Memory and Sampling in Contextual Judgment

James Tripp

Thesis submitted for the degree of Doctor of Philosophy

Department of Psychology

University of Warwick

October 2013

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank Gordon Brown for his supportive supervision of this thesis. His advice and comments were invaluable and tremendously useful. I am also indebted to other members of the Psychology Department at Warwick for their support and stimulating discussion. In particular, I thank Silvio Aldrovandi and Karen Melrose.

I am extremely thankful to the Psychology Department at Warwick for a Departmental Studentship which financially supported my study and allowed me the pleasure of teaching undergraduate students. Additional funding for testing participants was provided by a grant held by Gordon Brown.

Finally, I wish to thank my wife, Nancy, for her relentless support, and my Mother-in-Law for lovingly taking care of my amazing son, Joshua. Their compassion and kindness is always an inspiration.

Declaration

I hereby declare that the research reported in this thesis is my own work unless otherwise stated. Initial analyses of, and the data from, Experiment 1 and Experiment 2 were submitted in partial fulfillment of the requirements for a BSc (Hons) in Psychology at the University of Warwick. Subsequent model based analysis of the data from Experiment 1 was carried out for this thesis. Experiment 2 is included in this thesis for completeness and is part of a paper (under revision) which forms Chapter 3.

Abstract

This thesis investigates the interaction of memory and decision making in relative and retrospective judgment. Theories of memory and decision making are often rigorously tested using a variety of data sets, and the resulting theories can be applied to a large selection of psychological phenomena. In Chapter 1 I argue that theoretical development in relative and retrospective judgment is in contrast often very specialized. Theories of relative and retrospective judgment cannot easily be applied to other memory and decision making phenomena. Another approach is to take broad models or principles from the wider literature and apply them to relative and retrospective judgment. I suggest that the SIMPLE model of memory (Brown, Neath, & Chater, 2007) and the decision by sampling model (DbS; Stewart, Chater, & Brown, 2006) can in combination offer a comprehensive and unifying account of relative judgment. In Chapter 2 I find that both relative and retrospective judgments are consistent with range-frequency theory. I also find evidence for range effects which are inconsistent with decision by sampling. Chapter 3 investigates the role of similarity in these relative judgments using the distance based sampling model (Qian & Brown, 2005). The results show no evidence for distance based sampling. A combined SIMPLE and DbS model (SDbS) is applied to data from previous studies in Chapter 4. SDbS and range-frequency theory can account for the data – including range effects - equally well. In Chapter 5 I use an incentivized free recall task to elicit atypical serial position curves in three experiments. SIMPLE is shown to be able to fit the effect of output position which appears important in decision making behavior. Overall, this thesis suggests that SDbS is a candidate model for unifying retrospective and relative judgment with the wider memory and decision making literature.

Chapter 1

Memory and Decision Making

How are our decisions influenced by our memory? It is self-evident that past experiences influence our judgments. We choose to go to our work every day assuming that the experience will be similar to that of previous days. People walk confidently across a zebra crossing because they expect the traffic to stop and allow them to cross the road. Although these examples are trivial, they demonstrate that past events are accessed from a memory system when a judgment is made. Given these obvious examples, it is surprising that research into decision making has generally failed to capitalize on the extensive research on memory in psychology (Johnson & Weber, 2009; Weber & Johnson, 2006). There are some exceptions to this generalization and I consider them below.

In this chapter it is argued that formal models of memory have the potential to offer a unifying framework for theories of decision making. First, it is shown that the memory literature can contribute to the understanding of decision-making behavior. Second, the use of memory processes in accounts of retrospective and relative judgment is discussed. Theoretical accounts of both retrospective and relative judgment generally acknowledge the importance of the memory system. Third, a broad decision-making theory called decision by sampling (Stewart et al., 2006) is outlined. Finally, the SIMPLE (Brown et al., 2007) and MINERVA2 (Hintzman, 1984) models of memory and their contribution to decision making are examined. This thesis investigates the possibility that a combined model of memory and decision making can predict both relative and retrospective judgments. The aim of this chapter is to motivate this investigation.

Memory can be conceived as implicit or explicit. Implicit memory influences past experiences without conscious awareness (Schacter, Chiu, & Ochsner, 1993). An example of implicit memory is the process of reading words which occurs automatically without requiring conscious access to every past reading event. In contrast, explicit memory is not automatic and is involved in the recollection of specific past events. Explicit memory can be further divided into autobiographical or episodic memory. Autobiographical memory is the synthesis of past events into a *cohesive narrative* focused on the self (Fivush, 2011) whereas episodic memory is the specific recollection of *discrete episodes* in the past (Tulving, 2002).

Distinguishing between different types of recall is useful when examining decision-making behavior. Consider the priming effect, where a previous stimulus can influence the identification of a degraded cue even when participants cannot recall seeing the stimulus (Tulving & Schacter, 1990; Tulving, Schacter, & Stark, 1982). The implicit memory system is thought to underpin the priming effect and the same memory system is used in explanations of social stereotyping (Banaji, Hardin, & Rothman, 1993), attitude formation (Greenwald & Banaji, 1995) and consumer choice (Coates, Butler, & Berry, 2006). Studies in neuroscience also support the distinction between explicit and implicit memory (Rugg et al., 1998).

Metacognition about memory processes influences decisions and judgments. Metacognition is knowledge of one's own cognitive processes (Flavell, 1979; Fleming & Dolan, 2012). Consider the speed with which items are recalled from memory and the influence this information has on judgment. The "availability heuristic" proposed by Tversky and Kahneman (1973) suggests that the accessibility of items in memory is in itself a source of information used by people when forming judgments. For example, N. Schwarz et al. (1991) asked participants to recall either 6 or 12 examples of their own assertiveness and then rate their assertiveness. Participants who gave 12 examples rated themselves as less assertive than those who gave 6 examples. In other words, if it was difficult to recall the required number of examples then the participant considered themselves to be unassertive.

These two examples illustrate the importance of the memory literature in theoretical accounts of decision-making behavior. In the first example, superficially distinct decision-making behaviors were investigated using a common concept, implicit memory. In the second example, knowledge about how accessible items are in memory influenced decision making. Taken together these examples illustrate a crucial point: explanations of decision-making behavior can benefit from incorporating extensively studied memory processes.

Understanding recall processes is important because people appear to make judgments based on small samples drawn from memory and the immediate environment. Findings from the decision-making literature suggest that people behave like “naïve intuitive statisticians” who fail to consider the limitations of making judgments based on small samples (Fiedler & Juslin, 2006; Juslin, Winman, & Hansson, 2007). For instance, people seem to poorly adjust for sample size in judgments and behave as though a small sample is more representative of the underlying population than it actually is (Fiedler & Juslin, 2006; Lindskog, Winman, & Juslin, 2013). These findings suggest that the retrieval of items from memory is an important component of decision making.

Decision Making

In this section, the role of memory processes in previous theories of retrospective and relative judgment will be reviewed. In the former, the memory system influences the accessibility of past events used in summary judgments and produces a ‘snapshot’ of the past. In the latter, the contribution and formulation of the memory system thought to underpin relative judgment varies considerably. I will argue that theories from both literatures are limited in scope, contain vague conceptions of memory and are generally isolated from recent research into episodic memory.

Retrospective Judgments

We must often form summary judgments of past events. For example, consider your last holiday. Upon your return colleagues and relatives may ask how enjoyable it was. Here the experience of a complex event which occurred over a period of time must be distilled into a single hedonic judgment. The retrospective judgment literature relates momentary and summary judgments whilst assuming an influence of selective accessibility in memory. In this section the key findings and theories from the retrospective judgment literature are reviewed.

There are two reasons to review hedonic judgments here. Firstly, hedonic experience modulates memory formation. Experiences which are either very pleasant or unpleasant are more likely to be recalled (Berridge & Kringelbach, 2011; Hamann, Ely, Grafton, & Kilts, 1999) suggesting that pleasure is a distinct dimension in memory. Secondly, studies examining retrospective judgment often use a hedonic scale (Kahneman, 2000). This second point is particularly important because studies of both relative and retrospective judgment have used hedonic scales and the phenomena may be linked (Brickman & Campbell, 1971).

The goal of research into retrospective judgment is to explain the relationship between momentary experiences and summary judgments. In one study patients recorded their pain every 60 seconds during a colonoscopy (Redelmeier & Kahneman, 1996). At the end of the procedure the patients were asked to judge retrospectively how much pain they had experienced. The surprising finding was that duration of the procedure was a poor predictor of the retrospective judgment. The researchers called this effect "duration neglect'". This effect is particularly surprising because the duration of the procedure varied considerably (4-67 minutes). Instead of the total duration, the pain reported at its peak and in the final few minutes was correlated with the remembered pain. Several other studies have found similar results with responses to pleasurable or aversive film clips (Fredrickson & Kahneman, 1993), annoying sounds (Schreiber & Kahneman, 2000), and pain from vice grips (Ariely, 1998).

A serious limitation of most explanations of duration neglect is that they are not explicit about the underlying memory process. Fredrickson and Kahneman (1993) argued that participants create a 'snapshot' of the previous event which is a selective representation of the experience. According to this explanation, the high correlation between peak/end and summary judgments are due to these events being over-represented in the 'snapshot'. This explanation describes the data and offers an explanation based on distortions in memory. However, the exact process that produces these distortions and how this explanation relates to broader research on memory is unclear. A later explanation argues that prototypical representations of experiences are created and form the basis of judgment (Kahneman, 2000). Again, memory processes are invoked but the nature of these memory processes and the relation of these phenomena to other psychological processes is unclear. In chapter 2 I examine the link between memory processes and summary judgments.

Other psychological factors such as familiarity and variance can negate duration neglect. Morewedge, Kassam, Hsee, and Caruso (2009) asked participants to give overall ratings of either a familiar or unfamiliar sound. In their second experiment participants judged the pleasantness of a long or short constant sound which was either familiar (a rotary telephone ring) or unfamiliar (a synthesized beep). As expected, the length of the sound did not influence judgments of the unfamiliar sound. However, the familiar sound was judged as less pleasant when it was longer. In other words, participants were influenced by the length of the sound if it was familiar. In another study, Ariely (1998) varied the duration and intensity of heat produced by a thermode attached to the arm of the participant. Duration influenced retrospective judgments and had a greater effect when the intensity of the heat varied. Both of these findings illustrate that familiarity and variation influence judgments.

Duration neglect in retrospective judgment can be disrupted by segmenting a continuous experience. Ariely and Zauberman (2000) manipulated the structure of a series of annoying sounds. The series of sounds were either played one after another without a gap between the sounds or played with a small period of silence between each sound. In other words, the sounds formed either a series of discrete events or a single flowing event. The participants rated each sound for its annoyance and then gave an overall rating. The peak and end ratings predicted overall ratings in the continuous sound condition. However, the average rating was a better predictor in the segmented sound condition. Interestingly, the initial rating influenced overall decisions when the volume of the sound increased and then decreased. The study demonstrates the influence of segmentation and the initial experience in overall ratings.

These effects are similar to findings from the memory literature. For example, improved recall performance is often seen for stimuli that are more isolated along a temporal dimension in comparison to other stimuli. Also, many studies from the memory literature over the past century have observed improved recall performance for items at the start (primacy) and end (recency) of a series when compared to items in the middle (e.g., Lewandowsky & Murdock, 1989; Murdock, 1962; Ward, Tan, & Grenfell-Essam, 2010). These findings are predicted by formal models of memory such as SIMPLE (Brown et al., 2007).

Including a formal model of memory into an account of retrospective judgment may bring several advantages. First, the retrospective judgment literature may be integrated into the memory literature. Currently retrospective judgments and episodic memory are described in largely disparate literatures. Second, the unexplained influences of experience isolation, familiarity, and apparent recency and primacy effects in retrospective judgment may be accounted for by using theories developed in the memory literature (such as SIMPLE). The current theoretical accounts of these effects in retrospective judgment are limited in application to other broader theoretical frameworks. Thirdly, other phenomena could be investigated under the same framework. Chapter 4 examines the application of a combined memory and judgment model which can be applied to a range of memory and decision-making phenomena.

Relative Judgment

Judgments of items are often made in relation to other similar items. For example, how expensive is a £1 pint of milk? Here a subjective judgment is being made about an objective price. Intuitively, our judgment will be influenced by the prices of other pints of milk drawn from the environment and memory. If the average price of milk is £3 then we might expect a person to judge a £1 pint of milk as not very expensive. On the other hand, if the average price is 50p then that same £1 pint of milk will be judged as more expensive. In other words, altering the contextual information available to the individual may influence the subjective judgment of a stimulus. Research in relative judgment investigates the relationship between the objective values of contextual stimuli (e.g., prices of other pints of milk) and subjective judgments of a stimulus (e.g., a £1 pint of milk).

This section on relative judgment is separated into three parts. First, research demonstrating the influence of context on judgment and the basic methodology used in relative judgment studies is introduced. Second, the highly influential adaptation level theory (ALT; Helson, 1964a) and range-frequency theory (RFT; Parducci, 1995) are discussed, and the role of memory processes in these accounts is reviewed. Third, the multi-domain decision by sampling model (DbS; Stewart et al., 2006) is introduced. It is argued that a broad account based on DbS may offer an account integrating the relative judgment literature with broader findings from the memory and decision-making literature.

Studying Context Effects. As described above, contextual stimuli influence subjective judgments in many domains. In a seminal study, Parducci (1965) asked participants to rate the size of a series of squares on a seven point Likert scale ranging from 1 (*very small*) to 7 (*very large*). All the participants saw the same sized squares. The presentation frequency of each square in the context was manipulated. If judgments were based on the objective size of a square then the frequency of other squares should have no influence on 'largeness' ratings. Parducci found that the frequency of other squares influenced the subjective judgments that participants gave to each square. For example, the same square was judged as larger if most of the stimuli were smaller in comparison to the judged square. In other words, Parducci's study demonstrated that context and judgment are interlinked. This link between contextual items and both judgment and decision making has been shown in economic (Stewart, Chater, Stott, & Reimers, 2003; Ungemach, Stewart, & Reimers, 2011), social (Wedell, Parducci, & Geiselman, 1987) and psychophysical judgments (for a review, see Wedell, Hicklin, & Smarandescu, 2007).

Here I focus on data from the psychophysics literature for several reasons. First, relative judgments have been studied in psychophysics for the last 70 years. This research has investigated several key phenomena and produced several theories. These theories often form the theoretical basis of research into other areas of psychology where relative judgment takes place, such as price perception (for a review see Mazumdar, Raj, & Sinha, 2005) and social judgment (e.g., Wedell et al., 2007).

Second, studies in psychophysics often aim to describe how objective magnitudes are transformed into subjective impressions (for a review see Murray, 1993). The method for doing this is highly systematic. The context of a stimulus is varied along a single physical dimension (e.g., weight or size) and a corresponding judgment along a psychological dimension (e.g., heaviness or largeness) is recorded. The key to theoretical development in this literature is finding the mathematical functions which relate objective properties to subjective impressions. This approach shows some similarity to the methodology of retrospective judgment in which the aim is to relate past events which occur along either a physical (e.g., sound intensity) or psychological (e.g., pain) continuum to the formation of a single subjective impression. This commonality between the two literatures provides a suitable starting point for the studies into subjective judgment and memory presented in this thesis.

Third, each of the theories of relative judgment include some form of memory. Reviewing the contribution of memory to theories of relative judgment is important because it motivates the inclusion of memory in a unifying account of relative and retrospective judgment, and it presents another common psychological process in both literatures. For these reasons, reviewing data from psychophysics is necessary for motivating the investigation carried out in this thesis.

Adaptation Level Theory. According to ALT, responses are made relative to a single internal reference point (Helson, 1964b). This reference point is called the 'adaptation level' (AL). For example, according to ALT the satisfaction derived from a £5 payment depends on the distance of £5 from an internal standard (such as the mean payment). In the original formulation of the theory the AL is derived from the combined intensities of past and present stimuli. In other words, ALT assumes that due to the assimilation of past and present stimuli there is a stimulus intensity to which an organism is adapted (the AL) and gives a neutral response. If participants are asked to rate an item along a 7-point Likert scale then a stimulus at the AL will be given a rating of 4. Responses to other stimuli are a function of the deviation of the stimulus from the AL.

The first formulation of ALT explicitly includes memory processes in the calculation of the AL. The ALT was first applied to responses to surfaces of varying color and luminance (Helson, 1938). Helson found that several phenomena from the vision literature could be predicted using a single unified theory. In this formulation the AL is a weighted mean of previous and present stimuli

$$AL = P^{P_w} F^{F_w} B^{B_w} \tag{1}$$

where the activation of stimuli in the past, P, foreground, F, and background, B, are weighted by three independent parameters, P_w F_w B_w. In this formulation the AL can be composed entirely of past stimuli drawn from memory. A key limitation to this approach is that the exact memory processes are not made explicit.

Later formulations of ALT calculated the AL as the mean of the stimuli. ALT was later extended to responses along psychophysical scales as a frame of reference theory (Michels & Helson, 1949). Responses were predicted based on the deviation of the stimulus intensity from the mean,

$$ALT_i = \bar{x}_s - S_i \tag{2}$$

where the response to a stimulus, ALT_i, is the deviation of stimulus intensity, S_i, from the mean intensity of the stimuli set, $\bar{x}_s$. Once again the memory processes underpinning these judgments are not explicit. Numerous variations of the above formulation have been successfully applied in several areas of psychophysics (Appley, 1971). In each application, a single pooled contribution of memory is considered or implicitly inferred.

In the price perception and income literature the single reference point theories are highly influential (Mazumdar et al., 2005). Within this literature the influence of a price on a customer is a function of its deviation from a reference price. Some models consider previous prices and memory accessibility, but these accounts are largely descriptive (Briesch, Krishnamurthi, Mazumdar, & Raj, 1997). Other models such as RFT have been applied to reference price effects (Niedrich, Sharma, & Wedell, 2001).

Range-frequency Theory. Volkmann (1951) noted that stimulus ratings could be predicted based on the range position of the stimulus along a dimension of interest. According to the range principle,

$$R_i = (S_i - S_{min})/(S_{max} - S_{min}) \tag{3}$$

where the response to a stimulus, R_i, is the distance of the stimulus, S_i, from the smallest stimulus, S_{min}, divided by the range of stimulus intensities, $S_{max} - S_{min}$. Consider an example in which a participant is told to expect a payment between £1-

£6 or £4-£9. Then the participant is given £5 and is asked to judge their satisfaction with the payment. The range-based prediction for the £5 payment is .8 (high satisfaction) for the £1- £6 range and .2 (low satisfaction) for the £4-£9 range. A higher satisfaction response is predicted in the £1-£6 range because £5 is higher up in the expected range. According to the range principle response categories (e.g. 1-7 on a Likert scale) are evenly distributed across the range of expected stimuli types.

The frequency principle predicts responses based on the relative rank of the stimulus. According to Parducci (1965) the frequency principle reflects a participant's tendency to use each response category with equal frequency. For example, if there are seven stimuli and seven response ratings then the first stimulus will be given a rating of 1, the second stimulus a rating of 2, and so on. So the predicted response is the relative rank of the stimulus,

$$F_i = (r_i - 1)/(N - 1) \tag{4}$$

where F_i is the rank of the stimulus, r_i, minus 1 divided by the number of stimuli, N, minus 1.

In RFT responses are predicted as a weighted average of the range and rank position of a stimulus within the comparison set (Parducci, 1965, 1995),

$$RFT_i = wR_i + (1 - w)F_i \tag{5}$$

where the prediction of the range-frequency model, RFT_i , is the weighted average of the range, R_i, and rank, F_i, predictions as described in Equations 3 and 4. The *w* parameter is a free parameter allowed to vary between 0 and 1 which controls the weight given to the range or rank principles. When w approaches 1 the responses become closer to those based on only the range position of the stimulus.

Parducci (1965) noticed that the predictions of the range and frequency principles would be different if some stimuli occurred more frequently than others.

Consider the above £5 example again with an expected distribution of payments of either (a) £3 £4 £5 £6 £6 £6 or (b) £3 £3 £3 £4 £5 £6. Distribution (a) is negatively skewed and distribution (b) is positively skewed.

How do the predictions for these distributions differ? Let us consider a £5 payment and a response between 0 (*unsatisfied*) to 1 (*satisfied*). In both payment distributions the response predicted by the range principle is .67. This is because the distance of the £5 payment to the smallest and largest payment is the same in both distributions. The frequency principle predicts different responses for each distributions. For the positively skewed distribution (£3 £3 £3 £4 £5 £6) the frequency principle predicts a satisfaction rating of .8. For the negatively skewed distribution (£3 £4 £5 £6 £6 £6) most of the payments are higher than £5: the frequency principle predicts a satisfaction rating of .4. The difference in predictions is due to the difference in frequency of each payment.

In a series of studies Parducci demonstrated that RFT can predict responses to skewed distributions (e.g., Parducci, 1956, 1965, 1968; Parducci & Haugen, 1967). Parducci recorded responses from skewed distributions of stimuli and, as predicted, responses were influenced by frequency manipulations. Responses were predicted using the weighted average of the range and frequency principles as outlined in Equation 5.

An advantage of RFT is its ability to predict findings that cannot be predicted by either the range principle only or ALT. Parducci (1965) argued that if the AL is the mean of the stimuli (see Equation 2) then altering other properties of the stimulus distribution, such as the frequency of items, should not influence judgments. He asked participants to rate the largeness of squares on a seven-point scale from 1 (*very small*) to 6 (*very large*). All participants rated the same 9 squares. The sizes of the squares were on a non-linear scale. He manipulated the frequency of square sizes so that some squares were more frequently presented than others. These manipulations produced 13 stimulus distributions. Here I will focus the bimodal and unimodal shaped distributions (see Figure 1 and Figure 2).

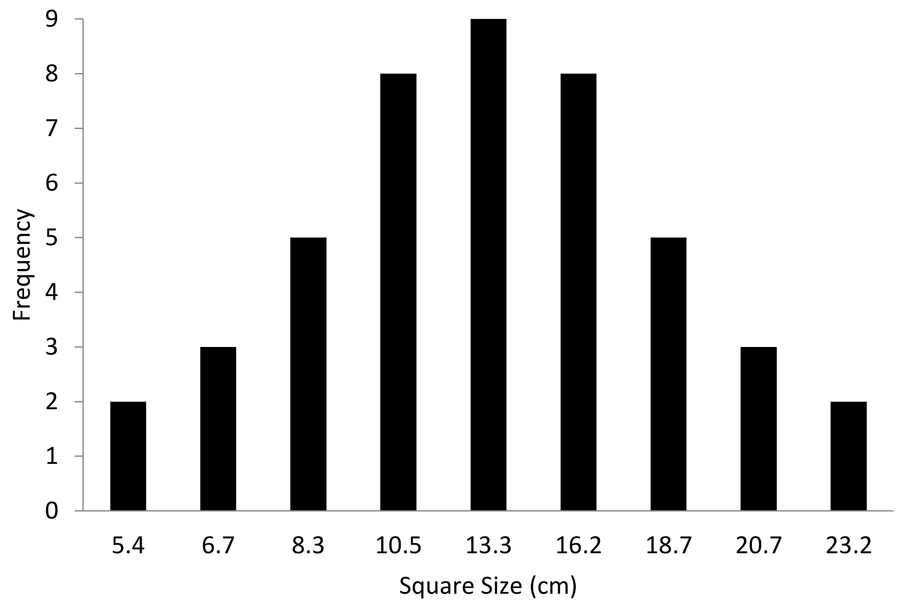

Figure 1. Unimodal distribution of square sizes presented in Parducci (1965)

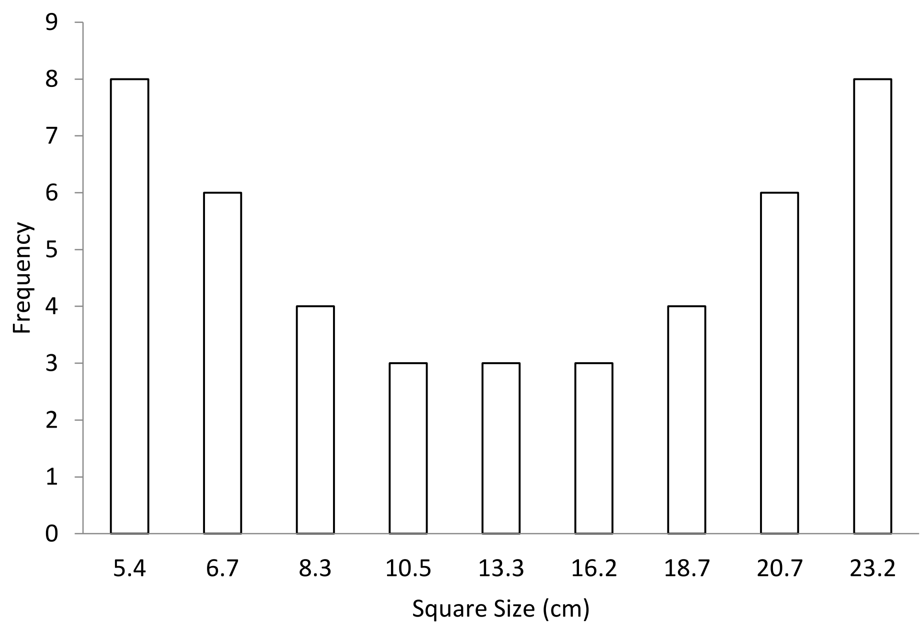

Figure 2. Bimodal distribution of square sizes presented in Parducci (1965)

The predictions of the range principle and ALT the distributions in Figure 1 and Figure 2 are the same. The position of each square in the range of the experienced squares is the same in both distributions. According to the range model (see Equation 3) 13.3 cm square should be given the same rating in both distribution conditions. The mean square size for both distributions is 14 cm. If the AL is equal to the mean of the stimuli (see Equation 2) then ALT predictions that the square will be given the same response rating in both distributions. The difference between the two distributions is the frequency of each stimulus type. Due to this difference in frequency the relative rank of the stimuli will differ. For example, the rank position of squares of size 8.3 cm will be higher in the bimodal distribution because more of the squares are a smaller size.

Parducci (1965) found that the responses to the distributions differed. Figure 3 shows the mean ratings of each stimulus type from Parducci (1965) for the distributions in Figure 1 and Figure 2. The results are inconsistent with both the ALT and the range principle. In other words, the rank of a stimulus within the stimuli set contributed to the subjective rating of the stimulus.

Parducci (1965) found that participant responses were best predicted by a combination of the range and frequency principles. The predictions of RFT with an equal weighting on the range and frequency predictions (w = .5) are plotted on Figure 3. In both cases, the RFT predictions closely match the qualitative pattern of the data. Figure 4 shows the range principle (w = 1), frequency principle (w = 0) and RFT (w = .5) predictions for the U-shaped distribution used by Parducci (1965). Neither the range principle or frequency principle predictions fit the average responses well. The equally-weighted RFT predictions perform much better. In other words, the range principle or frequency principle alone are poor predictors of participant responses in comparison to the weighed mean in RFT.

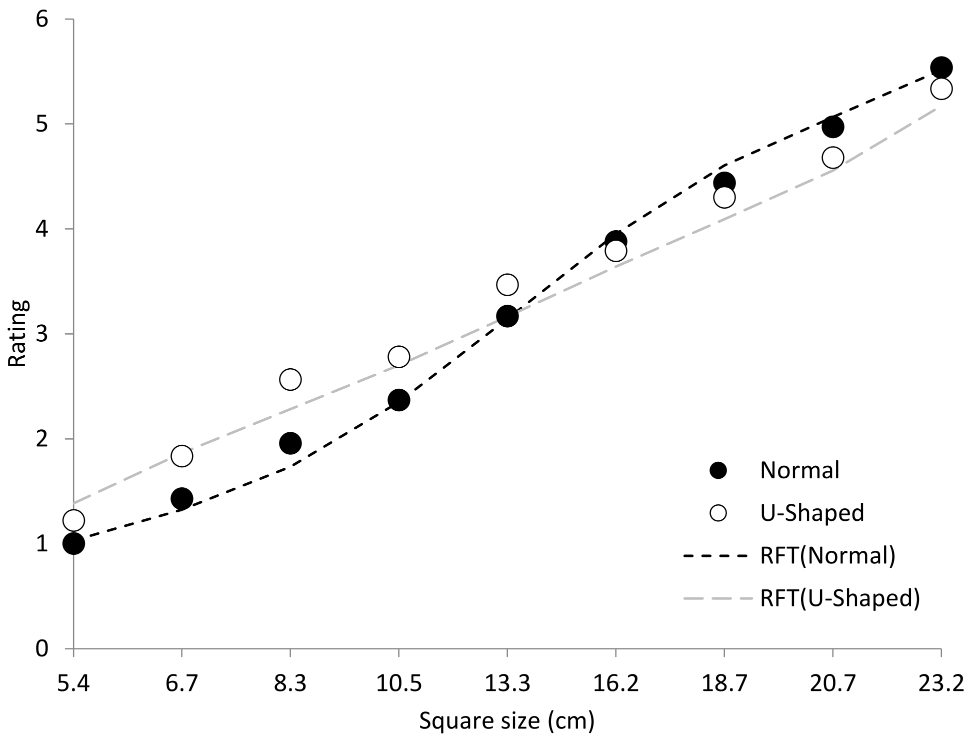

Figure 3. Mean judgments of each stimulus type and predictions of RFT from Parducci (1965)

Similar manipulations with a diverse array of stimuli have supported RFT. The model has been applied to psychophysical (Parducci, Perrett, & Marsh, 1969) (Parducci, Perrett, & Marsh, 1969), hedonic (Parducci, 1968) and social (Pettibone & Wedell, 2007; Wedell et al., 1987) judgments. For example, Wedell, Parducci, and Roman (1989) asked students to grade hypothetical results from a 100-point exam using an A-F scale. The frequency of low and high marks was varied. RFT predicted the grades that the students allocated to the grade marks. All of these studies support RFT over ALT and either a range principle only or frequency principle (rank) only.

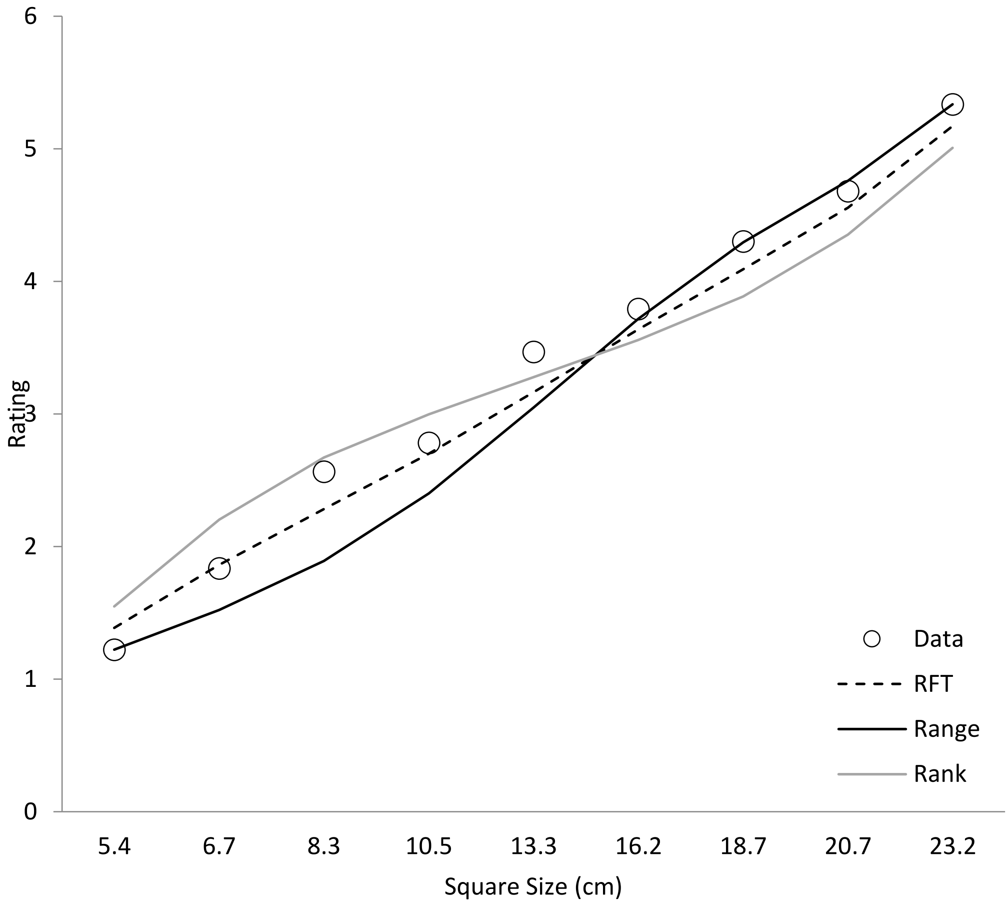

Figure 4. Range principle, frequency principle (relative rank) and RFT predictions of responses to the U-shaped distribution of squares from Parducci (1965)

Of particular interest to this thesis is the possibility of applying RFT to summary judgments. RFT predicts a higher average rating for stimuli that are negatively skewed. In a seminal study, Parducci (1968) showed that RFT may be applied to summary judgments. He asked participants to take part in a pseudo gambling experiment. Participants were shown the distribution of monetary payments they could expect to receive in the experiment. They then chose one of three cards, received a payment equal to the amount shown on the underside of their chosen card, and then rated their satisfaction with the payment. The card task was repeated for every payment in the distribution shown to the participant at the start of the experiment. Participants were either given a positively or negatively skewed distribution of payments with the same total and mean payment. To ensure that the mean and total payments were the same the absolute range of the payments were different. The average reported satisfaction was higher for participants when the payment distribution was negatively skewed (items in the upper portion of the payment range were most frequent).

Smith, Diener, and Wedell (1989) extended RFT to summary judgments. Following the findings from Parducci (1968), Smith et al. (1989) asked participants to rate the happiness they would expect a waitress to experience after receiving a series of either positively or negatively skewed hypothetical tips. After rating all the tips the participants were asked to rate how happy the waitress would be with the distribution of tips shown to the participant. These summary ratings were higher when the distribution of hypothetical tips was negatively skewed, which is consistent with the average higher ratings for negatively skewed distribution reported by Parducci (1968). These findings led Parducci (1995) to suggest that RFT is a theory of both online and summary happiness. These overall skew effects are examined in Chapter 2 as they represent a link between retrospective and relative judgment.

However, the initial formulation of RFT was unable to account for stimulus or category manipulations. Parducci and Wedell (1986) replicated the square size estimation experiment whilst varying the number of categories and unique stimuli. Increasing the number of rating categories that participants could use decreased the effect of frequency (i.e., skewness) on judgments – the category effect. Decreasing the number of unique stimuli also decreased the effect of frequency on judgments – the stimuli effect.

To account for these effects a more explicit account of memory was added to RFT. The initial formulation of RFT suggested that memory was involved in judgments (Parducci, 1965; Parducci et al., 1969; Parducci & Perrett, 1971) but specific memory processes were not included in the model. Wedell and Parducci (1985) extended RFT to accommodate stimuli and category effects by including rudimentary memory processes. Central to the extended model is the assumption that the distribution of stimuli is recreated from memory for each judgment. The presentation frequency of a stimulus is retrieved from memory by accessing each presentation event serially and is performed for all the stimulus types in parallel. If little time is allowed for search in memory then the recalled frequency of each stimulus type was assumed to be similar and the recalled distribution more uniform. Wedell and Parducci (1985) argues that increasing the number of response categories decreases the time available to search memory.

When Parducci and Wedell (1986) increased the number of stimulus types they reduced the frequency of each stimulus type. If searching for the presentation of each stimulus in memory occurs in parallel for each stimulus type then increasing the number of stimulus types reduces the time required to build the distribution from memory (i.e., the recalled distribution is closer to the actual distribution when each stimulus type is less frequent). However, it is unclear how this formulation of memory may be applied to wider memory phenomena (e.g., primacy, recency and isolation effects). Similar accounts overcome this limitation by incorporating memory retrieval and a need for consistency in responses (i.e., responses to the same stimulus should be the same) also account for these effects (Haubensak, 1992; Parducci & Wedell, 1986). But these other similar are still unable to predict some phenomena in the memory literature.

In summary, incorporating memory processes into models of relative judgment improves the ability of these models to predict responses. Price perception models using a single reference price can predict a wider range of behavior if the influence of previous prices is considered. RFT without explicit memory processes is unable to fully account for judgment behavior. In these cases the memory processes are highly specific to relative judgment. At present, there is no model which can account for the full range of context effects, remain compatible with the present findings in memory, and be applied to the retrospective judgment literature.

In other words, memory processes appear to be important in relative judgment but the current models are generally isolated from the broader memory literature. The focus of this thesis is the development of a model that encapsulates a contemporary understanding of episodic memory and is able to unite apparently disparate decision-making phenomena. In the next section a model of decision making which has been applied directly to relative judgment is introduced.

Decision by Sampling. The decision by sampling (DbS; Stewart et al., 2006) model of judgment and choice predicts responses based on simple cognitive processes. In DbS a judgment is formed based on the number of stimuli higher and lower than a stimulus. Using the number of stimuli above and below a stimulus allows the participant to calculate the relative rank of the stimulus

$$DbS_i = \frac{N_{lower}}{N_{lower} + N_{higher}} \quad (6)$$

This relative rank model is identical to Parducci's frequency principle (see Equation 4) if we assume that all of the experimental stimuli are in the memory sample that participants recall.

There are three main differences between DbS and RFT. First, the range position of a stimulus does not directly influence the predictions of DbS. The formulation of DbS in Equation 6 is purely the relative rank of a stimulus within the stimuli. The implications of this are discussed below. Second, DbS assumes that rudimentary cognitive processes such as long term and working memory are central to decision making. The popular formulation of RFT (see Equation 5) does not explicitly incorporate cognitive processes. Third, decisions are made dynamically in DbS. The RFT model is largely descriptive and assumes that stimuli are categorized before the first judgment is made. In contrast, DbS suggests that decisions are made on a decision by decision basis.

DbS assumes that decisions are formed on the basis of a limited sample which is recalled from memory. Stewart et al. (2006) make two assumptions about this sampling process. Their first assumption is that a sample of around six values is kept in working memory - the limit of around six or seven items in working memory is a robust finding in psychology (Baddeley, 1994; Miller, 1956). Their second assumption is that the sampling process is stochastic (i.e., random). According to DbS, the participants draw on past distributions of values and the current environment to form a sample within working memory: Judgments of a stimulus are predicted by the relative rank of the stimulus within that sample.

Figure 5 illustrates how the range, rank, RFT and DbS models differ. Consider an example where a person judges the attractiveness of a £7 price. In this example, each model predicts a different response to the £7 price. The range, rank and range-frequency models predict responses based on other prices in the present environment (£5, £8, and £9). This immediate environment is the other experimental stimuli presented either sequentially (such as weights) or simultaneously (as in social judgments such as Brown et al, 2009). The range-only model predicts response based on the highest and lowest value stimuli (in this case (7-5)/(9-5) = .5). Rank predicts responses based on the number of items lower and higher than the judged item ((2-

1)/(4-1) = .33). Range-frequency theory predicts that responses will be between the range and rank model predictions.

Decision by Sampling predicts responses based on a sample from memory which can include items from the present and past environments (see Figure 5). In DbS responses are based on rank based comparison. In the figure the DbS predictions may be based on the current environment – the same as the rank only model (.33) – or include both the present and past environments (in this case, with both past and present items the predictions is (5-1)/(8-1) = .57).

The central difference between the models is that the assumption in DbS that sample are drawn from memory. We assume that a rank based process underpins judgment. In the original formulation of DbS this was assumed to be a stochastically drawn set of stimuli combining both the present and past stimuli. The sampling from the past and present stimuli is depicted in the 'sampling' box in figure 5.

In this thesis I investigate the sampling processed which may underpin judgment. The process underpinning sampling from memory may be stochastic and the stimuli are evaluated using a rank based processes (see lower right portion of figure 5). On the other hand, other processes may mediate this sampling procedure. Other factors include sampling based on the distance of other stimuli from the items being judged (see chapter 3) and the local distinctiveness of item amongst its neighbors (see chapter 4). The GEMS and SIMPLE models offer cohesive accounts of these processes which may mediate the sampling of items before rank comparison takes place. The main focus of this thesis is the extent to which other cognitive processes such as those found in memory research can explain the apparent range effects reported in the relative judgment literature.

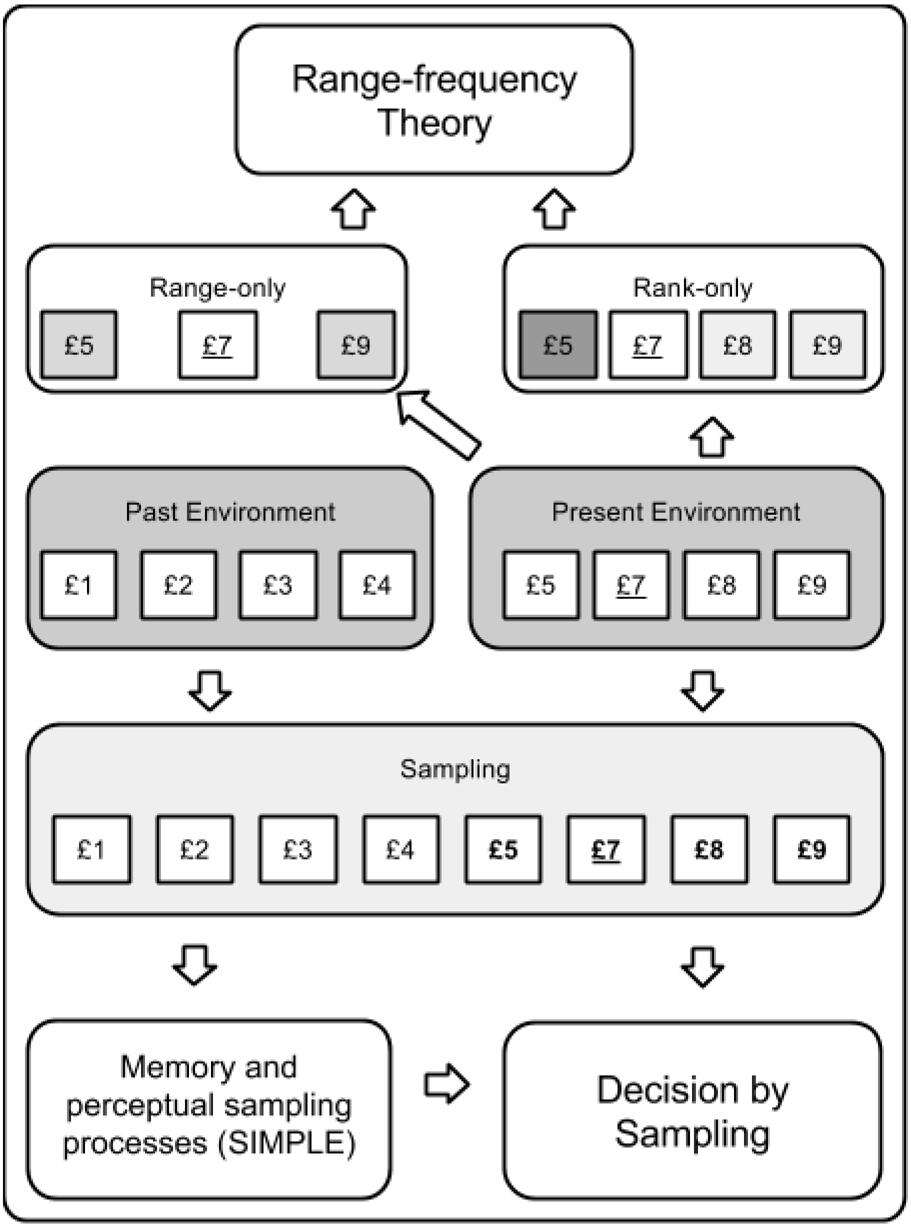

Figure 5. Illustration outlining how RFT, Range-only, Rank-only and DbS use different features of the enviroment to predict judgments.

Consider a simple example that illustrates the different predictions of DbS and RFT. Imagine going to a supermarket to buy milk. You have shopped there before and the price is normally quite low. However, this time most of the milk is expensive. The distributions of prices in the environment and memory for this example are given in Figure 6. In the environment (the supermarket) higher milk prices are more frequent. In long-term memory most of the milk prices are low.

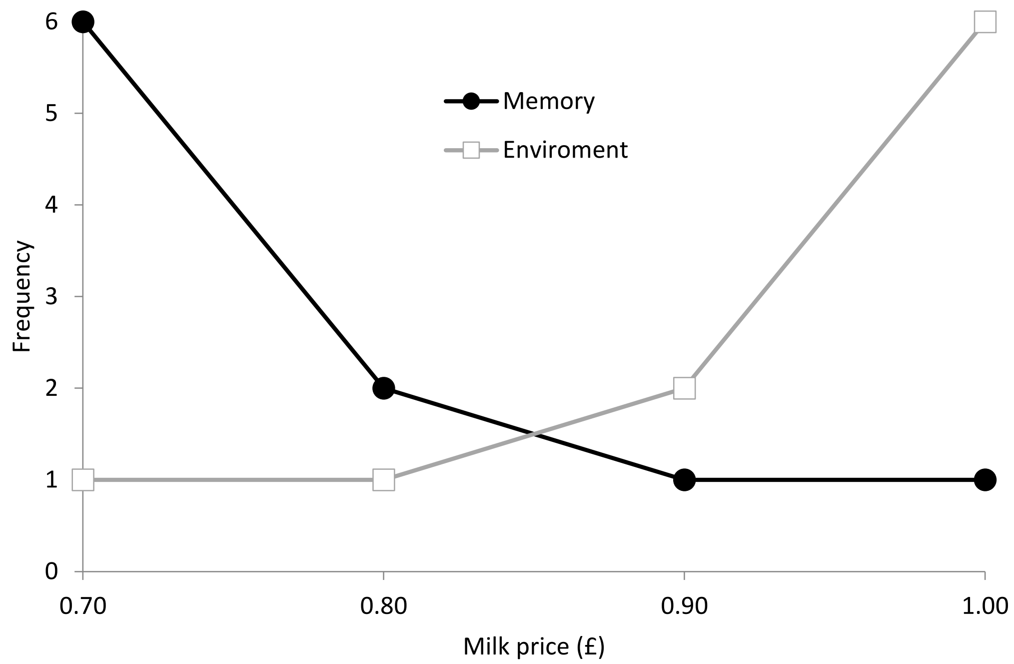

Figure 6. Example price distributions in memory and the environment

Let us assume that a buyer will rate a higher price as more expensive. How will the predictions of DbS differ from the frequency principle (excluding memory) and RFT? DbS predicts ratings of expensiveness based on the relative rank of a price within a 6-item sample drawn from both long-term memory and the environment. As the sample combines both memory and the environment we would expect as many 70p prices as £1 prices. In this case, MATLAB has produced a random sample of 70p, 70p, 80p, 90p, £1 and £1. Figure 7 shows the difference in predictions for each price from DbS, the frequency principle (w = 0) and RFT (w = .5).

DbS matches our intuitions about judgment. In DbS previously experienced prices are included in judgments. It is unlikely that a consumer will suddenly forget the prices from last week and rely solely on what is available in the immediate environment. Consider the expensiveness of milk costing 80p and 90p as shown in Figure 7. The immediate environment is dominated by milk costing £1 whereas past experience is dominated by milk costing 70p (see Figure 6). If responses are based only on the environment then milk costing 80p or 90p should appear relative inexpensive as predicted by RFT and the frequency principle. One the other hand, if past experience influences decisions then milk costing 80p and 90p should be much more expensive because past experience is dominated by lower values, as predicted by DbS. This example demonstrates a crucial difference between DbS and previous models of relative judgment: A psychologically plausible memory mechanism is central to DbS.

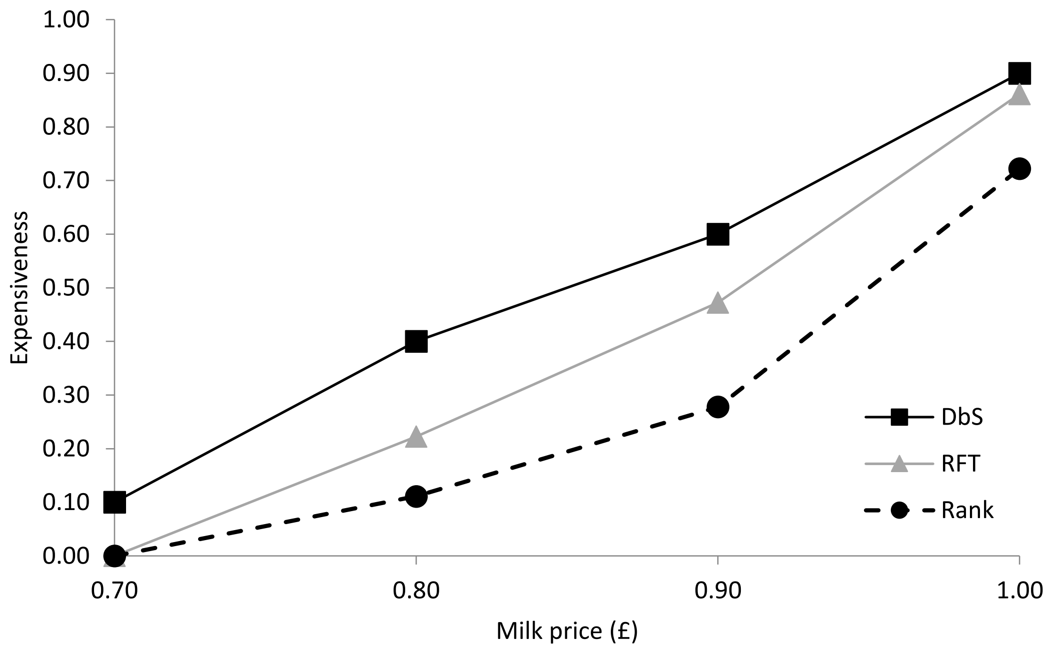

Figure 7. Predicted milk price expensiveness ratings from the RFT, DbS, and rank-only models

The DbS approach can explain two features of decision making described by Kahneman and Tversky (1979). The first feature is that people treat losses and gains unequally. Participants appear to lose more utility for a loss when compared to a gain of the same financial value. Kahneman and Tversky (1979) investigated the difference between economic losses and gains. Participants in their study were split into Sellers, Choosers and Buyers. Every participant had to choose between going home with an attractive mug or a pre-specified amount of money. Sellers were given a mug and asked how much they would exchange it for. Choosers chose between gaining the mug or money. Buyers indicated the price for which they would purchase the mug. The median price for the Sellers was $7.12. For the Choosers it was $3.12 and the Buyers were willing to pay $2.88. The difference between the three groups was the framing of the decision. For sellers the cash was in exchange for the loss of the mug. For both the Choosers and Buyers the cash was in exchange for the gain of the mug. The experimenters concluded that losses had a larger subjective impact than corresponding gains – loss aversion. Loss aversion has been observed in several other scenarios (Neale, Huber, & Northcraft, 1987; Thaler, 1980). The non-linear relationship between a loss and a gain was described by Kahneman and Tversky's (1979) value function. The second feature is that high and low probabilities are differently weighted in decision making. Kahneman and Tversky (1979) asked participants to choose between lottery options. They found that people act as though very low probability events are more probable than they actually are, and treat high probability events like as less probable than they actually are.

DbS offers a process-based account of this behavior. Stewart et al. (2006) examined the frequency of cash credits and cash debits in a sample of the UK population and the frequency of phrases describing probability from the British National Corpus. The credits and debits both followed power law distributions such that smaller payments and debits were most common. If these distributions are taken to accurately represent the distribution in memory that people draw upon when making decisions then a sharper drop in subjective value is expected for losses compared to gains. In other words, DbS predicts the value function from prospect theory based on samples drawn from memory. The probability phrases showed that phrases like “Never” and “Always” were generally more frequent than phrases which representing equal probability. Placing these items on a probability scale between 0 and 1 allowed the relative rank of each probability to be calculated. The function showed a remarkable similarity to the probability weighting function in Kahneman and Tversky (1979). Taken together, these findings suggest that DbS can predict similar effects to those described by prospect theory. Crucially, in DbS these behaviors are caused by the use of rudimentary cognitive tools (i.e., ordinal comparison) and distributions of values drawn from memory.

DbS can also predict social judgments. For example, Wood, Brown, and Maltby (2012) examined the link between the subjective risk associated with drinking alcohol and the relative rank of alcohol consumption. In their first study, participants reported (a) their alcohol consumption, (b) their perceived chance of developing illness and (c) their subjective distribution of alcohol intake throughout the UK. The relative rank of the participant within their recalled distribution of alcohol consumption was a significant predictor of their perceived chance of developing illness. In a second study Wood et al (2012) experimentally manipulated the distribution of alcohol consumption which participant used to make their judgments. Participants were shown different distributions of weekly alcohol consumption. The distributions were shown one at a time and participants were asked to give the probability of long-term illness for each value in the distribution.. In both studies relative rank was a better predictor than the distance from the distribution average. Other studies in other domains of social judgment have shown similar findings (Brown, Gardner, Oswald, & Qian, 2008; Melrose, Brown, & Wood, 2012). This work suggests that distributions accessed from long-term memory do influence social judgments.

However, DbS appears unable to predict range effects in the relative judgment literature. As reviewed above, RFT combines two principles. The frequency principle is the relative rank of the stimulus and is identical to DbS. The range principle is the range position of the stimulus. According to the range principle the rating of a stimulus will increase as the distance of the stimulus from the lowest and highest value increases. Parducci (1965) demonstrated that a weighted average of both principles predicts responses better than either in isolation. If we assume that all of the stimuli in an experiment are present in a sample from memory then DbS predictions are based on just the relative rank of an item and so DbS appears unable to produce range-based effects. Consider the data and predictions in Figure 4. An equal compromise between the range and frequency (rank) principles fits responses from a U-shaped distribution better than rank alone.

Is it possible for DbS to predict these effects predicted by the range principle? The participants in Parducci (1965) were shown 45 squares one after another. In the formulation of DbS above (which is the same as the frequency principle) we may assume that all of the stimuli are equally accessible in memory. For example, the 23rd square will be given a neutral rating (.5) because 22 smaller squares and 22 larger squares are considered in the judgment. This seems unlikely. A large literature has examined systematic distortions in memory for serially presented items (e.g., Hogan, 1975; Laming, 2010; Murdock, 1962; Rundus & Atkinson, 1970). Stewart et al. (2006) assumed that the sampling processes from memory was stochastic (i.e., random) and they acknowledged that this was a simplifying assumption which was most likely wrong. One possibility which will be examined in this Chapter 4 is that DbS may be able to account for range effects if the accessibility of items in memory is calculated using a formal model of memory (such as SIMPLE).

This section has reviewed the relative judgment literature and illustrated several limitations of past findings and theories. In retrospective judgments the isolation and serial position of an experience alters summary judgments. These effects have been studied in the memory literature. In relative judgment, RFT cannot predict a wide range of findings from the memory and decision-making literature (e.g., primacy, recency and the economic behavior described by prospect theory). Including some memory processes accounts of relative judgment generally improves the predictive ability of the models. However, models from both the relative and retrospective judgment literature are generally isolated from other literatures. An advantage of DbS over these accounts is that it has already been widely applied to judgments and decisions from the social and economic literatures, and explicitly incorporates episodic memory. The next section considers two models of memory which could be combined with DbS to unify these literatures.

Memory

SIMPLE

In this section the SIMPLE model of memory is discussed. To foreshadow, the SIMPLE memory may be able to account for both peak and end effects in the peak-end literature, and apparent range effects in the relative judgment literature. SIMPLE (Brown et al., 2007) has been applied to free and serial recall data. In both paradigms the participants are shown a series of items one after another. After being shown all the items they are asked to recall as many as they can. In serial recall the items must be recalled in the order that they were presented in. In free recall the items can be recalled in any order. Key results from both paradigms are predicted by the SIMPLE model (e.g., Brown et al., 2007; Lewandowsky, Duncan, & Brown, 2004; Lewandowsky, Nimmo, & Brown, 2008).

Two assumptions are central to the SIMPLE model. First, memories of events are located within a multidimensional memory space. In principle, multiple dimensions can be included in the model. Second, the probability of recalling an item from memory decreases as its confusability with other items increases.

Let us consider a simple example to illustrate the SIMPLE model. Imagine you are buying a pack of biscuits and trying to recall previous biscuit prices. In this example price is the dimension of interest. The prices when you last visited the store were:

50p 60p 70p 80p £1 £1.50 £3

and you want to recall the previous prices before you decide the expensiveness of the other biscuits in the store.

First, the similarity of each price to the others is calculated,

$$\eta_{ij} = e^{-c|M_i - M_j|} \tag{7}$$

where the similarity of stimuli i and j is an negative exponential function of the absolute differences in magnitude along the dimension of interest, $|M_i - M_j|$. The exponential function transforms external magnitudes into psychological distances (Nosofsky, 1986; Shepard, 1957). Applied to the biscuit price example Equation 7 gives us a matrix of similarity between items. This matrix of similarities is shown in Table 1. The £3 biscuit has a much lower overall similarity to all of the other prices. Intuitively, we would expect the £3 price to be more easily discriminated and, according to SIMPLE, more likely to be recalled.

Table 1

Similarity of the prices of packs of biscuits to one another

	Price						
Price	50	60	70	80	100	150	300
50	1	.61	.37	.22	.08	.01	0
60	.61	1	.61	.37	.14	.01	0
70	.37	.61	1	.61	.22	.02	0
80	.22	.37	.61	1	.37	.03	0
100	.08	.14	.22	.37	1	.08	0
150	.01	.01	.02	.03	.08	1	0
300	0	0	0	0	0	0	1

Note: Shading indicates similarity. The *c* parameter was set to 0.05 for this illustration.

The *c* parameter reflects the relationship between psychological distance and similarity. The relationship between the *c* parameter and similarity in psychological space is shown in Figure 8. In the figure the transformed distance between the 50p price and the other prices is shown for multiple *c* parameter values. Increasing the *c* parameter reduces the similarity of one price to another and changes the impact of distance on similarity. From a psychological viewpoint the *c* parameter modulates the steepness of the relationship between distance and confusability.

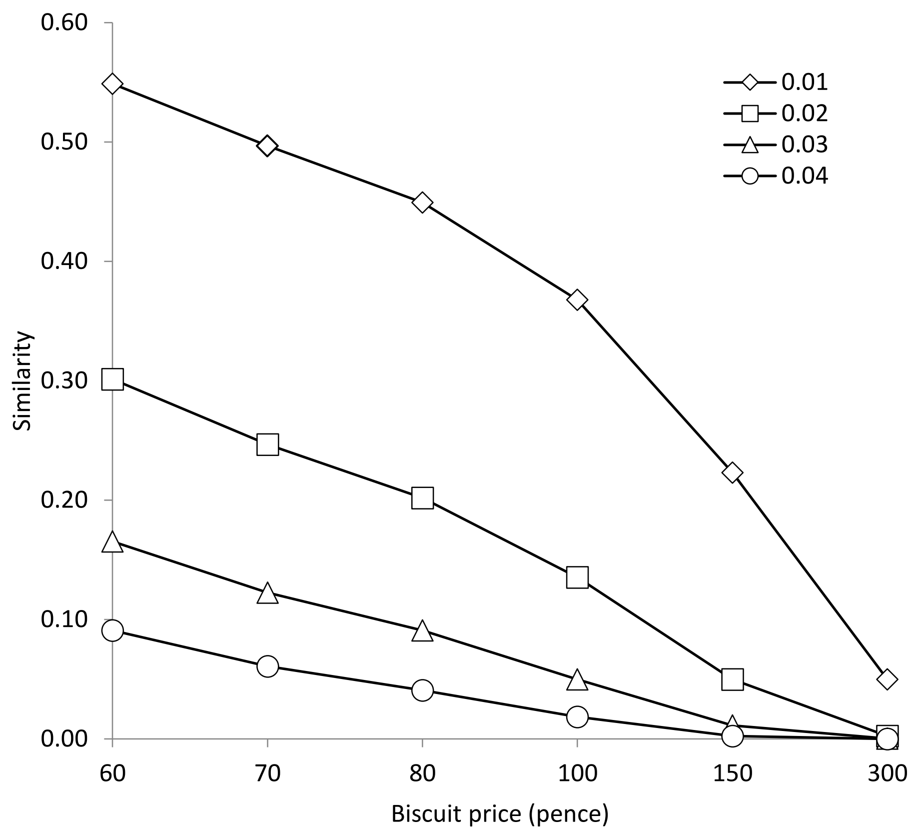

Figure 8. The effect of the *c* parameter on the similarity of 50p to other prices

In the simplest formulation of SIMPLE the probability of recalling a price is given by

$$D_i = \frac{1}{\sum_{k=1}^{n}(\eta_{i,j})} \tag{8}$$

where the discriminability of the price, D_i, is 1 divided by the sum of the similarity of a price to all the other prices. The discriminability of each price is shown in Figure 9.

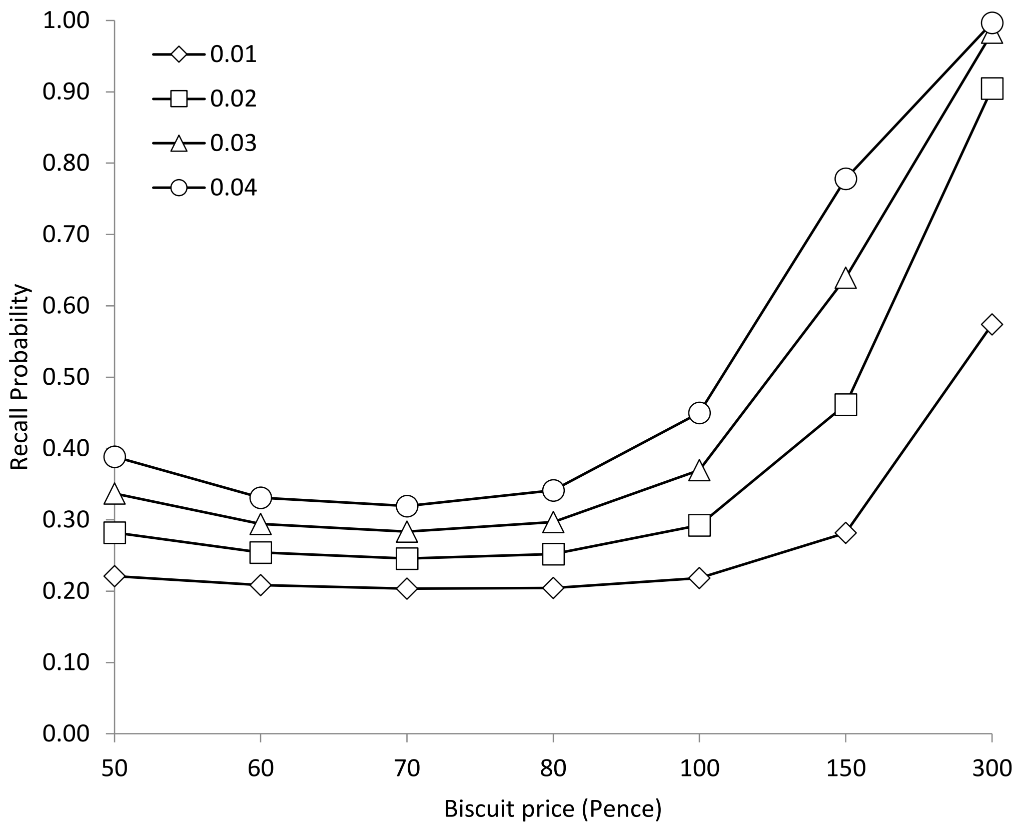

Figure 9. The impact of the *c* parameter on the discriminability of biscuit prices

The recall probabilities in Figure 9 match our earlier intuition. The £3 price is most likely to be recalled and the 70p price is least likely. Applied to DbS, the higher recall probability of £1.50 and £3 would make them more likely to be in the sample drawn from memory.

The SIMPLE model is able to predict key findings from the memory literature. In studies of free and serial recall items at the start and end of a series are much more likely to be recalled than items in the middle. These are the primacy – the first few items – and recency – the last few items - effects, which have been observed in serial learning, free recall, recognition memory and other paradigms (e.g., Crowder, 1976; Healy, Havas, & Parker, 2000; Lansdale, 1998; Murdock, 1974). In SIMPLE recall is like a discrimination task. Along the temporal dimension the first and last items are more easily discriminated from other items because they have no temporal neighbors. The successful application of SIMPLE to serial position effects in free recall typically involves two additional parameters, *t* and *s*, which control an omission threshold that is necessary to fit the free recall data. To ensure that the above example was as simple as possible these parameters were not considered, but the parameters are used and explained in applications of SIMPLE presented in the later chapters.

In this thesis I examine the application of SIMPLE to judgments phenomena for at least two reasons. First, SIMPLE predicts that locally distinctive events are the most likely be recalled. Along the temporal dimension these are the first and last items in a series, and along other dimensions (e.g., pain and pleasure) these may be peak and troughs of a series. These predictions appear to match the peak and end effects observed in the summary judgment literature briefly reviewed above. Second, SIMPLE is a multidimensional model. Relative judgment stimuli are altered along a single dimension. But experiences in contributing to retrospective judgments unfold over time and vary in intensity. In other words, they vary along two dimensions. In SIMPLE the discriminability of an item in memory space can be multidimensional. These reasons are why SIMPLE rather than other models such as MINERVA 2 (see below) is applied to judgments in this thesis.

MINERVA-DM

MINVERA-DM (MDM; Dougherty, Gettys, & Ogden, 1999) is a model of memory applied to decision-making behavior. The decision-making model extends a model of memory called MINERVA2 (Hintzman, 1984, 1988) to judgments of likelihood and frequency. In MDM recall behavior depends on the similarity of a probe to traces in ‘secondary memory’. This subsection has three aims. First, the

MINERVA2 model is introduced. Second, the MDM model and its applications are discussed. Third, the potential applicability of MDM versus SIMPLE to relative and retrospective judgments is considered.

Consider the following example, which illustrates MINERVA2. In this example a participant is shown the same item twice and is later asked to identify the item they saw. Let the item be a pint of milk. In the model this item is represented by a vector of feature values such as

[1 1 -1 0]

where 0 represents an irrelevant feature, 1 is excitory and -1 is inhibitory. The meaning of each vector value is arbitrary and in our case the first two values represent the packaging and the last two values represent the shape of the bottle.

Every time a memory item is presented to the participant a memory trace is encoded into secondary memory. The accuracy of the memory trace depends on the value of the *L* parameter. Each non-zero feature value has an *L* probability of entering the memory trace and a 1-*L* probability of being a 0 in the trace. The resulting memory traces after two presentations of the milk bottle with the *L* parameter set to 0.5 are shown in Figure 10. Roughly half of the features in the representation of the milk bottle are present in each of the traces.

Memory Event	1	1	-1	0
T1	0	1	0	0
T2	0	1	-1	0

Figure 10. Memory event encoded into secondary memory on two occasions with a learning parameter (*L*) of 0.5

Recall is predicted by the response of all the traces in secondary memory to a probe from primary memory. The underlying principle of MINERVA2 is that all of the traces in memory respond to a probe simultaneously in the form of an "echo". Recall behavior is a function of the *intensity* and *content* of the echo. In turn, the content and intensity are influenced by the similarity of the probe to the traces in memory. If a probe is more similar to the traces then the echo and content will be stronger. Similarity is calculated as

$$S_i = \sum_{i=1}^{N} P_i T_{ij} / N \tag{9}$$

where the similarity of a trace, S_i, is the sum of each feature in the probe multiplied by the corresponding feature i, of trace j, $P_i T_{i,j}$, divided by the number of non-zero features in either the trace or the probe, N. The similarity of the trace to the probe is transformed into the activation by cubing the similarity

$$A_i = S_i^3 \tag{10}$$

where the activation of a trace, A_i, is the cube of the similarity of the probe to the trace, S_i^3. This allows traces that are more similar to the probe to dominate the echo. The intensity of the echo is the sum of the activation of all of the traces.

$$I = \sum_{i=1}^{M} A_i \tag{11}$$

Let us apply these equations to our milk bottle example (see Figure 10). The participant is shown a set of items which contains the original milk bottle and is asked to rate the familiarity of the items. The original milk bottle is the retrieval cue to which the traces in memory respond.

As shown in Figure 11, the milk bottle (probe) and the first memory trace T_1 share only the second out of three non-zero features which gives a similarity of .33

and an intensity of .04. The probe and the second trace T_2 share two features out of three non-zero features resulting in a similarity of .67 and an activation of .3. Cubing the similarity of the trace give a lower activation of traces less similar to the probe (trace one) and a higher activation of the more similar trace (trace two). The intensity of the echo from the traces given the probe is .34 which is the sum of the activations from each memory trace.

Probe	1	1	-1	0
T_1	0	1	0	0
$P_1T_{1\ i}$	0	1	0	0
N_1	1	1	1	0

$\sum P_1T_{1,j} = 1$

$\sum N_1 = 3$

$S_1 = 1/3 = 0.33$

$A_1 = S_1{}^3 = 0.04$

Probe	1	1	-1	0
T_2	0	1	-1	0
N_2	0	1	1	0
$P_2T_{2,j}$	1	1	1	0

$\sum P_2T_{2,j} = 2$

$\sum N_2 = 3$

$S_2 = 2/3 = 0.67$

$A_2 = S_2{}^3 = 0.3$

$I = 0.04 + 0.3 = 0.34$

Figure 11. Example calculations of similarity and activation in MINERVA2

The intensity of the echo increases as the similarity of the probe to all the traces in secondary memory increases. In the MDM model estimates of frequency and likelihood are a function of the intensity of the echo. Increasing the L parameter increases the correspondence of the trace and the original memory event such that low value of L result in poor recognition of a stimuli (probe) which has been previously encoded into a trace.

In MDM the content of the echo does not influence judgments. The content of the echo represents a reproduction of the original event and the intensity the strength of the recall. The content is calculated as

$$C_i = \sum_{i=1}^{M} A_i T_{ij} \tag{12}$$

where the content of the echo, C_i, is the sum of all the features in all of the traces weighted by the trace activation.

MDM predicts frequency judgments as a function of the echo intensity. The intensity of the echo depends on the similarity of the probe to all of the traces in memory. If an item is encoded in secondary memory many times, more of the traces in memory will be highly similar to the probe. A trace is created in secondary memory every time an item is presented, so more traces which are similar to the probe suggests that the probe has been shown to the memory system many times before. In other words, if an echo is very intense then the probe was more frequent. The model was applied successfully to frequency judgments by Hintzman (1988).

There are several reasons why SIMPLE might be more suitable than MDM for modeling relative and retrospective judgments. Firstly, SIMPLE models predict both primacy effects because in the model time is log compressed which decreases the confusability of the final items of a series relative to the first or last items. Secondly, SIMPLE predicts recall based on distinctiveness and can be extended to multiple dimensions. The ability of the peak and end of an experience to predict retrospective judgments could reflect the distinctiveness of the end experience along the temporal dimension and the distinctiveness of the peak experience along an intensity dimension. SIMPLE is more able to model these properties than MINERVA2. Thirdly, SIMPLE may offer a way to predict the range effects observed in relative judgment. Changes in the range position of a stimulus alter the distance of the stimulus to its neighbors. In other words, range manipulations may alter the distinctiveness of an item. Deriving the recall probability of a stimuli using SIMPLE may allow DbS to predict range effects (for example, see Brown & Matthews, 2011).

Summary of Remaining Chapters

The remaining chapters of this thesis were originally written to be separate academic papers. They have remained in this format. To highlight their purpose in the thesis as a whole each chapter has an “Introduction to Chapter” section. Here I provide a brief overview of the contribution of each chapter within this thesis.

In Chapter 2 I present experiments that investigate the influence of range position on relative and retrospective judgments. I show that the range position of a stimulus can influence both types of judgment and that these effects are predicted by RFT at an individual level. These range effects are a challenge for DbS.

The model comparison in Chapter 3 examines the extent to which these range effects can be attributed to distance based sampling. Exemplar models of memory predict that contextual items may be weighted by their distance from the item being judged. The generalized exemplar model of sampling (GEMS; Qian & Brown, 2005) incorporates distance based sampling: RFT as a special case of the GEMS model. Individual level model comparison using data from 5 previous studies shows that RFT offers a better account of the data and that relative judgment effects may not be attributed to distance based sampling.

In Chapter 4 I develop the combined SIMPLE and DbS model (SDbS) and fit it to the data from 5 previous studies. The analysis shows that SDbS and RFT can fit the qualitative pattern of the data equally well. I suggest that the SDbS should be favored when considering the independent empirical support of DbS and SIMPLE.

The experiments in Chapter 5 used monetary incentives to elicit atypical free recall behavior. Both recency (Kahneman, 2000) and output order (Johnson, Häubl, & Keinan, 2007) seem to influence decision making and judgment. Participants recalled highly incentivized items in the first output position producing atypical serial position curves. I demonstrate that SIMPLE can be used to fit and investigate the resulting data.

Finally, Chapter 6 summarizes these findings and considers them in relation to the wider research area. I discuss directions for future work to investigate an integrated and unifying account of relative judgment and decision making.

Chapter 2

Negatively Skewed Distributions Are More Satisfying

Introduction to Chapter

In the previous chapter I outlined the Decision by Sampling (DbS) model of decision making. This model can predict many economic and social decision-making phenomena. The aim of this thesis is to investigate if this model can account for relative judgment phenomena? However, DbS appears unable to predict range effects which are widely reported in the relative judgment literature. As an initial step, I investigate the size and reliability of range effects at the level of individual participants.

In this chapter I present three experiments that investigate a link between range effects in relative and retrospective judgment. These experiments demonstrate that (a) range effects are present at an individual level, (b) range effects are present in retrospective judgments. Model based analysis of the experimental data compares adaptation level theory, range-frequency theory, the range principle and the frequency principle at an individual level.

The findings presented here lay the foundation for later chapters. I go beyond previous work by examining range effects at an individual level with sequentially presented stimuli. The effect of range on retrospective hedonic judgments of sequential stimuli links the relative judgment and peak-end studies literatures. Demonstrating range effects at an individual level motivates the model comparisons presented in Chapters 3 and 4.

Abstract

How does the structure of a series of payments influence recipient satisfaction? One hypothesis is that each payment will be compared with a single “standard” or “reference” payment (e.g., the average payment). Applications of cognitive models of judgment such as range-frequency theory predict in contrast that the entire payment distribution will be influential. Three experiments examined satisfaction with a series of payments. Most payments were relatively high in the experienced distribution (negatively skewed) or relatively low in the experienced distribution (positively skewed). The total and average payment was held constant. Experiment 1 found that average satisfaction with individual payments was higher when the payments were negatively skewed, extending earlier findings with model-based analysis at the individual level. We compared range-frequency theory with the range and frequency principles. Experiment 2 examined satisfaction with whole sequences of payments and found that a negatively skewed sequence was more satisfying than a positively skewed sequence. Experiment 3 replicated the effect with dissatisfaction judgments, and found that effects of payment skew on satisfaction with overall sequences could not be explained by memory distortions due to sampling from real-world payment distributions. It is concluded that negatively skewed payment distributions are more satisfying, as predicted by range-frequency theory.

What effect does payment structure have on payment satisfaction? Employees often receive regular payments for work (e.g., weekly or monthly wages) punctuated by occasional larger amounts (e.g., monthly or yearly bonuses). Under such a system the majority of payments received are at the lower end of the range of experienced payments – i.e., the payment distribution is positively skewed. However, recipients' overall evaluations of an experienced sequence of payments may be adversely affected under such conditions. Cognitive models of context-based judgment suggest that occasional high payments may overshadow the more frequent lower payments. Intuitively, it may be dissatisfying to receive, on the majority of occasions, payments that are at the lower end of the range of payments ever received (Parducci, 1968). If overall amount of payment received is held constant, would people be more satisfied with a negatively skewed distribution of payments, in which relatively high payments occur most of the time, even if the overall amount of pay was the same, as Parducci suggested?

Here we apply cognitive models of context-based judgment to satisfaction with different payment structures. We develop and extend work by Parducci (1995) and others to examine (a) whether and under what conditions negatively skewed payment structures will be more positively evaluated and (b) whether selective memory for particular payments is responsible for such effects. We go beyond previous work in using likelihood-based model fitting at the level of individual participants (to allow model-based analysis of individual differences) and in examining both satisfaction with individual payments and satisfaction with whole sequences of payments.

The structure of the rest of this paper is as follows. First we outline the differing predictions made by various cognitive models of contextual judgment for the effects of payment structure skew on satisfaction. We then review previous studies that have examined preferences for different payment structures. Three experiments are then presented. Experiment 1 replicates an earlier finding that negatively skewed payment distributions lead to greater satisfaction (Parducci, 1968), and uses individual-level model fitting to demonstrate that there are individual differences in sensitivity to the skewness of payment distributions. Experiment 2 finds that the preference for negatively skewed payment distributions holds when whole sequences of payments, rather than individual payments, are evaluated. Finally, Experiment 3 finds that the negative-skew preference remains when dissatisfaction rather than satisfaction is elicited, and that memory bias in the recall of payments does not explain the preference.

Contextual Models of Judgment

A central assumption underpinning the present research is that the context (here, the payment distribution) within which a payment is received will influence the satisfaction judgment associated with its receipt. Within cognitive psychology, there are several different models of how judgments are made within a context (Vlaev, Chater, Stewart, & Brown, 2011). Here we review the most relevant models of contextual judgment. Each model is described and its predictions for the evaluation of different payment structures are outlined.

Reference Level Model

One possibility is that each payment is compared to some average, “typical”, or “reference level” amount. The idea that subjective judgments of payments involve comparison of each to a single reference point can be seen as an application of the adaptation level theory (ALT; Helson, 1947, 1964a) and is shown in Equation 13,

$$J_i = k(S_i - \bar{S}) \tag{13}$$

where the judgment J_i of a stimulus S_i depends on its distance from the average payment $\bar{S}$. A constant k scales responses to fit within the range of possible responses.

Within the income satisfaction literature several studies have suggested that satisfaction with one's income depends on its relation to an average or "reference" income (Clark & Oswald, 1996; Luttmer, 2005), but these studies relate to across-individual comparisons rather than the within-individual comparisons that form the focus of present paper.

The predictions of the reference-level model for positively and negatively skewed payment distributions depend on the reference payment to which all payments are compared. If the reference payment is the mean payment then the single reference point model predicts a neutral overall response to sequences with the same mean (Brickman & Campbell, 1971).

Relative Rank Model

Alternatively, people may evaluate payments according to their relative rank within the distribution of expected or experienced payments. This approach is consistent with the decision by sampling model (DbS; Stewart et al., 2006) according to which economic quantities such as payments are evaluated by counting up the number of higher and lower quantities that are present in a mental comparison sample. Several studies support a relative rank account of judgments of wages (Brown et al., 2008) and life satisfaction (Boyce, Brown, & Moore, 2010), such that people gain satisfaction from an income to the extent that it ranks higher than others rather than (or as well as) its absolute amount, but such accounts, like the reference-level model, have not generally been applied to within-individual judgments of payments.

Relative rank of the *i*th payemnt F_i is given by

$$F_i = \frac{r_i - 1}{N - 1} \tag{14}$$

where r_i is the rank position of the *i*th item and N is the total number of items. The relative rank model predicts no difference in overall responses if the number of items in the series is the same — distributional changes leave rank information unchanged. Consequently, a pure relative rank model predicts no effect of the skewness of the payment structure (although see Brown & Matthews, 2011).

Range Model

The third type of model assumes that the position of a payment within the range of experienced or expected payments will influence satisfaction with the payment. Changing the skew of a payment structure alters the range position of the payments. In a bonus style payment structure (i.e., a positively skewed structure) most payments are in the lower portion of the range. In a negatively skewed payment structure most of the payments are in the upper portion of the range. Consequently, average responses based on the range position of the payments are higher on average in a negatively skewed payment structure (Parducci & Wedell, 1986). Formally, a range based prediction is given by the equation below.

$$R_i = \frac{S_i - S_{min}}{S_{max} - S_{min}} \tag{15}$$

where R_i is the range based judgment of stimulus S_i, given the smallest $S_{\min}$ and largest $S_{\max}$ stimulus. The range based prediction is thus the distance of a payment from the smallest payment divided by the total range of payments. An effect of range on subjective judgments has been reported in the salary literature (Highhouse, Luong, & Sarkar-Barney, 1999; Rynes, Schwab, & Heneman III, 1983).

Range-frequency Theory

Range-frequency theory (RFT; Parducci, 1965) combines rank and range based predictions. RFT predictions are a weighted compromise between the range based and rank based responses as shown in the below equation.

$$J_i = wR_i + (1 - w)F_i \tag{16}$$

where R_i and F_i are as in Equation 14 and Equation 15.

The relative weighting of range and rank influence is specified by the w parameter. When w equals 1 then predictions are as for the range model. When w equals 0 then predictions are as for the relative rank model. Predictions based solely on relative rank (w = 0) are the same for two differently skewed payment structures. However, predictions based on the range position of stimuli will lead to higher average responses in a negatively skewed payment structure because there are more payments in the upper portion of the range in a negatively skewed compared to a positively skewed structure. In RFT we would expect higher average responses in the negatively skewed payment structure as the weighting parameter approaches 1.

There is considerable support in the wider judgment literature for RFT. Studies examining judgments of drink sweetness (Riskey, Parducci, & Beauchamp, 1979), hypothetical tips (Wedell & Parducci, 1988), payments (Parducci, 1968, 1995), squares of different sizes (Parducci, 1982) and many other types of stimuli (e.g., Birnbaum, Parducci, & Gifford, 1971; Parducci, Calfee, Marshall, & Davidson, 1960; Wedell & Parducci, 1988; Wedell et al., 1987) show response patterns that are predicted by RFT but are inconsistent with either rank only or range only models.

The models described here predict different average responses to skewed payment structures. The mean comparison and rank based models predict no effect of skewness on either response to individual items in the distribution or summary judgments. The range only and range frequency models (with $w > 0$) predict a preference for negatively skewed payment structures; these model predictions will be compared in Experiment 1.

Previous Findings

Skew effects have been studied under a variety of conditions. Some studies ask participants to rate their satisfaction with every payment in a series they experience (e.g., Parducci, 1968). Other studies, such as preferences for lotteries, ask participants to provide a single summary judgment (e.g., Garrett & Sobel, 1999). Furthermore, comparisons can be made within the participant's experienced payments or across group payments.

Lottery Preferences

One possibility is that a judgment of satisfaction with a wage distribution is akin to judgment of a range of risky outcomes (i.e., a lottery), with the different payments playing the role of different outcomes and the relative frequency of a given payment being its probability. Both humans (Burke & Tobler, 2011a; Golec & Tamarkin, 1998; Symmonds, Wright, Bach, & Dolan, 2011) and animals (Caraco & Chasin, 1984; Coombs & Pruitt, 1960) prefer lotteries with positively skewed outcomes when the mean expected gain is held constant. Moreover, brain regions such as the insula appear responsive to lottery skewness (Burke & Tobler, 2011b; Wu, Bossaerts, & Knutson, 2011).

To be consistent with the lottery preference literature, a payment distribution offering a low probability of a large payment and a high probability of a small payment would be preferable to the reverse, assuming that the total income is the same. Such preferences would be consistent with the underweighting of low probabilities described in prospect theory (Kahneman & Tversky, 1979), but inconsistent with both the intuitions mentioned earlier and the psychological literature reviewed below.

Responses to payment structures may in any case be quite different to preferences for lotteries, for at least two reasons. First, the judgment of risky prospects, where the outcome is not under the control of the person making a judgment or choice, seems quite different from a case where payments are earned and are assumed to reflect effort. Second, several cognitive biases (such as loss aversion) characterize anticipated feelings prior to an outcome rather than the reactions that are actually experienced following an outcome or outcomes (Gilbert, Morewedge, Risen, & Wilson, 2004; Kermer, Driver-Linn, Wilson, & Gilbert, 2006). People are subject to “affective forecasting errors” such that for example they overestimate the intensity of the negative feelings they will experience when they suffer a loss. Moreover, described and experienced outcomes are often differently evaluated (Hertwig, Barron, Weber, & Erev, 2004). Therefore people’s preferences for positively and negatively skewed outcomes in described lotteries may not relate to their satisfaction with experienced distributions of probabilistic payments.

Across-individual Comparisons

A large literature has examined the effects of income distribution (inequality) on various measures of wellbeing (Wilkinson & Pickett, 2010). Given that real-world income distributions are almost invariably positively skewed, research that has examined the effects of income inequality on happiness (e.g., Alesina, Di Tella, & MacCulloch, 2004; Hagerty, 2000) is relevant to across-individual comparisons and to preferences for greater or lesser amounts of positive skew. It is therefore of limited relevance to the current study though the payment of others can influence judgments of one's own payments (for a review see Gerhart & Rynes, 2003).

Within-individual Comparisons

Most relevant to the present investigation are the few studies that have directly examined satisfaction with skewed payment structures. A seminal paper by Parducci (1968) gave participants either a positively or negatively skewed payment structure. The participants received a payment after selecting one of three cards. After each payment the participant rated their satisfaction with that payment. The average of the satisfaction ratings was higher when the payment structure was negatively skewed. This finding is important because ratings along hedonic scales such as satisfaction or well-being may indicate levels of utility (Oswald & Wu, 2010; Sandvik, Diener, & Seidlitz, 1993). However, the study did not examine participants' satisfaction with the overall sequence of payments that they received (e.g., by asking them for a summary satisfaction judgment at the end of the study). Retrospective judgment of an overall sequence of payments might be the most relevant to real-world applications (e.g., when workers are looking back on a series of payments and evaluating their resulting satisfaction, perhaps when deciding whether to move jobs or request a raise). The study also did not examine individual differences (i.e., to discover whether all participants' response patterns were individually best described by RFT) and does not allow us to determine whether the range-only model, or the range-frequency compromise, best fit the data. Experiment 1 below replicates and extends Parducci's important study to answer these additional questions.

Smith et al. (1989) found that summary evaluations were higher for negatively skewed payment structures. Smith et al. asked participants to rate the happiness they would experience with a series of six hypothetical tips and then to rate their overall happiness with the distribution of tips. The tips were either positively or negatively skewed. The mean of the payments ratings and the summary judgment were both higher when the distributions of tips were negatively skewed. These results, like those of Parducci (1968; see also Parducci, 1995) are consistent with the expectation that negative skewed payment distributions will be preferred, but used only hypothetical payments which were not experienced sequentially.

Overview

Three experiments are reported here. In all of the experiments the participants receive a negatively or positively skewed series of payments and reported their satisfaction with the outcome of the payment series. In Experiment 1 the participants additionally rated their satisfaction with each payment in the payment structure. RFT is fit to each participant's data and compared to competing models. In Experiment 2 participant satisfaction with the overall sequence of experienced payments is examined without the presence of trial-by-trial ratings. In Experiment 3 we give participants monetary payments and ask participants to recall the payment structure they received.

Experiment 1[1]

In Experiment 1, following Parducci (1968), we asked participants to rate their satisfaction with each of a series of credit payments drawn from either a positively or negatively skewed payment structure. We extended Parducci (1968) in two ways. Firstly, we compared the fit of RFT at both the individual and group level. Secondly, we asked participants to rate their satisfaction with the outcome of the

[1] As noted in the Declaration, this experiment is not new to this thesis and has already formed part of a dissertation. It is reported here for completeness and as part of the manuscript (currently under revision) that forms this chapter

overall sequence of payments. RFT predicts that (a) mean credit satisfaction will be higher in the negatively skewed condition and (b) the RFT will best fit individual responses. Following Parducci (1968) we predict that (c) the outcome of a negatively skewed payment structure will be more satisfying.

Method

Participants. The sample consisted of 40 undergraduates from the University of Warwick separated into two groups of 20. Each participant received five candies as payment for taking part in the experiment.

Materials. The materials used in the experiment were two histograms, two decks of cards, a satisfaction scale and some paper tokens. The two histograms depicted the skewed payment structures used in the experiment (see Figure 12) which replicated the payment structures used by Parducci (1968). One deck contained cards with the payment values from the negatively skewed payment structure, and the other deck contained cards with the payment values from the negative skewed payment structure. Each card had a squared pattern on one side and a credit value (e.g., "20p") on the other. The satisfaction scale was a seven point Likert scale ranging from 1 (v*ery dissatisfied*) to 7 (*very satisfied*). The scale was printed out and placed in front of participants. Paper tokens depicting credit payments were 1cm x 1cm in size.

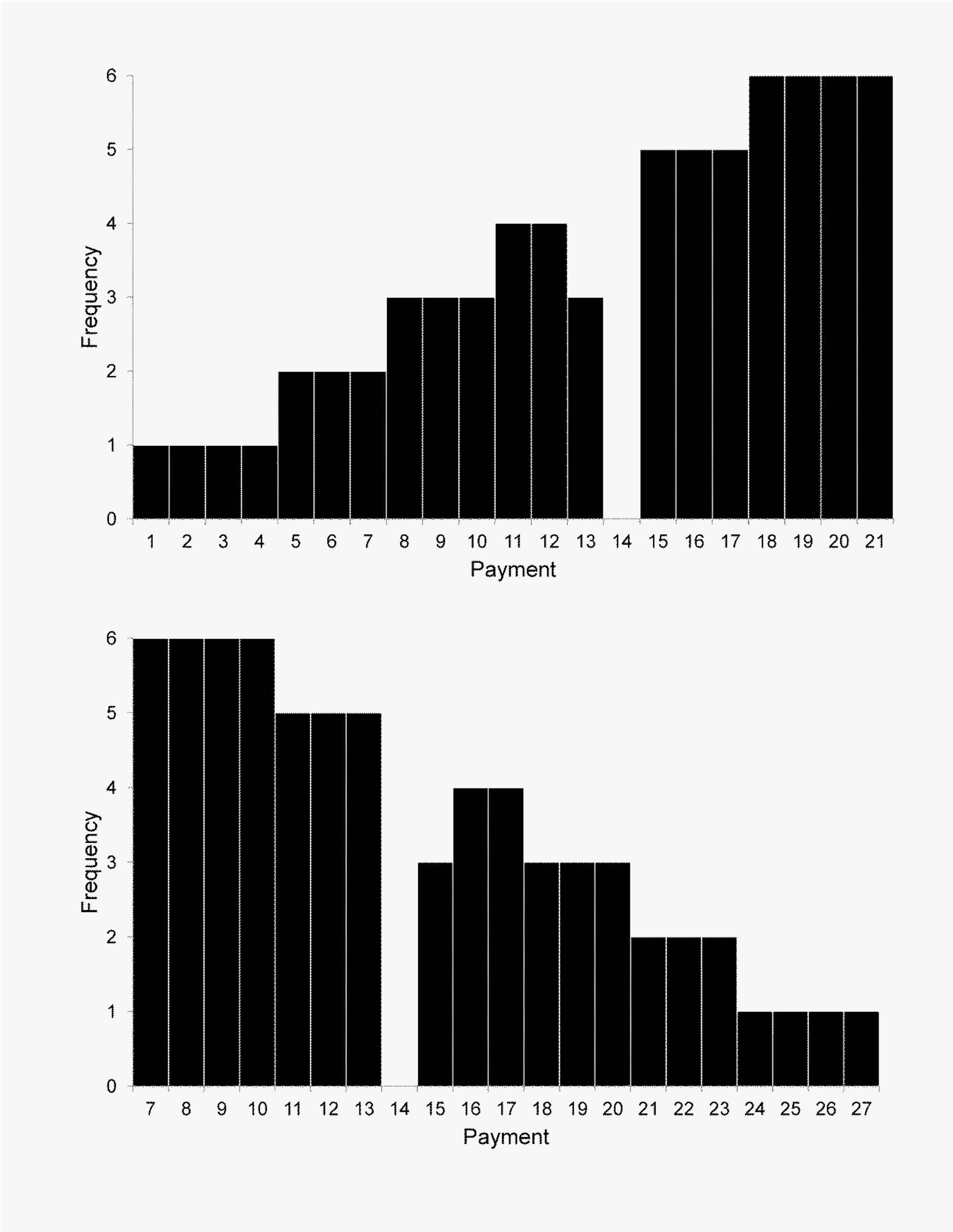

Figure 12. Histograms depicting negatively (a) and positively (b) skewed payment structures shown to participants in Experiment 1

Procedure. There were three phases in the experiment. First, participants were given instructions and performed a practice trial. Then the participants completed 69 experimental trials in which they received a credit payment and rated their satisfaction with it. Finally, participants traded all their credit payments for five candies and then rated their satisfaction with the overall amount of payment. Here we describe each phase in the order experienced by the participants.

At the start of the experiment, participants were seated at a table with three cards placed face down on it. A printed satisfaction scale was placed below the cards. Three sets of candies were placed above the cards. The sets of candies contained four, five and six candies which were labelled: ‘0-599’, ‘600-1199’ and ‘1200+’. The participants were told that they would trade the total number of credits they won in the experiment for candies at the end of the experiment – e.g., 800 credits would give the participants five candies at the end of the experiment. The participant was shown either the negatively or positively skewed payment structure, as shown in Figure 12, depending on which experimental group the participant was in. The participant was told that they could expect to receive the payments depicted by the histogram in the experiment. The histogram was removed after the participant had examined it. From this point the procedure was the same as the one used in Parducci (1968).

Next, the participant performed a practice trial. The participant chose one of the three cards in front of them. All three cards had a value of 14 on the other side (i.e., the mean of the payment distribution). The participant then turned over their chosen card, saw the 14 value on the underside of the card, and then given a 14-credit token. After receiving the credit token, the experimenter asked the participant to rate their satisfaction with the credit payment on the seven point Likert scale.

The participant then completed 69 experimental trials. First, the three cards from the preceding trial were collected and put in the bottom of the deck of cards. Then three new cards from the top of the deck were dealt face down. In each trial the credit value printed on the face down side of all three cards was (unbeknownst to the participant) the same. The deck contained cards with values from the payment structure shown to the participant at the start of the experiment. In each trial, the participant turned over one card, received a credit token equal to the value shown on the upturned card, and rated their satisfaction with the credit payment. The participant completed 69 trials and received every payment in the payment structure.

At the end of the experiment, the participant was told they had won 966 credits in total. The 966 credits allowed the participant to exchange the candies for the middle prize of 5 candies presented at the start of the experiment. The participant gave the experimenter all of the credit tokens that they had received in exchange for five candies. The experimenter then asked the participant to rate on the seven point Likert scale their satisfaction with the overall amount of candies they had received.

Results

The aims of Experiment 1 were to (a) replicate the findings of Parducci (1968), (b) to compare the fit of RFT to the fit of competing models at an individual level and (c) to examine whether the overall payments gained by participants were more satisfying if they were the results of a negatively skewed payment structure.

The data of three participants were removed from the analysis due to a low correlation between the satisfaction rating given to credit payments and the value of the credit payments. The Spearman correlation coefficients between payments and satisfaction ratings for these participants were all less than .6 which suggests that these participants had misunderstood the task.

Are payments from a negatively skewed payment structure more satisfying, on average, than payments from a positively skewed payment structure? The participants rated their satisfaction with each of the 69 payments they experienced in the experimental trials. The average of these satisfaction ratings was significantly higher when the payments were part of a negatively skewed payment structure (M =

4.49, *SD* = 0.31) than when they were part of a positively skewed payment structure (*M* = 3.77, *SD* = 0.59) skewed, $t(35) = 4.66$ $p < .001$. This difference replicates the finding of Parducci (1968) that negatively skewed distributions are more satisfying on average in comparison to positive skewed distribution.

We next examined whether the overall outcome of a negatively skewed payment structure is more satisfying than the same outcome from a positively skewed payment structure. The participants rated how satisfied they were with the candies they received in exchange for the payments they were given in the experimental trials. The reported satisfaction with the outcome of the positively (*M* = 4.88, *SD* = 0.86) and negatively (*M* = 4.94, *SD* = 0.90) skewed payment structures were not significantly different, $t(35) = 0.41$ $p = .69$. Further analysis found a significant correlation between satisfaction ratings for the final credit payment and satisfaction ratings with the outcome of the payments structure, $R_s(37) = .39$, $p = .02$. These findings suggest that participants were influenced more by their recent responses than by the skew of the payment structure when making overall judgments.

Model Comparison. Model-fitting was undertaken to determine whether RFT fit the satisfaction ratings for credit payments better than did competing models. We answer this question by comparing RFT to rank only and range only models. Model fitting and comparison was carried out at both the group, consistent with previous work, and individual levels.

For model comparison we used a maximum likelihood method. Each participant's response is assumed to be the same as a model's prediction plus normally distributed noise. The parameters of the model (i.e., the *w* parameter in RFT) and the standard deviation of the normally distributed noise were allowed to vary freely.

We varied the weighting parameter in RFT to compare the rank only ($w = 0$) and range only ($w = 1$) models to RFT ($0 \leq w \leq 1$). First, we compare the model fits to the group level data (i.e., average response across the participants). Then we fit the models to each participant's data individually to compare the model and examine individual differences in the model fits. Lower values of -2lnL indicate a higher probability of the model given the participant's response and, consequently, a better fit of the model to the data. Both range and rank models are nested within RFT, and the difference between the model approximates a χ^2 distribution. If the -2lnL is more than 3.84 lower we can conclude that the fit of the RFT model is significantly better than the comparison model because 3.84 is the critical value of the χ^2 distribution with 2 degrees of freedom.

To replicate previous model based analysis we fit each model to the average responses of the participants in each skew condition. RFT fit the group level data (-2lnL = 131.9, $w = .42$) significantly better than did the rank only (-2lnL = 252.41), $\chi^2(1, N = 2) = 120.51$, $p < .001$, or range-only (-2lnL = 200.48), $\chi^2(1, N = 2) = 68.58$, $p < .001$, models. These findings are consistent with previously reported comparisons of RFT. The best fitting RFT prediction is shown in Figure 13.

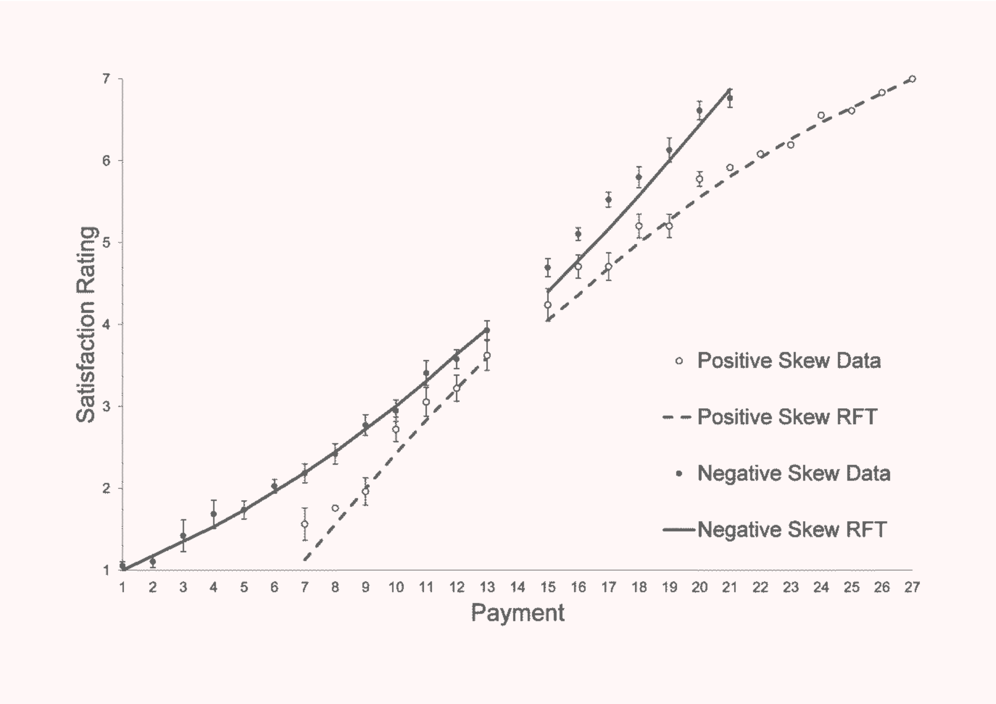

Figure 13. The average rating given to credit payments in Experiment 1

Range, rank and RFT. Is RFT significantly better than its competitors at an individual level? We fit the models to each participant's data (for individual fit statistics see Table 2 and Table 3). The satisfaction measure was ordinal so a prior distribution of standard deviations was used (see Appendix A for a detailed explanation). RFT is significantly better when the -2lnL for the RFT model is at least 3.84 lower than competing model. We have marked the participants significantly better fit by RFT on the tables using bold characters. Overall, RFT best fit individual responses.

Table 2

Fit statistics for participants in the negative skew condition

Range-only		Rank-only		RFT		
SD	-2lnL	SD	-2lnL	*w*	*SD*	-2lnL
0.87	**161.98**	**0.43**	**91.8**	**.74**	**0.34**	**72.38**
0.73	**144.59**	**0.6**	**124.2**	**.57**	**0.39**	**86.85**
0.88	**158.95**	**0.6**	**123.13**	**.66**	**0.46**	**102.59**
1.09	180.86	0.43	93.12	.91	0.42	90.87
1.72	207.97	0.64	125.4	1	0.64	125.4
0.74	**149.42**	**0.79**	**155.06**	**.47**	**0.55**	**120.57**
0.79	**153.25**	**0.74**	**143.4**	**.54**	**0.54**	**115.77**
0.54	**118.66**	**0.9**	**167.45**	**.28**	**0.46**	**102.99**
0.82	**155.2**	**0.74**	**143.37**	**.55**	**0.55**	**117.89**
0.78	**151.97**	**0.61**	**125.87**	**.61**	**0.44**	**97.25**
0.62	128.64	1.22	192.42	.06	0.61	128.11
1.26	**190.28**	**1.26**	**186.47**	**.54**	**1.11**	**177.6**
1	**175.06**	**0.36**	**74.05**	**.87**	**0.32**	**67.85**
0.63	**139.24**	**0.53**	**119.19**	**.55**	**0.3**	**67.55**
1.07	167.41	1.36	186.47	.14	1.07	166.72
0.99	**179.58**	**0.85**	**160.21**	**.63**	**0.74**	**147.48**
1.02	**194.3**	**1.08**	**201.45**	**.43**	**0.91**	**182.32**
0.8	165.99	1.49	239.19	0	0.8	165.99
0.93	**185.59**	**1.03**	**195.75**	**.42**	**0.83**	**172.59**

Note: Bold typeface indicates that RFT performed significantly better than either the range or rank only models

Table 3

Fit statistics for participants in the positive skew condition

Range-only		Rank-only		RFT		
SD	-2lnL	SD	-2lnL	*w*	*SD*	-2lnL
2.57	**268.96**	**1.6**	**218.31**	**1**	**1.6**	**218.31**
0.64	**124.66**	**0.7**	**138.57**	**.45**	**0.45**	**90.86**
0.56	**116.7**	**0.84**	**161.78**	**.3**	**0.47**	**99.98**
1.81	224.06	1.26	187.32	1	1.26	187.32
1.56	211.62	0.83	150.92	1	0.83	150.92
1.16	197.94	0.51	113.08	.98	0.51	113
1.58	218.29	0.51	110.23	1	0.51	110.23
1.01	**197.63**	**1.01**	**196.35**	**.51**	**0.88**	**180.37**
0.65	131.53	1.25	184.84	.1	0.64	130.23
1.29	185.91	0.77	140.91	.87	0.75	139.3
0.95	**168.35**	**0.94**	**161.28**	**.54**	**0.76**	**144.67**
0.86	153.66	1.74	211.53	0	0.86	153.66
0.68	**139.45**	**0.86**	**154.33**	**.41**	**0.52**	**114.34**
1.31	**200.14**	**1.19**	**188.13**	**.65**	**1.11**	**182.52**
0.99	**179.71**	**0.39**	**84.75**	**.87**	**0.36**	**79.01**
0.94	159.3	1.54	200.91	0	0.94	159.3
0.59	130.94	1.07	195.72	.12	0.58	128.62
0.99	**174.52**	**0.36**	**78.04**	**.86**	**0.33**	**70.53**

Note: Bold typeface indicates that RFT performed significantly better than either the range or rank only models

What is the relationship between individual satisfaction and the range-rank compromise? RFT predicts that skew effects – higher mean satisfaction in the negative skew and lower mean satisfaction in the positive skew condition - should be correlated with the weighting parameter estimates. This is because the skew manipulation alters the number of items in the upper and lower portion of the range whilst keeping the number of items constant. Predictions based on relative rank alone (w=0) will be the same for both skew conditions. However, as the w parameter approaches 1 the model predictions are increasingly based on the range-position of the payments. When the payments are negatively skewed items –i.e., are most frequent in the upper portion of the range – the mean predicted response increases as *w* approaches 1. When payment are positively skewed – i.e., items are most frequent in the lower portion of the range – the mean predicted response decreases as *w* approaches 1. We would expect *w* to positively correlate with mean satisfaction in the negatively skewed condition and negatively correlate with mean satisfaction in the positively skewed condition.

For the participants whose data were best fit by RFT, there was a significant correlation between *w* and mean satisfaction in both the positively, $R_s(12) = -0.70$, $p = .01$, and negatively, $R_s(17) = 0.95$, $p < .001$, skewed conditions. The fact that participants showed a strong relationship between the range-rank compromise and their satisfaction responses supports RFT.

Discussion

The first finding from Experiment 1 was that the average satisfaction with individual payments was higher when payments came from a negatively skewed distribution. This finding replicates Parducci (1968), and is consistent with the predictions of RFT. Model-fitting demonstrated that RFT accounted for the data better than did range-only, or rank-only models at both group and individual level.

However the main finding, that participants preferred a negatively skewed payment distribution, must be qualified; participants' stated satisfaction with the distribution as a whole was not affected by skew. Instead, retrospective judgments depended on the final credit payment response as predicted by the peak-end rule (Langer, Sarin, & Weber, 2005; Tversky & Kahneman, 1981). The significant positive correlation between the final and overall response found in Experiment 1 is consistent with this explanation. However, the complete absence of a skew effect is inconsistent with the skew preference literature reviewed above and with evidence of averaging in summary hedonic evaluations (e.g., Miron-Shatz, 2009).

Furthermore, the results are consistent with at least two other interpretations. Firstly, if participants neglected the skew of the distribution and instead compared the candies won to three possible outcomes, then the relative rank or range models would predict a neutral response because all participants were given the middle of the three possible prizes. Secondly, the use of online responses may have distorted the effect of distribution skew on summary responses. Online ratings in some studies have reduced or removed the effect of stimuli structure (Ariely, 1998; Ariely & Carmon, 2000; Ariely & Zauberman, 2000, 2003). Participants may (having already stated their satisfaction with the individual payments they received) interpreted the task demands such that they produced an evaluation of different aspects of the experiment in their summary evaluation. We address this possibility directly in the next experiment by replicating Experiment 1 without the online ratings.

Experiment 2[2]

Experiment 1 found no skew effect on the participants' retrospective satisfaction with the outcome of a skewed payment structure. Two explanations consistent with RFT were that the use of online ratings or the relative rank of the outcome negated the effect of skew. To test these possibilities, the design of Experiment 1 was almost identical to that of Experiment 2 except that the online ratings used in Experiment 1 were replaced by a task in which the participants merely entered the credit payment they received on each trial. This was done to ensure attention to the individual rewards. As before, participants provided a

[2] As noted in the Declaration, this experiment is not new to this thesis and has already formed part of a dissertation. It is reported here for completeness and as part of the manuscript (currently under revision) that forms this chapter

summary evaluation of their satisfaction with the overall distribution of payments they received. Based on previous findings reviewed above, we predict that the participants will be more satisfied with an outcome from a negatively skewed payment distribution.

Method

Participants. The sample consisted of 12 students from Kings College in Wembley and 12 from the University of Warwick. Each cohort was divided equally into two experimental groups and each participant was paid five candies for their participation.

Materials. The stimuli and instructions were presented to participants using a Java program viewed within an internet browser (see procedure). The histograms depicting the skewed payment structures were identical to those used in Experiment 1.

Procedure. Participants were seated in front of a computer showing instructions for the experiment and told that they would receive credits which could later be exchanged for candies. The number of credits needed for the candies was the same as in Experiment 1. The candies and labels were displayed in front of the computer keyboard throughout the experiment.

Participants gave their consent by ticking a box on the screen and then the histogram used in Experiment 1 was displayed onscreen. The histogram displayed either a positively or negatively skewed payment distribution (as shown in Figure 12) for several seconds before participants could click on the screen and continue with the experiment. Next, participants were shown three cards and asked to choose one by clicking on it. After the choice was made, a payment (number of credits) was shown.

The credits received were identical to those of Experiment 1, starting with a practice trial of 14 credits. Unlike in Experiment 1, participants had to enter the number of credits using the keyboard instead of reporting a satisfaction rating after clicking on a card.

After completing all the trials the participant were told the number of credits they had received and exchanged the credits for candies. The next screen asked participants to judge how satisfied they were with the candies by selecting a value on a seven point Likert scale similar to the one used in Experiment 1.

Results and Discussion

The results supported the hypothesis that negatively skewed distributions of payments would be more satisfying than positively skewed payment distributions. The payment value shown onscreen and payment value entered by the participant were correlated (p's $< .001$for all participants) which confirmed that each participant attended to each payment. In contrast to the findings of Experiment 1, satisfaction with the overall received payment was higher in the negative skew condition (M = 5.36, SD = 1.36) when compared to the positive skew condition (M = 3.83, SD = 1.47), $t(21) = 2.59$, $p = .02$. The higher satisfaction ratings reported by participants in the negative skew condition support the summary judgment predictions from Parducci (1995) and match the findings reported by Smith et al. (1989).

Taken together the results of Experiments 1 and 2 support the conclusion that negatively skewed distributions of payments are more satisfying than are positively skewed distributions, even when the overall amount of payment is constant. Next, we turn to the question of whether the preference for negatively skewed distributions might reflect bias in memory for the experienced payments.

Experiment 3

The theories of relative judgment outlined above assume that, when making summary judgments, participants accurately recall the distribution of payments they have received. This assumption is hard to reconcile with literature from memory pointing out distortions in recall such as the ubiquitous primacy and recency effect reported in the episodic memory literature. However it is possible that participants' memory for a distribution is distorted by their prior beliefs about the distribution that payments typically follow. Experiment 3 was designed to test this possibility, as well as to examine whether the preference for negatively skewed distributions would remain when dissatisfaction, rather than satisfaction, judgments were elicited.

More specifically, participants will have considerable experience of positively skewed payment distributions in their real-world experience. For detailed analysis of the skewness of real world payment distributions see Stewart et al. (2006). If they (a) treat the satisfaction judgment as a task that involves reconstructive inference, such that they infer the true payment distribution from a combination of the payments they remember and their prior beliefs, and (b) have prior beliefs that payment distributions are more likely to be positively skewed, an apparent preference for negative skew might result.

This could occur if the best-fitting positively skewed (e.g., lognormal) distribution for the negatively skewed stimuli (cf. Figure 12) had a higher mean than the best-fitting positively skewed distribution for the positively skewed stimuli. This assumes that people distort the stimuli they have experienced based on their experience of a positively skewed distribution of payments outside of the experimental setting. In fact it is the case that the relevant estimates of the mean are 14.7 and 14.0 for the negatively and positively skewed distributions respectively. In the light of such possibilities, and given the likelihood of memory distortions more generally, such that the more “distinctive” payments are more likely to be remembered at the point of summary evaluation (Brown et al., 2007), it was considered important to determine whether participants were able to recall the experienced payments accurately.

Experiment 3 differed from Experiment 2 in three ways. Firstly, at the end of the experiment participants recalled the payments they had received. The skew effect reported in Experiment 2 may have been a result of the participants misremembering the payment distribution. Secondly, credit payments were replaced by direct monetary payments. Thirdly, the participants rated either their satisfaction or their dissatisfaction with the total payment to investigate any possible effect of framing on overall satisfaction evaluations.

Method

Participants. The sample consisted of 84 undergraduates from the University of Warwick separated into four groups of 21. Each participant was paid £4.16 as a result of their participation.

Design. The three factors manipulated between participants were the skewness of payment distribution (positive or negative), the type of scale used by participants when judging their total income (satisfaction vs. dissatisfaction), and the position of the recall task in the procedure (before or after provision of an overall satisfaction rating). We recorded the value that the participants typed into the computer in each trial, the participant’s satisfaction/dissatisfaction with their total payments, and the payments that the participants recalled at the end of the experiment. As in previous experiments, the order of payment permutation was random.

Materials. The presentation of the stimuli was controlled by a computer program written in Blitz3D. The structure of the program and the instructions given to the participants were the same as those used in Experiment 2. The values in the histogram presented to the participants and the credit payments were half those used in Experiment 2. Two Likert scales were used. The dissatisfaction scale ran from 1 (*very satisfied*) to 7 (*very dissatisfied*). The satisfaction scale ran from 1 (*very dissatisfied*) to 7 (v*ery satisfied*).

Procedure. The procedure was separated into four phases. First, a series of instructions was displayed which included a histogram depicting the payment structure. Second, the participant performed a practice trial to familiarize themselves with the procedure. Third, the participant carried out 69 experimental trials with monetary payments. Finally, the total of the monetary payments was displayed and participants judged their satisfaction/dissatisfaction with the total payment and recalled as many monetary payments as they could remember. Throughout the experiment the participant sat next to the experimenter in a laboratory cubicle.

Participants were first given a series of instructions. The participants were told they would be given a monetary payment in each experimental trial and that they would receive the total of those payments at the end of the experiment. Onscreen instructions informed them that in the experiment they would select one of three cards which would then be replaced with a payment value which they would type into the computer program. Then a histogram depicting the skewed payment structure they could expect to receive in the experiment trials was displayed. The skew of the payment structure depended on the skew condition that participants were in. The histogram was replaced with a screen introducing the practice trial when the participant pressed any key on the keyboard.

The practice trial familiarized the participants with the procedure. The patterned side of three cards was displayed onscreen. When the participants clicked on one of the cards the patterned side of that card was replaced with a card outline containing the text “14p”. The participants typed in the value displayed in the card outline and then pressed enter. The participants were asked if they had any questions about the procedure to ensure they understood the experiment. Any questions the participants had were addressed by the experimenter. After the experimenter answered any questions the participants then proceeded to the experimental trials.

Participants received monetary payments in each of the 69 experimental trials. The procedure of the experimental trials was the same as the practice trials. The text displayed within the outline of the card was one value from either the negatively or positively skewed payment structure, depending on the skew condition. All of the 69 payments depicted in the histogram were shown to each participant.

Participants were then informed of their total payment: A screen informing participants that they had received a total of £4.16 was displayed. After participants pressed any key the next screen was displayed. The participants then performed a recall task either before or after a rating task. In the recall task the participants wrote on a piece of paper as many of the monetary payments from the experimental trials as they could remember. They were given a 2 minute time limit. In the rating task the participants were shown either a satisfaction or dissatisfaction scale onscreen and onscreen instructions asked them to click on the scale to indicate their satisfaction with the total payment.

Results and Discussion

We first examined whether the greater satisfaction for negatively skewed distributions (a) replicated the findings of Experiment 2, and (b) survived the shift to judgments of dissatisfaction rather than of satisfaction. Three factors were manipulated between participants and factored into an ANOVA analysis: the skew of the payment structure (positive or negative), when the recall task was performed (before or after overall ratings), and type of response scale (satisfaction/dissatisfaction). Overall judgments made on a dissatisfaction scale were reversed to allow comparison of all the responses in a single ANOVA analysis.

A 2 x 2 x 2 ANOVA found a significant main effect of skew, $F(1,76) = 13.65$, $MSE = 22.11$, $p < .001$, $\eta^2 = 0.15$. As expected, participants were more satisfied with the same total payment when it was the outcome of a negatively ($M = 5.31$ $SD = 1.2$) rather than a positively ($M = 4.29$, $SD = 1.29$) skewed payment distribution. Neither the type of scale, $F(1,76) = 0.01$, $MSE = 0.01$, $p = 0.94$, $\eta^2 = 0$, nor the position of the recall task, $F(1,76) = 0.01$, $MSE = 0.02$, $p = 0.92$, $\eta^2 = 0$, had a significant main effect on overall judgments. There was no significant interaction between payment structure skew and scale type, $F(1,76) = 0.27$, $MSE = 0.44$, $p = .6$, $\eta^2 = 0$, payment structure skew and the position of the recall task, $F(1,76) = 3.29$, $MSE = 5.33$, $p = .08$, $\eta^2 = 0.04$, or scale type and the position of the recall task, $F(1,76) = 0.27$, $MSE = 0.44$, p = .6, $\eta^2 = 0$. No significant three way interaction was found, $F(1,76) = .2$, $MSE = .32$, $p = .7$, $\eta^2 = 0$. The only main effect found in the analysis was skew which replicated the findings reported in Experiment 2 and is consistent with previously reported findings.

Next we examined accuracy of the recalled distributions. The skewness of the payment structures recalled by each participant was calculated and entered into a 2 x 2 x 2 ANOVA which examined the effects of the payment structure skew (positive or negative), the type of scale used for overall judgments (satisfaction or dissatisfaction) and the position of the recall task (before or after overall judgments).

The skewness of the recalled payment structure in the positively (*M* = 0.46, *SD* = 0.32) and negatively (*M* = -0.7, *SD* = 0.32) skewed payment structures were significantly different, $F(1,76) = 200.46$, *MSE* = 33.38, $p < .001$, $\eta^2 = 0.73$. This difference suggests that participants had correctly encoded the skew of the distribution in memory. No significant main effect of either the position of the recall task, $F(1,76) = 1.15$, *MSE* = 0.19, $p = .29$, $\eta^2 = 0.02$, or the type of scale used, $F(1,76) = 1.65$, *MSE* = 0.28, $p = .2$, $\eta^2 = 0.02$, was found. There were no significant interactions.

Then we examined whether participants accurately recalled the payment structure that they received in the experimental trials. In the negative skew condition there was a significant difference between the skewness of the experienced (-0.55) and recalled (*M* = -0.76, *SD* = 0.43) payment structure, $t(41) = -3.12$, $p = .003$. In the positive skew condition there was no significant difference between the skewness of the recalled (*M* = 0.5, *SD* = 0.38) and experienced (0.55) payment distribution, $t(41) = -0.85$, $p = 0.402$. These results show that participants recalled a payment structure that was more extremely skewed than the payment structure they experienced. Memory models such as SIMPLE predict that recalled distributions should be less skewed than the experienced distribution (as in the positive skew condition).

The mean of the recalled distributions was significantly smaller than the experienced mean of 7.5 in both the negatively skewed (M = 7.12, SD = 0.7), $t(41) = -3.5$, $p < .001$, and positively skewed (M = 7.20, SD = 0.65), $t(41) = -3.01$, $p = .004$, payments structures. There was no difference between the recalled mean based on the skewness of the payment structure, $t(82) = 0.51$, $p = .61$. Though the mean was underestimated by the participants, the skewness of the distribution did not influence the mean of the recalled distribution.

In summary: Experiment 3 confirmed that the total payment from a negatively skewed payment distributions is more satisfying, and extended the previous findings to show that (a) the effect remains when dissatisfaction rather than satisfaction is judged, and (b) that memory distortions due to previously experienced positively skewed payment distributions are unlikely to be responsible for the effect.

General Discussion

Three experiments showed that negatively skewed payment distributions are judged to be more satisfying. This effect holds when the average satisfaction with individual payments is examined (Experiment 1; Parducci, 1968) and is estimated to hold for about 90% of participants based on individual-level model-fitting (Experiment 1). It is also seen when a whole series of payments must be given a summary evaluation (Experiments 2 and 3), and applies when dissatisfaction rather than satisfaction judgments are elicited (Experiment 3).

The results have both theoretical and practical implications. At a theoretical level, the findings challenge reference-level models according to which payments are judged relative to some mean or reference level payment. Instead, the results appear consistent with an interpretation in terms of RFT (e.g., Parducci, 1968, 1995). Crucially, we extended earlier work by examining these models at an individual level. RFT best predicted the responses of a majority of participants.

These results also shed light on the role of the memory system in overall judgments. Our findings suggest that people do not sample from real-world distributions when forming subjective judgments about skewed distributions in an experimental setting. Instead, when a single judgment was made the total payment was given a higher satisfaction rating. However, satisfaction with the total payment was correlated with recent satisfaction judgments. One interpretation of this findings is that people recall their most recent judgment when deciding their overall satisfaction with a distribution, which is consistent with findings from the peak-end literature (Redelmeier & Kahneman, 1996) and the widely reported recency effects from the memory literature (for examples, see Brown et al., 2007). This conclusion is however tentative as it relies on correlational analysis and not experimental manipulation. At a practical level, the findings suggest that payment satisfaction can be altered by manipulating only the skew of the payment structure. The design of incentive schemes typically focus on the relation between productivity and payment. A typical bonus style scheme in which the highest payments are the least frequent is an example of this. Our findings suggest that restructuring these schemes so that highest payments are the most frequent may increase the satisfaction of the recipient. Existing schemes ignore the possible effects of distribution which may underline the effects of incentive schemes. We have demonstrated that payments can be more satisfying without cost to the payer due to distribution effects which are predicted by a simple contextual judgment model.

Chapter 3

Distance-based Sampling

Introduction to Chapter

In the previous chapter I found evidence of range effects in both relative and retrospective judgments. These findings are hard to reconcile with Decision by Sampling (DbS; Stewart et al., 2006). In the present chapter I investigated the role of similarity in relative judgment. Models of memory predict that similarity influences recall performance in a cued recall task: Items that are more similar to the cue are more likely to be recalled. One could argue that relative judgment tasks are similar to cued recall. The judged item is the cue in the recall of other items. If this is the case then we would expect to see an effect of similarity in judgment.

I use the generalized exemplar model of sampling (GEMS; Qian & Brown, 2005) to examine similarity effects in relative judgment. Within GEMS the Range-frequency theory (RFT; Parducci, 1965) is a special case. In GEMS the frequency principle is altered to produce distance based sampling effects. These effects are set by the γ parameter in the model. Similarity effects are modeled when items closest to the judged item are most heavily weighted in judgment. Using GEMS I directly examine the contribution of the similarity effects predicted by the memory literature beyond the range and frequency principles. I do this by comparing the performance of RFT with and without the distance based weighting of contextual items.

Abstract

How do individuals use contextual information when making subjective magnitude judgments about items, for example, when judging the subjective largeness of a square? According to the highly influential Range-Frequency Theory (RFT; Parducci, 1965) all contextual items are equally weighted when making subjective judgments about any given stimulus. Studies over the past 60 years which investigate subjective judgments of items have consistently supported RFT. The model predicts that a judgment is a weighted average of the position of a stimulus within the range of similar items and the relative rank of the stimulus amongst the contextual stimuli. However, findings from the psychophysics of prices directly challenge the equal impact assumption. These findings support an alternate hypothesis which states that the impact of an item on judgment depends on the distance of the item from the stimuli being judged. Furthermore, this "distance weighting" hypothesis is consistent with the predictions of exemplar models of memory. In this chapter I assess the performance of RFT at an individual level in comparison to (a) its range-only and rank-only components, and (b) the generalized exemplar model of sampling (GEMS) in which contextual information about similar items is weighted by the distance of an item from the judged stimulus. I go beyond previous work to examine the limitations and strengths of RFT at an individual level, and address the distance based sampling hypothesis directly using data from several previous studies (N=370). Furthermore, I examine individual level uncertainty in the weighting of the range and frequency components of RFT. The results are discussed in relation to recent findings suggesting that people make their subjective judgments based on both real world distributions and experimental stimuli.

People often make subjective magnitude judgments. For example, how long is this chapter? Such judgments may depend upon a comparison set. This chapter is short compared to longer review articles which may be 50 pages in length. On the other hand, it is long in comparison to short reports which can be four pages in length. Intuitively, we would expect that a subjective judgment (e.g., of length) will be made relative to a set of contextual items (e.g., other papers). If the set of contextual items is changed then the subjective judgment may also change. Here lies the key theoretical question: What is the relationship between the changes in contextual items along an objective dimension (e.g., number of papers) and the subjective judgment given to an item (e.g., the length of a paper)?

Studies over the past 60 years have systematically investigated this relationship. In typical studies a participant is presented with contextual stimuli and then asked to judge each of a set of stimuli along a subjective dimension. Early findings suggested that judgments are made relative to a single neutral reference point (Helson, 1947). This adaptation level theory (ALT) was widely applied to judgments of psychophysical stimuli (for a review see Appley, 1971; Helson, 1964a) after success predicting responses in color vision (Helson, 1938; Helson & Jeffers, 1940), and remains influential in, for example, the price perception literature (Briesch et al., 1997; Mazumdar et al., 2005).

However, later findings from the psychophysics literature strongly support and alternative: range-frequency theory (RFT; Parducci, 1965, 1995). Parducci et al. (1960) presented participants with a series of numerals on a sheet of paper and asked them to judge the size each numeral. He found that judgments were influenced by the median and midpoint of the series. However, changing the mean had no effect on judgments. These findings were inconsistent with the predictions of ALT but are predicted by RFT (Parducci, 1965). Subsequent studies show that RFT can be applied to a diverse range of judgments (e.g., Birnbaum et al., 1971; Parducci, 1968; Parducci & Wedell, 1986) and contextual conditions (Parducci, 1982, 1995; Parducci & Perrett, 1971). Recent studies suggest that the model can be extended to social judgments (Wedell, Santoyo, & Pettibone, 2005), categorization (Wedell, 2008) and price perception (Niedrich et al., 2001).

Here I examine the strengths and limitations of RFT by addressing two questions. First, does RFT predict the judgments of individuals? Previous work established that RFT can predict the average response of participants. Establishing the individual level performance of the model and individual differences in fits is particularly important given recent work which suggests that people sample from real world distributions when making judgments about stimuli (Juslin et al., 2007). Second, are contextual items weighted by their distance from a judgment stimulus when a judgment is made? Consider an intuitive example of buying a TV set where one needs to judge the attractiveness the prices of the TV. We might expect people to pay more attention to TV sets with a similar price. In other words, nearby prices are most heavily weighted in judgment. Evidence from psychophysics and predictions from exemplar models of memory suggest that contextual items are weighted by their distance from the items being judged. In RFT, either contextual items are equally weighted in judgment (frequency principle) or only stimuli at the ends of the contextual range (range principle) influence judgment.

The paper is divided into two sections. Each section addresses one of the above questions. In the first section, RFT is outlined and individual level analysis is motivated. Then I report the results of individual level model comparison comparing the rage-only and rank-only component of the model to RFT. Uncertainty arising from the maximum likelihood methodology is addressed by carrying out bootstrap analysis on the parameter controlling the weight given to the range and rank components. In the second section, I note evidence from price psychophysics supporting the use of distance based sampling in judgment. I discuss models of memory which predicts distance based sampling and then compare RFT to the Generalized Exemplar Model of Sampling (GEMS; Qian & Brown, 2005) which incorporates distance based sampling in judgments.

Range-frequency Theory

RFT predicts judgments based on two principles. According to the range principle, judgments depend on the position of the judged item within the range of experimental stimuli (see Equation 15).

Parducci (1965) argues that this principle applies because participants subdivide the response categories equally across the range of stimuli. For example, payments which are in the upper portion of a payment range will be given a higher satisfaction rating.

The frequency principle predicts judgments based on the number of items lower and higher than the judged stimulus (see Equation 14).

. According to Parducci (1965) the frequency principle is a result of participant's tendency to use the response categories with equal frequency in the experiment. Intuitively, we might expect a payment to be most satisfying when most of the expected payments are lower. The frequency principle is consistent with findings from financial decision making (Stewart, 2009; Stewart et al., 2006).

Predictions of RFT are a weighted average of the range and frequency principles,

$$RFT_i = wR_i + (1 - w)F_i. \tag{17}$$

When the weight parameter, *w*, equals 1 predictions are based on the range position of the items and when *w* = 0 predictions are based on relative rank only. In previous studies the best fitting *w* was between 0 and 1, which gives predictions somewhere between range-only and rank-only and supports RFT (e.g., Parducci, 1956; Parducci, 1965; Wedell et al., 1989).

Individual level modeling allows us to examine individual variation in the weight of the Range and Frequency principles in judgment. The Decision by Sampling model (DbS; Stewart et al., 2006) predicts judgments based on the relative rank of an item in a sample drawn from memory. If a sample contains only the experimental stimuli then DbS predicts the same responses as the rank principle. On the other hand, people may draw a combined sample of experimental stimuli and items from a real-world distribution of similar items, as appears to be the case in financial judgments (Stewart et al., 2003). These judgments based on a combined sample from memory may be closer to the range predictions.

I ran a simulation to investigate the influence of combined environment and experiment sampling on the contextual judgments of items. Consider an experiment where participants are shown a bimodal distribution of payments: £2 £3 £4 £12 £13 £14. They are asked to rate their satisfaction with each payment on a scale ranging from 0 (*dissatisfied*) to 1 (*satisfied*). In this simulation I assume that the participant's prior distribution of payments from the real world is normally distributed. When making a judgment about a payment the participant calculates the relative rank of the payment within a memory sample. This memory sample consists of the experimental stimuli and N items from the prior distribution.

The results of the simulation are shown in Figure 14. When N=0 then predictions are the relative rank of each item within the experimental stimuli. This is expected because the sample from memory contains only the experimental stimuli. However, the simulation predicts a non-zero value of the *w* parameter as more items from the real world distribution are added to the sample. These variations in N may result from between participant differences in memory performance. If people differ in how much they are sampling from prior distributions then I would expect variation in the best fitting *w* parameter at the individual level.

This first series of model fits will examine RFT at an individual level. The performance of both the range-only and rank-only models will be compared to RFT. Model comparison will be carried out using data from five previous studies (N=370) which ask participants to respond to different stimuli.

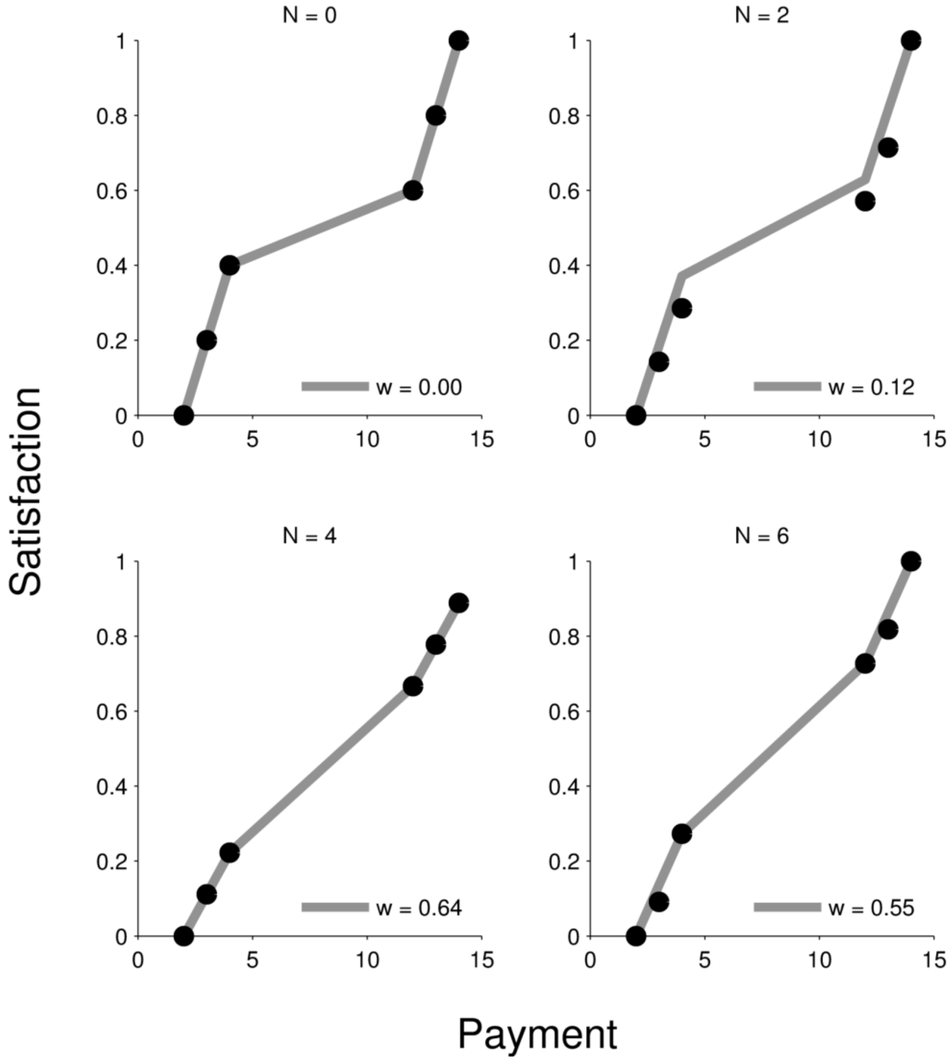

Figure 14. RFT fits of DbS predicted responses given N items sampled from a prior normal distribution

Method

Previous Data. I fit data from five previously published studies. In each study participants made subjective judgments of stimuli which varied along a single dimension (e.g., monetary value). When making each judgment the participant was aware of the entire distribution (see Table 4 for the distribution values and participant numbers) and made a subjective judgment for every stimulus. Next I consider the method of each study in details and outline the key variables.

Brown, Gardner, Oswald and Qian (2008). Participants were asked to rate a series of hypothetical wages. They were told that these wage values were prospective wages for their first job after finishing their university degree. Each participants was shown six wage distributions. The distributions were shown one at a time and the wage values for a given distribution were shown simultaneously. Wages were written on labels and the participant's task was to place the label for each payment on a 1 to 7 satisfaction scale.

In the model fitting reported here I kept the parameters of the model constant for all of a participant's responses. Each participant responded to all 11 wages in each of the six wage distributions (a total of 66 responses).

Melrose, Brown and Wood (2012). Participants were asked to give the probability of a statement describing a symptom from the diagnostic statistics manual being true. They were shown a series of 11 values. Each value represented a different person and was the number of days per month that person had experienced a symptom. Participants were asked: "For each person, please indicate in the form of a percentage how likely you think the following statement is to be true where 100% is certain to be true". The participant was shown all of the distribution values on a piece of paper and then assigned percentages to each value.

This task was repeated for six symptoms. The symptoms were feeling (1) depressed, sad, blue, tearful (2) tired and having no energy, and (3) worthless or excessively guilty about things you have or have not done, or anxiety, and experiencing (1) excessive anxiety about a number of events or activities (2) irritability, and (3) muscle tension. Each participant gave 66 responses. Model parameters were held constant for each participant across all six symptoms.

Wood, Brown and Maltby (2012a). Participants were asked to state their perception of the risk (%) of someone developing an alcohol related disorder when consuming a given number of alcohol units for 20 years. The participants were shown either a unimodal or bimodal distribution of alcohol units on a piece of paper.

This task was repeated eight times. The alcohol disorders were (a) any alcohol-related illness, (b) any serious psychological difficulties as a direct result of alcohol use, (c) a dependency on alcohol and (d) cirrhosis of the liver. Participants responded to each disorder for males and females. Model parameters were held constant across all eight responses. Each participant gave 44 responses.

Wood, Brown and Maltby (2012b). Participants read two vignettes in which 11 hypothetical people volunteered different amounts either time or money to help the participant. Participant then rated their gratitude for each value of time/money from 1 (*not at all grateful*) to 10 (*extremely grateful*). Model parameters were held constant across the 22 responses from each participant.

Maltby, Wood, Vlaev, Taylor and Brown (2012). Participants were shown either a bimodal or unimodal distribution of 11 exercise durations on a piece of paper. The participants rated on a 1 to 10 scale: (1) how healthy it was for the individual, (2) the extent of the health benefits, (3) how sufficient it would be to avoid the serious health problems associated with low levels of exercise. Participants gave 33 responses. As with previous fits, the model parameters were held constant across a participant’s responses to all three questions.

Table 4

Summary of previous studies

Paper	Stimulus	Distribution	Judgment	N	Stimuli values										
Wood, Brown and Maltby (2011a)	Units of alcohol	Unimodal	Chance of negative consequence (%)	40	5	19	26.5	30	32	33	34	36	39.5	47	61
		Bimodal		41	5	6	8	11.5	19	33	47	54.5	58	60	61
Brown, Gardner, Oswald and Qian (2008)	Payments (£)	Positive Skew	Satisfaction	24	17.2	17.6	18.1	18.7	19.5	20.3	21.4	22.7	24.3	26.1	28.4
		Negative Skew			17.2	19.5	21.3	22.9	24.2	25.3	26.1	26.9	27.5	28	28.4
		Unimodal			17.2	20	21.5	22.2	22.6	22.8	23	23.4	24.1	25.6	28.4
		Bimodal			17.2	17.4	17.8	18.5	20	22.8	25.6	27.1	27.8	28.2	28.4
		Low Range			14.3	17.1	18.6	20	21.4	22.8	25.9	26.8	27.5	28	28.4
		High Range			17.2	17.6	18.1	18.8	19.7	22.8	24.2	25.6	27.1	28.5	31.3
Melrose, Brown and Wood (2012)	Symptom occurance	Unimodal	Truth of DSM statement (%)	26	3	10	12	13	14	16	18	19	20	22	29
		Bimodal		26	3	4	6	8	10	16	22	24	26	28	29
Wood, Brown and Maltby (2011b)	Time/ money	Unimodal	Gratitude	38	9	23	27	30	33	36	39	42	45	49	63
		Bimodal		40	9	12	15	19	23	36	49	53	57	60	63
Matlby. Wood, Vlaev, Taylor and Brown (2012)	Exercise Duration	Unimodal	Health of exercise duration	68	36	92	108	120	132	144	156	168	180	196	252
		Bimodal		67	36	48	60	76	92	144	196	212	228	240	252

Model Fitting. The goal of the model fitting was to find the RFT parameters which maximized the likelihood of the model. The likelihood of the model was calculated assuming a normal distribution for each response. Here we assume that response have a normally distribution noise. This normal distribution had a mean equal to the prediction of a model for a response. The standard deviation of this normal distribution was kept constant across the responses of a participant but otherwise allowed to vary.

I recorded the probability of the data given the assumed normal distributions. This gave us the likelihood of the model given the data which was then transformed into a negative log likelihood (-2lnL). As the likelihood of the model increases the -2lnL decreases. A function minimization algorithm in MATLAB (`fminsearch`) altered the *w* and SD parameters to minimize the -2lnL. This gave us the estimates of the most likely parameters given the data. The model parameters and standard deviation were held constant for each participant across a participant's responses to different questions or stimuli distributions.

To compare the model fits the maximum likelihood of the range-only (w = 1), rank-only (w = 0) and RFT (0 < w < 1) models were calculated for each participant. The predictions of each model were scaled to the data using the following equation,

$$J_i = bM_i + a \tag{18}$$

where a is the participants lowest response, b is the size of the response range, and M_i is the model prediction for stimuli i.

Results

***W* Parameter Estimates**. I examined the between study variation in *w* parameter estimates from the Range-Frequency model. These individual *w* parameter estimates were analyzed using a Kruskal Wallis test which revealed a significant difference between studies: $\chi^2(4) = 113.17$, $p < .001$.

A Tamahane's T2 post-hoc test showed that the *w* parameter estimates were significantly lower (all p's < .05) for participants from Maltby, Wood, Vlaev, Taylor, and Brown (2012), M= .15 , compared to participants from Brown et al. (2008), M= .36, Melrose et al. (2012), M = .53, Wood, Brown, and Maltby (2011), M= .61, and Wood, Brown, et al. (2012), M= .52. Estimates for participants from Brown et al. (2008) were significantly higher than participants from Maltby et al (2012) and significantly lower than the other studies (all p's < .05).

Model Comparison. Likelihood ratio tests were used to compare the fits of the range-only, rank-only and RFT models. By varying the *w* parameter the RFT model produces rank-only ($w = 0$) and range-only ($w = 1$) predictions. In other words, the range-only and rank-only models are specific instances of the more general RFT. Consequently, the difference in negative log likelihood (-2lnL) estimates between the general and specific model approximates a χ^2 distribution with 1 degree of freedom. For one model to significantly better fit the data the difference in -2lnLs must be at least 3.85 lower ($p < .05$). Individual level parameter estimates and model fits are shown in Appendix B.

Overall RFT best fit a minority of participants. Table 5 shows the percentage of participants for whom RFT performs significantly better than either the range-only or rank-only models. There are notable between study differences in the performance of RFT. Most of the participants in the Brown et al (2008) study were significantly better fit by RFT. In contrast, less than 2% of participants were significantly better fit by RFT in the Wood et al. (2011) and Maltby et al. (2012)

studies. These cross study variations have not been examined at an individual level before.

Table 5

Percentage of participants for whom the likelihood ratio test favors RFT

Study	Vs Range	Vs Rank	Both
Wood, Brown & Maltby 2011a	69	73	43
Brown, Gardner, Oswald and Qian (2008)	100	92	92
Melrose, Brown and Wood (2012)	58	58	23
Wood, Brown and Maltby (2011b)	8	33	1
Matlby. Wood, Vlaev, Taylor and Brown (2012)	55	0	0
Overall	51	37	19

Most surprising is the poor performance of RFT overall. Only 19% of the participants across all of the studies were significantly better fit by RFT. RFT performed either equally well or worse than the range-only or rank-only model for 81% of the participants in the analysis.

Bootstrap Analysis. I ran individual level bootstrap analysis to quantify the uncertainty in the above *w* parameter estimates. Bootstrap analysis was carried out as follows. First, RFT was fit to an individual's responses. This fit of the model is the most likely given the data and provides the maximum likelihood estimates of the *w* and SD parameters. Then a bootstrap sample was created by drawing one sample for each response from a normal distribution. These samples were drawn using the `normrnd` function in MATLAB. The normal distribution had a mean equal to the RFT model prediction of a response given the best fitting *w* and SD estimates from the fit of the model to the data.

For each participant 1000 bootstrap samples and corresponding parameter estimates were generated. For each participant there were therefore 1000 *w* parameter estimates. Parameter values were considered credible if they were within the 95% confidence interval of these bootstrap *w* parameter estimates. The mean and 95% confidence intervals of these estimates for each study are shown in Figure 15.

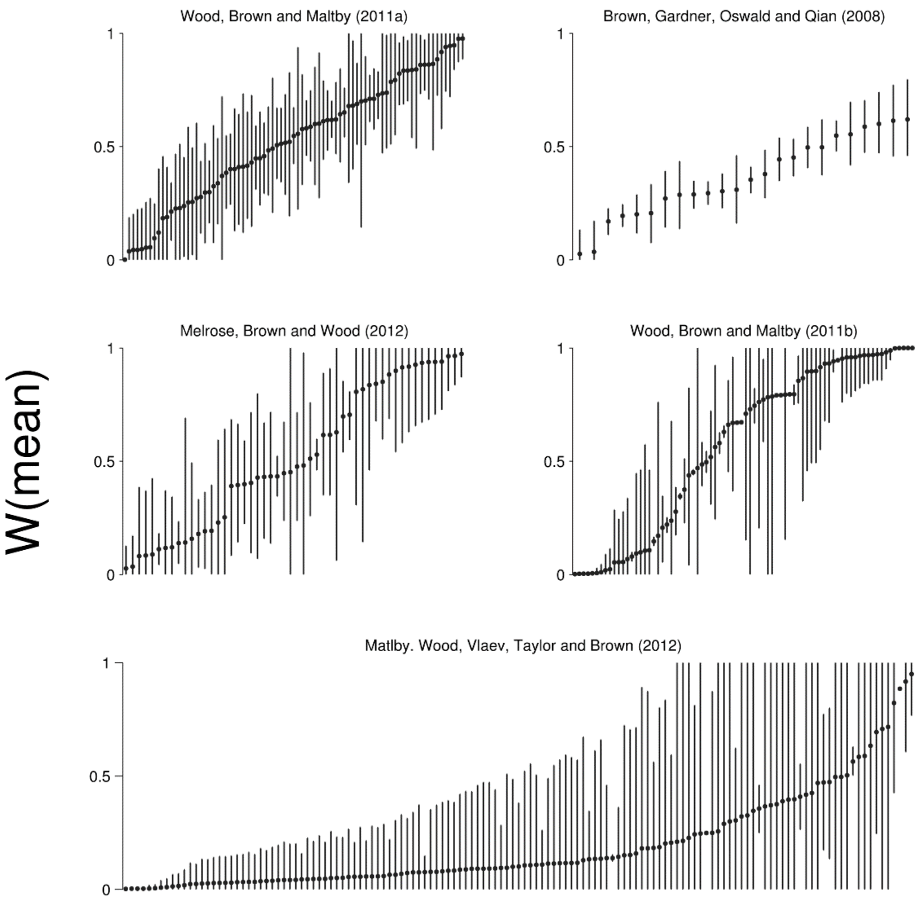

Figure 15. Credible *w* parameter estimates based on bootstrap analysis

Of particular interest are those participants for whom parameter values of exactly 0 and 1 are not within the confidence interval. For these participants, the RFT is the only credible model based on bootstrap sampling. The percentage of participants for whom RFT is the only credible model is 40% overall, notably larger than the percentage from the likelihood ratio tests. Consistent with the likelihood ratio tests reported above, the performance of RFT varies depending on the study (see Table 6). The participant data from Brown et al. (2008) and Wood, Brown, et al. (2012) offer the greatest support for RFT. The data from Maltby et al. (2012) generally support the range-only model. Once uncertainty in parameter estimates is considered RFT appears more credible, though still for a minority of participants, and considerable cross study variation remains.

Table 6

Percentage of participants for whom each model is credible based on 95% confidence interval of bootstrap parameter estimates

Study	Rank-only	Range-only	Range-Frequency
Wood, Brown & Maltby 2011a	20	22	49
Brown, Gardner, Oswald and Qian (2008)	8	0	92
Melrose, Brown and Wood (2012)	27	37	35
Wood, Brown and Maltby (2011b)	29	41	35
Matlby. Wood, Vlaev, Taylor and Brown (2012)	92	26	26
Overall	48	28	39

Discussion

The results of the first series of model comparisons found mixed support for RFT. Less than half of the participants were significantly better fit by RFT. In some studies, such as Maltby et al. (2012), far fewer participants were significantly better fit by RFT only. These findings at an individual level are quite different to the strong support of the model at group level reported in previous studies (e.g., Parducci, 1968, 1982; Parducci & Haugen, 1967). I consider this to be an important limitation of the model.

Individual estimates of the *w* parameter varied considerably within each study as predicted by memory based models of relative judgment. In the simulation above I demonstrated that individual memory performance could produce individual variation in the *w* parameter estimates across participants. One interesting possibility is that memory performance could influence extent to which responses are closer to the range-only model. If the variation is due to recall from real world distributions, as suggested by DbS, then I would expect poor recall of real world distributions to produce rank based responses. RFT does not make any specific predictions about individual variability in the *w* parameters, though some individual variation may be expected.

Cross study differences may be consistent with DbS. Previous findings from the DbS literature suggest that people sample from real-world distributions when making decisions. For example, payments are often positively skewed: Larger payments are less frequently than small payments (Stewart et al., 2006). The five studies I examined above ask participants to make judgments about very different stimuli (see Table 4). These stimuli may have different real world distributions. If people are sampling from these different prior distributions when we would expect variation based on study. Though individual difference such as working memory (number of items in working memory) may mediate the effect of these prior distributions (see Figure 14 for an example).

There is some precedent in the RFT literature of cross study variation due to methodology (Parducci & Wedell, 1986) and stimuli (Parducci, 1995). Parducci and Wedell (1986) found that *w* parameter estimates differ depending on (a) the number of response categories, and (b) the number of unique stimuli. These effects were observed using frequency manipulations in which some stimuli are more frequent than others. Parducci accounted for these effects by extending RFT to include sampling from memory. In this process based model participants search through their memory system to build a distribution of stimuli (Parducci & Wedell, 1986; Wedell & Parducci, 1985). Task limitations reduce the effectiveness of this search to produce distorted (often flat) distributions which gave RFT predictions consistent with their findings.

However, the response categories and number of stimuli were the same or similar in the studies used in the current model comparison. In both Maltby et al. (2012) and Wood et al. (2011) participants responded on a seven point ordinal scale to distributions with the same number of stimuli. Yet the *w* estimates were significantly different in both of these studies. Also, the distributions analysis above was created by manipulating the relative spacing of items in the stimuli range so the distributions were flat. The process based RFT model (Wedell & Parducci, 1985) is applied specifically to frequency manipulations and it is unclear how the model would be extended to the spacing manipulations such as those examined above. Some memory models, for example SIMPLE (Brown et al., 2007), do predict different recall performance for distributions such as those examined above (Brown & Matthews, 2011) and I investigate this combined SIMPLE + DbS model in Chapter 4.

In summary, the analysis above documents a notable limitation of RFT. For the majority of participants RFT was not the best fitting model. Instead, the range or rank only models better fit participant responses. There was considerable variation in the best fitting parameters between and within the studies. The poor performance of RFT at the individual level may reflect individual differences in sampling. Some participants may be influence by the distance of items from the judged item (as examined below) or recall items differently (see chapter 4). These variations may allow us to distinguish between RFT and other models of relative judgment. In the next section I compare the RFT with a competing account using the same data and again at the individual level.

Distance-based Sampling

Are contextual items equally weighted when people subjectively judge items? According to RFT judgments are the weighted average of the rank position and range position of the judged stimulus. The rank principle (which Parducci refers to as the frequency principle) assumes that all of the contextual items are equally weighed when a judgment is made. In contrast, the range principle predicts judgment based on the position the item relative to the highest and lowest item. For the range principle only the endpoints influence judgment: the intermediate items have no impact. So RFT predicts either equal weighting of contextual items or weighting of only the most extreme items.

However, empirical findings and theories of memory predict that distance from the items being judged will influence judgments. Recent work from the psychophysics of prices suggest that items closer or further from the judged item have more impact on judgments than the other contextual items. This distance based weighting hypothesis directly challenges the assumption in RFT that items are equally weighted or only the most extreme items influence judgment.

Here I address this tension directly by comparing RFT against the generalized exemplar model of sampling (GEMS; Qian & Brown, 2005). GEMS incorporates both RFT and the distance based sampling hypothesis as special cases of the GEMS model which allows us to directly compare RFT with and without distance based sampling.

Intuitively, we may expect distance based weighting in relative judgment. One example is judging the satisfaction of a wage. According to the frequency principle a person will weight all of the contextual items equally. In other words, the wages of everyone the person knows will contribute equally to a satisfaction judgment. However, this seems somewhat simplistic. Instead, the wages of ones colleagues may have more influence on judgment than the wage of the company director and the office intern (i.e., those wages closer to the judged wage are more heavily weighted in judgment). Other examples are the attractiveness of TV prices or happiness with school grades. In all of these cases we would intuitively expect the weight of contextual items in judgment to depend on their distance from the judged item.

Exemplar models of memory can predict similarity based sampling. According to DbS people draw upon a sample from memory when making a judgment. One possibility is that the weight of an item is proportional to the recallability of that item from memory, and items closer in magnitude to the judged stimulus are more likely to be in a recalled sample.

This task is analogous to a cued recall experiment: the participants is shown a cue (the stimuli to be judged) and attempts to recall other items (the contextual stimuli). In exemplar models of memory, such as MINERVA2 (Hintzman, 1984), PROBEX (Juslin & Persson, 2002) and SIMPLE (Brown et al., 2007), items similar to the cue are most likely to be recalled. Table 7 shows the prediction of SIMPLE in a cued recall task. I calculated the SIMPLE cued recall predictions for eight items along a single dimension. As we would expect SIMPLE predicts that items matching the cue are most likely to be recalled: these probabilities are in bold. Also, SIMPLE predicts that similar items are more likely be recalled. The table shows that recall probability decreases as distance from the cue increases.

Table 7

Similarity in cued recall as predicted by SIMPLE

	Position							
Cue	1	2	3	4	5	6	7	8
1	**.50**	.22	.12	.08	.06	.05	.05	.04
2	.17	**.36**	.17	.10	.07	.05	.05	.04
3	.09	.15	**.30**	.15	.09	.06	.05	.05
4	.06	.09	.14	**.28**	.14	.09	.06	.05
5	.05	.06	.09	.14	**.28**	.14	.09	.06
6	.05	.05	.06	.09	.15	**.30**	.15	.09
7	.04	.05	.05	.07	.10	.17	**.36**	.17
8	.04	.05	.05	.06	.08	.12	.22	**.50**

Note: shading indicates similarity.

Distance based sampling is implemented in the generalized exemplar model of sampling (GEMS; Qian & Brown, 2005). The rank-only and range-only models are a special case of RFT. Varying the *w* parameter allows the model to produce rank-only (w =0), range-only (w=1) and RFT ($0 < w < 1$) predictions. In the same way, RFT is a special case of the GEMS model. In the GEMS model,

$$J_i(x) = wR_i + (1-w)\left[0.5 + \frac{\sum_{j=1}^{i-1}\left(x_i - x_j\right)^y - \sum_{j=i+1}^{N}\left(x_i - x_j\right)^y}{2\left(\sum_{j=1}^{i-1}\left(x_i - x_j\right)^y + \sum_{j=i+1}^{N}\left(x_i - x_j\right)^y\right)}\right] \tag{19}$$

judgments are a weighted average of two components. As with RFT, the first term is the range position of the stimulus (see Equation 18), and when *w* equals 1 the model produces range-only predictions (see Equation 17).

In GEMS the γ parameter controls distance based sampling in the model. The value of the γ parameter determines the distance based weighting of items. Consider an example where a participant is judging the fourth of 11 equally spaced items. The relative weight of the other items in judgment is shown in Figure 16. In the figure darker items carry more weight in a judgment compared to lighter items. When $\gamma = 0$ all of the items carry the same weight and the model produces rank-only predictions. Increasing the γ parameter reduces the impact of nearby items. Instead, items are furthest from the judged stimulus are given the most weight. The opposite occurs when the γ is negative: Items closest to the stimulus carry the most weight. This latter case is comparable to the recall probabilities of similar items shown in Table 7.

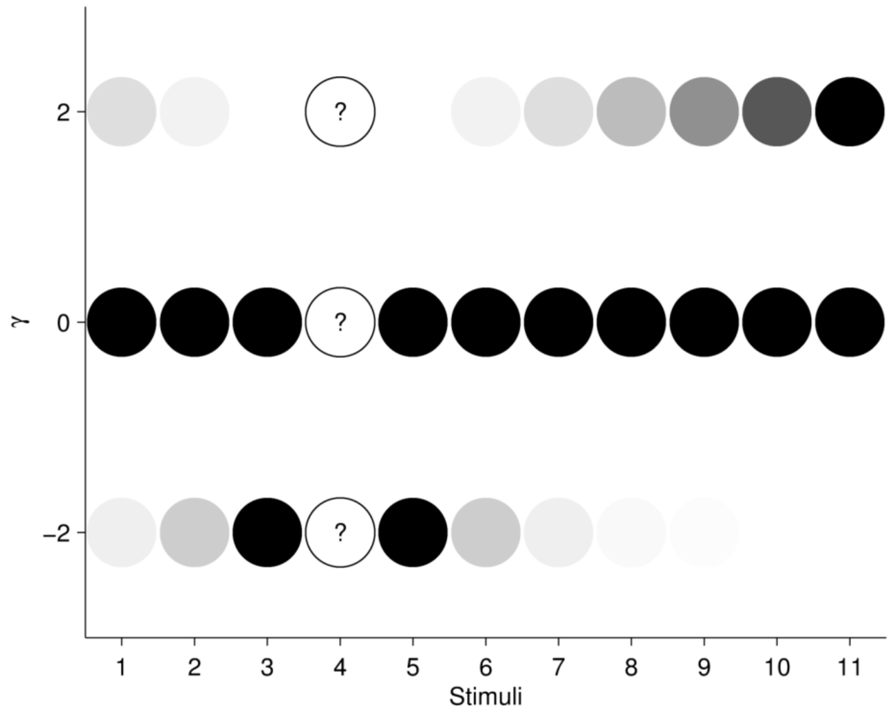

Figure 16. The relative weight given to other stimuli when judging the third stimulus for different values of γ. Note: Darker circles indicate a greater relative weight

The corresponding model predictions are shown in Figure 17. When the model predictions are rank-only (γ=0) the model predicts a straight line because the all of the items have an equal impact on judgment. Focusing on the fourth item, the predicted response is lower when γ > 0 and higher when γ <0. This is because the ratio of ratio of higher and lower items is different when the γ parameter varies due to the distance based weighting depicted in Figure 16. For instance, when γ >0 then the total weight of the items larger than the fourth item is greater than those which are smaller.

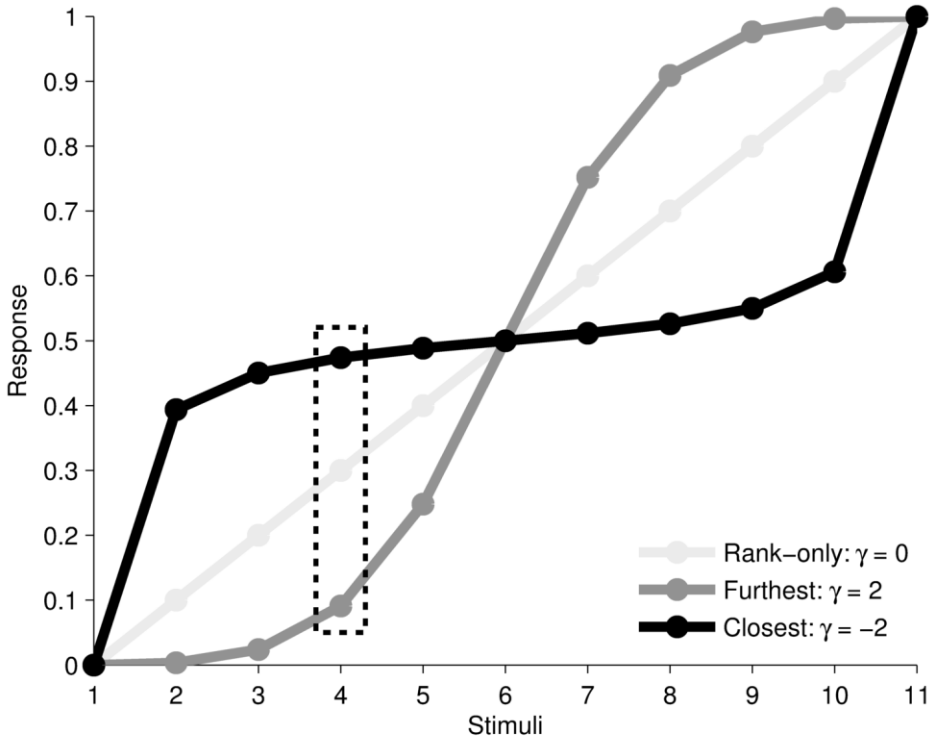

Figure 17. Predictions of relative judgments from the GEMS model based solely on the rank component (*w*=0) for varying values of the γ parameter. These predictions are for 11 equally spaced stimuli (e.g., square sizes or payments).

The distance based sampling hypothesis is implemented in the second term and controlled by the γ parameter. When γ equals 0 this component is reduces to the rank-only model (see Equation 18). Consequently, when $\gamma = 0$ and w is allowed to vary between 0 and 1 then the model produces RFT predictions.

Several studies have shown distance based sampling in relative judgment. Qian and Brown (2005) presented participants with a series of hypothetical holiday prices and then asked them to rate the attractiveness of each price. In Experiment 1, the prices were shown on labels of paper and the ten participants placed the prices along a 1 (*extremely unattractive*) to 7 (*extremely attractive*) scale. Individual and group level fits of the GEMS model found evidence for distance based sampling - the model fitted the data best when $\gamma \neq 0$ suggesting that people weight prices by their distance from the judged price. For example, some participants appeared to attend to only the items closest to the price being judged. . In Experiment 2, the prices were shown sequentially on a computer. GEMS fitting again suggested that the items closest to the judged stimulus were more heavily weighted in judgment. Group level model analysis of the data found that GEMS performed significant better than the RFT model.

In a subsequent study, Brown et al. (2008) asked participants to rate their expected satisfaction with a series of hypothetical wages. The cover story was that the values were the wage of their first job after university. Each wage was on a paper label and, like the above experiment, participants were asked to place all of the wage labels on a 1 (*least satisfied*) to 7 (*most satisfied*) scale. Brown et al reported group level model fitting comparing the GEMS and RFT models. RFT did perform better than the GEMS model. However, the γ parameter of the GEMS model did not equal 0 even though RFT is nested within the GEMS model. These results were not significant.

Here I examine the evidence supporting distance-based sampling in contextual judgment. I fit both the GEMS and RFT model to the same data used in section one. To compare the models I held γ at 0 and allowed w to vary to find the RFT prediction that best fit the data. Then the w parameter was fixed and γ was allowed to vary to contribution of distance based sampling to the model fit. The parameters were fixed to avoid possible parameter trade-off between the w and γ parameters. The focus of this analysis is to examine the extent to which the GEMS model better predicts responses over and above RFT. The models were again fit using the maximum likelihood estimation technique outlined above.

Results

Bootstrap analysis. To examine the uncertainty in the parameter estimates I carried out bootstrap analysis on the γ parameter. The parameter estimates examined above are the most likely parameter estimates given the data. But other γ parameter values may also be credible. In our case, the γ parameter value tells us if the predictions are consistent with the distance based sampling hypothesis. The hypothesis is supported when $\gamma \neq 0$.

The bootstrap analysis used the same method as was outlined in the first series of model fits. A total of 1000 bootstrap samples per participant were generated based on the GEMS model parameters which best fit an individual's data. Then the GEMS model was fit to these bootstrap samples and the resulting bootstrap parameter estimates were recorded.

From these bootstrap γ estimates I calculated the 95% confidence intervals shown in Figure 18. The qualitative pattern of the figure suggests that a γ parameter of 0 was credible for most participants. For the majority of participants a γ parameter of 0 was credible, though there is some between-study variation. Taken together, the qualitative pattern of the bootstrap analysis suggests that a pure RFT model without distance based sampling best is most credible for most of the participants. Next I compare the models directly to establish if RFT outperforms the distance based sampling model (GEMS).

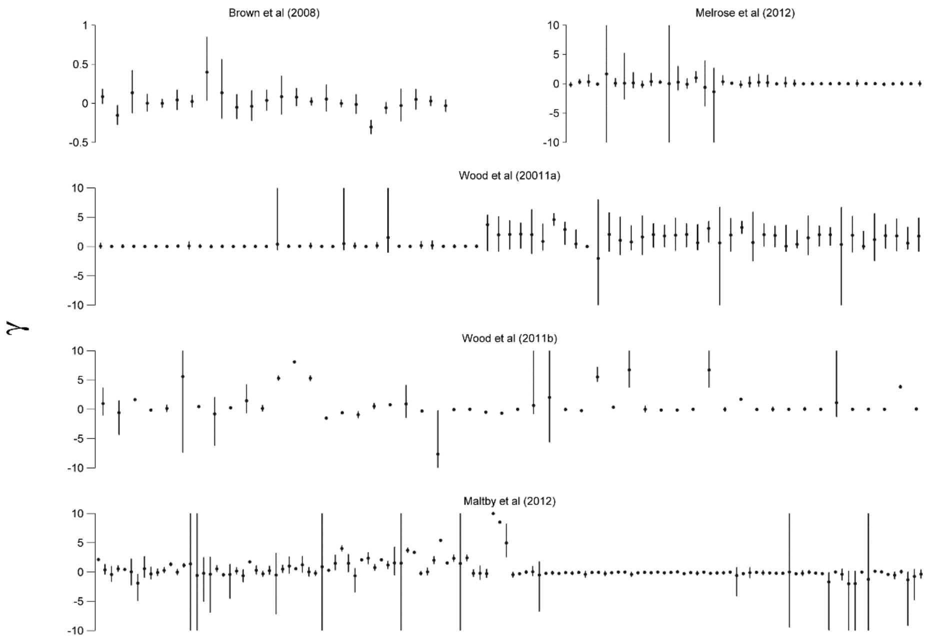

Figure 18. Credible γ parameter estimates based on bootstrap analysis

Model Comparison. In the comparison reported here the GEMS models has one more parameter, γ, than RFT. We would expect the GEMS model to better fit the data based on this additional parameter alone (for a discusion of model complexity see Myung, 2000). In the model fitting, lower -2lnL values correspond to better model fits. To penalize these fits a penalty must be added to the model fit statistic (-2lnL) that depended on the 2 parameters in RFT (*w*, SD) and 3 parameters in GEMS (*w*, γ, SD). Individual fit statistics and model comparison are shown in Appendices C and D.

I penalized the fit (-2lnL) of each model by calculating the Akaike Information Criterion (AIC) and the Bayes information Criterion (BIC). Both criteria penalize the fit of each model based on the number of parameters. The AIC is calculated as

$$AIC = -2lnL + 2k \tag{20}$$

where k is the number of model parameters. When applied to our model comparison the AIC penalty is 4 for the RFT model and 6 for GEMS. In the studies reported here the number of data points is low (11) per person. Burnham and Anderson (2002) recommend using a modification of the AIC for small samples,

$$AIC_c = -2lnL + 2k\left(\frac{N}{N-K-1}\right) \tag{21}$$

where N is the number of data points and k is the number of parameters. In our case, AIC_c increases the -2lnL by 5.5 for RFT and 8.25 for GEMS. The BIC is an alternative to the AIC first proposed by G. Schwarz (1978) motivated by Bayesian theory. The BIC is calculated as,

$$BIC = -2lnL + k\ln N \tag{22}$$

where k is the number of parameters and N is the number of data points. BIC penalizes RFT by 9.59 and GEMS by 14.39. Of these three information criteria the BIC most harshly penalizes model complexity. I calculated the AIC, AIC_c and BIC for each participant individually.

First, I compared the models by calculating the percentage of the participants for whom GEMS better fit the data. Burnham and Anderson (2002) suggest that a difference in the criterion of greater than 3 suggests that the data favor one model over the other. In the model comparison I fit RFT to the data and then GEMS, allowing γ to vary. Given the present method the GEMS model could only do better than or as well as the RFT model.

I counted up the number of participants for whom the -2lnL, AIC, AICc and BIC were more than 3 points lower for the GEMS model as compared to the RFT model. The percentage of participants significantly better fit by GEMS is shown in Table 8.

Table 8

Percentage of participants better fit by the GEMS model (criterion difference > 3)

Paper	-2lnL	AIC	AICc	BIC	N
Wood, Brown and Maltby (2011a)	4	2	2	1	81
Brown, Gardner, Oswald and Qian (2008)	13	8	8	4	24
Melrose, Brown and Wood (2012)	4	2	2	0	52
Wood, Brown and Maltby (2011b)	7	5	3	3	86
Maltby, Wood, Vlaev, Taylor and Brown (2012)	23	15	15	14	135
% of total sample	12	8	7	6	378

As expected fewer people were significantly better fit by the model as the criterion more harshly penalized the GEMS model for complexity. Overall, less than 10% of the participants were significantly better fit by the GEMS model once the criteria were applied. GEMS did perform best when fit to the data from Maltby et al. (2012), but this was for 15% of the participants. The analysis based on the difference in criteria do not support the distance based weighting hypothesis.

Next, I calculated the relative posterior probability of the models given the data. The differences in BIC estimates for the two models (ΔBIC) can be interpreted as the log of the Bayes factor for these two models (Lewandowsky & Farrell, 2010). In Bayesian statistics the Bayes factor,

$$B = \frac{p(M_1|y)}{p(M_2|y)} \tag{23}$$

is the ratio of the probabilities of each model given the data. We can calculate the Bayes factor from the BIC differences as:

$$B = \exp\left(-\frac{1}{2}\Delta BIC\right). \tag{24}$$

This Bayes factor can then be used to calculate the Bayesian model weights,

$$BIC_w = \frac{exp\left(-\frac{1}{2}\Delta BIC_M\right)}{\sum_i exp\left(-\frac{1}{2}\Delta BIC_i\right)}. \tag{25}$$

where BIC_w is the Bayesian model weight, ΔBIC_M is the difference in BIC between model M and the best model, and BIC_i is an array of differences between the BIC of each model and the best model. The BIC_w can be interpreted as the posterior probability of a given model assuming that RFT and GEMS are the only models which can explain the data.

A similar approach can be taken with AIC differences. The AIC and BIC differ in their theoretical underpinnings so the AIC_w cannot be interpreted in terms of posterior probabilities. However, Burnham and Anderson (2002) suggest that AIC_w can be interpreted as giving the weight of evidence for the a model given the data. The calculations for AIC_w are the same as Equation 27 with ΔAIC replacing ΔBIC.

To compare the models, I calculated the AIC_w and BIC_w for each participant. The average AIC_w and BIC_w for participants in each study is shown in Table 9. In all of the studies the criterion weights favor RFT over GEMS. Consistent with the above analysis the data do not support the distance based weighting hypothesis and offer unanimous support for RFT.

Table 9

Average criterion weights from individual level model fits

	AICw		BICw	
Paper	GEMS	RFT	GEMS	RFT
Wood, Brown and Maltby (2011a)	.34	**.66**	.15	**.85**
Brown, Gardner, Oswald and Qian (2008)	.41	**.59**	.23	**.77**
Melrose, Brown and Wood (2012)	.32	**.68**	.14	**.86**
Wood, Brown and Maltby (2011b)	.34	**.66**	.25	**.75**
Maltby, Wood, Vlaev, Taylor and Brown (2012)	.46	**.54**	.33	**.67**

Discussion

In this second series of model fits the data do not support the distance based weighting hypothesis. The fits of the GEMS model were consistent with the heavier weighting of items furthest from the judged item. However, bootstrap analysis found that a purely RFT account was credible for many of the participants. Direct model comparison of RFT with and without distance based sampling strongly supports the RFT model. Based on these findings, I conclude that distance based sampling does not influence judgments. Instead, RFT offered a better account of the data.

General Discussion

In this chapter I investigated at an individual level (a) the performance of RFT when compared to the range-only and rank-only models, and (b) the evidence for distance based weighting in subjective judgments. The findings from the first series of model comparisons show that RFT performs better than the range-only and rank-only models for a minority of participants, and the performance of the model differs considerably across studies. The second series of model comparisons found little evidence of distance based sampling in judgment. When compared with the GEMS model the data strongly favored RFT across all of the studies.

This chapter presented an investigation of RFT at an individual level. A large body of work has shown that RFT accurately predicts the average responses of participants in subjective judgment tasks (Parducci, 1995). One might assume that RFT would also fit individual level responses equally well. However, RFT did not fit all of the participant data when compared with the range-only and rank-only models. If responses of individual participants are generally not as predicted by RFT, then this represents a considerable limitation of the model.

These individual differences may have theoretical applications. Individual level analysis requires a given model to fit many more data points compared to group level analysis of average responses. Consequently, a given model needs to account for a richer set of data when analysis is at individual level. The poor performance of RFT at the individual level (as above) leads one to wonder if other models can better capture the response patterns in this data.

A variety of models predict subjective judgments based on previous responses. For example, a variety of models predict judgments based on an internal referent based on the responses to a previous trial (for reviews see Laming, 1997; Petzold & Haubensak, 2004; Sarris, 2004). Unfortunately the sequence of judgments was not recorded in the above data set so it is unclear how these models can be evaluated using the above data.

Other models predict responses based on recollection of items from memory. According to the multiple-standard model responses are a combination of previous responses and the recollection of extreme values from the memory system (Petzold & Haubensak, 2004). Similarly, the consistency model predicts that participants try to be consistent with previous stimulus-response pairs drawn from memory (Haubensak, 1992).

In the data analyzed above the participants were shown the stimuli simultaneously and were able to see their previous judgments throughout the experiment. Given this methodology it is unclear how these memory based models could be applied to this data set. However, models could apply to the current data if recallability is a function of the local distinctiveness of a given item within a stimuli set (Brown & Matthews, 2011). I apply the SIMPLE model to these data in the next chapter.

Chapter 4

Decision by Sampling and SIMPLE

Introduction to Chapter

In the previous chapter I demonstrated that range-frequency theory (Parducci, 1965) can predict relative judgments using individual level data from five previous studies. The effect of context on judgment could not be attributed to distance based weighting of contextual items – an effect predicted by exemplar models of memory.

In this chapter I investigate another type of memory effect in contextual judgment. According to the SIMPLE model of memory (Brown et al., 2007) the relative spacing of items influences recall: Items which are more distinct along a dimension of interest are more likely to be recalled. Previous work suggests that Decision by Sampling (Stewart et al., 2006) can predict range effects if the probability of a contextual item entering a memory sample is derived from SIMPLE (Brown & Matthews, 2011). I investigate this interesting possibility and whether the range component is necessary at all in this combined model.

Abstract

It is well established that subjective judgments of the magnitudes of stimuli are strongly influenced by the magnitudes of other stimuli in the context of judgment. According to one recent model of context-based judgment, decision by sampling, the subjective magnitude of an item is given by its relative ranked position within the comparison context. According to the longer-established range-frequency theory, in contrast, the subjective magnitude of an item is given partly by its relative ranked position within the comparison context and is also influenced by the "range position" of the item – i.e., by where it stands in relation to the highest and lowest contextual values. Here I examined whether the rank-only process assumed by decision by sampling can account for effects that have previously been attributed to range. It was shown by model comparison based on the results of five previously published experiments that when the relative distinctiveness of contextual items is taken into account, as predicted by an independently-motivated distinctiveness model of memory (SIMPLE), a rank-only process can account for apparent range effects. It was concluded that a rank-only model can account for data that have hitherto been taken to implicate range-based processes.

Subjective magnitude judgments are heavily influenced by the relevant context of comparison. Thus the judged magnitude of a simple psychophysical stimulus, such as a square, brightness, or number, is influenced by the range and distribution of other to-be-judged stimuli (Parducci, 1963, 1995; Parducci et al., 1960; Wedell & Parducci, 1985). Contextual effects are also seen in judgments of everyday stimuli. Here I investigate the processes underlying such contextual effects and, in particular, attempt to adjudicate between two widely-used models: the decision by sampling model (DbS; Stewart et al., 2006) and range frequency theory (RFT; e.g., Parducci, 1963, 1965, 1995).

According to DbS, judgments of the magnitudes of stimuli within a comparison context depend only on the relative ranked position of the stimulus within the remembered or experienced context of judgment. According to RFT, in contrast, the position of a stimulus with respect to the highest and lowest stimuli in the context (its range position) also matters. Thus the models differ in terms of whether they predict that the *distribution* of stimuli within a fixed range will affect judgments of individual items within the context. A rank-only model such as DbS must, by definition, predict no effect of distribution per se - if a stimulus is (e.g.) the 3rd largest in a context, that is all that can matter; the distribution of other stimuli can have no effect and the mean judgment of a set of contextual stimuli must always be .5 (on a 0-1 scale). However, a number of studies — reviewed below — have found effects of skew, such that the mean judgment of a set of negatively skewed stimuli, in which most of the items are clustered near the upper end of the range, is higher than the mean judgment of a set of positively skewed stimuli, in which most of the items are clustered near the lower end of the range (e.g., Parducci, 1968). Effects of skew are well described by RFT, suggesting that a rank-only account may be inadequate. Here I examine whether a purely rank-based approach can account for apparent range effects when the relative distinctiveness of contextual items, as independently determined by a memory model (Brown et al., 2007), is taken into account.

The structure of this paper is as follows. First, I review models of relative judgment. I motivate and then introduce the combined SIMPLE and decision by sampling model (SDbS; Brown & Matthews, 2011) of judgment. Then I report results from two model comparisons. The first directly compares the performance of RFT and SDbS. The second compares SDbS with and without a range component.

Decision by Sampling

According to DbS, items are evaluated based on the number of smaller and larger items in a sample of items present in working memory (WM) memory at the time of judgment. Samples can be drawn from a combination of previously experienced items and experimental stimuli. Each to-be-judged item is compared, ordinally, with each sample item present in WM. Consider the problem of determining the expensiveness of a cup of coffee costing £1.50. According to DbS, one might call to mind two occasions on which a lower price is or has been charged for a cup of similar-quality coffee (N_{lower} = 2), and three higher prices (N_{higher} = 3). That is, the present price is the *i*'th most expensive out of *n*, where *i* = 3 and *n* = 6. The resulting estimate of the expensiveness of the present price, M_i, is according to DbS simply:

$$M_i = \frac{i-1}{n-1} = \frac{N_{lower}}{N_{lower} + N_{higher}} = 0.4 \quad (26)$$

DbS has been successfully applied both to key economic phenomena (Stewart, 2009; Stewart et al., 2006; Stewart & Simpson, 2008; Ungemach et al., 2011) and to real life social judgments (Maltby et al., 2012; Melrose et al., 2012; Wood et al., 2011).

However, as noted above DbS appears unable to predict a key finding in the relative judgment literature. Findings from psychophysics show that subjective judgments appear to be influenced by the distance of the judged item from the most extreme items in the immediate environment. In a seminal study, Parducci (1965) asked people to rate the 'largeness' of a series of squares. Each participant was shown the all of the squares they would rate. The responses of the participants were best fit by RFT, according to which judgments are a weighted average of the relative rank of the item (similar to DbS) *and* the distance of the item from the highest and lowest items in the experiment. The relative rank of stimuli was unable to capture the qualitative pattern of the data. Numerous studies have found range effects in subjective judgment (e.g., Smith et al., 1989; Wedell & Parducci, 1985; Wedell et al., 1987; Wedell et al., 1989).

Range-frequency Theory

According to RFT judgments are a weighted average of quantities representing the *range* and *frequency* principles. According to the frequency principle,

$$F_i = (r_i - 1)/(N - 1) \tag{27}$$

the judgment of stimuli i is the rank of the item within the contextual items, r_i , divided by the number of context items, N. The prediction of the frequency principle depends on the number of items in the immediate context which are higher and lower on the objective dimension of interest. For example, the frequency principle predicts that the subjective "tallness" of a building will be high if most of the buildings are comparatively smaller. The predictions of the frequency principle are the same as those of DbS and are typically attributed to participants trying to use each response category with equal frequency throughout an experiment.

According to the range principle, judgments also depend on the position of the judged item relative to the lowest and highest item,

$$R_i = \frac{S_i - S_{min}}{S_{max} - S_{min}} \quad \textbf{(28)}$$

where S_{min} is the lowest item and S_{max} is the highest item. The range principle predicts that the same stimulus will be judged as higher on the subjective scale if it is close to the highest items and far from the lowest item.

The final judgment is assumed by RFT to be a weighted compromise of the range and frequency principles:

$$RFT_i = wR_i + (1 - w)F_i \quad \textbf{(29)}$$

where RFT_i depends on the value of a weighting parameter, *w*. As the *w* parameter approaches 1 the predictions of the model become closer to those of the range principle alone.

Later formulations of RFT incorporated memory processes into the model of judgment formation. Parducci and Wedell (1986) found greater context effects on judgment when there were more stimuli or fewer categories. Wedell and Parducci (1985) presented a model predicting such effects based on an incomplete recollection of the stimuli distribution. Parducci and Wedell (1986) extended the model and showed it could predict responses to spacing-based manipulations of the stimuli distribution (see below). These accounts retain the range and frequency principle from the original theory. However, the predictions are different because the recalled distribution differed based on a sample drawn from memory. These samples are drawn solely from the stimuli presented in an experiment.

RFT has received empirical support from a wide range of domains (e.g., Niedrich et al., 2001; Parducci, 1965, 1968; Pettibone & Wedell, 2007; Riskey et al., 1979; Smith et al., 1989; Wedell & Parducci, 1985; Wedell & Parducci, 1988; Wedell et al., 1989; Wedell et al., 2005). Here I focus on range effects, as they are key to discriminating between RFT and DbS.

Range effects are widely reported in the relative judgment literature. Volkmann (1951) first suggested that the range position of an item within a context may influence social judgments. Subsequent work by Parducci and colleagues show that a combination of the frequency and range principles predicts responses better than relative rank alone (Parducci, 1965, 1968, 1982; Parducci et al., 1960; Parducci & Wedell, 1986). Wedell and colleague showed similar effects of range position in social judgments (Pettibone & Wedell, 2007; Wedell et al., 1989). In these studies the stimulus distributions were skewed by changing the frequency of different stimuli.

Range effects are also present in overall judgments of a series of stimuli. According to RFT the average judgment in a negatively skewed distribution will be higher than a positively skewed distribution because most of them are in the upper portion of the range (Parducci, 1968, 1995). In studies examining satisfaction and happiness judgments the skew of the distribution does influence overall and mean judgments (Parducci, 1968; Smith et al., 1989). The only difference between the positively and negatively skewed distributions in these studies was the frequency of the items. The number of items was the same. If the number of stimuli is the same then the frequency principle predicts the same average judgments. Consequently, these skew effects on overall judgments must be attributed to range effects.

Relatively recent studies show range effects by instead manipulating the spacing between the stimuli along the dimension of interest. In these studies each stimuli is presented once. For example, an equally spaced distribution can be skewed by moving most items close together in the upper portion of the range. These manipulations produce individual level range effects in health (Melrose et al., 2012), payment (Brown et al., 2008) and social (Maltby et al., 2012; Wood et al., 2011) judgments. All of these range effects appear inconsistent with a pure relative rank principle as assumed by the frequency principle and DbS.

Is the inability of DbS to predict these ubiquitous range effects a serious limitation of the model? Next, following Brown and Matthews (2011), I report two series of model comparisons to determine whether a model resulting from the combination of a memory model (Brown et al, 2007) with the rank-only DbS model can predict range effects. In the first I compare the combined model with the RFT. Based on a preliminary group level analysis reported by Brown and Matthews (2011) on the data of just one experiment, we expect a combined model to fit the data as well as RFT at a group level. Here I report individual level analysis of data from a series of experiments. In the second I compare the combined SDbS model with and without the range principle from RFT. If adding a range component to the model does not improve the model fit then this would suggest that the combined model best predicts the range effects in the data.

SIMPLE and Decision by Sampling

The initial application of DbS to economic decision making assumed that sampling from memory is stochastic and unbiased (Stewart et al., 2006; Stewart &

Simpson, 2008). They acknowledged that this was a simplifying assumption which is most likely wrong. Instead, sampling from memory is often biased. For example, in free recall items presented first and last in a series are most likely to be recalled (e.g., Murdock, 1962). DbS is unable to predict range effects if the sample drawn from memory is completely unbiased and is comprised of experimental stimuli. However, if the probability of recalling items is calculated using a formal model of memory then DbS may be able to predict range effects.

Brown and Matthews (2011) derived the probability of an item entering a memory sample from SIMPLE (Brown et al., 2007). In Brown and Matthews (2011) the predictions were based on the magnitude each stimulus along a dimension of interest (wage value). According to the SIMPLE model of memory the recall of an item from memory is a discrimination task. Stimuli which are most unlike other stimuli along a dimension of interest are more likely to be recalled or attended to. The model can be applied to any dimension of interest, such as time (as in the case of primacy and recency effects) or stimulus magnitude.

In SIMPLE the confusability of items in psychological space is an exponential function of the distance between them,

$$\eta_{i,j} = e^{-cd_{i,j}} \tag{30}$$

where $d_{i,j}$ is the distance between stimuli i and j along a dimension of interest, and c is a free parameter. The discriminability of an item, D_i, is assumed to be inversely proportional to its summed similarity to all of the other items in the stimuli set,

$$D_i = \frac{1}{\sum_{k=1}^{n}\left(\eta_{i,k}\right)} \tag{31}$$

where n is the number of comparison stimuli. The probability of recalling an item from memory is the discriminability of an item after considering possible omissions,

$$P_i = \frac{1}{1 + e^{-s(D_i - t)}} \tag{32}$$

where P_i is the probability of recalling stimulus i given its discriminability D_i and the free s and t parameters. The t parameter is the threshold and s is the slope of the transforming function.

As noted above, range effects occur in distributions where each item is shown once and the relative spacing of items in the distribution is manipulated. Consider the three panels in Figure 19. Each circle represents a single stimulus along a dimension of interest. In panel (a) the stimuli are equidistant. Both the range and frequency principles predict the same response (as shown above and below the stimuli). In panel (b) most of the stimuli are grouped together in the upper portion of the stimuli range: The distribution is negatively skewed. The frequency based predictions in panels (a) and b are the same because the number of stimuli has not changed. However, the range principle predicts higher ratings for the items because their positions in the stimulus range have changed. DbS with unbiased sampling from memory produces frequency only (relative rank) predictions.

However, according to SIMPLE some of these items are more likely than others to be present in WM and hence contribute to the context of judgment. In SIMPLE, recall is akin to a discrimination task. Changing the distance between the items in the stimuli range alters their discriminability. The item in the middle of the items that are grouped together in Figure 19 is less distinct than are two lower value stimuli. The color of the circles in panel *c* corresponds to the recall probability of these items calculated using SIMPLE (*c*= 0.22, *s*=5.67, *t*=0.91).

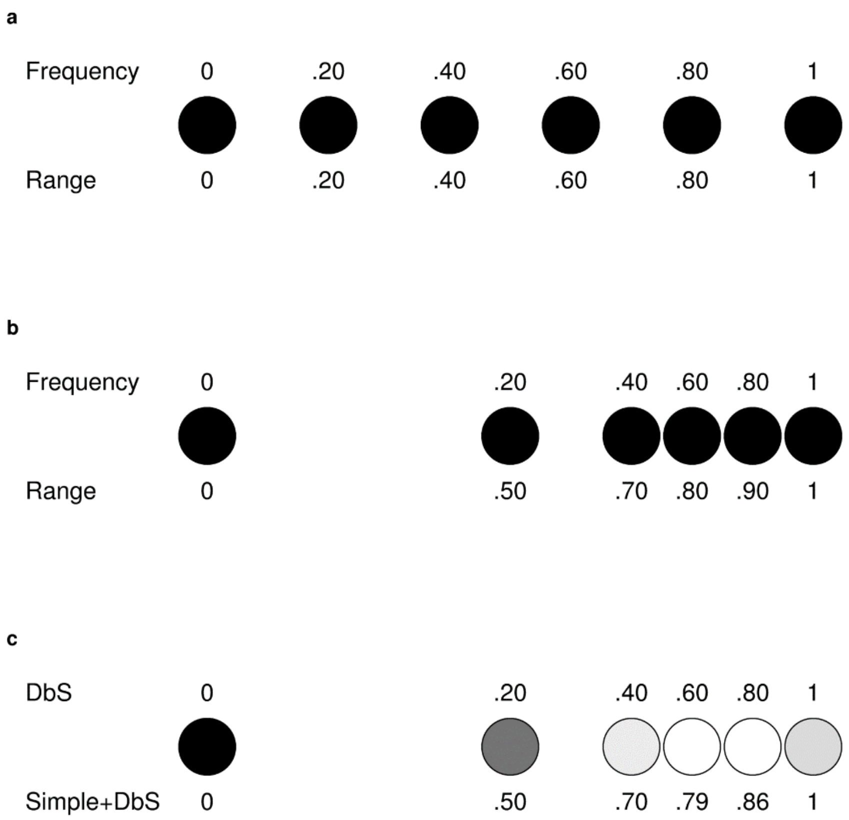

Figure 19. The predictions of the range and frequency principles when the distribution of stimuli is (a) equally spaced and (b) negatively skewed, and (c) decision by sampling predictions for the same distributions

We can calculate the SIMPLE and DbS (SDbS) prediction by summing the probabilities of items higher and lower than the item being judged,

$$SDbS_i = \frac{\sum_{j=1}^{i-1} p_j}{\sum_{j=1}^{i-1} p_j + \sum_{j=i+1}^{N} p_j} \tag{33}$$

where p is the probability of recalling each item. The bottom values in panel c show the predicted responses from the combined SDbS model. In this case the SDbS model can produce predictions which are similar to those made by the range principle.

Here I examine the ability of this combined SIMPLE and DbS model to predict range effects at an individual level. First, the combined model is compared to RFT. Then I extend the combined model to include the range principle and evaluate the additional contribution of the range component.

Range-frequency Theory and SDbS

Here I compare the combined SIMPLE and DbS model (SDbS) with RFT. Both models are fit using a maximum likelihood method (Myung, 2003).

Method

Previous Data. I fit data from five previously published studies. In each study participants made subjective judgments of stimuli which varied along a single dimension (e.g., monetary value). When making each judgment the participant was aware of the entire distribution (see Table 4 for the distribution values and participant numbers) and made a subjective judgment for every stimulus. For some studies (e.g., Brown et al, 2008) the participants were shown a series of different distributions one after another. For other studies the participant gave multiple responses for each distribution. Next I consider the method of each study in details and outline the key variables. In each description the number of distribution and the number of responses is outlined (also see Table 4).

Brown, Gardner, Oswald and Qian (2008). Participants were asked to rate a series of hypothetical wages. They were told that these wage values were prospective wages for their first job after finishing their university degree. Each participant was shown six wage distributions. The distributions were shown one at a time and the wage values for a given distribution were shown simultaneously. Wages were written on labels and the participant's task was to place the label for each payment on a 1 to 7 satisfaction scale.

In the model fitting reported here I kept the parameters of the model constant for all of a participant's responses. Each participant responded to all 11 wages in each of the six wage distributions (a total of 66 responses).

Melrose, Brown and Wood (2012). Participants were asked to give the probability of a statement describing a symptom from the diagnostic statistics manual being true. They were shown a series of 11 values. Each value represented a different person and was the number of days per month that person had experienced a symptom. Participants were asked: "For each person, please indicate in the form of a percentage how likely you think the following statement is to be true where 100% is certain to be true".

This task was repeated for six symptoms. The symptoms were feeling (1) depressed, sad, blue, tearful (2) tired and having no energy, and (3) worthless or excessively guilty about things you have or have not done, or anxiety, and experiencing (1) excessive anxiety about a number of events or activities (2) irritability, and (3) muscle tension. Each participant gave 66 responses. Model parameters were held constant for each participant across all six symptoms.

Wood, Brown and Maltby (2012a). Participants were asked to state their perception of the risk (%) of someone developing an alcohol related disorder when consuming a given number of alcohol units for 20 years. The participants were shown either a unimodal or bimodal distribution of alcohol units on a piece of paper.

This task was repeated eight times. The alcohol disorders were (a) any alcohol-related illness, (b) any serious psychological difficulties as a direct result of alcohol use, (c) a dependency on alcohol and (d) cirrhosis of the liver. Participants responded to each disorder for males and females. Model parameters were held constant across all eight responses. Each participant gave 44 responses.

Wood, Brown and Maltby (2012b). Participants read two vignettes in which 11 hypothetical people volunteered different amounts either time or money to help the participant. Participant then rated their gratitude for each value of time/money from 1 (*not at all grateful*) to 10 (*extremely grateful*). Model parameters were held constant across the 22 responses from each participant.

Maltby, Wood, Vlaev, Taylor and Brown (2012). Participants were shown either a bimodal or unimodal distribution of 11 exercise durations on a piece of paper. The participants rated on a 1 to 10 scale: (1) how healthy it was for the individual, (2) the extent of the health benefits, (3) how sufficient it would be to avoid the serious health problems associated with low levels of exercise. Participants gave 33 responses. As with previous fits, the model parameters were held constant across a participant's responses to all three questions.

Model Fitting. For both models the fit was calculated as the probability of the data using a normal distribution with a mean equal to a model prediction and a standard deviation which was allowed to vary freely. The model parameters and standard deviation were held constant for each participant.

The model parameters were varied using the `fminsearch` algorithm in MATLAB. In RFT the w parameter was allowed to vary between 0 and 1. For the SDbS model the c and s parameters were allowed to vary freely and the t parameter was restricted to between 0 and 1, as specified by the SIMPLE model[3].

---

[3] Initial fits found that the best fitting t parameter was often close to 1. This was quite unlike previous model fits of the SIMPLE model and a prior distribution was used to penalize such high t values (see Appendix A for more information).

The predictions of each model were scaled to the response range of the participant. The model predictions were between 0 and 1, and the following equation was used to scale the predictions to the response range of the participant,

$$J_i = bM_i + a \quad (34)$$

where a is the participant's lowest response, b is the size of the response range, and M_i is the model prediction for stimuli i.

Results

Model Parameters. Both models were fit to the data from all five previous studies. The average best fitting model parameters are shown in Table 10. The *w* parameter fits are the same as those presented in Chapter 3 and the SIMPLE parameter estimates are consistent with previous SIMPLE parameter estimates. Individual fit statistics are in Appendix E and model comparisons are in Appendix F.

Table 10

Means of the best fitting model parameters across participants

	RFT		SDbS			
Paper	Noise	*w*	Noise	*c*	*s*	*t*
Wood, Brown and Maltby (2011a)	8.45	.52	8.38	0.83	19.02	.60
Brown, Gardner, Oswald and Qian (2008)	0.36	.36	0.35	2.54	15.88	.62
Melrose, Brown and Wood (2012)	8.02	.53	7.92	0.65	24.70	.62
Wood, Brown and Maltby (2011b)	0.20	.59	0.21	0.97	34.69	.56
Matlby. Wood, Vlaev, Taylor and Brown (2012)	1.11	.12	1.05	4.68	20.42	.50

Log likelihood comparison. To investigate the fit of the model to the date I examined the difference in -2lnL between the models. As suggested by Burnham and

Anderson (2002) I considered a model to better fit the data if the -2lnL was 3 lower than the other model.

The percentage of participants for whom either RFT, SDbS or neither model better fit the data is shown in Table 11. For most of the participants there was no significant difference between the -2lnLs for the models. For most of the remaining participants SDbS was more probable than RFT (see third column of Table 11). This pattern is consistent with the group level model fit reported by Brown and Matthews (2011) and suggests that the DbS model can predict the range effects observed in individual level data.

Table 11

Summary of model fits based on -2lnL

	% of participants			
Paper	RFT	SDbS	Neither	N
Wood, Brown and Maltby (2011a)	2	20	**78**	81
Brown, Gardner, Oswald and Qian (2008)	21	38	**42**	24
Melrose, Brown and Wood (2012)	0	21	**79**	52
Wood, Brown and Maltby (2011b)	0	12	**88**	86
Maltby, Wood, Vlaev, Taylor and Brown (2012)	2	3	**95**	135

Model Complexity. The analysis presented above does not consider the complexity of the models being compared. Model complexity is an important topic in model selection (for a review see Myung, 2000). Ideally, when faced with multiple models which fit data equally well one would follow the principles of Occam's razor and choose the simplest model.

In our case the combined SDbS model is the most complex. RFT has two free parameters: the standard deviation of the likelihood function and the *w* parameter.

SDbS has four free parameters: the standard deviation of the likelihood function and *c*, *s* and *t*. Consequently, SDbS is the more complex based purely on the number of free parameters.

To penalize the model fits for model complexity I calculated the AIC, AICc and BIC for each participant (see Equations 22, 23, and 24 detailing each information criterion). I then counted up the number of participants for whom the combined model outperformed RFT after the complexity penalties had been applied. As shown in Table 12 a minority of participants were best fit by the combined model after the complexity penalties were applied.

Table 12

Percentage of participants for whom the penalized fits are at least 3 points lower for the SDbS model

	% of participants			
Paper	AIC	AICc	BIC	N
Wood, Brown and Maltby (2011a)	7	7	4	81
Brown, Gardner, Oswald and Qian (2008)	17	17	0	24
Melrose, Brown and Wood (2012)	12	12	4	52
Wood, Brown and Maltby (2011b)	7	3	3	86
Maltby, Wood, Vlaev, Taylor and Brown (2012)	1	1	1	135

However, what is the relative support for each model across *all* participants once complexity is incorporated into the model fitting? The AIC and BIC weights were calculated for each participant. The BIC weight can be interpreted as the posterior probability of each model if we assume that RFT and SDbS are the only candidate models. The AIC weight can be interpreted as the relative evidence for each model. The calculation and interpretation of these weights is described in more detail in Chapter 3 (see Equation 27). These weights are based on the performance of each model after being penalized for complexity for all participants and are between 0 and 1. The support for the model increases as the weight approaches 1. The average criterion weights are shown in Table 13 and suggest that the data support RFT.

Table 13

Average criterion weights for RFT and SDbS

	AICw		BICw	
Paper	RFT	SDbS	RFT	SDbS
Wood, Brown and Maltby (2011a)	**.73**	.27	**.89**	.11
Brown, Gardner, Oswald and Qian (2008)	**.63**	.37	**.92**	.08
Melrose, Brown and Wood (2012)	**.72**	.28	**.90**	.10
Wood, Brown and Maltby (2011b)	**.77**	.23	**.87**	.13
Maltby, Wood, Vlaev, Taylor and Brown (2012)	**.86**	.14	**.95**	.05

Discussion

The results of the first series of model comparisons show that the combined SIMPLE and DbS model can predict range effects at an individual level. For most participants both RFT and SDbS were equally likely given the data. This findings confirms that RFT and SDbS can produce the same predictions as previously reported by Brown and Matthews (2011). For many participants the combined model was *more* likely than RFT.

One interpretation of these findings is that the RFT is the best model based on the individual level analysis of the data. However, the complexity penalties ignore the theoretical scope of the models under consideration. One may consider the ability of a model to fit the data *and* the theoretical contribution of a model. In our case the -2lnLs reported above show that the predictions of the two models are indistinguishable when fit to the data from the majority of participants. The combined SIMPLE and DbS model offers a unifying account of decision making and judgment phenomena in social, psychophysical and economic domains using rudimentary cognitive processes. Ifboth models perform equally well and SDbS can account for a wider selection of phenomena then one may argue that SDbS should be preferred. In the next section, I adopt an alternative approach.

SDbS and the Range Principle

In the combined SIMPLE and DbS model (SDbS) judgments are purely based on the relative rank of an item within a sample drawn from memory. The absence of a range based component is inconsistent with other memory based models of relative judgment such as the consistency model (Haubensak, 1992) and the multiple standards model (Petzold & Haubensak, 2001, 2004). In these models the range position of an item is a major component of judgment.

In this first series of model comparisons the RFT and SDbS models were equally probable given the data. One interpretation of these findings is that SDbS is able to predict range effects and a range component is unnecessary in the SDbS model. However, it may be the case that there are range effects in the individual level data which are not accounted for in the combined model.

In a second series of model comparisons I examined the contribution of a range component in the combined model. It may be the case that judgments are a combination of the relative rank of an item in a memory sample *and* the position of the item relative to the highly discriminable range endpoints. I address this possibility directly by comparing the fit of SDbS with and without the range principle from RFT,

$$J_i = wR_i + (1 - w)SDbS_i \tag{35}$$

where judgments are the weighted average of the range principle and SDbS. For ease of reference I refer to this model as SDbS + Range. As the *w* parameter increases the range component is more heavily weighted. If SDbS captures all of the range effects in the data then SDbS should perform equally well when compared to SDbS+Range.

Method

To examine the contribution of the range principle I manipulated the *w* parameter. First, the *w* parameter was fixed to 0 which produced SDbS only predictions. Then the *w* parameter was allowed to vary freely between 0 and 1.When w>0 the model predictions are a combination of SDbS and the Range principle.

Results and Discussion

***W* Parameter Estimates**. Do the *w* parameter estimates differ between studies? A Kruskal Wallis test revealed that the *w* parameter estimates differed significantly between studies, $\chi^2(4) = 64.89$, $p < .001$. The results of post-hoc Tamhane T2 analysis are shown in Table 14. Notably, the *w* parameter estimates appear lower than those reported in Chapter 3 which is consistent with the notion that the combined SDbS model predicts range effects.

Table 14

Average w parameter estimate from model fits of the SDbS+range model

	Mean *w* estimate
Wood, Brown and Maltby (2011a)	.23[a]
Brown, Gardner, Oswald and Qian (2008)	.22[a]
Melrose, Brown and Wood (2012)	.35[ab]
Wood, Brown and Maltby (2011b)	.39[b]
Maltby, Wood, Vlaev, Taylor and Brown (2012)	.11[c]

Note: means not sharing same superscript are significantly difference $p < .05$

Model comparison. Does the SDbS model perform better with a range component? I first calculated the -2lnL of the model with SDbS only (w=0). Then I allowed the w parameter to vary. For individual level parameter estimates see Appendix G and for model comparison statistics see Appendix H.

The difference in -2lnLs is the improvement in model performance due to the range component of the model. We can examine the statistical significance of this difference using a likelihood ratio test because SDbS is nested within the generalized SDbS+Range model. The percentage of participants significantly better fit by the SDbS+Range model is shown in Table 15. Participant data are significantly better fit by SDbS+Range model if the difference in -2lnL is greater than 3.85. The 3.85 value is chosen because the difference between nested models is approximately a χ^2 distribution with K degree of freedom, where K is the difference in parameters between the models (Lewandowsky & Farrell, 2010). In our case, there is one parameter which differs between the models (w). The critical value of χ^2 with 1 degree of freedom for $p < .05$ is 3.85. Table 15 shows that for the majority of participants the range component does not significantly improve the performance of the model.

Table 15

Percentage of participants better fit by the SDbS+Range model

Paper	%	N
Wood, Brown and Maltby (2011a)	0	81
Brown, Gardner, Oswald and Qian (2008)	33	24
Melrose, Brown and Wood (2012)	6	52
Wood, Brown and Maltby (2011b)	2	86
Maltby, Wood, Vlaev, Taylor and Brown (2012)	15	135

An alternative approach to likelihood ratio tests is to calculate the criterion weights. The weight based comparison of SDbS with and without the range principle is shown in Table 16 and strongly favors the SDbS model. In summary, the results show that SDbS captures the range effects previous attributed to the range principle.

Table 16

Criterion weight based comparison of SDbS with and without a range based component

	AICw		BICw	
Paper	SDbS	+range	SDbS	+range
Wood, Brown and Maltby (2011a)	**.70**	.30	**.85**	.15
Brown, Gardner, Oswald and Qian (2008)	.44	**.56**	**.61**	.39
Melrose, Brown and Wood (2012)	**.70**	.30	**.97**	.03
Wood, Brown and Maltby (2011b)	**.73**	.27	**.98**	.02
Maltby, Wood, Vlaev, Taylor and Brown (2012)	**.67**	.33	**.94**	.06

Bootstrap analysis. The parameter estimates calculated above are the most likely parameter estimates. To examine the uncertainty in these estimates I carried out bootstrap analysis on the *w* parameter. The mean and 95% CI of these estimates is shown in Figure 20. For most participants a *w* parameter value of 0 is within the 95% CI; for these participants the SDbS model alone is credible when parameter uncertainty is considered.

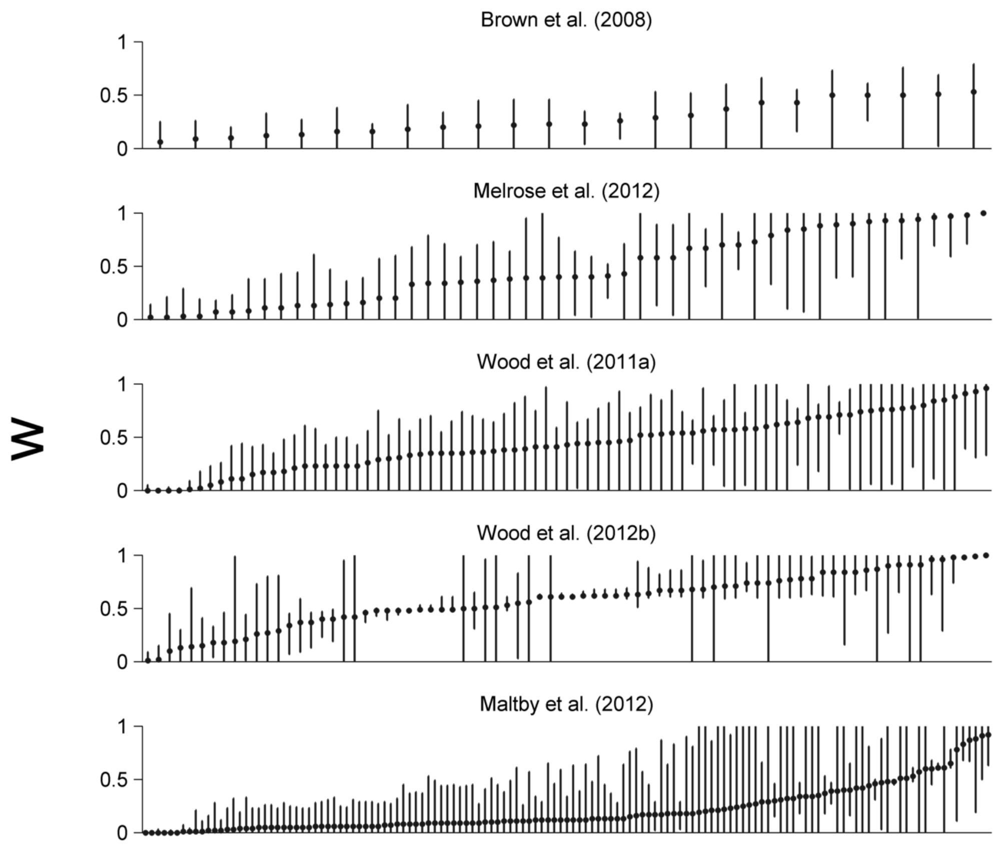

Figure 20. The *w* bootstrap estimates for the SDbS + range model

General Discussion

The aim of this chapter was to investigate the extent to which DbS can predict range effects in subjective judgments. I applied a model combining DbS and the SIMPLE model of memory to the individual level data analyzed in Chapter 3. In the first model comparison the SDbS model was compared to RFT. I found that for the majority of participants the models were equally likely, but model comparison favored RFT when the fit were penalized for model complexity. In the second model comparison I compared the performance of SDbS with and without the range principle from RFT. The results show that for many participants the range component was not necessary.

An advantage of the SIMPLE and DbS model (SDbS) is that it offers a unifying framework for understanding apparently disparate phenomena. Independently, these models can account for many psychological phenomena. SIMPLE can predict key memory phenomena such as serial position effects (Brown et al., 2007) and interference based forgetting (Lewandowsky et al., 2004). DbS can account for social decision making (Maltby et al., 2012; Moore, Wood, Brown, & Shepherd, 2012; Wood, Boyce, Moore, & Brown, 2012; Wood, Brown, et al., 2012).

In this thesis I have shown that DbS can account for many key relative judgment phenomena. Chapter 2 showed range effects do occur in overall judgments. Chapter 3 found range effects at an individual level in studies manipulating item spacing. In this chapter the findings show that range effects can be accounted for using the combined SIMPLE and DbS model.

However, there are limitations to the results presented here. First, many range effects occur when the frequency of stimuli is manipulated. SIMPLE and DbS has not yet been extended to examine these frequency effects[4]. Second, the SIMPLE model may be restricted to only accounting for serial position effects. A major confound in the serial position curve is output position. Items recalled first are more

[4] Preliminary analysis suggests that range effects from frequency manipulations are consistent with a combined SDbS model.

likely to be recalled irrespective of their input (serial) position. In the next chapter I address this confound directly.

Chapter 5

The Cost of Recall

Introduction to Chapter

In this chapter I continue the examination of SIMPLE, using economic incentives, to investigate the relationship between serial position effects and output interference. Both of these are important in the study of judgment and decision making: recency effects may play a role in the peak-end effects (Fredrickson & Kahneman, 1993) and forgetting due to output interference is an important component of query theory (Johnson et al., 2007).

First, I investigate the extent to which recall is influenced by output order effects. To do this I carry out three free recall experiments. The novel methodology associates input position with monetary payments. Consistent with previous literature I establish that output position appears to influence recall performance. I illustrate recall behavior with atypical serial position curves. Second, I examine if SIMPLE can be extended to account for atypical recall behavior. I fit a modified version of the SIMPLE model to this atypical data.

Our implementation of SIMPLE has several advantages over other previously reported implementations. First, as in previous chapters I fit the model at an individual level. Second, the SIMPLE architecture is used to compare the performance of two versions of the model. In one, forgetting during output is due to the passage of time alone; in the other, output interference occurs. Finally, I incorporate the observed output order of items from participant data into the model.

Abstract

How accessible are items in memory? The ubiquitous serial position effects in free recall suggest that items at the start and end of a list are more accessible in memory than those in the middle. However, these serial position effects do not provide an unambiguous measure of memory strength. Items in the first and last input positions are the most likely to be recalled overall. Here I report three free recall experiments in which input position was associated with a monetary incentive. Participants received differential monetary incentives for recalling items from particular portions (thirds) of the presented list. Experiment 1 found that items associated with higher payments were recalled earlier. Experiment 2 shows that this effect survives rapid presentation. Experiment 3 found that recall strategies can be altered when incentives are shown just before recall. Formal modeling using the SIMPLE model of memory investigates the extent to which these recall effects can be attributed to output interference or time based forgetting. I go beyond previous work by fitting the models at an individual level and incorporating the observed output order of recalled items.

Here I investigate the interaction of input and output position in a free recall task. In a series of experiments I use a novel paradigm in which I pay people for recalling items. Payment depends on the input positions of the recalled words in a series. The payment for words in the first, second and final third of the word list is manipulated. Then I use a formal model to examine whether the effect of output position on recall performance can be attributed to (a) the passage of time or (b) output interference. I extend and adapt a previous implementation of the SIMPLE model (Lewandowsky et al., 2004) and conduct model fitting at an individual level using maximum likelihood estimation (Myung, 2003).

The plan for this paper is as follows. First, I briefly review the literature on the effect of input and output position in free recall. Second, I consider previous work that has attempted to control the order of recall in both free and serial recall tasks. This is followed by a report of three experiments using financial incentives to direct recall behavior. Then the results of individual level model fit of SIMPLE are presented. I use the SIMPLE architecture to compare the performance of time-based forgetting and output interference based models.

Input Position Effects

In studies of free recall the input position of an item seems to influence overall recall performance. In the free recall paradigm participants are shown a series of items one after another and are then asked to recall as many of these items as they can in any order. The task was first introduced by Kirkpatrick (1894) and subsequently has become a key paradigm in memory research (for reviews see Bhatarah, Ward, & Tan, 2008; Greene et al., 1986; Murdock, 1974; Roberts, 1972).

Free recall tasks are particularly useful for researchers because they allow the examination of the relationship between the input position of the item (position in the list) and number of items correctly recalled. Many studies have found that items at the start and end of a series are more likely to be recalled when compared to items in the middle (e.g., Deese & Kaufman, 1957; Murdock, 1962; Murdock, 1974). These input position effects have been separated into the primacy - first few input positions– and recency – final few input positions – effects.

Primacy effects are robust. Studies manipulating the number of items in a series have shown primacy effects across list lengths (Murdock, 1962; Postman & Phillips, 1965; Ward et al., 2010). Primacy in recall has been attributed to the rehearsal of items in the first input positions throughout subsequent encoding. Initially, primacy was taken as evidence of a dual-store model of memory (Atkinson & Shiffrin, 1971; Raaijmakers & Shiffrin, 1981). In one free recall study supporting this interpretation, Rundus and Atkinson (1970) asked participants to rehearse items out loud. The average number of rehearsals was highest for items in the first output positions. In all but the final input positions a higher number of rehearsals was associated with better recall performance. These results suggest that rehearsal of items moved items in an early input position into long-term memory.

However, Tan and Ward (2000) argue that primacy may be due to *when* these items are rehearsed. Instead, rehearsal may create instances of items in early input positions that are closer to the point of recall (e.g., Brodie & Murdock, 1977; Laming, 2010). Tan and Ward (2000) found that items rehearsed closer to the point of recall were more likely to be recalled. If rehearsals strengthen the memory trace then the number of rehearsals should be the determining factor in recall performance. Tan and Ward (2000) suggest instead that the primacy effect can be explained using a similar mechanism as explains the recency effect.

Output Position effects

Recency effects may be due to the output position of items. In the free recall paradigm items can be recalled in any order. A robust finding in the literature is that items presented last are general recalled first (Beaman & Morton, 2000; Bhatarah et al., 2008; Farrell, 2010; Hogan, 1975; Ward et al., 2010). Intuitively, we might expect recall performance to be best for items recalled first.

Here I investigate two sources of these output position effects. In most models of memory the recall of an item decreases as a function of time alone. In trace decay accounts this is due to a degradation in the memory impression of an event. Other models, such as SIMPLE, predict that as time passes items become more easily confused with one another and are less likely to be recalled correctly for that reason (Brown et al., 2007). A second source may be output interference. The act of recalling an item creates additional noise or degrades the memory representation of other items. In effect, the outputting of an item interferes with the recall of the next. Both output interference predict that recall performance will decrease as a function of output position. I later apply individual level model fitting to the data from Experiments 4, 5 and 6, comparing both mechanisms within the SIMPLE model of memory.

Crucially, output position effects can confound the serial position curve. Based purely on the frequency of recall a researcher may conclude that items in the first and last input positions are more accessible in memory. However, this does not consider the output order of items. It may the case that items at the end and middle of a list are equally likely to be recalled when they are recalled in the first output position. The recency effect may be entirely or partially due to output position alone.

Directing Recall

In serial recall the experimenter can determine output order. Cowan, Saults, Elliott, and Moreno (2002) asked participants to start recall at specific input positions. They found that primacy effects could be attributed to output interference. Output interference may however depend on the modality of input and output, as shown by Harvey and Beaman (2007). Furthermore, Oberauer and Lewandowsky (2008) found that in delayed recall the degradation in performance was best explained by output interference.

In free recall items can be recalled in any order. Unlike in serial recall, the experimenter does not specific the order in which items are recalled. In many ways this type of recall more closely resembles everyday recall tasks. For instance, when shopping the order of recalling the items one needs to buy may not be important. As we might expect recency effects are most notable in free recall tasks where people recall the final list items in the first output positions. In an early study, Dalezman (1976) showed participants 15 words and used written instructions to direct their recall. He asked participants to recall either the first (input: 1-5), middle (input: 6-10) or last (input: 11-15) items, and then to recall any other items. Items recalled first were more likely to be recalled correctly.

Here I further develop the directed recall paradigm by associating input positions with financial incentives. Many previous studies have used incentive to influence recall behavior. These studies usually use incentives to improve recall by associating items with incentives or increasing the number of rehearsals for a given item (e.g., Cuvo, 1974; Eysenck & Eysenck, 1982; Haines & Torgesen, 1979; Hartley & Walsh, 1980; Hill, Storandt, & Simeone, 1990; Kunzinger III & Witryol, 1984; Rundus & Atkinson, 1970; Wilson, Witryol, & Hust, 1975). Here I use economic incentives by associating them with input positions.

Experiment 4

The methodology of the experiment is similar to the methodology used by Dalezman (1976). Participants were shown 15 item word lists. One item was presented per second. Recall was incentivized for a specific third of the list. The main difference between the work here and the work carried out by Dalezman (1976) was the use of financial incentives rather than directive instructions.

Method

Participants. Seventy five participants took part in the study. Participant payment varied depending on their payment schedule and recall performance.

Materials. Words from the Toronto word pool (Friendly, Franklin, Hoffman, & Rubin, 1982) were presented using a computer program written in Blitz3D. The words presented were randomly selected from the 1000 word noun subset of the Toronto word pool. Each word was two syllables in length.

Design. I manipulated the payment schedule between participants. There were five payment schedules (see Table 17). Payment depended on the presentation position of the words that were recalled in the increasing, decreasing and middle high conditions. The participant was paid for each word correctly recalled. The payment for each participant varied depending on the payment schedule. In the none condition the participant received a £3 payment irrespective of their performance and in the same condition the payment was two pence per word correctly recalled. In the remaining conditions the payment for each word varied. Table 17 shows the payment associated with input position in each condition. The recall accuracy and order of recall were recorded.

Table 17

Payment schedule used in Experiment 4 and Experiment 5

Payment Schedule	Serial Position		
	1 - 5	6 - 10	11 - 15
Increasing	1	2	3
Decreasing	3	2	1
Middle High	2	3	1
None		£3 for the experiment	
Same	2	2	2

Note: payments are in pence unless otherwise specified

Procedure. The participants were seated beside the experimenter and in front of a computer monitor in an experimental cubicle. Each participant was shown a series of instructions followed by 19 experimental trials.

Onscreen instructions informed the participant that they would be presented with a series of words which would be presented one after another after which they would be asked to recall the presented words. A second screen told the participants that their payment in the task might depend on the input positions of the words they recalled. The next screen differed depending on the participant's payment schedule. In all but the none payment schedule condition a bar chart depicting the payment schedule was shown. Those in the none condition were told they would receive £3 from taking part irrespective of their recall performance. Participants were then given the opportunity to ask the experimenter any questions.

Each participant completed 19 experimental trials. The bar chart or text reminding the participant of their payment schedule was displayed at the start of every trial until the participant pressed the space bar. Then a series of 15 words was shown onscreen one after another. Each word was displayed for 1000 ms. After all

15 words had been shown, an onscreen prompt told participants that they had one minute to write down as many words as they could recall on a piece of paper. The participant told the experimenter when they had recalled as many as they could. The experiment then pressed the space key to display the words presented on one screen and marked on the piece of paper the correct responses. In all but the none condition the experimenter told the participant how much money they had earned based on their performance and payment schedule.

Results and Discussion

To examine the effect of monetary incentives I analyzed three measures of recall: serial position curves, probability of first recall, and lag recency. Each of these measures is important for examining the overall performance and the output dynamics of recall.

Input Position Effects. To examine the effect of payment schedule I performed a repeated measures ANOVA on the probability of recalling items in each serial position. There was no significant main effect of payment schedule, $F(1,4) = 0.41$, $p = .80$, $\eta_p^2 = 0.02$. However, there was a significant main effect of serial position, $F(14,980) = 28.67$, $p < .001$, $\eta_p^2 = 0.29$, and a significant interaction between serial position and payment schedule, $F(56,980) = 2.45$, $p < .001$, $\eta_p^2 = 0.12$. These results show that the payment manipulation influenced the probability of recall with different effects depending on the serial position of an item.

A repeated measures ANOVA on the data from the same and none conditions revealed no significant effect of payment schedule, $F(1,1) = 0.03$, $p = .88$, $\eta_p^2 = 0.001$, and no significant interaction of payment schedule and serial position, $F(14,420) = 1.31$, $p = .20$, $\eta_p^2 = 0.04$. There was a significant main effect of serial position, $F(14,420) = 25.80$, $p < .001$, $\eta_p^2 = 0.46$, as shown in the serial position curve in Figure 21. Thus recall probability of words was similar when participants were paid £3 for the experiment compared to a uniform incentive of 2 pence per word.

To investigate the effect of differing incentives (increasing, decreasing, middle high) I ran a repeated measures ANOVA which revealed no significant main effect of payment schedule, $F(2,40) = 0.67$, $p = .52$, $\eta_p^2 = 0.32$, a significant main effect of serial position, $F(14,560) = 10.66$, $p < .001$, $\eta_p^2 = 0.21$, and a significant interaction between serial position and payment schedule, $F(28,560) = 3.11$, $p < .001$, $\eta_p^2 = 0.13$. The interaction is shown in Figure 21, where items associated with the highest payment were generally more likely to be recalled.

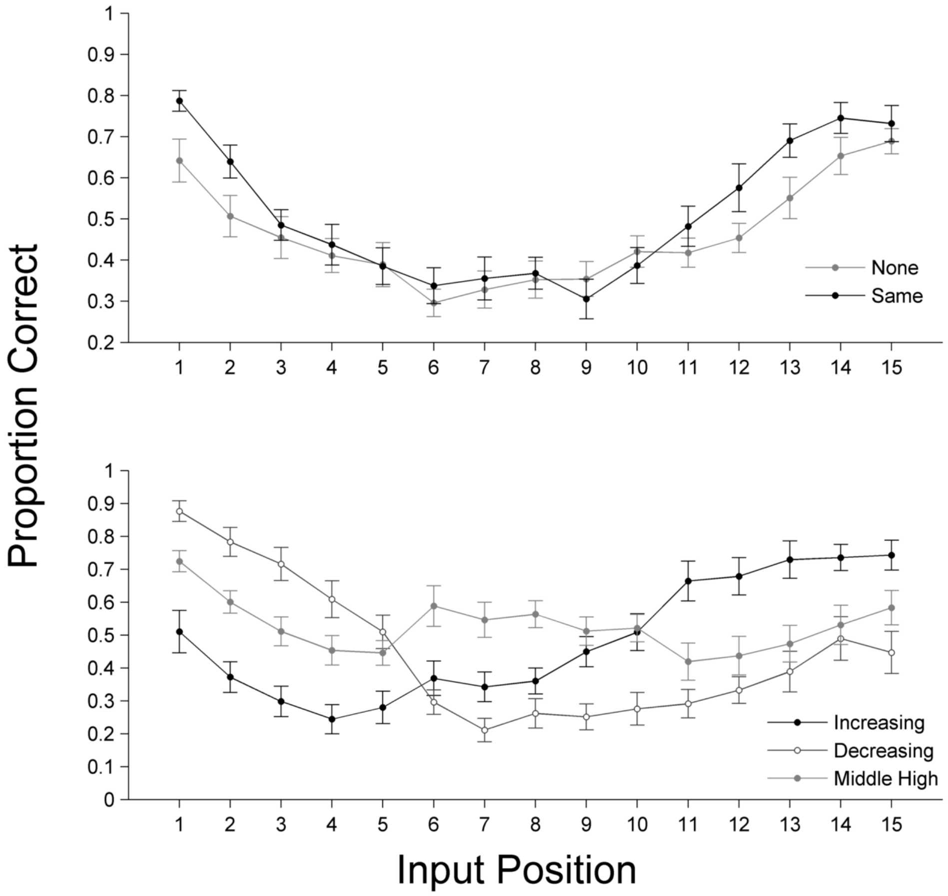

Figure 21. Serial position effects in Experiment 4

Effect of Incentive. To examine the effect of incentive I calculated the average proportion of items recalled in each third of the word list. A 3 (third of list) x 5 (payment schedule) repeated measures ANOVA on the average proportion correct found a significant main effect of list third, $F(2,140) = 39.20$, $p < .001$, $\eta_p^2 = 0.36$, no significant main effect of payment schedule, $F(4,70) = 1.53$, $p = .20$, $\eta_p^2 = 0.08$, and a significant interaction between payment schedules and third of list, $F(8,140) = 18.32$, $p < .001$, $\eta_p^2 = 0.51$. To explore this interaction I carried out one-way ANOVAs for each payment condition. In all of the conditions except middle high there was a significant main effect of third (all $p < .05$).

Table 18 shows the mean proportion correct in each third of the word list. The results of the post-hoc analysis are shown in the table. Thirds within each condition which share a superscript are not significantly different ($p > .05$). Participants recalled more of the items which were highly incentivized. However, performance was similar across thirds in the middle high condition.

Table 18

Mean proportion of items recalled in each third of the word list

	Third of list		
	First	Middle	Last
None	$.48^a$ *.18*	$.35^b$ *.13*	$.55^a$ *.12*
Same	$.55^a$ *.11*	$.35^b$ *.14*	$.65^a$ *.13*
Increasing	$.34^a$ *.17*	$.41^a$ *.13*	$.71^b$ *.15*
Decreasing	$.70^a$ *.16*	$.26^b$ *.14*	$.39^c$ *.19*
Middle High	.55 *.11*	.55 *.15*	.49 *.17*

Note: standard deviations are in italics.

Probability of First Recall. Are highly incentivized items the first to be recalled? We would expect this to be the case if people are trying to maximize their payment. To examine which items are recalled first I calculated the probability of first recall (see Figure 22). In other words, this measure is the proportion of items in each input position which were recalled first.

I ran a repeated measures ANOVA on the first recall probability of items in each serial position. The ANOVA showed no significant main effect of payment schedule, $F(4,70) = 0.47$, $p = .79$, $\eta_p^2 = 0.03$, a significant main effect of serial position, $F(14,980) = 22.80$, $p < .001$, $\eta_p^2 = 0.27$, and a significant interaction between serial position and payment schedule, $F(56,980) = 3.84$, $p < .001$, $\eta_p^2 = 0.18$.

Figure 22 shows the effect of incentives on the probability of first recall. In the same and none conditions the participants were most likely to first the last items presented to them. This is similar to the findings of previous studies (Bhatarah et al., 2008; Ward et al., 2010). When items are associated with a financial incentive participants tend to start recall with high incentive items.

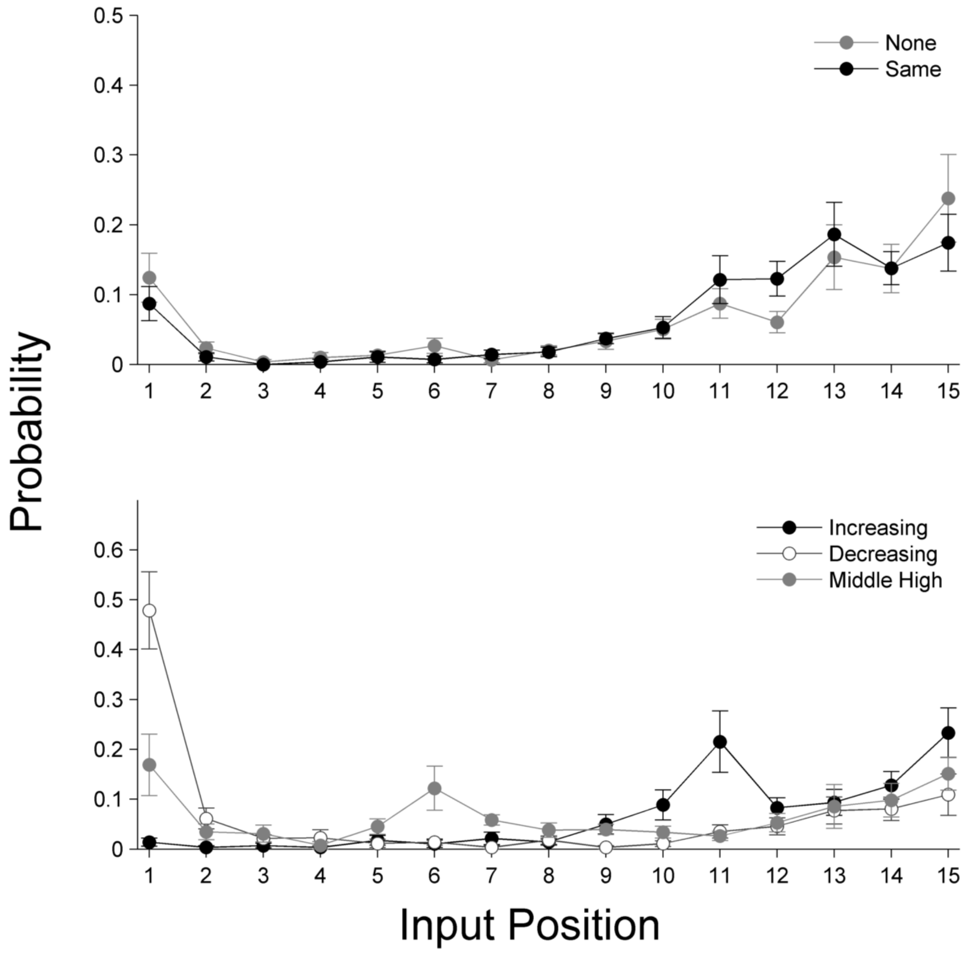

Figure 22. Probability of recalling items in output position 1 in Experiment 4

To examine the influence of incentives on the first recall probability of items in each third of the list I ran a 3 (third of list) x 5 (payment schedule) repeated measures ANOVA. This found a significant main effect of serial position, $F(2,140) = 63.86$, $p < .001$, $\eta_p^2 = 0.47$, no significant main effect of payment schedule, $F(1,70) = 0.46$, $p = .77$, $\eta_p^2 = 0.02$, and a significant interaction between payment schedule and third of list, $F(8,140) = 11.27$, $p < .001$, $\eta_p^2 = 0.39$.

To explore this interaction I carried out one-way ANOVAs on the proportion of correct recall for each payment condition. In all of the conditions except middle high there was a significant main effect of third. Thirds within each condition which do not share a superscript are significantly different ($p < .05$).

Table 19

Mean probability of first recall for items in each third of the word list

	Third of list		
	First	Middle	Last
None	$.03^{a}$	$.03^{a}$	$.14^{b}$
	.03	*.03*	*.05*
Same	$.02^{a}$	$.03^{a}$	$.15^{b}$
	.02	*.02*	*.04*
Increasing	$.01^{a}$	$.04^{b}$	$.15^{c}$
	.01	*.04*	*.05*
Decreasing	$.12^{a}$	$.01^{b}$	$.07^{a}$
	.06	*.01*	*.06*
Middle High	.06	.06	.08
	.06	*.04*	*.05*

Note: standard deviations are in italics.

Lag Recency. Previous studies have examined the conditional response probabilities of output position in free recall (Kahana, 1996). When people recall items it is informative to calculate how people move between input positions during recall. For instance, people may "jump" erratically between input positions when recalling items. On the other hand, people may recall items in the sequence that they were presented in (e.g., recalling the 4^{th} presented item after the 3^{rd} presented item).

I calculated these conditional response probabilities for each condition separately as shown in Figure 23. As we can see in the figure the most likely lag is +1 for all of the participants irrespective of condition. In other words, participants generally recalled items in input position order and this pattern of recall was not influenced by the incentives associated with input position.

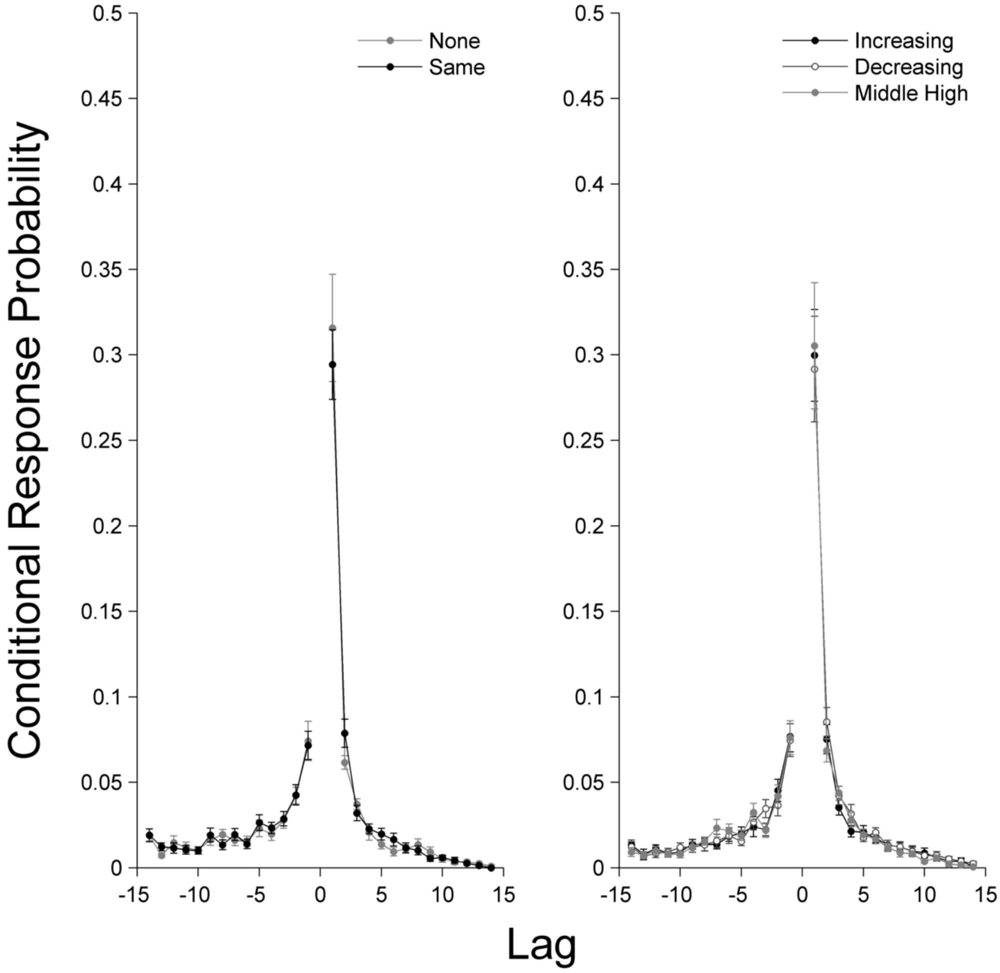

Figure 23. Lag recency of items in Experiment 4. Note: a lag recency of +1 indicates recall of a pair of items in input position order

Taken together, the results show that economic incentives can alter recall behavior. Items associated with the highest incentives were the most likely to be recalled. The first item in the highest incentive position of the list was more likely to be recalled first. One recall was initialized the items were generally recalled in input order.

Experiment 5

The results of Experiment 4 show that financial incentives can influence recall behavior. One aim of this paper is to examine the extent to which output position influences recall performance. In Experiment 4 each word was displayed for 1000 ms and participants knew which words were associated with the highest monetary incentives before encoding. It may the case that items were selectively rehearsed. In Experiment 5 the items were shown for only 500 ms to decrease rehearsal time.

Method

Participants. A total of 50 participants took part in the experiment with ten participants in each condition. As in Experiment 4 the participant payment depended on the payment schedule and the recalled words.

Procedure. Unlike Experiment 4 each word was presented for 500 ms. In all other ways the procedure of Experiment 5 was the same as that of Experiment 4.

Results and Discussion

To examine the influence of financial incentives I calculated the serial position curve, probability of first recall and lag recency. These are the same as presented in Experiment 3.

Input Position Effects. A repeated measure ANOVA using all of the data revealed a significant main effect of serial position, $F(14,630) = 22.69$, $p < .001$, $\eta_p^2 = 0.33$, no significant main effect of payment schedule, $F(4,45) = 0.76$, $p = .56$, $\eta_p^2 = 0.01$, and a significant interaction between payment schedule and serial position, $F(56,630) = 5.41$, $\eta_p^2 = 0.33$. As with Experiment 4, associating payments with words based on their serial position influenced how often the items were recalled. I next examined the uniform and differential payment schedules separately.

A repeated measures ANOVA examining responses from participants with the same and none payment schedules showed no significant main effect of payment schedule, $F(1,18) = 1.37$, $p = .23$, $\eta_p^2 = 0.07$, or significant main interaction between serial position and payment schedule, $F(14,252) = 1.22$, $p = .26$, $\eta_p^2 = 0.06$. However, there was a main effect of serial position, $F(14,238) = 13.02$, $p < .001$, $\eta_p^2 = 0.43$, as shown in Figure 24. As expected both recency and primacy effects were present in the data. Thus a small uniform incentive did not influence recall performance.

Do differential incentives direct recall performance even when the encoding duration is reduced? A repeated measures ANOVA showed a significant main effect of serial position, $F(14,378) = 11.27$, $p < .001$, $\eta_p^2 = 0.29$, no significant main effect of payment schedule, $F(1,27) = 0.65$, $p = .53$, $\eta_p^2 = 0.05$, and a significant interaction between payment schedule and input position, $F(28,378) = 9.38$, $p < .001$, $\eta_p^2 = 0.41$. As shown in Figure 24, the participants were much more likely to recall items associated with the highest incentive.

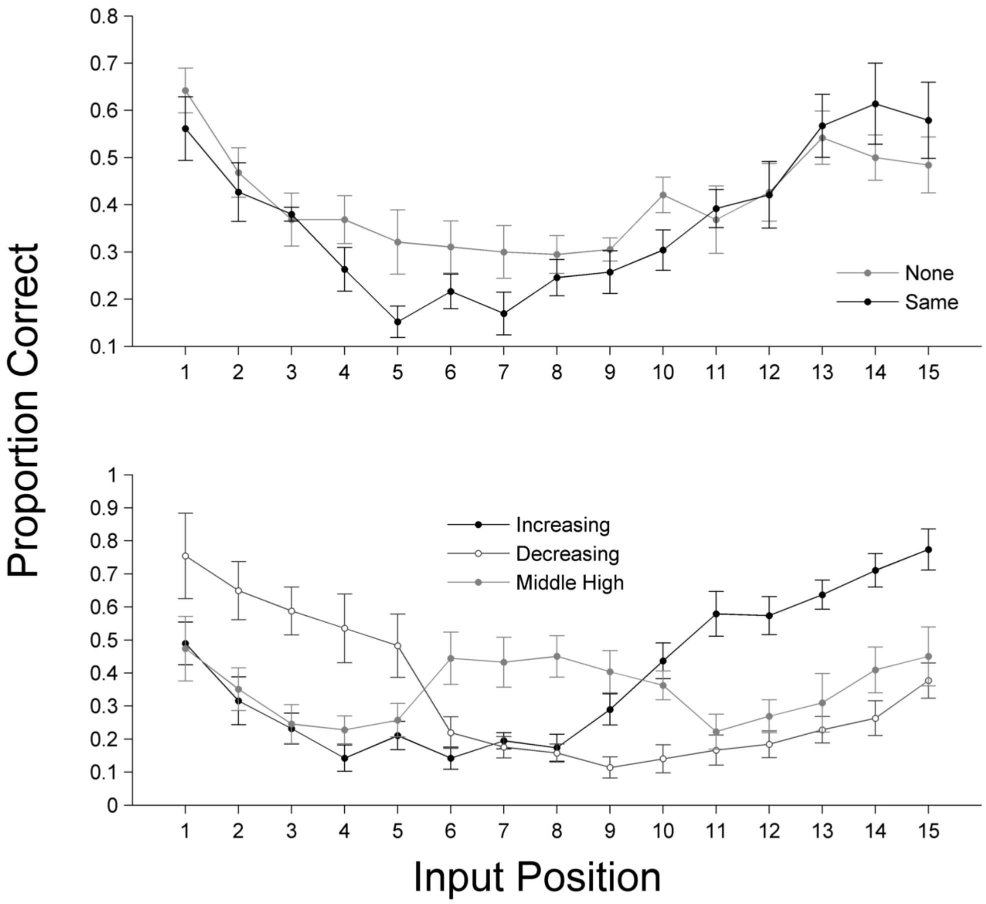

Figure 24. Serial position effects in Experiment 5

Effect of incentive. To examine the overall effect of incentive on the proportion of items recalled I calculated the average proportion correct for each third of the word list. A 3 (third of list) x 5 (payment schedule) repeated measures ANOVA found a significant main effect of list third, $F(2,90) = 21.14$, $p < .001$, $\eta_p^2 = 0.32$, no significant main effect of payment schedule, $F(4,45) = 0.76$, $p = .56$, $\eta_p^2 = 0.06$, and a significant interaction between payment schedules and third of list, $F(8,90) = 9.68$, $p < .001$, $\eta_p^2 = 0.46$. To explore this interaction I carried out one-way ANOVAs on the proportion of correct recall for each payment condition. In the same, decreasing and increasing payment schedule conditions there was a significant main effect of third. However, the effect of third was not statistically significant in the none condition (p = .06) and not significant in the middle high condition.

Table 20 shows the mean proportion correct in each third of the word list. The results of the post-hoc analysis are shown in the table. Thirds within each condition which do not share a superscript are significantly different (p < .05). These results clearly show that differential incentives can influence recall performance.

Table 20

Mean proportion of items recalled from each third of the word list

	Third of list		
	First	Middle	Last
None	.43 *.10*	.33 *.04*	.46 *.05*
Same	.38[a] *.10*	.23[b] *.04*	.50[a] *.05*
Increasing	.28[a] *.15*	.25[a] *.04*	.66[b] *.05*
Decreasing	.58[a] *.18*	.21[b] *.04*	.34[c] *.05*
Middle High	.32 *.10*	.40 *.04*	.33 *.05*

Note: standard deviations are in italics.

Probability of First Recall. Next I examined the first recall probability items in each input position. If incentives influence recall then I would expect the high incentive items to be recalled first. A repeated measures ANOVA on the first recall probability of items based on their serial position showed a significant main effect of serial position, $F(14,630) = 12.66$, $p < .001$., $\eta_p^2 = 0.22$, no significant main effect of payment schedule, $F(4,45) = 1.37$, $p = .26$, $\eta_p^2 = 0.11$, and a significant interaction between payment schedule and serial position, $F(56,630) = 2.77$, $p < .001$, $\eta_p^2 = 0.17$. The general pattern of the results was similar to that found in Experiment 4 (see Figure 25). Participants appear most likely to recall items associated with a high incentive first.

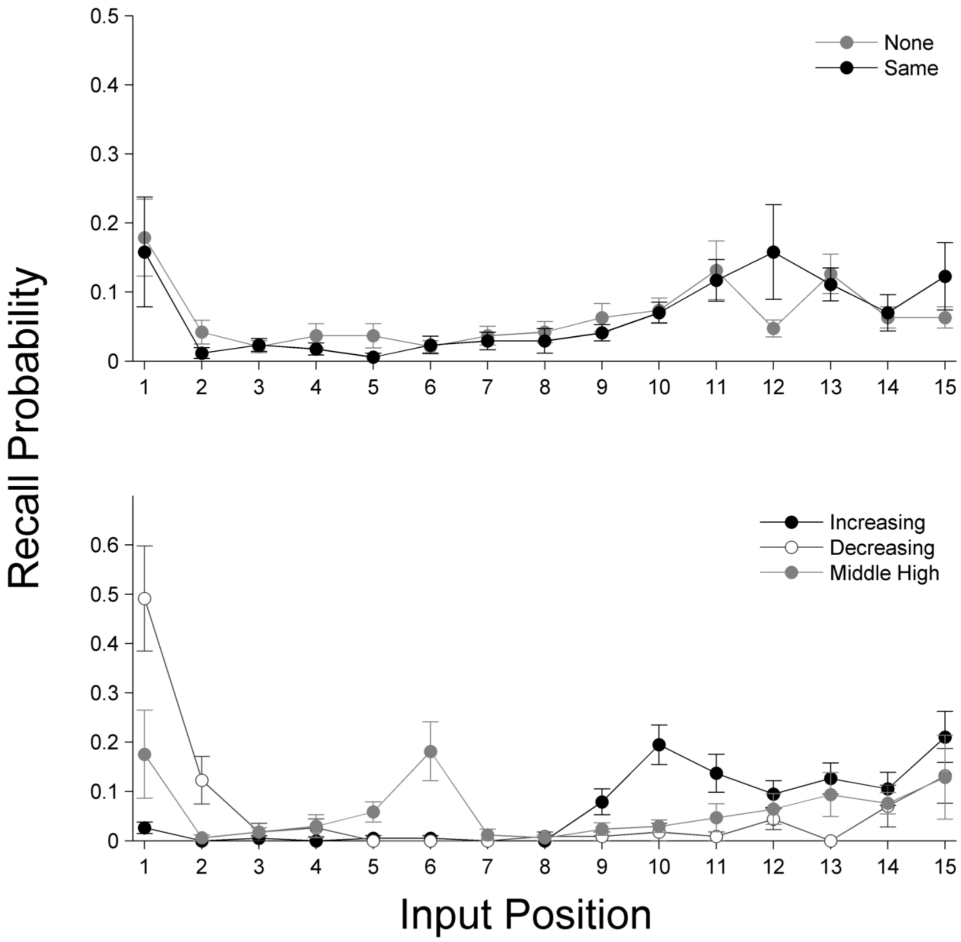

Figure 25. Probability of recalling items in output position 1 in Experiment 5

To examine the overall influence of incentive I calculated the average first recall probability for each third of the list. A 3 (third of list) x 5 (payment schedule) repeated measures ANOVA found a significant main effect of third of list, $F(2,90) = 14.65$, $p < .001$, $\eta_p^2 = 0.25$, no significant main effect of payment schedule, $F(4,45) = 1.37$, $p = .26$, $\eta_p^2 = 0.11$, and a significant interaction between payment schedules and third of list, $F(8,90) = 3.87$, $p = .001$, $\eta_p^2 = 0.26$. To explore this interaction I carried out one-way ANOVAs on the probability of first recall for each payment condition. In all of the conditions except middle high and same there was a significant main effect of third. Thirds within each condition which do not share a superscript are significantly different ($p < .05$). These results show that incentives influenced which item was recalled first.

Table 21

Mean proportion of items recalled in each third of the word list

	Third of list		
	First	Middle	Last
None	.06 *.05*	.05 *.02*	.09 *.04*
Same	.05^{ab} *.05*	.04^{a} *.02*	.12^{b} *.06*
Increasing	.01^{a} *.01*	.06^{b} *.04*	.14^{c} *.04*
Decreasing	.12^{ac} *.06*	.00^{b} *.03*	.07^{bc} *.05*
Middle High	.06 *.06*	.05 *.04*	.08 *.06*

Note: standard deviations are in italics.

Lag Recency. The probability of first recall shows us that people have a tendency to start their recall with the first high incentive item that they are shown. The serial position effects show that people recall more of the high inventive items. However, I have not examined the order in which successive pairs of items are recalled. I examined this by calculating lag recency. In Figure 26 shows that people generally recalled items in their serial position order because a lag recency of +1 is most common.

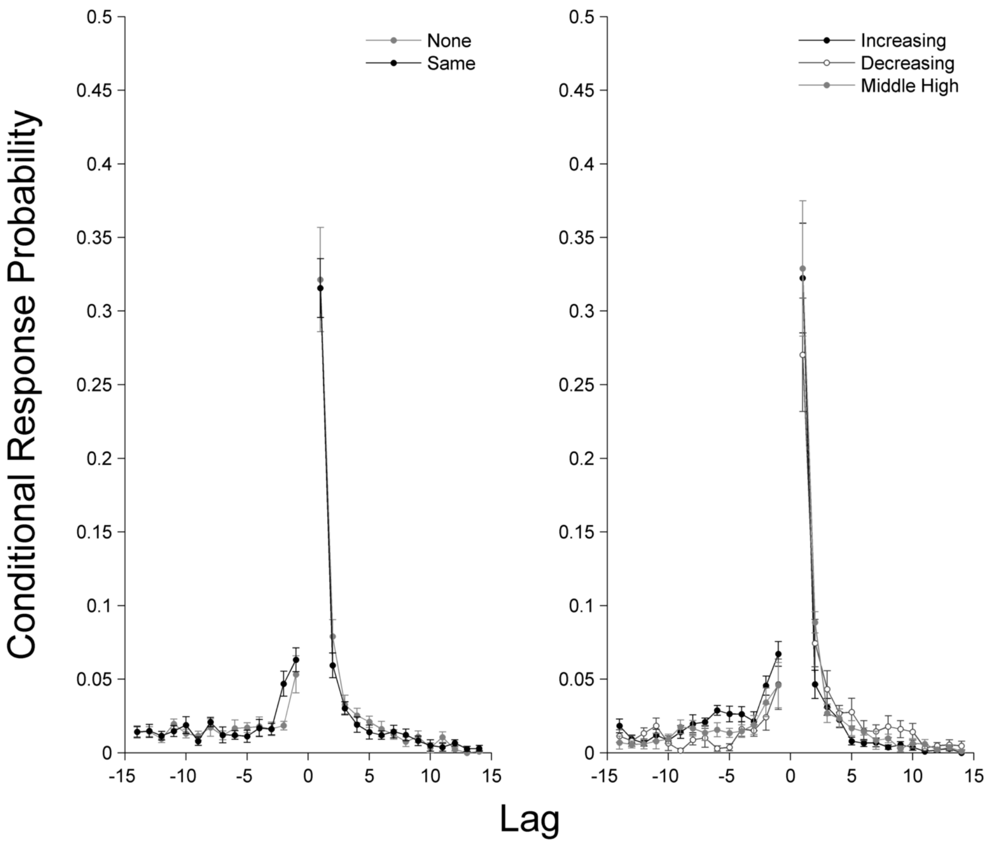

Figure 26. Conditional response probability in Experiment 5

Experiment 6: Postcued Recall Schedule

Experiments 4 and 5 showed that participants are able to selectively retrieve items in order to maximize their payment. Items associated with a high incentive were more likely to be recalled correctly and in the first output position. In both of these experiments the payment schedule was displayed immediately before the words were shown to the participant. The recall performance of high incentive items may have been due to increased attention on those items.

Experiment 6 was designed to examine how well people could selectively recall items when they had no knowledge of the payment schedule until the point of recall (i.e., after encoding). The order of the payment schedules shown at recall was random which will prevent participants from selectively rehearsing or attending to only the high incentive items.

Method

Participants. Nineteen participants took part and were paid according to their recall performance.

Materials. The materials were the same as those used in Experiment 4 and Experiment 5.

Design. The payment schedule was manipulated within participants. In the payment schedules the participant was paid 6 pence per word correctly recalled in the first, middle, or last five serial positions (see Table 22)[5]. The order in which the participant experienced each payment schedule was random.

Table 22

Payment schedules shown to participants in Experiment 6

	Serial Position of Word		
Payment Schedule	1 - 5	6 - 10	11 - 15
First	6	0	0
Middle	0	6	0
Last	0	0	6

[5] Preliminary experiments found that participants became confused when the payment schedules used in Experiment 4 and Experiment 5 were post-cued. Due to this we used a simpler payment schedule in Experiment 5.

Procedure. The procedure of Experiment 6 differed from Experiment 5 in two ways. First, the payment schedule was shown after the 15 words had been displayed onscreen. Second, participants experienced a total of 21 trials consisting of seven trials per payment schedule.

Results and Discussion

In Experiment 6 participants were shown a payment schedule just before initiating recall. The aim of the analysis presented here is to investigate if people altered their behavior depending on the payment schedule.

Input Position Effects. A repeated measures ANOVA revealed a significant main effect of payment schedule, $F(2,28) = 4.67$, $p = .02.$, $\eta_p^2 = 0.25$, a significant main effect of serial position, $F(14,196) = 9.02$, $p < .001.$, $\eta_p^2 = 0.39$, and a significant interaction between payment schedule and serial position, $F(28,392) = 11.88$, $p < .001.$, $\eta_p^2 = 0.46$. The interaction between payment and serial position is shown in Figure 27.

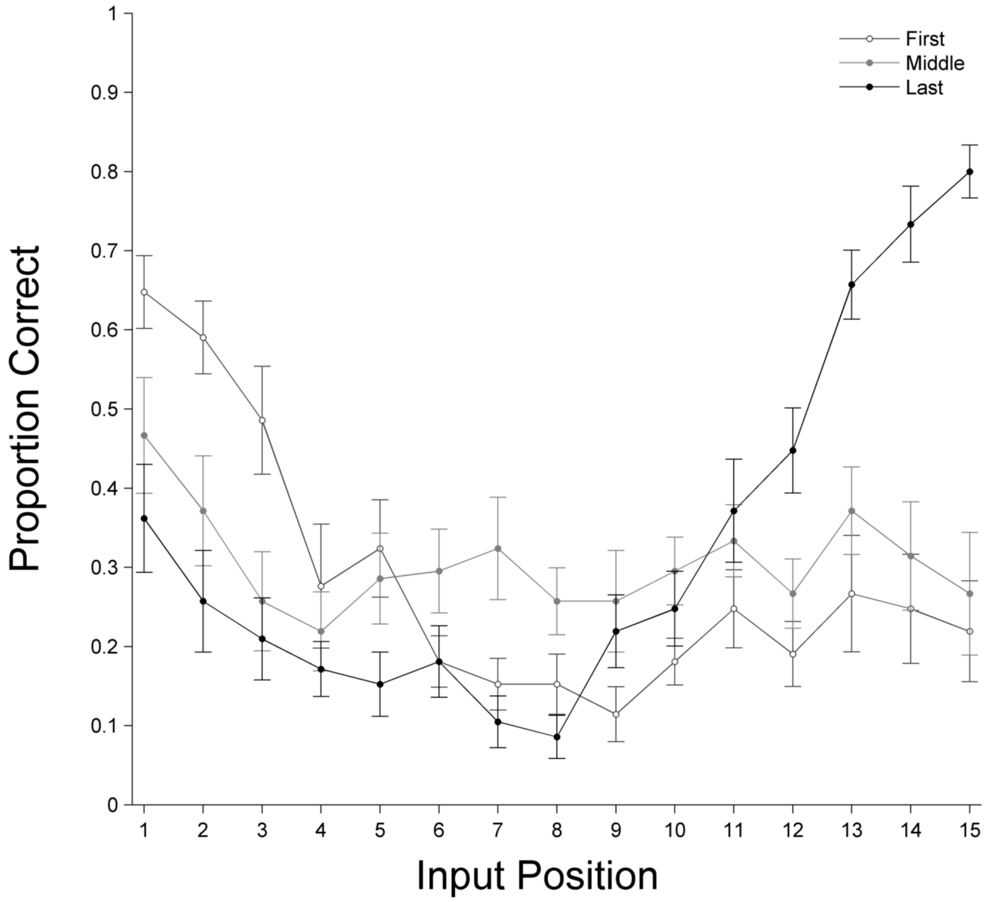

Figure 27. Recall performance of participants in differing incentive conditions

Effect of incentive. To examine the overall effect of incentive on recall performance I calculated the average proportion of items recalled in each third of the word list. A 3 (third of list) x 3 (payment schedule) repeated measures ANOVA was carried out using the average proportion of correct items recalled. There was a significant main effect of list third, $F(2,56) = 11.37$, $p < .001$, $\eta_p^2 = 0.45$, a significant main effect of payment schedule, $F(2,56) = 4.67$, $p = .02$, $\eta_p^2 = 0.25$, and a significant interaction between payment schedules and third of list, $F(4,56) = 26.98$, $p < .001$, $\eta_p^2 = 0.66$. To explore this interaction I carried out one-way ANOVAs for each payment condition. In the first and last conditions there was a significant main effect of third (all *p's* < .05).

Table 23 shows the mean proportion correct in each third of the word list. The results of the post-hoc analysis are shown in the table. Thirds within each condition which share a superscript are not significantly different ($p < .05$). These results show that incentive did influence recall. In the first and last conditions the participant recalled more of the items which were highly incentivized.

Table 23

Mean proportion of items recalled in each third of the word list

	Third of list		
Payment Schedule	First	Middle	Last
First	.46[a] *.17*	.16[b] *.06*	.23[b] *.19*
Middle	.32 *.17*	.29 *.13*	.31 *.17*
Last	.23[a] *.16*	.17[a] *.12*	.60[b] *.09*

Note: standard deviations are in italics.

Probability of First Recall. A 3 (payment schedule) x 15 (serial position) repeated measures ANOVA showed a significant main effect of payment schedule, $F(2,36) = 7.28$, $p = .002$, $\eta_p^2 = 0.29$, a significant main effect of serial position, $F(14,252) = 13.15$, $p < .001$, $\eta_p^2 = 0.42$, and a significant interaction of payment schedule and serial position, $F(28,504) = 6.93$, $p < .001$, $\eta_p^2 = 0.28$. This interaction is shown in Figure 28.

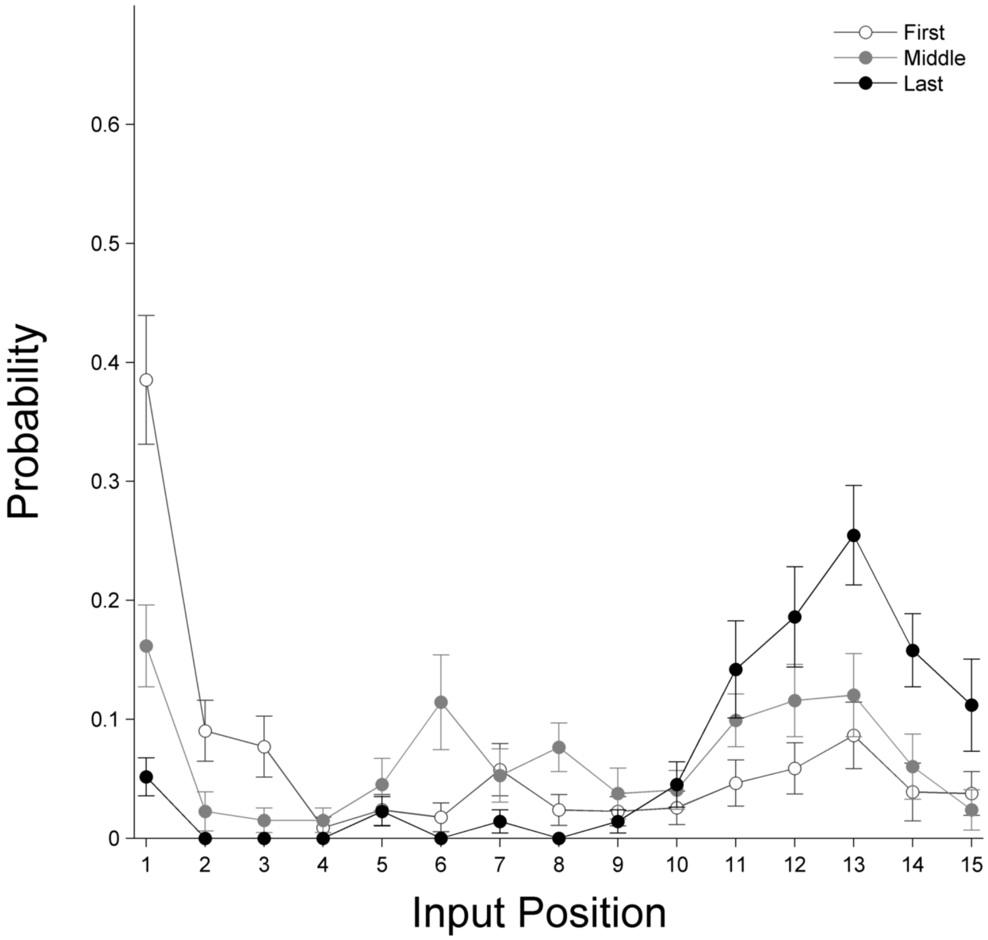

Figure 28. Probability of first recall for items in Experiment 6

I examined the number of items that participant recalled in the first output position. A 3 (third of list) x 3 (payment schedule) repeated measures ANOVA found a significant main effect of third of list, $F(2,36) = 23.91$, $p < .001$, $\eta_p^2 = 0.57$, no significant main effect of payment schedule, $F(2,36) = 7.28$, $p = .002$, $\eta_p^2 = 0.29$, and a significant interaction between payment schedules and third of list, $F(4,72) = 24.47$, $p < .001$, $\eta_p^2 = 0.58$. To explore this interaction I carried out one-way ANOVAs on the average number of items correctly recalled in the first output position in each third of the word list. In all of the conditions except middle high and same there was a significant main effect of third. The average first recall probability of items in each third is shown in Table 24. Thirds within each condition which do not share a superscript are significantly different ($p < .05$). As we might expect, the first items recalled were generally those associated with a high incentive.

Table 24

Average number of items recalled in the first output position for each third of the word list

	Third of list		
Payment Schedule	First	Middle	Last
First	0.74[a] *0.38*	0.19[ab] *0.16*	0.36[b] *0.40*
Middle	0.36 *0.24*	0.43 *0.38*	0.58 *0.42*
Last	0.11[a] *0.14*	0.11[a] *0.14*	1.20[b] *0.19*

Note: standard deviations are in italics.

Lag Recency. The lag recency effects were similar to those in Experiment 4 and Experiment 5. As shown in Figure 29, participants were most likely to recall items in forward neighboring input positions.

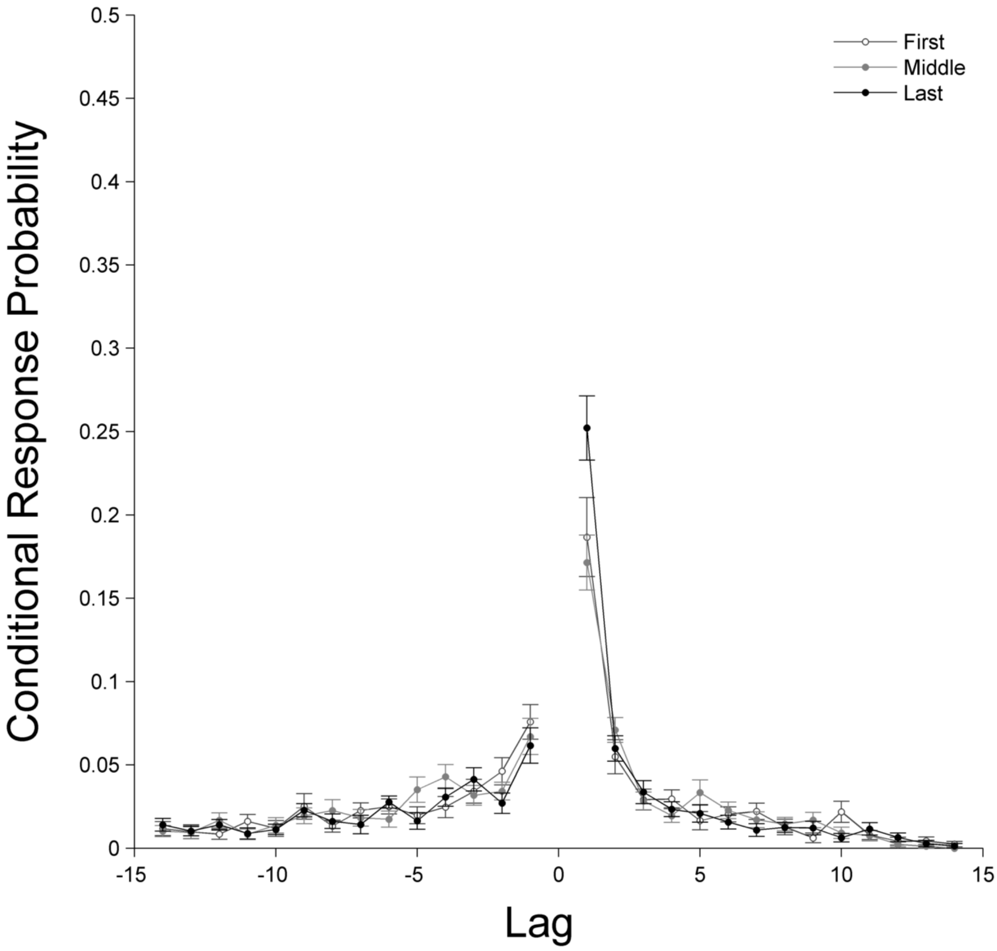

Figure 29. Lag recency observed in Experiment 6

Model Based analysis

The above findings show that output position does influence recall performance. Items associated with high incentives were the most likely to be recalled first and the most likely to be recalled overall. This effect persisted when rehearsals were reduced by either increasing the presentation rate or associating items with incentives after encoding.

Yet there remains a key theoretical question: Is the advantage of items in the first few output positions due to output interference or time-based forgetting? In this study I have incentivized recall to produce atypical serial position effects. The

SIMPLE model offers an account of memory which may be extended to decision making (see Chapter 4 and Brown & Matthews, 2011). However, previous applications of the SIMPLE model have focused on typical serial position effects. Here I examine if SIMPLE can be extended to the atypical serial position effect reported above. An advantage of the SIMPLE model is that it allows us to examine the extent to which decreases in recall performance can be attributed to output interference or time based forgetting.

According to event-based models, the decrease in recall performance over output position cannot be attributed to the passage of time (Murdock, 1995; Neath & Crowder, 1990). One mechanism proposed by these event-based models is output interference (Lewandowsky & Farrell, 2008; Oberauer & Lewandowsky, 2008) in which the recollection of items from memory interferes with subsequent recall attempts.

Time-based models predict a decrease in recall performance due to the passage of time. As more items are recalled by participants the time between encoding and retrieval increases. In some models memory are assume to decay with the passing of time (Baddeley, 1992; Page & Morris, 1998). Other accounts, such as SIMPLE (Brown et al., 2007), predict poorer recall performance due to a decreased distinctiveness of items along a temporal dimension.

Here I use an implementation of SIMPLE to compare output interference and time based forgetting accounts using the data collected from Experiments 4, 5 and 6. This implementation is similar to that of Lewandowsky et al. (2004). The model predicts a decrease in recall performance as a function of output position due to either (a) output interference or (b) the passage of time. Unlike Lewandowsky et al.

(2004) I fit the model to individual level data and use the observed output position of items for each participant to weight the predictions of the model.

Before reporting the results of the model fitting I will outline the model. First, I detail the output interference model in which time has no effect on the recall probability of items. Instead, the memory representations of items becomes less discriminable with each recalled item. Second, the time based forgetting model is described. In this model the time between item presentation and recall is log compressed (as detailed in Brown et al., 2007). With each output position the presented items are more distant in time. When logged the result is less local distinctiveness between presented items and lower recall probabilities. Finally, I use data from a participant in Experiment 4 to demonstrate how the observed output position of items is used to weight the model predictions.

An advantage of using SIMPLE as a common framework is that the model can generate predictions due to both output interference and time-based Forgetting. SIMPLE allows recall to be a function of the local distinctiveness of an item along any particular dimension. In the implementation applied here the temporal and input position dimensions are used. Predictions can be made based on a weighted average of these dimensions,

$$Distance_{i,j} = w\left|\log TD_i - \log TD_j\right| + (1-w)\left|IP_i - IP_j\right| \tag{36}$$

where the distance in psychological space between items *i* and *j* is a function of their distance from the point of recall, *TD*, and input position, *IP*. The weight of these dimensions is controlled by the *w* parameter. When *w* = 1 the predictions are based purely on the temporal distance. If *w* = 0 then the predictions are based solely on the input position of the item. The predictions are based on a weighted average of the two dimensions when 0<*w*<1.

Output Interference

In SIMPLE recall is a discrimination task. In principle this discrimination can be based on any particular dimension. Following Lewandowsky et al. (2004) the dimension of interest in the output interference model is the input position of the item. The psychological distance between items is the absolute difference between input positions,

$$I_{i,j} = \left|IP_i - IP_j\right| \quad (37)$$

given the input positions (IP) of items *i* and *j*. When applied to the 15 word list used in the above experiments we get the absolute distances shown in Table 25.

Table 25

Absolute distance of items along input position dimension

	Input														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
3	2	1	0	1	2	3	4	5	6	7	8	9	10	11	12
4	3	2	1	0	1	2	3	4	5	6	7	8	9	10	11
5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	10
6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9
7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8
8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7
9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6
10	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5
11	10	9	8	7	6	5	4	3	2	1	0	1	2	3	4
12	11	10	9	8	7	6	5	4	3	2	1	0	1	2	3
13	12	11	10	9	8	7	6	5	4	3	2	1	0	1	2
14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	1
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0

These absolute distances are transformed into similarities using the following equation,

$$\eta_{i,j} = e^{-cI_{i,j}}. \tag{38}$$

In the output interference model we assume that the confusability of the presented words depends on their input position. The confusability of item is controlled by the *c* parameter in Equation 40. The value of *c* given output position *n* is,

$$c_n = ko^{n-1} \tag{39}$$

where *k* is a constant and *o* is a value between 0 and 1. When *o* is 1 there is no output interference because the value of *c* is the same in all output positions. If $o < 1$ then *c* decreases with each output position.

As a result items are more confusable in later output positions. When $k = 1.5$ and $o = .8$ the *c* parameter for the second output position is 1.2. The similarity of the items when $c = 1.2$ is shown in Table 26. The shading of the table cells correspond to the similarity of the items. The darkest cells are the most similar based on the absolute distance of the items in Table 25.

Table 27 shows the similarity of items when $c = 0.2$. As shown by the cell shading the items are much more similar in output position 10. In other words, it will be harder to discriminate between the items at later output positions.

Table 26

Similarity of items along the input position dimension in the second output position

	Input														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	.30	.09	.03	.01	0	0	0	0	0	0	0	0	0	0
2	.30	1	.30	.09	.03	.01	0	0	0	0	0	0	0	0	0
3	.09	.30	1	.30	.09	.03	.01	0	0	0	0	0	0	0	0
4	.03	.09	.30	1	.30	.09	.03	.01	0	0	0	0	0	0	0
5	.01	.03	.09	.30	1	.30	.09	.03	.01	0	0	0	0	0	0
6	0	.01	.03	.09	.30	1	.30	.09	.03	.01	0	0	0	0	0
7	0	0	.01	.03	.09	.30	1	.30	.09	.03	.01	0	0	0	0
8	0	0	0	.01	.03	.09	.30	1	.30	.09	.03	.01	0	0	0
9	0	0	0	0	.01	.03	.09	.30	1	.30	.09	.03	.01	0	0
10	0	0	0	0	0	.01	.03	.09	.30	1	.30	.09	.03	.01	0
11	0	0	0	0	0	0	.01	.03	.09	.30	1	.30	.09	.03	.01
12	0	0	0	0	0	0	0	.01	.03	.09	.30	1	.30	.09	.03
13	0	0	0	0	0	0	0	0	.01	.03	.09	.30	1	.30	.09
14	0	0	0	0	0	0	0	0	0	.01	.03	.09	.30	1	.30
15	0	0	0	0	0	0	0	0	0	0	.01	.03	.09	.30	1

Table 27

Similarity of items along the input position dimension in the tenth output position

	Input														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	.82	.67	.55	.45	.37	.30	.24	.20	.16	.13	.11	.09	.07	.06
2	.82	1	.82	.67	.55	.45	.37	.30	.24	.20	.16	.13	.11	.09	.07
3	.67	.82	1	.82	.67	.55	.45	.37	.30	.24	.20	.16	.13	.11	.09
4	.55	.67	.82	1	.82	.67	.55	.45	.37	.30	.24	.20	.16	.13	.11
5	.45	.55	.67	.82	1	.82	.67	.55	.45	.37	.30	.24	.20	.16	.13
6	.37	.45	.55	.67	.82	1	.82	.67	.55	.45	.37	.30	.24	.20	.16
7	.30	.37	.45	.55	.67	.82	1	.82	.67	.55	.45	.37	.30	.24	.20
8	.24	.30	.37	.45	.55	.67	.82	1	.82	.67	.55	.45	.37	.30	.24
9	.20	.24	.30	.37	.45	.55	.67	.82	1	.82	.67	.55	.45	.37	.30
10	.16	.20	.24	.30	.37	.45	.55	.67	.82	1	.82	.67	.55	.45	.37
11	.13	.16	.20	.24	.30	.37	.45	.55	.67	.82	1	.82	.67	.55	.45
12	.11	.13	.16	.20	.24	.30	.37	.45	.55	.67	.82	1	.82	.67	.55
13	.09	.11	.13	.16	.20	.24	.30	.37	.45	.55	.67	.82	1	.82	.67
14	.07	.09	.11	.13	.16	.20	.24	.30	.37	.45	.55	.67	.82	1	.82
15	.06	.07	.09	.11	.13	.16	.20	.24	.30	.37	.45	.55	.67	.82	1

The discriminability of item from the a given memory location is,

$$D_i | I_j = \frac{\eta_{i,j}}{\sum_{k=1}^{n} \eta_{i,k}} \tag{40}$$

where $\eta_{i,j}$ is calculated from Equation 40. Applying this equation to Table 26 we divide each similarity in a row by the row total.

Next we apply a thresholding function to calculate the recall probabilities,

$$P(R_i|D_i) = \frac{1}{1 + e^{-s(D_i - t)}} \tag{41}$$

where s and t are the parameters for the slope and threshold. Here s = 10 and t = .4. The resulting cued recall probabilities for output position two are shown in Table 28. Following Lee and Pooley (2012) given n cues the probability of recalling item i is

$$\theta_i = 1 - \prod_{j=1}^{n} \left(1 - P_{i,j}\right) \tag{42}$$

where $P_{i,j}$ is the probability of recalling item i given cue j. The overall probability of recalling a given item in the second output position is shown at the bottom of Table 28. In this case the model predicts a high recall probability for all of the items. Items in the first and last input position are most likely to be recalled.

Table 28

Cued recall probabilities and overall recall probability in output position two

	Input														
Cue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.95	.13	.03	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02
2	.09	.85	.09	.03	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02
3	.03	.09	.82	.09	.03	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02
4	.02	.03	.09	.80	.09	.03	.02	.02	.02	.02	.02	.02	.02	.02	.02
5	.02	.02	.03	.08	.80	.08	.03	.02	.02	.02	.02	.02	.02	.02	.02
6	.02	.02	.02	.03	.08	.80	.08	.03	.02	.02	.02	.02	.02	.02	.02
7	.02	.02	.02	.02	.03	.08	.80	.08	.03	.02	.02	.02	.02	.02	.02
8	.02	.02	.02	.02	.02	.03	.08	.80	.08	.03	.02	.02	.02	.02	.02
9	.02	.02	.02	.02	.02	.02	.03	.08	.80	.08	.03	.02	.02	.02	.02
10	.02	.02	.02	.02	.02	.02	.02	.03	.08	.80	.08	.03	.02	.02	.02
11	.02	.02	.02	.02	.02	.02	.02	.02	.03	.08	.80	.08	.03	.02	.02
12	.02	.02	.02	.02	.02	.02	.02	.02	.02	.03	.09	.80	.09	.03	.02
13	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.03	.09	.82	.09	.03
14	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.03	.09	.85	.09
15	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.03	.13	.95
	.97	.91	.88	.87	.87	.87	.87	.87	.87	.87	.87	.87	.88	.91	.97

However, the recall probability of items decreases as the output position increases. As described in Equation 41 the value of the *c* parameter depends on the output position of the item. Later output positions will have a lower *c* parameter value which decreases the local distinctiveness of all of the items. This captures the intuition that each additional recall decreases the ease with which later items can be recalled.

Table 29 shows the overall recall probability of each input and output position. To produce these values I applied the above calculations to every output position. As expected the output interference model predicts that the probability of recalling a word in any input position will decrease as a function of the output position. This effect is produced by decreasing the value of the *c* as output position increases.

Table 29

Recall probability of items based on input and output position

	Output														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.98	.97	.93	.86	.76	.65	.56	.49	.44	.42	.40	.40	.39	.39	.39
2	.95	.91	.84	.76	.67	.60	.54	.49	.46	.43	.42	.41	.40	.40	.40
3	.94	.88	.79	.69	.61	.55	.51	.48	.46	.44	.43	.42	.41	.41	.41
4	.94	.87	.77	.66	.58	.53	.49	.47	.45	.44	.43	.42	.42	.42	.41
5	.94	.87	.76	.65	.57	.51	.48	.46	.45	.44	.43	.43	.42	.42	.42
6	.94	.87	.76	.64	.56	.51	.48	.46	.45	.44	.43	.43	.43	.42	.42
7	.94	.87	.76	.64	.56	.50	.47	.45	.44	.44	.43	.43	.43	.42	.42
8	.94	.87	.76	.64	.55	.50	.47	.45	.44	.44	.43	.43	.43	.43	.42
9	.94	.87	.76	.64	.56	.50	.47	.45	.44	.44	.43	.43	.43	.42	.42
10	.94	.87	.76	.64	.56	.51	.48	.46	.45	.44	.43	.43	.43	.42	.42
11	.94	.87	.76	.65	.57	.51	.48	.46	.45	.44	.43	.43	.42	.42	.42
12	.94	.87	.77	.66	.58	.53	.49	.47	.45	.44	.43	.42	.42	.42	.41
13	.94	.88	.79	.69	.61	.55	.51	.48	.46	.44	.43	.42	.41	.41	.41
14	.95	.91	.84	.76	.67	.60	.54	.49	.46	.43	.42	.41	.40	.40	.40
15	.98	.97	.93	.86	.76	.65	.56	.49	.44	.42	.40	.40	.39	.39	.39

Time Based Forgetting

In the time based model there is no output interference. Instead, any effect of output position is due to the passage of time alone. In the time based forgetting model discrimination in memory is performed along the temporal dimension. The psychological distance between items is the log compressed absolute distance of items on the temporal dimension,

$$TI_{i,j} = \left|\ln TD_i - \ln TD_j\right| \tag{43}$$

given the temporal distance (TD) of items *i* and *j* from the point of recall. As a result of log transformation there will be less difference between the temporal distances of items as their distance from the point of recall increases.

As more items are recalled the distance between the presented items and point of recall increases. Also, participants correctly recall fewer items over time (for a review see Wixted & Rohrer, 1994). I implement this in the model by increasing the recall interval, *RI*, as a function of output position,

$$RI_n = kl^{n-2} \tag{44}$$

where n is output positions 2 to 15, k is the recall interval between the first and second output position, and *l* is a latency constant. When *l* = 1 the recall interval is equal to k and is constant across all output positions. Increasing the *l* parameter results in a larger temporal interval between output positions. This is illustrated in Figure 30 which shows the increase in recall interval as *l* varies between 1 and 1.2.

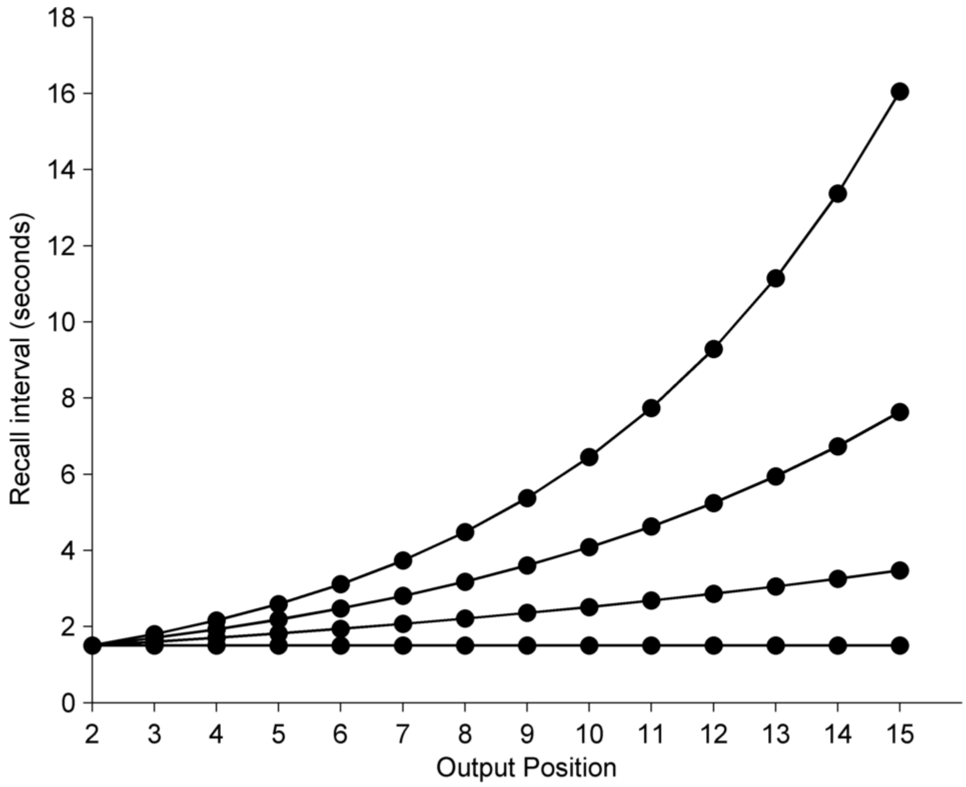

Figure 30. Change in recall interval as a function of the *l* parameters

Increasing the recall interval with each output position influence the predicted recall probability of all the items. For this example, I set *k* to 1.5 and *l* to 1.2. The presentation of the items is one per second and the retention interval is two seconds. In the SIMPLE model time is log compressed. As time passes the items all of the items appear more similar. The similarity of the items when $c = 7$ is shown in Table 30 and Table 31. The similarity of items increases as a function of the output position. This is because the temporal distance between presentation and recall for all items is larger.

Items that are recalled in later output positions are less likely to be recalled. The recall probability of each input and output position is shown in Table 32. These values were calculated using Equations 42, 43, and 44 with $t = .4$ and $s = 10$. In all of the output positions there is more recency than primacy. In other words, items in the final output positions are most likely to be recalled. This is due to the log transformation of the distances between the items. Also, the time between presentation and recall increases with output position. The log compression decreases the distinctiveness of items and makes them more confusable as time passes. The result is a decrease in the probability of recalling all items as output position increases. Note that in the time based forgetting model these output positions effects are purely due to temporal distance.

Table 30

Similarity of items at output position two

	Input														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	.66	.43	.27	.16	.09	.05	.03	.01	.01	0	0	0	0	0
2	.66	1	.65	.40	.25	.14	.08	.04	.02	.01	0	0	0	0	0
3	.43	.65	1	.63	.38	.22	.12	.07	.03	.01	.01	0	0	0	0
4	.27	.40	.63	1	.61	.35	.20	.10	.05	.02	.01	0	0	0	0
5	.16	.25	.38	.61	1	.58	.33	.17	.09	.04	.02	.01	0	0	0
6	.09	.14	.22	.35	.58	1	.56	.30	.15	.07	.03	.01	0	0	0
7	.05	.08	.12	.20	.33	.56	1	.53	.26	.12	.05	.02	.01	0	0
8	.03	.04	.07	.10	.17	.30	.53	1	.50	.23	.09	.03	.01	0	0
9	.01	.02	.03	.05	.09	.15	.26	.50	1	.46	.19	.07	.02	.01	0
10	.01	.01	.01	.02	.04	.07	.12	.23	.46	1	.42	.15	.05	.01	0
11	0	0	.01	.01	.02	.03	.05	.09	.19	.42	1	.37	.11	.03	0
12	0	0	0	0	.01	.01	.02	.03	.07	.15	.37	1	.31	.08	.01
13	0	0	0	0	0	0	.01	.01	.02	.05	.11	.31	1	.25	.04
14	0	0	0	0	0	0	0	0	.01	.01	.03	.08	.25	1	.17
15	0	0	0	0	0	0	0	0	0	0	0	.01	.04	.17	1

Table 31

Similarity of items from output position ten

	Input														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	.73	.52	.37	.25	.17	.11	.07	.05	.03	.02	.01	0	0	0
2	.73	1	.72	.50	.35	.24	.16	.10	.06	.04	.02	.01	.01	0	0
3	.52	.72	1	.70	.49	.33	.22	.14	.09	.05	.03	.02	.01	0	0
4	.37	.50	.70	1	.69	.47	.31	.20	.12	.08	.04	.02	.01	.01	0
5	.25	.35	.49	.69	1	.68	.45	.29	.18	.11	.06	.04	.02	.01	0
6	.17	.24	.33	.47	.68	1	.66	.43	.27	.16	.09	.05	.03	.01	.01
7	.11	.16	.22	.31	.45	.66	1	.64	.40	.24	.14	.08	.04	.02	.01
8	.07	.10	.14	.20	.29	.43	.64	1	.63	.38	.22	.12	.06	.03	.01
9	.05	.06	.09	.12	.18	.27	.40	.63	1	.61	.35	.20	.10	.05	.02
10	.03	.04	.05	.08	.11	.16	.24	.38	.61	1	.58	.32	.17	.08	.04
11	.02	.02	.03	.04	.06	.09	.14	.22	.35	.58	1	.56	.29	.15	.07
12	.01	.01	.02	.02	.04	.05	.08	.12	.20	.32	.56	1	.53	.26	.12
13	0	.01	.01	.01	.02	.03	.04	.06	.10	.17	.29	.53	1	.49	.23
14	0	0	0	.01	.01	.01	.02	.03	.05	.08	.15	.26	.49	1	.46
15	0	0	0	0	0	.01	.01	.01	.02	.04	.07	.12	.23	.46	1

Table 32

Probability of recall in each input and output position with time based forgetting

	Output														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.67	.63	.62	.61	.60	.59	.58	.56	.55	.53	.51	.50	.48	.46	.45
2	.63	.60	.60	.59	.58	.57	.56	.55	.54	.53	.52	.50	.49	.48	.46
3	.59	.57	.56	.56	.55	.55	.54	.53	.52	.52	.51	.50	.48	.47	.46
4	.57	.55	.54	.54	.53	.53	.52	.52	.51	.50	.49	.48	.48	.47	.46
5	.57	.54	.54	.53	.53	.52	.51	.51	.50	.49	.48	.48	.47	.46	.46
6	.58	.54	.54	.53	.53	.52	.51	.50	.49	.49	.48	.47	.46	.46	.45
7	.60	.56	.55	.54	.53	.52	.51	.50	.49	.49	.48	.47	.46	.46	.45
8	.64	.58	.57	.56	.55	.54	.52	.51	.50	.49	.48	.47	.46	.45	.45
9	.70	.62	.60	.59	.57	.56	.54	.53	.51	.50	.48	.47	.46	.45	.45
10	.77	.66	.65	.63	.61	.59	.57	.55	.53	.51	.49	.48	.47	.46	.45
11	.84	.73	.71	.68	.66	.63	.60	.58	.55	.53	.51	.49	.47	.46	.45
12	.91	.80	.78	.75	.72	.69	.65	.62	.58	.55	.53	.51	.49	.47	.46
13	.96	.88	.86	.83	.80	.77	.72	.68	.64	.60	.57	.54	.51	.49	.47
14	.99	.95	.93	.91	.89	.86	.82	.78	.73	.68	.64	.59	.55	.52	.49
15	1	.99	.98	.98	.97	.96	.94	.92	.88	.82	.76	.68	.61	.55	.50

Output Order

Our implementation of the Lewandowsky et al. (2004) model incorporates the observed output order of the recalled items. In the experiments above I recorded the output and input position of each item for every participant. For example, here I apply the output interference and time based forgetting models to data from just one participant, participant 23, in Experiment 4.

For each participant I calculate an output-input matrix. Table 33 shows this matrix for participant 23. This matrix tells us for each input position the proportion of the correct recall in for each input position. For example, when participant 23 recalled the first presented item is was recalled in the second output position 26% of the time.

Table 33

Proportion of correctly recalled items in each output position

	Output														
Input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.37	.26	.16	.05	.05	.11	0	0	0	0	0	0	0	0	0
2	.06	.28	.22	.11	.17	.06	.06	0	0	0	0	.06	0	0	0
3	.05	0	.16	.21	.05	.11	.11	.05	.16	0	.05	.05	0	0	0
4	.07	.07	0	.20	.20	0	.13	.13	0	.07	.07	0	.07	0	0
5	0	0	.25	.17	0	.25	0	.17	0	.08	.08	0	0	0	0
6	.08	.08	.08	0	0	0	.33	0	.25	0	.17	0	0	0	0
7	0	0	0	.38	0	.13	.13	.13	.13	0	0	.13	0	0	0
8	0	.11	0	0	0	.11	.22	.22	0	.11	.11	0	.11	0	0
9	0	0	0	.20	.30	.10	0	.10	.10	.10	0	0	0	.10	0
10	0	0	0	.08	.08	.31	0	.23	0	.23	.08	0	0	0	0
11	0	0	0	0	.25	0	.25	.13	0	.25	.13	0	0	0	0
12	0	0	0	0	.14	.29	0	.14	.14	.14	.14	0	0	0	0
13	.20	0	.30	.10	0	.10	.10	.10	0	.10	0	0	0	0	0
14	.25	.25	0	0	.08	.08	0	.25	0	.08	0	0	0	0	0
15	.20	.20	.13	0	.13	0	.20	0	.13	0	0	0	0	0	0

The output probabilities can be used to weight the predictions of either model. Both the output interference and time based forgetting models produce predictions for every each input and output position combination (see Table 29 and Table 32). These models predictions are weighted by the observed output probabilities (Table 33) by multiplying the corresponding probabilities together. When an item is not recalled in an input-output position then the predicted probability will be 0. Multiplying the corresponding probabilities gives the probability of the participant recalling an input-output combination *and* the model prediction.

The aim of this section is to compare the time-based forgetting and output interference accounts of the incentive effects reported above. Both models predict a decrease in recall performance as a function of output position. However, as I have shown above these models make quite different predictions. The time-based forgetting model predicts strong recency effects and a gradual decline in recall performance over time (see Table 32). The recency effect is due to the log transformation of the temporal dimension, whereas the gradual decrease in performance depends on the increase in interval between recalls. On the other hand, the output interference model predicts equal primacy and recency effects because the input position dimension is not log transformed (see Table 29).

Method

I compare three models: time-based forgetting, output interference, and both time-based forgetting and output interference. These models are all specific instances of the general model outlined above. Consequently, I can generate predictions from any of the models by changing the values of the parameters of the general model.

The general model has six parameters which I vary. The c, s, and t parameters are from the original formulation of SIMPLE (Brown et al., 2007). These parameters are allowed to vary in all of the model fitting below. The w parameter controls the “attentional weight” given to the temporal and input position dimensions.

Increases in confusability increasing with output position - which implements output interference - is determined by the *o* parameter. The rate of the increase in the durations between output positions – which contributes to time based forgetting - is controlled by the *l* parameter. Here I set the *l* parameter to 1.2 because higher values produce unrealistic increases in the inter output duration.

Table 34 shows the parameters which could vary in each model. For example, in the time-based forgetting model the attentional weight was entirely focused on the temporal dimension (w = 1), there was no output interference (o = 0) and the temporal interval between output positions gradually increased (l = 1.2). Setting the range of parameter values allowed us to compare the performance of each model.

Table 34

Range of parameter values allowed to vary in the model fitting function

Model	Parameter					
	c	*s*	*t*	*w*	*o*	*l*
Output interference	0 - 100	0 - 100	0 - 1	0	0 - 1	1
Time-based forgetting	0 - 100	0 - 100	0 - 1	1	0	1.2
Time-based forgetting and output Interference	0 - 100	0 - 100	0 - 1	0 - 1	0 - 1	1.2

To calculate the fit of the model I used a maximum likelihood estimation methodology (see Myung, 2003). In this method the `fminsearch` algorithm varied the parameter values within a pre-set range (see Table 34) to minimize the negative log likelihood (-2lnL).

These -2lnLs were calculated using the binomial probability function. The model produces a weighted input-output grid (for an example see Table 33). Taking the total of these probability predictions across output position gives us the overall predicted probability of recalling an item in each input position. Then the binomial probability function returns the probability of the observed number of correct recalls in each input position given the data.

Results

Experiment 4. In the model fitting I allowed the parameters to vary for each participant. The average best fitting parameter estimates across participants are shown in Table 35. To visualize the model fit I calculated the average of the individual level data and model predictions.

The average predictions are shown in Figure 31. The time-based forgetting model is unable to capture the strong primacy shown in the decreasing and same conditions. This is because the log transformation of the temporal dimension produces a larger recency than primacy effect. In the output interference model the position dimension is not log transformed allowing an equal primacy and recency effect to be predicted. The individual level fits are detailed in Appendices I and J.

Table 35

Average parameter estimates for model fits to data from Experiment 4

	Time-based			Output Interference				Both				
Schedule	*c*	*s*	*t*	*c*	*s*	*t*	*o*	*c*	*s*	*t*	*o*	*w*
None	8.30	29.22	.57	37.14	23.53	.63	.52	14.63	15.80	.53	.72	.50
Same	9.61	37.23	.50	35.47	43.45	.51	.76	9.64	29.78	.45	.91	.40
Increasing	22.94	11.13	.71	62.31	26.46	.64	.50	35.91	36.06	.56	.78	.66
Decreasing	4.82	32.47	.55	35.55	26.45	.49	.55	25.43	38.14	.47	.58	.35
Middle High	12.79	16.51	.45	18.70	30.94	.79	.61	23.45	18.73	.51	.56	.59
Overall	11.69	25.31	.56	37.83	30.17	.61	.59	21.81	27.70	.50	.71	.50

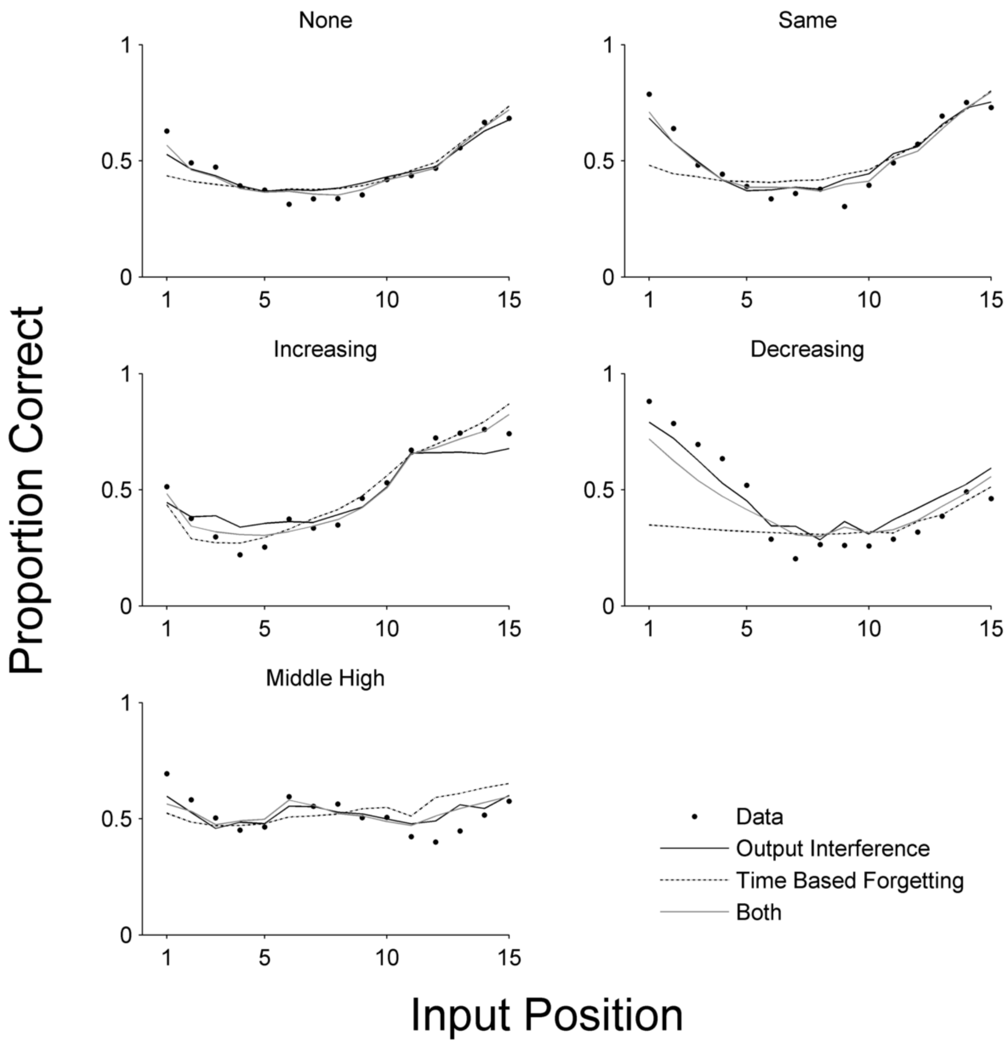

Figure 31. Mean data and model fit from Experiment 4

To compare the performance of the models I calculated two information criteria. The Bayes Information Criterion (BIC) and Akaike Information Criterion (AIC) penalize the models based on their complexity (Myung, 2000). In maximum likelihood estimation the -2lnL is measure of fit in which smaller values represent higher likelihood of the model given the data (better model fit). The BIC and AIC penalties add an additional value to the -2lnL as shown in Table 36. As a result of applying these penalties the time-based forgetting model will have an advantage because it has the fewest free parameters.

Table 36

Penalties applied by AIC and BIC based on the number of free model parameters (k) and number of observations

		Criterion	
Model	*k*	AIC	BIC
Time-based forgetting	3	6.00	8.12
Output interference	4	8.00	10.83
Time and output interference	5	10.00	13.54

Comparison of the models was carried out using the criterion weights. The AICw and BICw give an easily understood measure of the relative performance of a model given the data. The weights in a comparison will add up to 1. Higher values of the weights suggest that the data favor that model.

The criterion weights in Table 37 show that overall the data support the output-interference model. However, in the increasing and middle high conditions the data support the time-based forgetting model. As mentioned above, the time-based forgetting model can predict stronger recency effects which are seen in the increasing incentive condition. Despite task specific differences the data support the output-interference account.

Table 37

Criterion weights for each model based on the data collected in Experiment 4. Note: the highest weight in a comparison is indicated in bold

	AICw			BICw		
	Time	Output	Both	Time	Output	Both
None	.31	**.36**	.32	.37	**.36**	.27
Same	.13	**.53**	.35	.16	**.55**	.29
Increasing	**.38**	.27	.36	**.45**	.27	.28
Decreasing	.15	**.53**	.32	.18	**.56**	.26
Middle High	**.43**	.41	.16	**.48**	.41	.11
Overall	.28	**.42**	.30	.33	**.43**	.24

Experiment 5: Precued fast. Next I compared the models using data from Experiment 5 in which the items were presented at a faster rate. The average best fitting parameter estimates are shown in Table 38. The mean data and model predictions are shown in Figure 32. The time based forgetting model appears unable to capture the primacy effects in the decreasing and same conditions. This is because the temporal dimension is log compressed in the time based model and this produces large recency effects but small primacy effects. On the other hand, the position dimension in the output interference model is not log compressed and can produce symmetrical serial position curves.

Table 38

Average parameter estimates for model fits to data from Experiment 5

	Time-based Forgetting			Output Interference				Both				
Schedule	*c*	*s*	*t*	*c*	*s*	*t*	*o*	*c*	*s*	*t*	*o*	*w*
None	12.61	9.50	.82	31.63	22.63	.78	.64	14.06	19.20	.49	.80	.50
Same	8.04	25.49	.53	34.55	39.16	.60	.67	14.22	29.55	.46	.86	.70
Increasing	17.46	20.13	.64	40.33	27.99	.43	.58	26.71	19.63	.56	.78	.76
Decreasing	5.32	42.13	.46	50.22	29.43	.61	.49	30.59	40.88	.42	.48	.29
Middle High	3.90	50.81	.33	35.13	26.00	.61	.41	19.15	20.65	.44	.51	.46
Overall	9.47	29.61	.56	38.37	29.04	.61	.56	20.95	25.98	.48	.69	.54

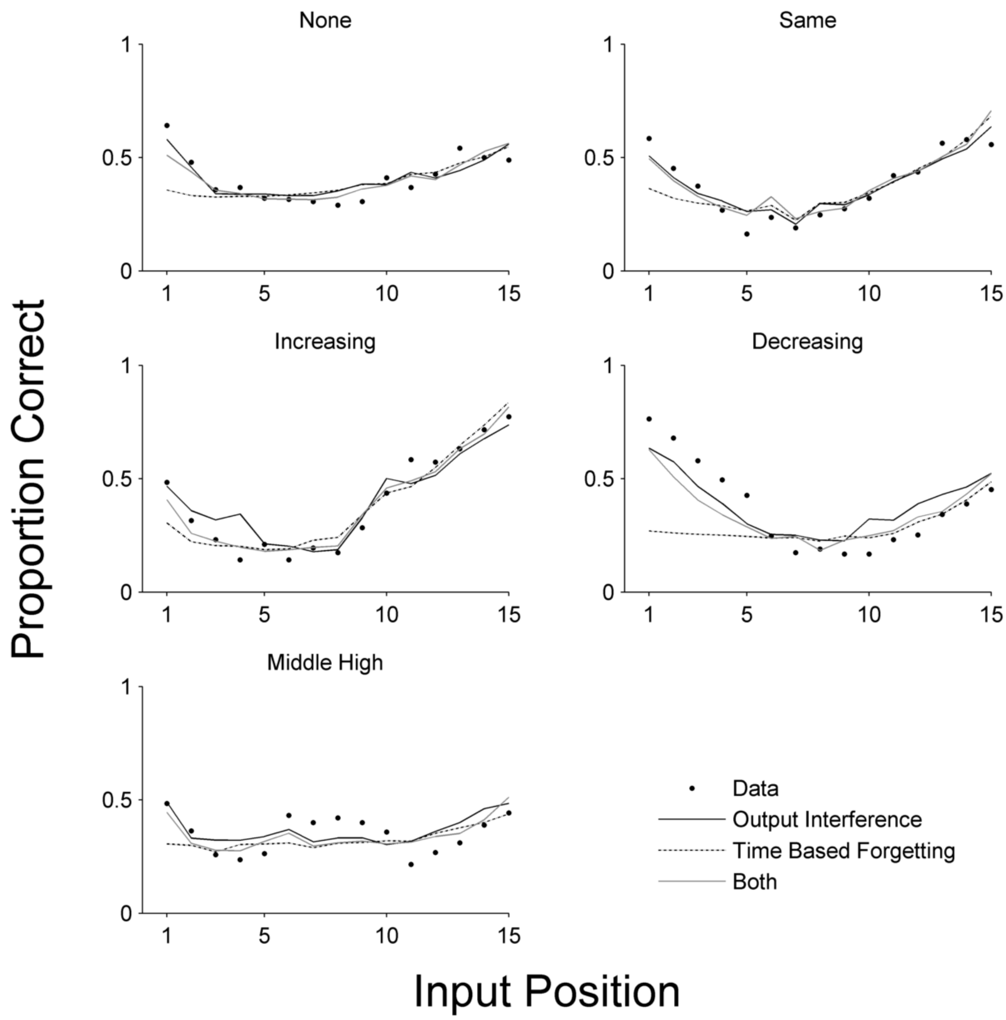

Figure 32. Mean data and model fit from Experiment 5

Again, I used criterion weights to gauge the extent to which the data supported each of the models. The criterion weights are shown in Table 39. These criterion weights suggest that both the time-based forgetting and output interference models are equally consistent with the data.

Table 39

Criterion weights for each model based on the fit of each model to data from Experiment 5

	AICw			BICw		
Schedule	Time	Output	Both	Time	Output	Both
None	**.34**	**.34**	.32	**.41**	.34	.25
Same	**.41**	.31	.28	.41	**.44**	.16
Increasing	**.42**	.33	.25	.41	**.44**	.16
Decreasing	.20	**.49**	.31	.44	**.45**	.11
Middle High	**.45**	.39	.17	**.45**	.44	.11
Overall	**.36**	**.37**	.27	**.42**	**.42**	.16

Experiment 6: Postcued. In this experiment the payment schedule was manipulated within participant. To incorporate this into the model fitting procedure I kept the model parameters constant across all of the responses of one participant. The average best fitting parameter values are shown in Table 40. These parameter values are generally consistent with the values in the previous model fits. However, for some participants the *s* parameter was at the highest possible value suggesting some difficulty in fitting the data. The comparative performance of the models is shown in Figure 33.

Table 40

Average parameter estimates across participants from Experiment 6

	Parameter				
Model	*c*	*s*	*t*	*o*	*w*
Time-based Forgetting	6.76	50.02	.47	-	-
Output Interference	16.16	21.47	.53	.59	-
Both	9.66	27.43	.41	.76	.70

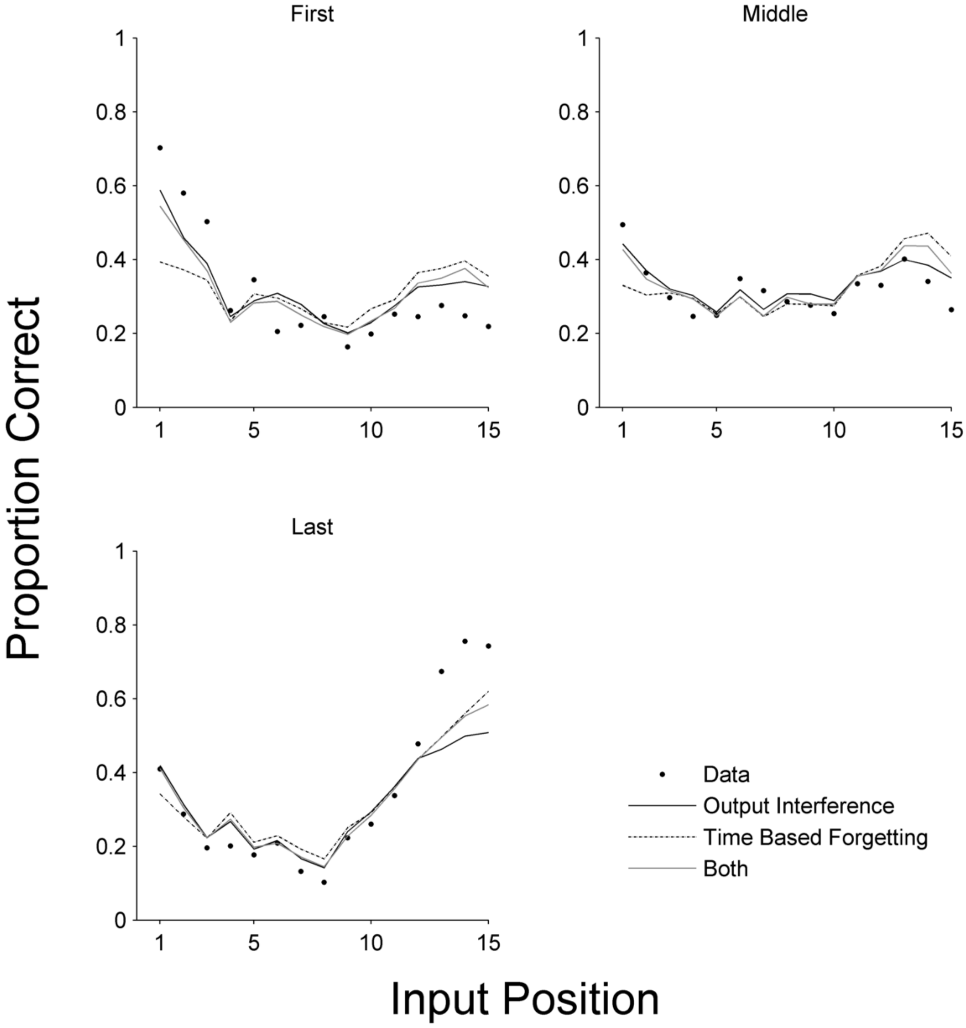

Figure 33. Mean data and model fit from Experiment 6

Which model is most likely given the data? To compare the models I again calculated criterion weights. Table 41 show us that the data favors the output interference model. As we would expect, this difference is greater in the AICw because AIC is a more lenient penalty of complexity (Lewandowsky & Farrell, 2010).

Table 41

Criterion weights for model comparison based on data from Experiment 6

		Model	
Weight	Time	Output	Both
AICw	.33	**.42**	.25
BICw	.40	**.43**	.17

General Discussion

In this study I first examined how monetary incentives can be used to direct free recall. Experiment 4 showed that items associated with monetary incentives are more likely to be recalled. This effect was also shown in Experiments 5 and 6. In all of these experiment the participants generally recalled high incentive items first and there was a strong sequential effect in recall (e.g., the fifth item was recalled after the fourth item).

Our financial incentive methodology has many advantages. First, the experimenter does not decide the order of recall. In both serial recall and directed free recall tasks the output order of items is prescribed by the experimenter. The findings may represent a more naturalistic strategy in recall. This is particularly pertinent given the debate encouraging more "naturalistic" approaches to memory research (Koriat & Goldsmith, 1996). Second, the approach offers a link between literature in behavioral economics and free recall. The use of incentives to direct behavior is widely applied in behavioral economics (Camerer, Loewenstein, & Rabin, 2011). In these tasks participants often adopt strategies to maximize their financial return.

Our findings suggest an advantage of the first and last items irrespective of output position. When items in the middle of the item list were highly incentivized the serial position curve was more flat despite a tendency to recall these items first. These findings are consistent with the local distinctiveness principles of the SIMPLE model.

There was a marked effect of output position in the primacy and recency effects. In the experiments the primacy and recency effects were amplified if the first or last few items was associated with the highest incentive. When recalled, these items were most likely to be recalled in the first few output positions. SIMPLE predicts large recency effect and small primacy effects.

Using model based analysis I showed that SIMPLE can be applied to the atypical data. I found that a pure output Interference account was slightly more supported by the data than time-based forgetting. although this varied depending on the incentive condition.

Chapter 6: General Discussion

Review of Rationale and Aims

This thesis has been concerned with the interplay between memory, judgment and decision making. A critical question running throughout has been: to what extent are memory processes involved in decision making? In some decisions memory has a minimal influence in decision making (Bechara & Martin, 2004). Yet in other areas memory and decision making are closely linked (Hinson, Jameson, & Whitney, 2003).

This thesis has focused on decision-making phenomena related to relative and retrospective judgments. Participants often make subjective judgments in response to a context of events. In relative judgment these events may be stimuli similar to the one being judged, as when the "largeness" of a square is judged with knowledge of other squares in an experiment (Parducci, 1965). A large literature from psychophysics has investigated these effects for the past 60 years (Helson, 1964a; Parducci, 1995). Retrospective judgments are summary responses to a series of events (for a review, see Kahneman, Wakker, & Sarin, 1997). In this case, the contextual events are spread out over time, as when a summary judgment of the discomfort of a series of sounds is made (Schreiber & Kahneman, 2000).

Theoretical accounts of both phenomena acknowledge the role of memory in the formation of judgments. Theories of relative judgment such as adaptation level theory (ALT; Helson, 1964a) and range-frequency theory (RFT; Parducci, 1965) incorporate memory into their models of judgment. The "snapshot" and "prototype" theories of retrospective judgment suggest that memory is central to peak end effects (Fredrickson & Kahneman, 1993; Kahneman, 2000).

All of these models, however, suffer from a potentially serious limitation: They are typically not explicit in their account of the relevant memory processes.

Rather, each of the models is designed to account for a specific phenomenon. In ALT it is the combination of multiple sources of information into a single adaptation level (Helson, 1938; Helson & Jeffers, 1940). In the “snapshot” theory of retrospective judgment the focus is on the impact of the peak and end components of an experience (Fredrickson & Kahneman, 1993). These ad hoc models of memory processes are not easily extended to phenomena in other areas of psychology such as serial position effects in memory (Murdock, 1962) or economic judgments (Kahneman & Tversky, 1979).

In this thesis I examined the role of memory processes in relative and retrospective judgments. To overcome the potential limitation of models with a narrow empirical focus I applied more general theories of memory and decision making. Specifically, I examined two models. The first model is the SIMPLE model of memory (Brown et al., 2007). In SIMPLE, recall is akin to a discrimination task along a dimension of interest. SIMPLE has been widely applied to various memory phenomena (Lewandowsky et al., 2004). Notably, it has been applied extensively to serial position curves. In retrospective judgments there is evidence of recency, primacy and distinctiveness (Langer et al., 2005; Miron-Shatz, 2009). All three of these memory effects can be accounted for in SIMPLE. The second model is Decision by Sampling (DbS; Stewart et al., 2006). In the model a decision is based on a small sample of items drawn from memory. The model has been successfully applied to both economic decision making (Stewart, 2009; Stewart & Simpson, 2008) and social judgments (Melrose et al., 2012; Moore et al., 2012; Wood et al., 2011). Both of these models are able to explain many phenomena in their respective literatures.

The combined SIMPLE and DbS model has several theoretical advantages. First, it has the potential to offer an internally consistent account of memory, decision making and relative judgment phenomenon. Second, the underlying mechanisms involved in making relative judgments can be made explicit. Third, the model will benefit from future theoretical development in a wider area of study. The components of the model (e.g., SIMPLE) are actively being examined within their specific field of study. For example, modifications of the SIMPLE model in response to new data in the memory literature can be applied to the combined model. For all of these reasons a combined model made up of SIMPLE and DbS is desirable. A key aim of this thesis has been to examine the extent to which this combined model can account for relative and retrospective judgment data.

In Chapter 2 I examined a link between relative and retrospective judgment. According to RFT, a negatively and positively skewed distribution of items should attract different subjective judgments (Parducci, 1965). This prediction results from the model's assumption that judgments will be determined by a weighted average of the range and rank position of each item. RFT predicts that the average judgment will differ as a function of skew: The average rating of stimuli will be higher in negatively skewed distributions. Parducci (1995) goes further and suggests that overall ratings (i.e., of the whole series) will differ. Smith et al. (1989) asked participants to rate the happiness they would expect a waitress to experience for a positively or negatively skewed distribution of hypothetical payments. The data matched Parducci's prediction: participants expected a waitress to be happier with a negatively skewed distribution of tips. These findings suggest a link between relative and retrospective judgments.

To investigate this link between relative and summary judgment I replicated and extended the classic experiment of Parducci (1968). Participants received a series of payments from either a negatively or positively skewed distribution. I extended Parducci (1968) by asking participants to rate their satisfaction after receiving all of the payments in the distribution. I found that satisfaction was higher for negatively skewed distributions. Like peak end studies (Redelmeier & Kahneman, 1996), the experiments in Chapter 2 asked participants to make hedonic judgments. These findings demonstrated that theories of relative judgments are applicable to summary judgments.

Model comparison confirmed that relative judgments effects were present at an individual level. I used maximum likelihood estimation to compare range-only, frequency-only and RFT models of judgment at an individual level using data from Experiment 1. One finding from the study was particularly relevant: None of the participants were best fit by the frequency principle. On its own the DbS model produces predictions which are the same as the frequency principle. It appears that the DbS model performs worse than the RFT.

These findings suggested that DbS alone does not predict the judgment effects found in Chapter 2. Also, DbS did not predict the overall effects on judgment reported in Chapter 2. Table 42 shows that DbS alone (i.e., the frequency principle) cannot account for the overall effect of negative skew on summary judgments. If we assume that the sample of items in memory consists of the experimental stimuli and these items are equally weighted, then DbS alone predicts no effect of skew at all. The effect of skew is due to the inclusion of the range principle in the model.

Table 42

Average rating predicted by RFT for the skewed distributions presented to participants in Chapter 2

	w		
	0	*.5*	*1*
	Decision by sampling	Range-frequency Theory	Range only
Negative Skew	.50	.58	.65
Positive Skew	.50	.43	.35

Summary ratings were influenced by the penultimate ratings. There was a strong correlation between the summary judgment and the last payment judgment in Experiment 1. RFT does not predict this finding. However, the finding is consistent with both the peak-end rule and SIMPLE. According to the peak-end rule the last experience of a series has a large impact on summary judgments. On the other hand, SIMPLE can predict this effect due to local distinctiveness along the temporal dimension.

However, there were findings which are difficult to explain in terms of the SIMPLE model. Experiment 3 required participants to recall as many payments as they remembered receiving. I calculated the skewness of this recalled distribution. The analysis shows that participants overestimated the skewness of the negatively skewed distribution and accurately recalled the skewness of the positively skewed distribution. SIMPLE predicts that the skewness of the distribution should be underestimated in both conditions. Yet it is also difficult to explain these effects in terms of the peak-end or range-frequency theories.

Chapter 2 showed a link between relative and retrospective judgments. One of the main findings was that the DbS model could not fit the individual level data. It may be the case that DbS cannot predict relative judgments. In Chapter 3 I examined whether similarity between items does influence relative judgment.

Many models of memory predict that similarity influences recall. In other words, when given a cue, items similar to the cue are more accessible in memory (Hintzman, 1984). Applied to relative judgment it may be the case that items similar to the stimuli being judged are more accessible in memory and therefore more heavily weighted in judgment. The effect of accessibility on judgment is well known (N. Schwarz et al., 1991; Tversky & Kahneman, 1973). The generalized exemplar model of sampling (GEMS) implements distance based sampling using the frequency principle. Items are weighted in judgment based on their distance to the judged item. In the GEMS model RFT is a special case. I used the GEMS model to compare RFT with and without distance based sampling at an individual level.

The analysis was divided into two sections. To compare the models I used data from five previously reported studies. In these studies stimuli were presented sequentially to the participants. In the first section I compared the RFT to the range principle and frequency principle. Surprisingly, likelihood ratio tests found that RFT best fit only 20% of participants overall, although bootstrap analysis showed that the RFT parameters ($0<w<1$) were within the 95% confidence intervals for the bootstrap samples for 40% of participants. In the second section I used the GEMS model to compare RFT with and without distance based sampling. The results of this model comparison strongly supported RFT without distance based sampling.

These findings have two implications for the thesis as a whole. First, the performance of the GEMS based model suggests that similarity has little influence on relative judgments. The absence of distance based sampling goes against previous reported findings in price psychophysics (Qian & Brown, 2005). This finding also suggests that we cannot easily apply a memory model in which similarity is a central component of recall (Hintzman, 1984). I conclude that there is little evidence for similarity based sampling in relative judgment. Second, the poor fit of RFT to the data is unexpected. One possibility is that with simultaneously presented items other cognitive processes may underpin judgment. One possible account is SIMPLE. Unlike other models of memory, SIMPLE can be applied to distinctiveness along any dimension of interest.

Chapter 4 examined the ability of a combined SIMPLE and DbS model to account for the relative judgment effects examined in Chapter 3. In Chapter 3 RFT performed poorly at an individual level. A meta-analysis of the data found range effects at an individual level, however. An important point is that a pure relative rank based account of the data did not outperform RFT. To examine if DbS can predict these effects when combined with a memory model I applied a combined SIMPLE and DbS model (SDbS) to the data.

Our findings show that RFT and SDbS fit the data equally well. Maximum likelihood estimation showed that for most participants neither model was most likely. Brown and Matthews (2011) also showed that both models can produce the same predictions, albeit on a more limited dataset and without comparing model fits at an individual level.

This result is particularly important because in SDbS responses are due to rank comparison. Range effects are due to the local distinctiveness of items. The results make a theoretical contribution because the latter process appears to underpin behavior in perception and memory (Brown et al, 2008).

However, when complexity penalties were applied to the model fits RFT outperformed the combined model. One interpretation of these findings is that the combined model should be rejected and RFT accepted. Such an approach does not consider the scope of the models under comparison. RFT can account several phenomena outside the domain of relative judgment, such as binary categorization (Wedell, 2008) and price perception (Niedrich et al., 2001). On the other hand, SIMPLE and DbS have been widely applied in many different fields (Brown et al., 2008; Brown et al., 2007; Stewart et al., 2006; Wood et al., 2011).

Further analysis found that SDbS accounts for range effects in the data. SDbS was compared with and without the range principle. The findings showed adding a range component did not improve the performance of the model. Taken together the model comparison indicates that the previous reported range effects can be captured by a local distinctiveness process also found in other areas of Psychology. Based on this analysis I concluded that SDbS should be preferred because it offers a comprehensive account of relative judgment based a rank comparison processes (as found in the economic literature) and local distinctiveness (as reported in the perception and memory literature). In other words, SDbS can produce the predictions of RFT and is supported by convergent evidence from other literatures.

In summary, Chapter 4 demonstrated SDbS can predict range effects. An advantage of SDbS is that it can be applied to memory phenomena which may influence decision making.

In Chapter 4 I argue that local distinctiveness offer a common explanation which links the memory, perception and relative judgment literature. Chapter 5 addresses a potential limitation of the SIMPLE model – that SIMPLE can only predict normal serial position curves and is unable to predict atypical recall performance. If local distinctiveness is a common mechanism underpinning perceptual and judgment tasks then SIMPLE must be able to account for the serial position curves resulting from widely reported phenomena such as output interference.

In Chapter 5 I investigated the relationship between serial position and output interference. Both of these are important for judgment and decision making. Recency may play a role in peak end effects, and forgetting due to output interference is an important component of theories such as Query Theory (Johnson et al., 2007). To investigate the relationship between serial position and output interference I used a novel methodology. In three experiments I used monetary incentives to direct free recall. Differential incentives were associated with each third of a 15 word list. The payment each participant received depended on the words they recalled. For example, a participant recalling three words from the first five words might receive three pence per word.

I found that associating payment with serial position influenced recall performance. Items associated with high payment were more likely to be recalled in earlier output positions. These items were also more likely to be recalled correctly. The effect of payment on recall was present when items were presented slowly (1s), quickly (0.5s) and when the payment schedule was presented after encoding. These results show that economic incentives can applied to direct free recall behavior.

Monetary incentives influenced the shape of the serial position curve. Associating a high incentive with the last few items resulted in a larger recency effect. When high incentives were associated with the first few items an increased primacy effect was observed. Interestingly, increasing the incentive associated with the middle items produced a flatter serial position curve.

Next, the SIMPLE model was used to investigate the impact of output position on recall performance. I used an extended version of SIMPLE to examine the evidence in the data of output interference and time based forgetting Lewandowsky et al. (2004). In both accounts later output positions are generally associated with decreased recall performance. According to time based forgetting, recall performance decreases as a function of time from initial presentation. This approach is consistent with trace decay theories of memory, as well as with temporal distinctiveness accounts such as SIMPLE. The version I implemented gradually increased the time between output positions to simulate the decrease in recall output observed in studies of memory (Wixted & Rohrer, 1994). On the other hand, output interference states that the act of recalling items will interfere with subsequent recall attempts.

Model comparison generally favored an output interference account. I compared both output interference and time-based forgetting at an individual level. The findings show that for most incentive manipulations the output interference model performed best. Incorporating output put interference in SIMPLE suggests that it may provide a process-level account of some of the key mechanisms theories of judgment such as Query Theory (Johnson et al., 2007).

Summary

This thesis examined the relationship between memory and decision making. I focused on relative judgment. It was shown that relative and retrospective judgment may be linked, and that SDbS is applicable to both phenomena. Taken together, the results suggest that memory and decision making are closely linked, and models in one area are directly applicable to the other.

Future Studies

In the first future study SDbS could be fit to frequency based manipulations. Many of the studies in the RFT literature altered the frequency of items in the range of stimuli (Parducci, 1968; Parducci et al., 1969; Riskey et al., 1979). Chapter 2 replicated one of these studies. In order to validate the conclusions of Chapter 5 SDbS and RFT could be fit data from frequency manipulations. The results of this fit would be particularly interesting from a memory perspective. In the RFT there is no speculation about the role of rehearsal in decision making. A possible application of the combined model would involve weighting an item in DbS based on its frequency in the model.

Another possible study could extend the model to sequential judgments. SIMPLE has been applied to serial position effects (Brown et al., 2007). In studies of both retrospective and relative judgment stimuli are sometimes presented sequentially. In SIMPLE one could model these effects using a weighted compromise of the temporal and magnitude position of a stimulus. The implementation of the model would be similar to the multidimensional approach adopted in Chapter 5.

Extending the model to the temporal dimension would allow more subtle manipulations of temporal effects in judgment. Studies in free recall show that items which are temporally isolated are more likely to be recalled (Lewandowsky et al., 2008). The combined model includes SIMPLE which predicts temporal isolation effects. Other models of relative judgment do not appear to predict temporal isolation effects (Haubensak, 1992; Petzold & Haubensak, 2004). Using temporal isolation I could directly test the combined SIMPLE and DbS model.

Finally, future work could examine the DbS component of the combined model. DbS assumes that values in the environment influence subjective judgments. For example, Ungemach et al. (2011) found that lottery choices were influenced by incidental payments at a supermarket, and Stewart (2009) showed a link between choices and distributions of economic quantities in the environment.

Model of relative judgment such as the consistency model (Haubensak, 1992) and RFT are restricted to the immediate context in decision making. I briefly suggested in Chapter 2 that individual differences in responses could result from a person sampling from prior distributions. We could elicit a participant's external distribution using a percentile task (Melrose et al., 2012; Moore et al., 2012; Wood et al., 2011). Participants would then carry out a standard relative judgment task. The extent to which people are sampling from a combination of their subjective distribution and the experimental stimuli could be estimated using simulations or individual level model analysis.

The aim of the present thesis was to investigate the relationship between memory and judgment. The findings suggest that relative judgment can be investigated using the same framework applied to economic decision-making, free recall and social judgment. The results of the studies above show that (a) relative and retrospective judgments are linked, (b) similarity does not influence relative judgment, (c) relative judgment can be predicted from a combined memory and decision making model, and (d) both recency and output interference effects can be investigated using this combined account. Taken together, I have demonstrated that memory and decision making are indeed intertwined. Utilizing theoretical developments in both areas has allowed us to contribute to the unification of disparate literatures.

References

Alesina, A., Di Tella, R., & MacCulloch, R. (2004). Inequality and happiness: Are Europeans and Americans different? *Journal of Public Economics, 88*(9), 2009-2042.

Appley, M. H. (1971). *Adaptation-level theory: A symposium*. New York: Academic Press.

Ariely, D. (1998). Combining experiences over time: The effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. *Journal of Behavioral Decision Making*, *11*(1), 19-45.

Ariely, D., & Carmon, Z. (2000). Gestalt characteristics of experiences: The defining features of summarized events. *Journal of Behavioral Decision Making, 13*(2), 191-201.

Ariely, D., & Zauberman, G. (2000). On the making of an experience: The effects of breaking and combining experiences on their overall evaluation. *Journal of Behavioral Decision Making*, *13*(2), 219-232.

Ariely, D., & Zauberman, G. (2003). Differential partitioning of extended experiences. *Organizational Behavior and Human Decision Processes, 91*(2), 128-139.

Atkinson, R. C., & Crothers, E. J. (1964). A comparison of paired-associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology, 1*(2), 285-315.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*(2), 82-90.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A. (1994). The magical number seven: still magic after all these years? *Psychological Review, 101*(2), 353-356.

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 65*(2), 272-281.

Beaman, C. P., & Morton, J. (2000). The separate but related origins of the recency effect and the modality effect in free recall. *Cognition, 77*(3), B59-B65.

Bechara, A., & Martin, E. M. (2004). Impaired decision making related to working memory deficits in individuals with substance addictions. *Neuropsychology, 18*(1), 152-162.

Berridge, K. C., & Kringelbach, M. L. (2011). Building a neuroscience of pleasure and well-being. *Psychology of Well-being, 1*(1), 1-26.

Bhatarah, P., Ward, G., & Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: The serial nature of recall and the effect of test expectancy. *Memory & Cognition, 36*(1), 20-34.

Birnbaum, M. H., Parducci, A., & Gifford, R. K. (1971). Contextual effects in information integration. *Journal of Experimental Psychology, 88*(2), 158-170.

Boyce, C. J., Brown, G. D. A., & Moore, S. C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological Science, 21*(4), 471-475.

Brickman, P., & Campbell, D. T. (1971). Hedonic relativism and planning the good society. In M. H. Appley (Ed.), *Adaptation-level theory: A symposium* (pp. 287-303). New York: Academic Press.

Briesch, R. A., Krishnamurthi, L., Mazumdar, T., & Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research, 24*(2), 202-214.

Brodie, D. A., & Murdock, B. B. (1977). Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior, 16*(2), 185-200.

Brown, G. D. A., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does wage rank affect employees' well‐being? *Industrial Relations: A Journal of Economy and Society, 47*(3), 355-389.

Brown, G. D. A., & Matthews, W. J. (2011). Decision by sampling and memory distinctiveness: Range effects from rank-based models of judgment and choice. *Frontiers in Psychology, 2*.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*(3), 539-576.

Burke, C. J., & Tobler, P. N. (2011a). Coding of reward probability and risk by single neurons in animals. *Frontiers in Neuroscience, 5*(121).

Burke, C. J., & Tobler, P. N. (2011b). Reward skewness coding in the insula independent of probability and loss. *Journal of Neurophysiology, 106*(5), 2415-2422.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: a practical information-theoretic approach*. New York: Springer-verlag.

Camerer, C. F., Loewenstein, G., & Rabin, M. (2011). *Advances in behavioral economics*. New York: Russell Sage Foundation.

Caraco, T., & Chasin, M. (1984). Foraging preferences: response to reward skew. *Animal Behaviour, 32*(1), 76-85.

Clark, A. E., & Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics, 61*(3), 359-381.

Coates, S. L., Butler, L. T., & Berry, D. C. (2006). Implicit memory and consumer choice: The mediating role of brand familiarity. *Applied Cognitive Psychology, 20*(8), 1101-1116.

Coombs, C. H., & Pruitt, D. G. (1960). Components of risk in decision making: Probability and variance preferences. *Journal of Experimental Psychology, 60*(5), 265-277.

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding serial recall. *Journal of Memory and Language, 46*(1), 153-177.

Crowder, R. G. (1976). *Principles of learning and memory*. Oxford, England: Lawrence Erlbaum Associates.

Cuvo, A. J. (1974). Incentive level influence on overt rehearsal and free recall as a function of age. *Journal of Experimental Child Psychology, 18*(1), 167-181.

Dalezman, J. J. (1976). Effects of output order on immediate, delayed, and final recall performance. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 597-608.

Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology, 54*(3), 180-187.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review, 106*(1), 180-209.

Eysenck, M. W., & Eysenck, M. C. (1982). Effects of incentive on cued recall. *The Quarterly Journal of Experimental Psychology, 34*(4), 489-498.

Farrell, S. (2010). Dissociating conditional recency in immediate and delayed free recall: A challenge for unitary models of recency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 324-347.

Fiedler, K., & Juslin, P. (2006). *Information sampling and adaptive cognition*. Cambridge, UK: Cambridge University Press.

Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology*, *62*, 559-582.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906-011.

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1594), 1338-1349.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, *65*(1), 45-55.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*(4), 375-399.

Garrett, T. A., & Sobel, R. S. (1999). Gamblers favor skewness, not risk: Further evidence from United States' lottery games. *Economics Letters*, *63*(1), 85-90.

Gerhart, B., & Rynes, S. L. (2003). *Compensation: Theory, evidence, and strategic implications*. London: SAGE Publications.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward The misprediction of regret. *Psychological Science, 15*(5), 346-350.

Golec, J., & Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy, 106*(1), 205-225.

Greene, R. L., Crowder, R. G., Sherry, D. F., Glenberg, A. M., Bradley, M. M., Kraus, T. A., . . . Toppino, T. C. (1986). Sources of recency effects in free recall. *Psychological Bulletin, 99*(2), 221-228.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27.

Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology, 78*(4), 764-771.

Haines, D. J., & Torgesen, J. K. (1979). The effects of incentives on rehearsal and short-term memory in children with reading problems. *Learning Disability Quarterly, 2*(2), 48-55.

Hamann, S. B., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience, 2*(3), 289-293.

Hartley, J. T., & Walsh, D. A. (1980). The effect of monetary incentive on amount and rate of free recall in older and younger adults. *Journal of Gerontology, 35*(6), 899-905.

Harvey, A. J., & Beaman, C. P. (2007). Input and output modality effects in immediate serial recall. *Memory, 15*(7), 693-700.

Haubensak, G. (1992). The consistency model: A process model for absolute judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 303-309.

Healy, A. F., Havas, D. A., & Parker, J. T. (2000). Comparing serial position effects in semantic and episodic memory using reconstruction of order tasks. *Journal of Memory and Language, 42*(2), 147-167.

Helson, H. (1938). Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination. *Journal of Experimental Psychology, 23*(5), 439-476.

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *The American Journal of Psychology, 60*(1), 1-29.

Helson, H. (1964a). *Adaptation-level theory*. New York: Harper & Row.

Helson, H. (1964b). Current trend and issues in adaptation-level theory. *American Psychologist, 19*(1), 26-38.

Helson, H., & Jeffers, V. B. (1940). Fundamental problems in color vision. II. Hue, lightness, and saturation of selective samples in chromatic illumination. *Journal of Experimental Psychology, 26*(1), 1-27.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15*(8), 534-539.

Highhouse, S., Luong, A., & Sarkar-Barney, S. (1999). Research design, measurement, and effects of attribute range on job choice: More than meets the eye. *Organizational Research Methods, 2*(1), 37-48.

Hill, R. D., Storandt, M., & Simeone, C. (1990). The effects of memory skills training and incentives on free recall in older learners. *Journal of Gerontology, 45*(6), P227-P232.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology. Learning, memory, and Cognition, 29*(2), 298-306.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16*(2), 96-101.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*(4), 528-551.

Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition, 3*(2), 197-209.

Johnson, E., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 461-474.

Johnson, E., & Weber, E. (2009). Mindful judgment and decision making. *Annual Review of Psychology, 60*, 53-85.

Juslin, P., & Persson, M. (2002). PROBabilities from EXemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science, 26*(5), 563-607.

Juslin, P., Winman, A., & Hansson, P. (2007). The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals. *Psychological Review, 114*(3), 678-703.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24*(1), 103-109.

Kahneman, D. (2000). Evaluation by moments: Past and future. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 693-708). New York: Cambridge University Press and the Russell Sage Foundation.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263-291.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, *112*(2), 375-406.

Kermer, D. A., Driver-Linn, E., Wilson, T. D., & Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychological Science, 17*(8), 649-653.

Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review, 1*(6), 602-609.

Koriat, A., & Goldsmith, M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences*, *19*(2), 167-187.

Kunzinger III, E. L., & Witryol, S. L. (1984). The effects of differential incentives on second-grade rehearsal and free recall. *The Journal of Genetic Psychology, 144*(1), 19-30.

Laming, D. (1997). *The measurement of sensation*. New York: Oxford University Press.

Laming, D. (2010). Serial position curves in free recall. *Psychological Review, 117*(1), 93-133.

Langer, T., Sarin, R., & Weber, M. (2005). The retrospective evaluation of payment sequences: Duration neglect and peak-and-end effects. *Journal of Economic Behavior & Organization, 58*(1), 157-175.

Lansdale, M. W. (1998). Modeling memory for absolute location. *Psychological Review*, *105*(2), 351-378.

Lee, M. D., & Pooley, J. P. (2012). Correcting the SIMPLE model of free recall. *Psychological Review, 120*(1), 293-296.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11*(5), 771-790.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *Psychology of Learning and Motivation, 49*, 1-48.

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. London: Sage.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, *96*(1), 25-57.

Lewandowsky, S., Nimmo, L. M., & Brown, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language, 58*(2), 415-428.

Lindskog, M., Winman, A., & Juslin, P. (2013). Naïve point estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 782-800.

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics, 120*(3), 963-1002.

Maltby, J., Wood, A. M., Vlaev, I., Taylor, M. J., & Brown, G. D. A. (2012). Contextual effects on the perceived health benefits of exercise: The exercise rank hypothesis. *Journal of Sport and Exercise Psychology, 34*(6), 828-841.

Mazumdar, T., Raj, S. P., & Sinha, I. (2005). Reference price research: review and propositions. *Journal of Marketing*, 84-102.

Melrose, K. L., Brown, G. D. A., & Wood, A. M. (2012). Am I abnormal? Relative rank and social norm effects in judgments of anxiety and depression symptom severity. *Journal of Behavioral Decision Making*, *26*(2), 174-184.

Michels, W. C., & Helson, H. (1949). A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data. *The American Journal of Psychology*, *62*(3), 355-368.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Miron-Shatz, T. (2009). Evaluating multiepisode events: Boundary conditions for the peak-end rule. *Emotion*, *9*(2), 206-213.

Moore, S. C., Wood, A. M., Brown, G. D. A., & Shepherd, J. (2012). Relatively drunk: subjective intoxication and estimated health consequences of alcohol consumption are conditional on the presence of less intoxicated individuals, not level of intoxication. *Addiction Science & Clinical Practice*, *7*(Suppl 1), A29.

Morewedge, C. K., Kassam, K. S., Hsee, C. K., & Caruso, E. M. (2009). Duration sensitivity depends on stimulus familiarity. *Journal of Experimental Psychology: General*, *138*(2), 177-186.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482-488.

Murdock, B. B. (1974). *Human memory: Theory and data*. New York: Lawrence Erlbaum Associates.

Murdock, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory & Cognition, 23*(5), 631-645.

Murray, D. J. (1993). A perspective for viewing the history of psychophysics. *Behavioral and Brain Sciences, 16*(1), 115-186.

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology, 44*(1), 190-204.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology, 47*(1), 90-100.

Neale, M. A., Huber, V. L., & Northcraft, G. B. (1987). The framing of negotiations: Contextual versus task frames. *Organizational Behavior and Human Decision Processes, 39*(2), 228-241.

Neath, I., & Crowder, R. G. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 316-327.

Niedrich, R. W., Sharma, S., & Wedell, D. H. (2001). Reference price and price perceptions: A comparison of alternative models. *Journal of Consumer Research*, *28*(3), 339-354.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39-57.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*(3), 544-576.

Oswald, A. J., & Wu, S. (2010). Objective confirmation of subjective measures of human well-being: Evidence from the USA. *Science, 327*(5965), 576-579.

Page, M. P. A., & Morris, D. (1998). The Primacy Model: A new model of immediate serial recall. *Psychological Review, 105*(4), 761-781.

Parducci, A. (1956). Direction of shift in the judgment of single stimuli. *Journal of Experimental Psychology, 51*(3), 169-178.

Parducci, A. (1963). Range-frequency compromise in judgment. *Psychological Monographs: General and Applied, 77*(2), 1-50.

Parducci, A. (1965). Category judgment: a range-frequency model. *Psychological Review, 72*(6), 407-418.

Parducci, A. (1968). The relativism of absolute judgments. *Scientific American, 219*(6), 84-90.

Parducci, A. (1982). Category ratings: Still more contextual effects! In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 89-105). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Parducci, A. (1992). Comment on Haubensak's associative theory of judgment. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 310-313.

Parducci, A. (1995). *Happiness, pleasure, and judgment: The contextual theory and its applications*. Mahweh, New Jersey: Lawrence Erlbaum Associates.

Parducci, A., Calfee, R. C., Marshall, L. M., & Davidson, L. P. (1960). Context effects in judgment: Adaptation level as a function of the mean, midpoint, and median of the stimuli. *Journal of Experimental Psychology, 60*(2), 65-77.

Parducci, A., & Haugen, R. (1967). The frequency principle for comparative judgments. *Attention, Perception, & Psychophysics, 2*(2), 81-82.

Parducci, A., Perrett, D. S., & Marsh, H. W. (1969). Assimilation and contrast as range-frequency effects of anchors. *Journal of Experimental Psychology, 81*(2), 281-288.

Parducci, A., & Perrett, L. F. (1971). Category rating scales: Effects of relative spacing and frequency of stimulus values. *Journal of Experimental Psychology, 89*(2), 427-452.

Parducci, A., & Wedell, D. H. (1986). The category effect with rating scales: Number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human Perception and Performance, 12*(4), 496-516.

Pettibone, J. C., & Wedell, D. H. (2007). Of gnomes and leprechauns: The recruitment of recent and categorical contexts in social judgment. *Acta Psychologica, 125*(3), 361-389.

Petzold, P., & Haubensak, G. (2001). Higher order sequential effects in psychophysical judgments. *Attention, Perception & Psychophysics, 63*(6), 969-978.

Petzold, P., & Haubensak, G. (2004). Short-term and long-term frames of reference in category judgments: A multiple-standards model. In C. Kaernbach, E. Schröger & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 45-68). Mahweh, New Jeresy: Lawrence Erlbaum Associates.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17*(2), 132-138.

Qian, J., & Brown, G. D. A. (2005). *Similarity-based sampling: Testing a model of price psychophysics.* Paper presented at the Proceedings of the 27th Annual Conference of the Cognitive Science Society

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*(2), 93-134.

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66*(1), 3-8.

Riskey, D. R., Parducci, A., & Beauchamp, G. K. (1979). Effects of context in judgments of sweetness and pleasantness. *Attention, Perception, & Psychophysics*, *26*(3), 171-176.

Roberts, W. A. (1972). Free recall of word lists varying in length and rate of presentation: A test of total-time hypotheses. *Journal of Experimental Psychology, 92*(3), 365-372.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392*(6676), 595-598.

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior, 9*(1), 99-105.

Rynes, S. L., Schwab, D. P., & Heneman III, H. G. (1983). The role of pay and market pay variability in job application decisions. *Organizational Behavior and Human Performance, 31*(3), 353-364.

Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well - being: The convergence and stability of self - report and non - self - report measures. *Journal of Personality, 61*(3), 317-342.

Sarris, V. (2004). Frame of reference models in psychophysics: A perceptual-cognitive approach. In C. Kaernbach, E. Schroger & H. Muller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 69-88). Mahweh, New Jersey: Lawrence Erlbaum Associates.

Schacter, D. L., Chiu, C.-Y. P., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16*(1), 159-182.

Schreiber, C. A., & Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *Journal of Experimental Psychology: General, 129*(1), 27-42.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*(2), 195-202.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika, 22*(4), 325-345.

Smith, R. H., Diener, E., & Wedell, D. H. (1989). Intrapersonal and social comparison determinants of happiness: A range-frequency analysis. *Journal of Personality and Social Psychology, 56*(3), 317-325.

Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice. *The Quarterly Journal of Experimental Psychology, 62*(6), 1041-1062.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1-26.

Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, *132*(1), 23-46.

Stewart, N., & Simpson, K. (2008). A decision-by-sampling account of decision under risk. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 261-276). London, UK: Oxford University Press.

Symmonds, M., Wright, N. D., Bach, D. R., & Dolan, R. J. (2011). Deconstructing risk: Separable encoding of variance and skewness in the brain. *Neuroimage, 58*(4), 1139-1149.

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1589-1625.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39-60.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1-25.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(4), 336-342.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

Ungemach, C., Stewart, N., & Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science, 22*(2), 253-260.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Science*, *15*(11), 546-554.

Volkmann, J. (1951). Scales of judgment and their implications for social psychology. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads* (pp. 279-274). New York: Harper & Row.

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1207-1241.

Weber, E., & Johnson, E. (2006). Constructing preferences from memory. In S. Lichtenstein & P. Slovic (Eds.), *The construction of preference* (pp. 397-410). New York: Cambridge University Press.

Wedell, D. H. (2008). A similarity-based range—frequency model for two-category rating data. *Psychonomic Bulletin & Review, 15*(3), 638-643.

Wedell, D. H., Hicklin, S. K., & Smarandescu, L. O. (2007). Contrasting models of assimilation and contrast. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology* (pp. 45-74). New York: Psychology Press.

Wedell, D. H., & Parducci, A. (1985). Category and stimulus effects: A process model for contextual memory in judgment. In G. d'Ydewalle (Ed.), *Cognition, information processing, and Motivation* (pp. 55-70). New York: North Holland.

Wedell, D. H., & Parducci, A. (1988). The category effect in social judgment: Experimental ratings of happiness. *Journal of Personality and Social Psychology, 55*(3), 341-356.

Wedell, D. H., Parducci, A., & Geiselman, R. E. (1987). A formal analysis of ratings of physical attractiveness: Successive contrast and simultaneous assimilation. *Journal of Experimental Social Psychology, 23*(3), 230-249.

Wedell, D. H., Parducci, A., & Roman, D. (1989). Student perceptions of fair grading: A range-frequency analysis. *The American Journal of Psychology, 102*(2), 233-248.

Wedell, D. H., Santoyo, E. M., & Pettibone, J. C. (2005). The thick and the thin of it: Contextual effects in body perception. *Basic and Applied Social Psychology, 27*(3), 213-228.

Wilkinson, R., & Pickett, K. (2010). *The spirit level: Why equality is better for everyone*. London, UK: Penguin.

Wilson, W. P., Witryol, S. L., & Hust, B. E. (1975). Differential incentive motivation in recency and primacy during children's free recall. *The Journal of Genetic Psychology, 127*(1), 131-145.

Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review, 1*(1), 89-106.

Wood, A. M., Boyce, C. J., Moore, S. C., & Brown, G. D. A. (2012). An evolutionary based social rank explanation of why low income predicts mental distress: A 17 year cohort study of 30,000 people. *Journal of Affective Disorders, 136*(3), 882-888.

Wood, A. M., Brown, G. D. A., & Maltby, J. (2011). Thanks, but I'm used to better: A relative rank model of gratitude. *Emotion, 11*(1), 175-180.

Wood, A. M., Brown, G. D. A., & Maltby, J. (2012). Social norm influences on evaluations of the risks associated with alcohol consumption: applying the rank-based decision by sampling model to health judgments. *Alcohol and Alcoholism, 47*(1), 57-62.

Wu, C. C., Bossaerts, P., & Knutson, B. (2011). The affective impact of financial skewness on neural activity and choice. *PloS one, 6*(2), e16838.

Appendices

Appendix A

In Chapter 3 I used a prior distribution of standard deviation (SD) parameters when fitting models to the data at an individual level. Here I explain what prior distribution was used and why it was required. First, I will outline how the models were fit to the data. Then I discuss why allowing the SD parameter to vary freely gave inappropriate model fits. Finally, I introduce the prior distribution of SD parameters used in Chapter 3.

The aim was to compare model predictions when they were closest to the observed data using a maximum likelihood estimation method. For each data point I recorded the probability of the data on a normal distribution with a mean equal to the prediction of each model. Consider the example in Figure A1, where the vertical line represents an observed response of 4 on a continuous scale. The predictions of the adaptation level and the range-frequency models are 3.8 and 4.4. In this example, the response is more probable on the normal distribution with a mean equal to the predictions of range-frequency model (solid black line) and the corresponding -2lnL is the lowest of the compared models.

Model fitting of the experiment used an interval method because the response scale was ordinal with discrete response options. The only ratings that participants can give are 1,2,3,4,5,6 and 7. A participant may wish to give a response of 4.25 to a stimulus and choose a rating of 4 instead. I incorporated this into the model fitting procedure. The likelihood of a model was given by the area of the normal distribution within a .49 interval of a response rating. As shown in Figure A2, more of the normal distribution representing the range-frequency model prediction of 3.8 is within the interval and the corresponding -2lnL will still be the lowest.

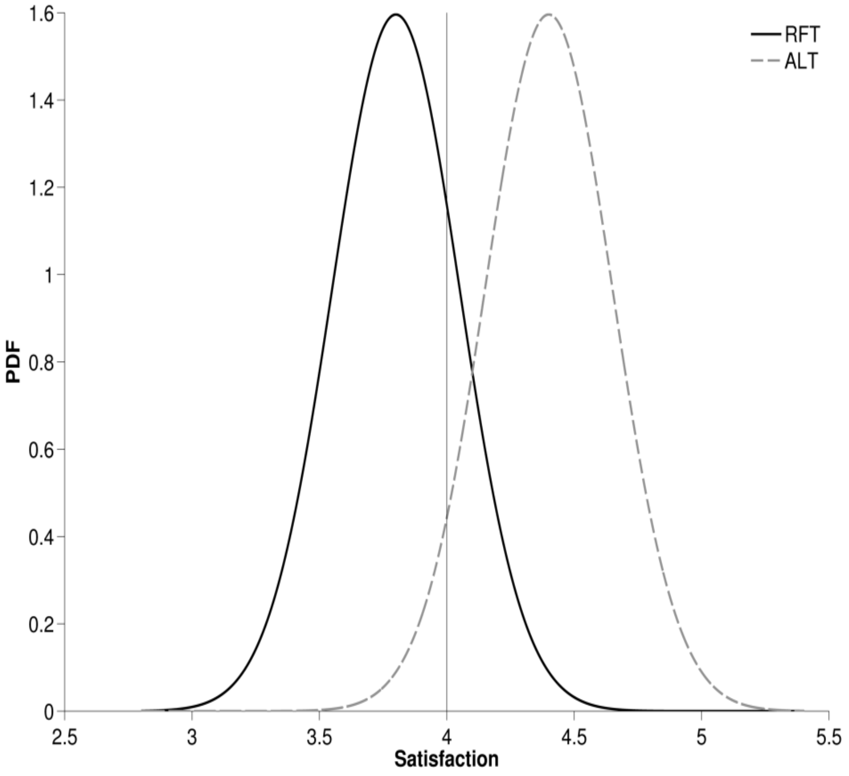

Figure A1. The relationship between a satisfaction response of four and two normal probability distributions (σ = 0.25) representing either range-frequency theory (RFT; solid line, μ = 3.8) or adaptation level theory (ALT; dashed line, μ = 4.4).

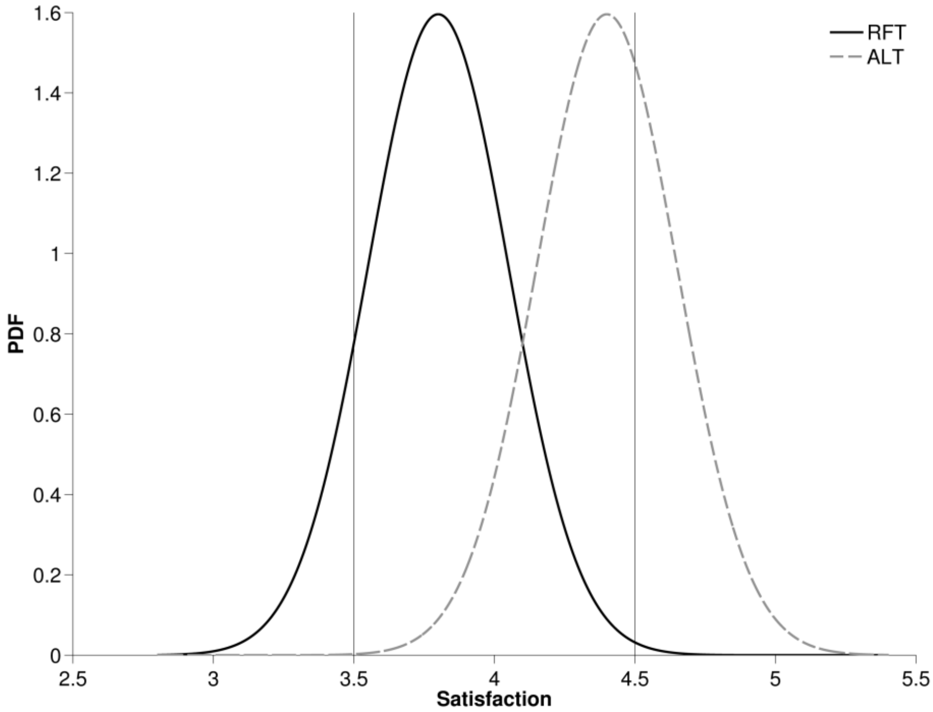

Figure A2. The +/- 0.49 interval used to calculate the -2lnL of ordinal data given a model prediction.

However, allowing an algorithm to freely vary the SD of the normal distribution can cause the measure of fit to be inappropriate. As shown in Figure A3 if the SD of a normal distribution is small enough then the entire distribution will fit within the response interval. If the predictions of both models are very close to the observed responses, as in Figure A3, then the -2lnL of the distribution area within the interval will incorrectly suggest that both models are equally probable given the data.

I resolved this issue by using a prior distribution of standard deviations. Tripp and Brown (unpublished) have used a similar method to fit the range-frequency model to individual data made on a continuous scale. Plots of this data show that these prior standard deviations can be fit using a log normal distribution (μ = -1.3, σ = 0.3) where an SD of 0.29 is most likely (see Figure A4). First, the models were fit by minimizing the -2lnL of the data given a normal distribution with a mean equal to a model prediction *and* the -2lnL of the SD parameter on the prior log normal distribution. Then, only the former -2lnL was used for model comparison.

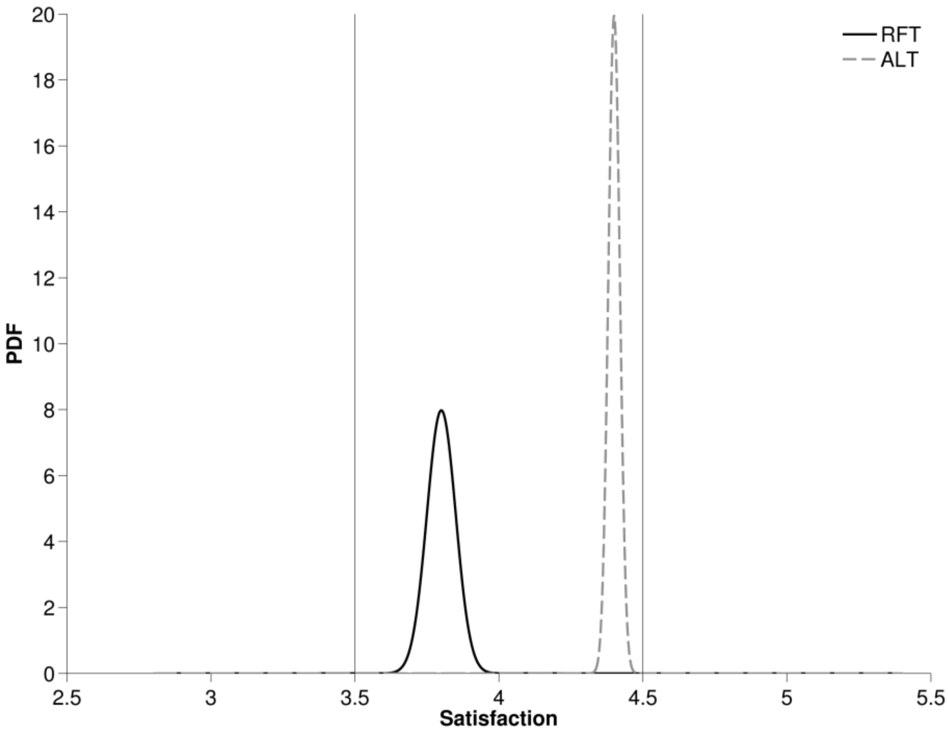

Figure A3. Graphical depiction of conditions in which two models may appear equally likely despite different predictions.

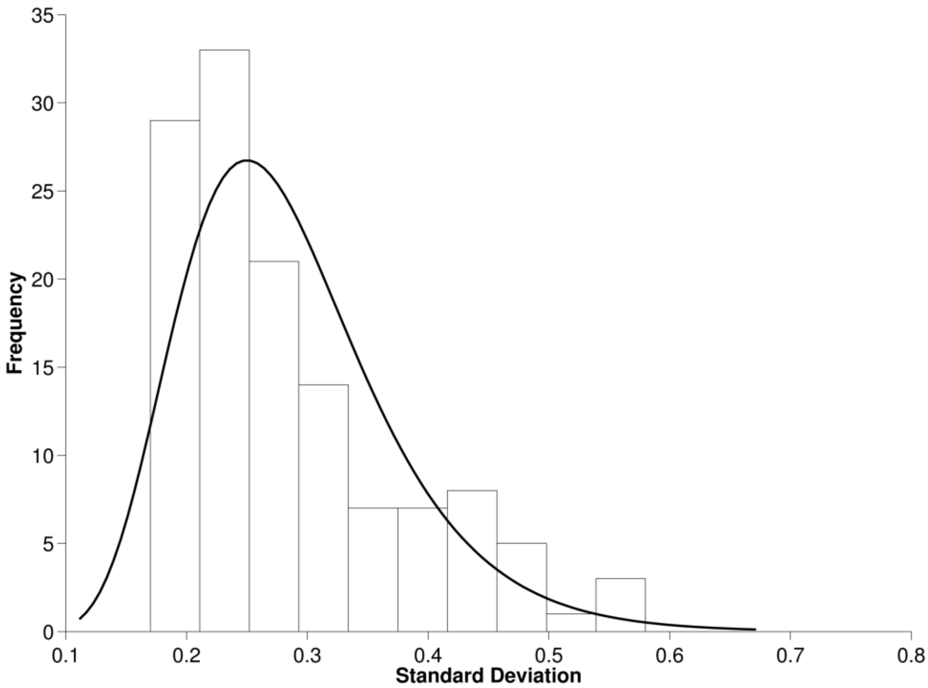

Figure A4. Histogram of standard deviation values from individual level fits of Brown et al. (2008) and Melrose et al. (2012), and the best fitting log-normal distribution (μ = -1.3, σ = 0.3).

Appendix B:

RFT, Range and Rank

Table B1

Statistics for the model fit of range, frequency and RFT to individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

Range SD	Range -2lnL	Frequency SD	Frequency -2lnL	RFT SD	RFT *w*	RFT -2lnL	*w* bootstrap Mean	*w* bootstrap 95% CI		Likelihood Ratio vs Range	Likelihood Ratio vs Rank
0.95	179.85	0.45	80.68	0.45	.00	80.68	.03	.00	.13	99.17	0.00
0.52	99.88	0.51	97.89	0.32	.49	34.95	.50	.41	.59	64.94	62.94
0.55	108.30	0.66	131.67	0.44	.60	80.21	.60	.47	.74	28.08	51.45
0.43	76.36	0.50	95.68	0.23	.55	-5.59	.55	.48	.61	81.95	101.28
0.59	118.45	0.29	23.30	0.17	.29	-49.96	.29	.24	.35	168.40	73.25
0.72	144.30	0.47	88.14	0.42	.27	72.29	.27	.14	.39	72.01	15.84
0.56	110.98	0.35	49.66	0.21	.35	-20.43	.35	.29	.41	131.41	70.08
0.59	118.24	0.71	142.23	0.51	.62	97.51	.61	.46	.77	20.73	44.72
0.63	127.32	0.75	148.78	0.56	.62	109.72	.62	.46	.80	17.60	39.06
0.74	147.71	0.55	109.06	0.49	.31	93.64	.31	.16	.46	54.06	15.41
0.59	118.36	0.59	117.87	0.43	.50	77.00	.50	.37	.62	41.36	40.87
0.77	152.43	0.46	84.68	0.43	.21	75.15	.21	.08	.33	77.29	9.53
0.56	111.01	0.62	124.77	0.43	.56	76.53	.55	.42	.70	34.49	48.24
0.52	101.76	0.46	84.53	0.28	.45	19.05	.45	.37	.53	82.72	65.49
0.70	139.63	0.24	-3.77	0.19	.17	-30.79	.17	.11	.23	170.42	27.03
0.75	148.51	0.54	104.84	0.48	.29	90.52	.29	.14	.43	57.98	14.32
0.67	134.41	0.23	-7.68	0.17	.20	-50.13	.19	.15	.24	184.53	42.45
0.62	123.48	0.47	88.18	0.36	.38	52.35	.38	.27	.48	71.14	35.83
1.01	188.48	0.58	115.03	0.58	.00	115.03	.04	.00	.17	73.45	0.00
0.62	123.32	0.35	47.86	0.25	.30	3.66	.30	.23	.38	119.65	44.20
0.51	98.51	0.61	122.65	0.39	.59	62.24	.59	.47	.70	36.27	60.40
0.55	109.12	0.48	90.73	0.32	.44	37.95	.44	.35	.54	71.18	52.79
0.61	121.98	0.31	33.09	0.21	.29	-21.07	.29	.23	.35	143.04	54.16
0.70	140.95	0.33	39.02	0.28	.20	20.33	.20	.12	.29	120.62	18.69

Table B2

Statistics for the model fit of range, frequency and RFT to individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

Range		Frequency		RFT			w bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
4.16	375.44	3.69	359.66	3.27	.40	343.83	.40	.22	.59	343.61	0.18
9.99	491.05	7.27	449.11	7.07	.19	445.45	.19	.00	.40	445.45	0.19
18.90	575.30	16.20	554.93	16.20	.00	554.93	.09	.00	.42	554.93	0.09
6.11	426.19	4.14	374.66	3.99	.19	369.95	.19	.03	.36	369.93	0.17
14.28	538.29	16.00	553.31	14.24	.87	537.92	.84	.46	1	537.46	0.38
15.84	551.99	13.82	533.97	13.77	.13	533.45	.16	.00	.49	533.45	0.16
19.90	582.06	20.44	585.59	19.71	.67	580.84	.63	.06	1	580.78	0.57
15.49	548.97	15.05	545.21	14.55	.42	540.78	.43	.07	.80	540.71	0.36
11.17	505.86	10.64	499.46	9.87	.43	489.50	.43	.16	.67	489.34	0.27
15.44	548.57	14.11	536.70	14.00	.22	535.61	.23	.00	.59	535.61	0.23
10.84	501.92	5.98	423.28	5.98	.00	423.28	.04	.00	.17	423.28	0.03
17.88	567.95	19.22	577.51	17.83	.85	567.59	.81	.31	1	567.28	0.50
9.97	490.85	9.93	490.38	9.68	.48	486.98	.48	.00	.98	486.98	0.48
9.81	488.72	9.90	489.95	8.98	.51	476.99	.51	.26	.76	476.73	0.25
10.86	502.14	8.48	469.50	8.48	.00	469.50	.08	.00	.37	469.50	0.08
7.42	451.86	10.92	502.83	7.35	.89	450.68	.88	.69	1	450.00	0.20
8.79	474.14	11.27	506.97	8.71	.86	472.94	.85	.60	1	472.34	0.25
8.55	470.58	7.15	446.90	7.15	.00	446.90	.08	.00	.39	446.90	0.08
9.74	487.70	4.48	385.33	4.48	.00	385.33	.03	.00	.13	385.33	0.03
6.10	426.10	6.01	423.98	5.41	.48	410.18	.48	.24	.71	409.94	0.24
8.64	471.95	8.31	466.75	7.81	.43	458.56	.43	.14	.72	458.42	0.30
15.58	549.74	14.34	538.79	14.19	.24	537.46	.25	.00	.64	537.46	0.25
13.76	533.34	13.04	526.29	12.51	.39	520.75	.39	.08	.68	520.67	0.31
9.62	486.17	5.85	420.58	5.60	.18	414.70	.18	.03	.33	414.67	0.15
9.98	490.94	9.43	483.53	9.00	.40	477.35	.40	.09	.72	477.25	0.31
9.63	486.19	9.10	478.83	8.03	.45	462.21	.45	.24	.67	461.97	0.21
3.90	367.12	6.09	425.69	3.27	.71	343.61	.71	.59	.81	343.02	0.11
6.29	430.00	6.76	439.64	5.96	.62	423.02	.62	.35	.91	422.67	0.27
3.19	340.52	6.45	433.43	3.19	1	340.52	.97	.84	1	339.68	0.13
7.62	455.40	9.18	479.99	7.62	.97	455.37	.92	.63	1	454.75	0.29
5.45	411.00	4.08	372.83	4.05	.11	371.98	.12	.00	.34	371.98	0.12
7.48	452.86	5.90	421.68	5.89	.09	421.28	.12	.00	.37	421.28	0.12

(continued)

Table B2

Statistics for the model fit of range, frequency and RFT to individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

Range SD	Range -2lnL	Frequency SD	Frequency -2lnL	RFT SD	RFT *w*	RFT -2lnL	*w* bootstrap Mean	*w* bootstrap 95% CI		Likelihood Ratio vs Range	Likelihood Ratio vs Rank
5.83	419.90	7.70	456.72	5.31	.70	407.78	.70	.54	.85	407.25	0.16
2.95	330.08	3.69	359.56	2.95	1	330.08	.93	.66	1	329.42	0.27
7.54	454.04	2.55	310.61	2.38	.11	301.66	.11	.04	.18	301.62	0.07
11.18	506.00	11.10	505.00	10.98	.43	503.54	.45	.00	1	503.54	0.45
7.52	453.58	3.31	345.47	3.13	.14	337.90	.14	.05	.23	337.85	0.09
5.52	412.69	5.10	402.44	4.76	.40	393.36	.40	.14	.66	393.22	0.25
6.49	434.21	9.60	485.91	6.49	1	434.21	.94	.71	1	433.50	0.23
8.91	476.03	8.03	462.35	8.03	0	462.35	.14	0	.69	462.35	0.14
6.02	424.17	11.71	512.11	6.02	1	424.17	.96	.81	1	423.36	0.15
10.93	502.92	14.94	544.21	10.92	1	502.92	.91	.58	1	502.34	0.33
14.77	542.74	15.45	548.65	14.77	1	542.74	.82	.14	1	542.59	0.67
3.32	345.88	2.80	323.25	1.84	.43	267.53	.43	.34	.52	267.18	0.09
2.38	301.57	3.01	332.89	2.38	1	301.57	.93	.67	1	300.90	0.26
9.82	488.90	12.75	523.29	9.82	1	488.90	.94	.68	1	488.21	0.25
12.21	517.58	14.58	541.00	12.21	1	517.58	.90	.54	1	517.04	0.36
8.24	465.64	8.89	475.76	7.79	.62	458.34	.62	.35	.89	457.99	0.26
3.12	337.35	3.40	348.91	1.66	.53	254.16	.53	.46	.60	253.70	0.07
4.71	391.94	10.64	499.38	4.71	1	391.94	.97	.87	1	391.07	0.10
5.51	412.51	8.50	469.74	5.37	.84	409.18	.84	.68	1	408.51	0.17
7.58	454.59	10.51	497.84	7.58	1	454.59	.94	.73	1	453.86	0.21

Table B3

Statistics for the model fit of range, frequency and RFT to individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
1.50	321.06	1.64	336.84	1.46	.69	316.32	.68	.40	.97	315.92	0.28
6.15	569.37	4.98	532.29	4.91	.19	529.64	.19	.00	.41	529.63	0.19
14.75	723.36	14.27	717.54	13.61	.43	709.18	.43	.15	.73	709.04	0.28
4.56	516.71	5.28	542.58	3.65	.58	477.66	.58	.46	.70	477.20	0.12
7.23	597.98	6.23	571.81	6.22	.09	571.49	.12	0	.40	571.49	0.12
5.41	546.94	5.11	536.67	4.33	.45	507.69	.46	.29	.61	507.40	0.17
11.29	676.31	9.55	646.89	8.89	.33	634.29	.34	.16	.54	634.13	0.18
8.59	628.16	9.84	652.20	7.69	.62	608.69	.62	.47	.78	608.23	0.15
4.01	494.16	7.79	611.12	4.01	1	494.16	.98	.87	1	493.29	0.10
14.72	723.04	16.61	744.32	14.43	.74	719.50	.73	.47	1	719.03	0.26
5.08	535.92	5.35	545.05	4.78	.58	525.21	.58	.33	.82	524.88	0.25
10.54	664.31	9.85	652.30	9.72	.28	650.08	.27	.00	.60	650.08	0.27
9.61	647.99	9.19	640.21	8.79	.41	632.20	.40	.13	.67	632.07	0.27
4.10	498.16	5.49	549.36	2.96	.62	440.65	.62	.54	.70	440.11	0.08
14.71	722.88	13.94	713.39	13.15	.41	703.14	.41	.16	.64	702.98	0.25
12.32	691.76	11.29	676.28	11.15	.25	674.11	.25	0	.59	674.11	0.25
10.82	668.91	9.52	646.29	9.38	.23	643.76	.23	0	.51	643.76	0.23
15.97	737.39	18.23	760.71	15.80	.80	735.54	.79	.51	1	735.03	0.28
7.07	594.06	9.56	647.08	6.47	.71	578.42	.71	.58	.84	577.84	0.13
7.83	612.04	8.72	630.94	6.80	.58	587.24	.59	.43	.75	586.80	0.15
3.80	484.79	7.02	592.67	3.68	.86	478.97	.86	.75	.97	478.22	0.11
4.90	529.56	7.57	605.99	4.19	.71	501.81	.71	.61	.82	501.20	0.10
6.21	571.07	9.71	649.81	5.44	.73	547.76	.73	.62	.83	547.13	0.11
7.78	610.75	11.25	675.80	7.71	.89	609.25	.89	.73	1	608.52	0.16
5.95	563.73	9.30	642.14	5.78	.84	558.43	.84	.70	.97	557.73	0.14
5.09	536.04	11.06	672.74	5.09	1	536.04	.98	.89	1	535.15	0.09
9.04	637.18	7.86	612.54	7.02	.38	592.83	.38	.23	.55	592.60	0.16
7.52	604.90	11.21	675.04	7.28	.82	599.23	.82	.68	.96	598.55	0.14
16.17	739.56	17.92	757.69	16.12	.86	739.02	.84	.49	1	738.53	0.35
6.98	591.69	8.24	620.81	5.97	.61	564.11	.62	.48	.75	563.62	0.13
8.62	628.94	8.14	618.81	7.60	.42	606.69	.42	.18	.65	606.51	0.23
4.99	532.53	8.23	620.71	4.87	.86	528.47	.86	.73	.99	527.74	0.13

(continued)

Table B3

Statistics for the model fit of range, frequency and RFT to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
5.24	541.40	10.27	659.61	5.23	.95	540.80	.95	.84	1	539.96	0.11
7.14	595.80	10.84	669.16	7.02	.86	592.83	.86	.71	1	592.12	0.15
6.29	573.34	6.81	587.47	4.98	.55	532.21	.55	.42	.67	531.78	0.12
17.70	755.47	17.11	749.56	16.48	.42	742.95	.41	.12	.73	742.83	0.29
5.08	535.84	6.66	583.33	4.32	.65	507.08	.65	.54	.76	506.54	0.11
10.16	657.75	14.16	716.27	10.16	1	657.75	.94	.72	1	657.03	0.23
10.97	671.26	15.02	726.53	10.96	.97	671.17	.94	.74	1	670.43	0.20
1.31	297.10	1.49	320.18	1.18	.62	279.11	.61	.44	.79	278.67	0.17
18.03	758.73	12.90	699.78	12.90	0	699.78	.06	0	.27	699.78	0.05
9.51	646.11	11.40	678.10	8.94	.69	635.18	.69	.51	.87	634.67	0.18
8.14	618.85	8.22	620.52	7.46	.51	603.34	.52	.29	.73	603.05	0.23
9.95	654.05	9.47	645.37	8.66	.44	629.73	.45	.24	.65	629.49	0.21
2.13	382.89	3.19	453.85	2.07	.83	378.20	.84	.69	.99	377.52	0.15
8.08	617.43	7.29	599.29	7.04	.32	593.17	.32	.07	.59	593.09	0.25
15.59	733.08	8.53	626.99	8.53	0	626.99	.04	0	.19	626.99	0.04
18.05	758.91	12.90	699.78	12.90	0	699.78	.04	0	.22	699.78	0.04
8.80	632.43	6.61	582.11	6.03	.30	566.00	.30	.16	.44	565.84	0.14
27.46	832.80	29.66	846.31	27.45	.93	832.73	.87	.48	1	832.24	0.38
3.30	460.10	3.48	469.15	3.15	.60	451.88	.60	.33	.85	451.55	0.27
10.16	657.81	10.51	663.77	9.84	.59	652.06	.60	.27	.91	651.79	0.33
8.72	630.91	7.37	601.28	7.23	.23	597.91	.23	0	.47	597.91	0.23
10.11	656.99	10.31	660.32	9.80	.56	651.39	.56	.22	.94	651.16	0.33
8.45	625.37	8.33	622.83	7.30	.49	599.59	.48	.29	.67	599.30	0.20
1.29	294.17	1.09	264.28	1.07	.18	262.44	.18	0	.46	262.44	0.18
5.28	542.64	5.27	542.21	4.98	.49	532.12	.49	.21	.80	531.92	0.28
7.42	602.47	7.48	603.93	6.24	.51	572.09	.51	.36	.67	571.73	0.15
6.73	585.28	5.61	553.17	5.48	.23	549.04	.24	.01	.45	549.03	0.22
19.59	773.30	13.40	706.49	13.40	0	706.49	.05	0	.23	706.49	0.05
23.68	806.73	24.42	812.16	23.54	.72	805.69	.70	.14	1	805.55	0.56
15.09	727.44	14.55	720.94	14.24	.37	717.18	.37	0	.72	717.18	0.37
18.23	760.68	11.49	679.48	11.49	0	679.48	.04	0	.20	679.47	0.04

(continued)

Table B3

Statistics for the model fit of range, frequency and RFT to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			w bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
16.75	745.74	17.89	757.32	16.43	.69	742.42	.68	.34	1	742.08	0.34
13.46	707.26	11.59	681.00	11.32	.26	676.80	.26	.02	.49	676.77	0.23
7.21	597.49	6.23	571.57	5.99	.30	564.79	.30	.07	.51	564.72	0.22
10.72	667.24	8.26	621.41	8.26	0	621.41	.05	0	.25	621.41	0.05
10.20	658.55	6.49	578.77	6.44	.09	577.41	.10	0	.25	577.41	0.09
10.93	670.67	12.62	695.99	10.65	.73	666.09	.74	.49	.99	665.60	0.24
8.60	628.52	7.69	608.89	6.83	.40	587.77	.40	.24	.56	587.52	0.16
10.12	657.09	7.82	611.69	7.37	.27	601.22	.28	.12	.44	601.10	0.16
14.90	725.13	17.04	748.83	14.83	.86	724.38	.84	.54	1	723.84	0.30
14.53	720.69	14.62	721.87	13.88	.52	712.73	.52	.19	.83	712.54	0.33
10.16	657.77	6.63	582.76	6.28	.21	573.03	.21	.09	.34	572.94	0.13
10.28	659.80	12.12	688.79	10.13	.79	657.25	.79	.55	1	656.70	0.24
9.63	648.36	9.81	651.59	8.23	.52	620.74	.51	.34	.67	620.40	0.17
9.65	648.77	10.97	671.25	9.00	.64	636.47	.64	.43	.84	636.04	0.21
7.92	613.90	7.32	600.03	6.04	.45	566.37	.45	.31	.59	566.06	0.14
15.91	736.69	18.68	765.01	15.91	1	736.69	.92	.58	1	736.11	0.34
8.20	620.16	9.83	651.88	7.80	.70	611.36	.70	.51	.90	610.86	0.20

Table B4

Statistics for the model fit of range, frequency and RFT to individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
0.34	24.38	0.36	26.00	0.30	.56	21.71	.56	.24	.92	21.47	0.32
0.00	18.00	0.31	28.60	0.00	1	18.00	1	.99	1	17.01	0.00
0.22	16.67	0.72	37.47	0.22	1	16.67	.96	.82	1	15.85	0.14
0.35	25.94	0.32	24.68	0.29	.42	22.52	.44	.04	.82	22.48	0.40
0.13	24.55	0.06	19.93	0.02	.15	19.38	.15	.13	.17	19.25	0.02
0.11	12.86	0.00	8.00	0.00	.01	8.00	.01	0	.01	8.00	0.00
0.54	31.29	0.36	26.48	0.24	.38	20.44	.37	.23	.51	20.22	0.15
0.17	15.95	0.31	26.37	0.17	1	15.95	.95	.76	1	15.19	0.19
0.79	37.67	0.96	40.35	0.78	.86	37.59	.79	.15	1	37.44	0.64
0.00	10.00	0.28	21.56	0.00	1	10.00	1	1	1	9.00	0.00
0.22	32.90	0.05	27.03	0.03	.08	25.97	.08	.06	.10	25.91	0.02
0.32	23.24	0.73	36.04	0.32	1	23.24	.93	.68	1	22.57	0.25
0.09	14.18	0.53	31.59	0.07	.94	13.61	.94	.89	.99	12.72	0.05
0.70	34.00	0.01	26.00	0.01	0	26.00	0	0	.01	26.00	0.00
0.23	20.76	0.63	34.85	0.23	1	20.76	.96	.80	1	19.96	0.16
0.07	14.81	0.21	19.85	0.07	.86	14.78	.86	.76	.96	14.02	0.10
0.35	25.85	1.06	41.64	0.35	1	25.85	.93	.67	1	25.18	0.26
1.29	44.26	2.20	47.62	1.29	1	44.26	.78	0	1	44.26	0.78
0.06	11.69	0.07	15.00	0.09	.48	11.80	.49	.39	.58	11.41	0.10
0.00	6.00	0.30	23.72	0.00	1	6.00	1	1	1	5.00	0.00
0.00	8.00	0.29	24.85	0.00	1	8.00	1	1	1	7.00	0.00
0.03	25.38	0.26	26.38	0.26	0	26.38	.07	0	.34	26.38	0.07
0.24	34.59	0.01	32.00	0.01	0	32.00	0	0	.01	32.00	0.00
0.09	27.12	0.04	25.24	0.04	0	25.24	.01	0	.05	25.24	0.01
0.24	18.83	0.16	15.88	0.04	.22	11.12	.22	.19	.25	10.94	0.04
0.00	20.00	0.28	27.24	0.00	.80	18.00	.80	.79	.80	17.21	0.00
0.07	11.32	0.30	25.28	0.05	.80	11.19	.80	.75	.84	10.44	0.05
0.06	17.15	0.18	23.29	0.06	1	17.15	.99	.95	1	16.20	0.04
0.08	15.30	0.25	23.98	0.08	1	15.30	.98	.91	1	14.40	0.07
0.49	31.50	0.75	37.34	0.49	1	31.50	.90	.49	1	31.01	0.41
0.11	10.24	0.15	11.62	0.04	.58	5.96	.58	.53	.63	5.43	0.05
0.11	6.84	0.17	13.58	0.00	.67	2.00	.67	.67	.67	1.33	0.00

(continued)

Table B4

Statistics for the model fit of range, frequency and RFT to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
1.28	45.75	2.43	49.04	1.28	1	45.75	.79	0	1	45.75	0.78
0.19	14.81	0.70	36.03	0.19	1	14.81	.97	.86	1	13.96	0.11
0.43	28.54	0.56	33.38	0.38	.67	26.78	.67	.37	.96	26.41	0.30
0.00	26.00	0.85	37.83	0.00	1	26.00	1	1	1	25.00	0.00
0.14	21.77	0.24	26.30	0.01	.34	20.00	.34	.33	.36	19.67	0.01
0.25	22.27	0.34	24.95	0.25	.66	20.50	.66	.45	.86	20.04	0.21
0.51	31.85	0.35	25.54	0.35	0	25.54	.09	0	.45	25.53	0.09
0.50	31.68	0.35	25.89	0.35	.04	25.88	.11	0	.46	25.88	0.11
0.30	30.32	0.24	28.21	0.24	0	28.21	.05	0	.25	28.21	0.05
1.12	40.83	0.53	32.00	0.53	0	32.00	.10	0	.46	32.00	0.10
0.11	16.18	0.08	16.40	0.11	1	16.18	.90	.49	1	15.69	0.40
0.19	13.92	0.16	14.44	0.02	.63	4.55	.63	.60	.66	3.95	0.03
0.34	23.15	0.35	25.92	0.32	.74	22.79	.71	.15	1	22.64	0.56
0.84	39.62	0.99	40.31	0.84	.82	39.57	.73	0	1	39.57	0.73
0.69	36.92	0.48	31.05	0.48	0	31.05	.11	0	.57	31.05	0.10
0.36	32.06	0.24	26.30	0.24	0	26.30	.06	0	.28	26.30	0.05
0.22	16.36	0.31	23.51	0.22	1	16.36	.90	.46	1	15.90	0.44
0.01	30.00	0.18	31.61	0.00	.67	30.00	.67	.67	.67	29.33	0.00
0.00	20.00	0.01	20.00	0.01	0	20.00	0	0	.01	20.00	0.00
0.06	9.76	0.15	18.58	0.00	.79	10.00	.79	.79	.79	9.21	0.00
0.46	31.56	0.45	31.12	0.42	.46	29.99	.47	0	1	29.99	0.47
0.25	24.02	0.16	17.88	0.16	0	17.88	.02	0	.11	17.88	0.02
0.20	14.32	0.43	27.69	0.20	1	14.32	.96	.78	1	13.53	0.17
0.00	20.00	0.01	18.00	0.01	0	18.00	0	0	.01	18.00	0.00
0.00	8.00	0.13	15.14	0.01	.45	8.00	.45	.44	.46	7.56	0.01
0.09	14.37	0.34	26.96	0.09	1	14.37	.97	.84	1	13.53	0.13
0.00	22.00	0.36	31.64	0.00	.95	22.00	.95	.94	.95	21.06	0.00
0.10	9.26	0.28	21.38	0.10	1	9.26	.97	.84	1	8.42	0.13
0.15	21.30	0.43	31.24	0.15	1	21.30	.97	.86	1	20.45	0.11
0.12	11.17	0.35	25.77	0.12	1	11.17	.96	.81	1	10.36	0.15
0.17	16.50	0.27	20.32	0.17	.77	15.66	.77	.59	.95	15.07	0.19
0.15	20.41	0.01	20.00	0.01	0	20.00	.01	0	.03	20.00	0.01

(continued)

Table B4

Statistics for the model fit of range, frequency and RFT to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
0.09	4.69	0.05	5.07	0.09	1	4.69	.91	.55	1	4.14	0.36
0.03	12.68	0.05	11.86	0.05	0	11.86	.02	0	.09	11.86	0.02
0.61	34.42	0.49	31.42	0.49	.06	31.40	.17	0	.76	31.40	0.17
0.20	19.10	0.37	27.42	0.08	.75	17.94	.75	.67	.82	17.27	0.08
0.38	26.74	0.30	22.65	0.29	.22	22.20	.24	0	.68	22.20	0.24
0.00	20.00	0.18	22.85	0.00	.79	20.00	.79	.79	.79	19.21	0.00
0.51	33.09	0.65	36.97	0.50	.82	32.84	.76	.20	1	32.63	0.56
0.29	19.70	0.19	16.21	0.09	.28	13.16	.28	.17	.39	12.99	0.10
0.36	24.50	0.17	16.72	0.15	.21	15.38	.21	.05	.35	15.33	0.15
0.30	21.95	0.36	26.14	0.30	1	21.95	.87	.32	1	21.62	0.54
0.22	17.72	0.13	12.95	0.04	.50	7.85	.50	.45	.54	7.40	0.05
0.00	30.00	0.26	31.92	0.26	0	31.92	.05	0	.28	31.92	0.05
0.09	9.86	0.25	17.84	0.09	1	9.86	.97	.86	1	9.00	0.12
0.25	15.72	0.24	18.27	0.18	.52	12.05	.52	.31	.72	11.74	0.21

Table B5

Statistics for the model fit of range, frequency and RFT to individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
0.25	29.60	0.03	11.27	0.03	0	11.27	.01	0	.05	11.27	0.01
0.43	44.65	0.27	32.02	0.27	.02	32.01	.07	0	.33	32.01	0.07
0.78	57.81	0.57	50.73	0.56	.16	50.37	.18	0	.56	50.37	0.18
3.99	74.08	4.87	74.94	3.99	1	74.08	.63	0	1	74.08	0.63
3.19	71.61	4.17	73.25	3.19	1	71.61	.71	0	1	71.61	0.71
6.19	74.88	7.33	75.48	6.19	1	74.88	.59	0	1	74.88	0.59
0.78	56.93	0.28	33.15	0.28	0	33.15	.04	0	.18	33.15	0.03
0.46	44.29	0.05	12.52	0.05	0	12.52	.01	0	.04	12.52	0.01
0.76	56.42	0.66	53.31	0.66	.18	53.13	.24	0	.81	53.13	0.24
0.43	48.08	0.49	47.88	0.42	.47	46.12	.47	0	.77	45.95	0.30
1.88	70.24	1.10	64.35	1.10	0	64.35	.14	0	.66	64.35	0.13
1.08	65.19	0.60	55.37	0.60	.11	55.09	.14	0	.46	55.09	0.14
1.00	61.43	0.38	45.26	0.38	0	45.26	.04	0	.23	45.26	0.04
0.51	52.04	0.28	39.99	0.28	.04	39.97	.06	0	.21	39.97	0.05
1.25	63.51	0.29	39.29	0.29	0	39.29	.03	0	.15	39.29	0.03
0.93	59.88	0.37	40.14	0.37	0	40.14	.04	0	.19	40.14	0.04
0.88	58.67	0.23	31.29	0.23	0	31.29	.03	0	.15	31.29	0.03
1.29	63.33	1.28	63.25	1.24	.48	62.83	.50	0	1	62.83	0.50
2.85	73.71	2.82	73.59	2.63	.46	73.23	.47	0	1	73.23	0.47
2.08	72.72	1.39	69.57	1.38	.09	69.52	.20	0	.83	69.52	0.20
2.52	74.19	1.71	71.88	1.67	.17	71.72	.25	0	.87	71.72	0.25
0.67	53.26	0.33	35.26	0.33	.05	35.20	.06	0	.22	35.20	0.06
0.94	59.76	0.26	30.62	0.26	.03	30.53	.04	0	.16	30.53	0.04
1.30	65.77	0.84	59.31	0.84	0	59.31	.12	0	.57	59.31	0.11
0.63	52.76	0.30	37.64	0.29	.08	37.46	.09	0	.28	37.46	0.09
0.88	58.78	0.63	51.83	0.63	.04	51.82	.11	0	.49	51.82	0.11
0.75	54.58	0.20	36.70	0.20	0	36.70	.03	0	.14	36.70	0.03
2.14	71.18	0.70	56.94	0.70	0	56.94	.06	0	.27	56.94	0.06
1.27	64.25	0.53	50.09	0.51	.12	49.44	.13	0	.34	49.44	0.13
1.53	67.27	0.63	53.86	0.63	0	53.86	.06	0	.27	53.86	0.06
2.59	70.37	2.19	68.74	2.18	.05	68.73	.21	0	1	68.73	0.21
1.47	66.00	0.47	45.17	0.47	0	45.17	.05	0	.23	45.17	0.05

(continued)

Table B5

Statistics for the model fit of range, frequency and RFT to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

Range		Rank		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
2.06	70.28	0.98	61.30	0.98	0	61.30	.09	0	.47	61.30	0.09
0.33	41.08	0.05	24.06	0.02	.14	23.40	.14	.12	.15	23.27	0.01
1.53	65.72	0.80	55.94	0.80	0	55.94	.09	0	.46	55.94	0.09
1.19	61.37	0.69	52.44	0.69	0	52.44	.06	0	.30	52.44	0.06
1.24	64.70	0.42	43.93	0.42	.04	43.84	.06	0	.21	43.84	0.06
1.50	64.27	1.47	64.24	1.44	.48	63.90	.50	0	1	63.90	0.49
0.20	45.94	0.01	46.00	0.01	0	46.00	0	0	.01	46.00	0.00
2.93	73.47	1.08	64.41	1.08	0	64.41	.09	0	.47	64.40	0.09
2.70	70.80	0.98	60.06	0.98	0	60.06	.09	0	.43	60.06	0.09
2.22	68.65	1.11	61.09	1.11	0	61.09	.09	0	.43	61.09	0.09
1.33	66.75	0.90	60.86	0.87	.17	60.53	.21	0	.59	60.53	0.21
1.97	69.11	0.32	35.73	0.32	0	35.73	.03	0	.13	35.73	0.03
2.63	71.35	1.08	62.10	1.08	0	62.10	.11	0	.50	62.10	0.11
1.87	68.56	0.41	45.58	0.41	0	45.58	.04	0	.17	45.58	0.04
1.31	62.70	0.23	37.11	0.23	0	37.11	.02	0	.11	37.11	0.02
2.21	69.04	0.51	50.80	0.51	0	50.80	.05	0	.21	50.80	0.04
2.47	67.67	1.74	64.17	1.74	0	64.17	.16	0	.71	64.17	0.16
2.19	68.46	2.39	68.91	2.18	.66	68.27	.58	0	1	68.27	0.58
0.40	50.41	0.15	45.18	0.15	0	45.18	.01	0	.06	45.18	0.01
0.89	58.63	0.29	42.29	0.29	0	42.29	.03	0	.14	42.29	0.03
1.02	61.60	0.31	37.54	0.31	0	37.54	.03	0	.15	37.54	0.03
0.42	45.04	0.33	39.37	0.26	.41	35.61	.41	.26	.55	35.34	0.14
0.75	56.98	0.30	34.15	0.30	.00	34.15	.05	0	.25	34.15	0.05
0.34	50.31	0.05	38.82	0.00	.25	38.00	.25	.25	.25	37.75	0.00
1.08	61.32	0.18	31.28	0.19	.08	31.14	.08	0	.15	31.13	0.07
2.03	66.30	0.79	55.78	0.79	0	55.78	.08	0	.37	55.78	0.08
3.71	71.76	3.76	71.70	3.65	.46	71.57	.50	0	1	71.57	0.49
1.14	60.26	0.52	45.60	0.52	0	45.60	.05	0	.27	45.60	0.05
0.68	55.21	0.36	40.14	0.35	.14	39.40	.14	0	.36	39.40	0.14
2.66	68.21	2.86	68.88	2.66	1	68.21	.72	0	1	68.21	0.72
0.50	49.45	0.45	45.16	0.42	.31	44.38	.30	0	.62	44.38	0.30
0.82	58.27	0.45	45.52	0.44	.05	45.46	.10	0	.38	45.46	0.10

(continued)

Table B5

Statistics for the model fit of range, frequency and RFT to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
0.38	47.59	0.01	34.00	0.01	0	34.00	0	0	.01	34.00	0.00
0.00	42.00	0.19	47.11	0.19	0	47.11	.03	0	.13	47.11	0.02
1.89	65.31	0.28	48.08	0.28	0	48.08	.02	0	.12	48.08	0.02
0.61	53.36	0.52	51.43	0.52	0	51.43	.13	0	.67	51.43	0.13
3.79	74.61	2.24	71.76	2.24	0	71.76	.25	0	1	71.75	0.25
0.63	51.89	0.26	29.48	0.26	0	29.48	.06	0	.28	29.48	0.06
1.27	67.52	0.80	60.98	0.80	0	60.98	.13	0	.61	60.98	0.13
2.42	71.63	2.23	70.89	2.12	.34	70.60	.40	0	1	70.59	0.40
0.43	49.23	0.49	50.46	0.43	.86	49.18	.82	.43	1	48.75	0.40
1.73	68.64	0.67	55.57	0.67	0	55.57	.09	0	.42	55.57	0.09
1.11	65.93	0.71	59.01	0.71	0	59.01	.15	0	.70	59.01	0.15
2.00	71.31	0.92	62.74	0.92	0	62.74	.12	0	.58	62.74	0.12
0.78	55.73	0.46	46.08	0.46	0	46.08	.08	0	.37	46.08	0.07
1.37	67.15	0.85	60.28	0.85	0	60.28	.15	0	.72	60.28	0.15
1.28	65.37	0.54	50.55	0.54	0	50.55	.06	0	.29	50.55	0.06
6.10	77.21	3.36	75.86	3.36	0	75.86	.32	0	1	75.86	0.32
1.33	65.71	0.78	57.28	0.78	0	57.28	.11	0	.52	57.28	0.11
0.99	61.32	0.62	52.52	0.62	0	52.52	.09	0	.44	52.52	0.09
3.73	73.58	2.33	70.47	2.33	0	70.47	.21	0	1	70.47	0.21
1.02	61.76	0.30	33.32	0.30	0	33.32	.04	0	.19	33.32	0.04
0.80	58.01	0.35	42.73	0.35	0	42.73	.04	0	.21	42.73	0.04
1.07	60.20	0.65	50.67	0.65	.02	50.67	.07	0	.31	50.67	0.07
4.63	75.15	2.75	72.68	2.75	0	72.68	.26	0	1	72.68	0.26
1.39	66.98	0.63	54.01	0.63	0	54.01	.08	0	.38	54.01	0.08
1.18	61.20	0.78	54.10	0.78	0	54.10	.08	0	.38	54.10	0.08
1.44	66.03	0.04	5.59	0.04	0	5.59	0	0	.02	5.59	0.00
1.11	62.26	0.04	7.18	0.04	0	7.18	0	0	.02	7.18	0.00
1.29	64.62	0.31	33.61	0.31	0	33.61	.03	0	.15	33.61	0.03
1.42	65.72	0.31	37.07	0.31	0	37.07	.03	0	.14	37.07	0.03
0.73	55.36	0.33	37.43	0.33	.11	36.97	.11	0	.26	36.97	0.11
1.74	68.55	0.66	53.64	0.66	0	53.64	.08	0	.35	53.64	0.08
1.21	63.01	0.40	48.03	0.40	0	48.03	.05	0	.23	48.03	0.05

(continued)

Table B5

Statistics for the model fit of range, frequency and RFT to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			w bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
2.09	68.50	1.49	65.25	1.49	0	65.25	.18	0	.89	65.25	0.18
1.29	64.15	0.43	46.23	0.43	0	46.23	.04	0	.20	46.23	0.04
2.17	70.39	0.99	61.00	0.99	0	61.00	.10	0	.48	61.00	0.10
2.12	70.18	1.15	63.60	1.15	0	63.60	.12	0	.57	63.60	0.12
2.39	72.50	1.14	65.65	1.14	0	65.65	.12	0	.59	65.65	0.11
1.37	65.32	0.16	17.81	0.16	0	17.81	.01	0	.07	17.80	0.01
4.32	74.49	3.57	73.66	3.57	0	73.66	.37	0	1	73.66	0.37
1.58	65.38	0.62	52.32	0.62	0	52.32	.08	0	.34	52.31	0.07
4.51	73.64	3.88	72.78	3.88	0	72.78	.35	0	1	72.78	0.35
0.48	48.72	0.24	33.04	0.24	0	33.04	.05	0	.21	33.04	0.05
1.75	64.63	1.56	62.60	1.56	0	62.60	.18	0	.87	62.60	0.18
1.06	60.60	0.63	51.51	0.63	0	51.51	.08	0	.39	51.51	0.08
1.97	69.82	0.98	61.18	0.98	0	61.18	.11	0	.55	61.18	0.11
1.60	66.60	0.88	58.31	0.88	0	58.31	.11	0	.55	58.31	0.11
9.60	77.53	7.87	77.26	7.87	0	77.26	.40	0	1	77.26	0.39
0.63	53.93	0.41	48.89	0.41	0	48.89	.04	0	.20	48.89	0.04
0.60	51.87	0.77	52.38	0.60	1	51.87	.92	.61	1	51.26	0.31
1.97	66.61	1.62	64.20	1.62	0	64.20	.23	0	1	64.20	0.23
1.10	63.42	0.86	59.43	0.86	.07	59.40	.19	0	.80	59.40	0.19
1.17	61.67	0.50	47.36	0.50	0	47.36	.05	0	.24	47.36	0.05
1.31	63.64	0.18	29.06	0.18	0	29.06	.02	0	.09	29.06	0.02
0.85	58.36	1.26	58.27	0.94	.72	58.17	.69	.25	1	57.92	0.45
0.16	37.55	0.01	38.00	0.01	0	38.00	0	0	.01	38.00	0.00
3.75	73.64	3.06	72.54	3.06	0	72.54	.33	0	1	72.54	0.33
0.00	42.00	0.21	45.57	0.00	.88	42.00	.88	.88	.89	41.12	0.00
4.23	73.43	3.72	72.61	3.72	0	72.61	.39	0	1	72.61	0.39
6.09	75.19	5.07	74.49	5.07	0	74.49	.42	0	1	74.49	0.42
0.69	52.88	0.31	37.06	0.31	0	37.06	.03	0	.16	37.06	0.03
6.46	75.10	5.95	74.76	5.95	0	74.76	.43	0	1	74.76	0.42
0.22	43.68	0.19	43.01	0.10	.56	40.84	.56	.50	.63	40.33	0.06
0.41	51.11	0.22	44.79	0.22	.36	42.59	.36	.25	.46	42.34	0.11
1.55	65.10	0.81	55.96	0.81	0	55.96	.10	0	.51	55.96	0.09

(continued)

Table B5

Statistics for the model fit of range, frequency and RFT to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

Range		Frequency		RFT			*w* bootstrap			Likelihood Ratio	
SD	-2lnL	SD	-2lnL	SD	*w*	-2lnL	Mean	95% CI		vs Range	vs Rank
2.70	68.14	2.43	66.59	2.43	0	66.59	.29	0	1	66.59	0.29
0.40	46.10	0.38	45.94	0.36	.47	44.43	.47	.14	.80	44.29	0.34
7.91	75.84	6.32	75.07	6.32	0	75.07	.38	0	1	75.07	0.37
5.66	74.42	4.88	73.66	4.88	0	73.66	.37	0	1	73.66	0.36
1.55	64.40	1.45	63.14	1.45	0	63.14	.30	0	1	63.14	0.30
0.35	49.72	0.41	53.06	0.35	1	49.72	.95	.77	1	48.96	0.18
0.00	48.00	0.01	50.00	0.01	0	50.00	0	0	.01	50.00	0.00

Appendix C:

RFT and GEMS Fit

Table C1

Statistics for the model fit of RFT and GEMS to individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

	RFT			GEMS			γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
0.45	.00	80.68	0.44	0	0.08	77.86	0.08	-0.01	0.18
0.32	.49	34.95	0.30	.49	-0.16	28.87	-0.15	-0.27	-0.03
0.44	.60	80.21	0.44	.60	0.12	79.35	0.13	-0.12	0.42
0.23	.55	-5.59	0.23	.55	0.00	-5.60	0.00	-0.10	0.12
0.17	.29	-49.96	0.17	.29	0.00	-49.96	0.00	-0.05	0.06
0.42	.27	72.29	0.42	.27	0.04	71.92	0.04	-0.08	0.17
0.21	.35	-20.43	0.21	.35	0.02	-20.77	0.02	-0.05	0.10
0.51	.62	97.51	0.49	.62	0.39	92.99	0.40	0.04	0.85
0.56	.62	109.72	0.55	.62	0.12	109.28	0.14	-0.19	0.56
0.49	.31	93.64	0.49	.31	-0.06	93.14	-0.05	-0.20	0.12
0.43	.50	77.00	0.43	.50	-0.04	76.82	-0.04	-0.22	0.16
0.43	.21	75.15	0.43	.21	0.04	74.77	0.04	-0.09	0.17
0.43	.56	76.53	0.43	.56	0.07	76.06	0.08	-0.14	0.35
0.28	.45	19.05	0.28	.45	0.08	17.27	0.08	-0.04	0.19
0.19	.17	-30.79	0.19	.17	0.02	-31.66	0.02	-0.03	0.07
0.48	.29	90.52	0.48	.29	0.05	90.12	0.05	-0.10	0.23
0.17	.20	-50.13	0.17	.20	0.00	-50.14	0.00	-0.05	0.04
0.36	.38	52.35	0.36	.38	-0.01	52.30	-0.01	-0.13	0.12
0.58	0	115.03	0.46	0	-0.30	84.45	-0.30	-0.39	-0.21
0.25	.30	3.66	0.24	.30	-0.06	1.29	-0.06	-0.13	0.01
0.39	.59	62.24	0.39	.59	-0.03	62.16	-0.03	-0.23	0.18
0.32	.44	37.95	0.32	.44	0.05	37.41	0.05	-0.08	0.18
0.21	.29	-21.07	0.21	.29	0.03	-21.78	0.03	-0.03	0.09
0.28	.20	20.33	0.28	.20	-0.03	19.70	-0.03	-0.11	0.04

Table C2

Statistics for the model fit of RFT and GEMS to individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
3.27	.40	343.83	3.25	.40	-0.19	342.97	-0.18	-0.53	0.24
7.07	.19	445.45	6.93	.19	0.29	442.79	0.31	-0.05	0.76
16.20	.00	554.93	16.12	.00	0.27	554.28	0.33	-0.38	1.52
3.99	.19	369.95	3.98	.19	-0.08	369.65	-0.06	-0.32	0.22
14.24	.87	537.92	14.24	.87	0.67	537.86	1.68	-10.00	10.00
13.77	.13	533.45	13.76	.13	0.05	533.43	0.08	-0.49	0.93
19.71	.67	580.84	19.70	.67	-0.34	580.75	0.10	-2.61	5.18
14.55	.42	540.78	14.55	.42	-0.01	540.78	0.14	-0.74	1.95
9.87	.43	489.50	9.83	.43	-0.22	488.96	-0.20	-0.72	0.49
14.00	.22	535.61	13.93	.22	0.29	535.05	0.38	-0.39	1.82
5.98	.00	423.28	5.81	.00	0.26	419.47	0.27	0.02	0.60
17.83	.85	567.59	17.83	.85	-0.36	567.56	0.02	-10.00	10.00
9.68	.48	486.98	9.68	.48	0.03	486.98	0.27	-1.09	3.01
8.98	.51	476.99	8.97	.51	-0.14	476.84	-0.08	-0.75	0.86
8.48	.00	469.50	8.04	.00	0.96	462.51	1.02	0.21	2.10
7.35	.89	450.68	7.31	.89	-0.84	449.88	-0.62	-3.77	3.86
8.71	.86	472.94	8.61	.86	-1.17	471.53	-1.38	-10.00	2.71
7.15	.00	446.90	7.08	.00	0.30	445.75	0.35	-0.21	1.38
4.48	.00	385.33	4.45	.00	0.09	384.39	0.09	-0.09	0.29
5.41	.48	410.18	5.40	.48	-0.18	409.80	-0.17	-0.69	0.49
7.81	.43	458.56	7.81	.43	0.04	458.54	0.11	-0.59	1.25
14.19	.24	537.46	14.17	.24	0.16	537.28	0.25	-0.47	1.67
12.51	.39	520.75	12.50	.39	0.10	520.68	0.18	-0.54	1.46
5.60	.18	414.70	5.60	.18	-0.03	414.65	-0.02	-0.24	0.23
9.00	.40	477.35	9.00	.40	0.03	477.34	0.10	-0.57	1.12
8.03	.45	462.21	8.02	.45	-0.02	462.21	0.01	-0.43	0.67
3.27	.71	343.61	3.27	.71	-0.02	343.56	-0.02	-0.16	0.14
5.96	.62	423.02	5.96	.62	0.00	423.02	0.01	-0.28	0.34
3.19	1	340.52	3.19	1	10.00	340.52	1.72	-10.00	10.00
7.62	.97	455.37	7.62	.97	-0.35	455.36	0.28	-10.00	10.00
4.05	.11	371.98	4.05	.11	0.00	371.98	0.00	-0.10	0.12
5.89	.09	421.28	5.89	.09	0.00	421.28	0.01	-0.12	0.14

(continued)

Table C2

Statistics for the model fit of RFT and GEMS to individual participant data from Melrose et al. (2012) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
5.31	.70	407.78	5.31	.70	0.02	407.75	0.02	-0.19	0.26
2.95	1	330.08	2.95	1	1.81	330.08	1.67	-10.00	10.00
2.38	.11	301.66	2.37	.11	0.01	301.46	0.01	-0.02	0.04
10.98	.43	503.54	10.98	.43	-0.01	503.54	0.03	-0.43	0.64
3.13	.14	337.90	3.13	.14	0.01	337.79	0.01	-0.04	0.06
4.76	.40	393.36	4.76	.40	0.01	393.34	0.02	-0.15	0.20
6.49	1	434.21	6.49	1	9.99	434.21	0.79	-10.00	10.00
8.03	0	462.35	7.99	0	-0.11	461.56	-0.10	-0.35	0.17
6.02	1	424.17	6.02	1	9.99	424.17	1.25	-10.00	10.00
10.92	1	502.92	10.92	1	10.00	502.92	-0.87	-10.00	10.00
14.77	1	542.74	14.77	1	-10.00	542.74	-1.10	-10.00	10.00
1.84	.43	267.53	1.83	.43	-0.01	267.40	-0.01	-0.07	0.05
2.38	1	301.57	2.38	1	2.99	301.57	1.30	-10.00	10.00
9.82	1	488.90	9.82	1	2.07	488.90	-0.68	-10.00	10.00
12.21	1	517.58	12.21	1	1.11	517.58	-2.03	-8.08	8.55
7.79	.62	458.34	7.79	.62	0.00	458.34	0.02	-0.27	0.36
1.66	.53	254.16	1.66	.53	0.00	254.14	-0.01	-0.07	0.06
4.71	1	391.94	4.71	1	6.08	391.94	0.20	-9.97	9.42
5.37	.84	409.18	5.37	.84	0.02	409.17	0.02	-0.38	0.48
7.58	1	454.59	7.58	1	1.85	454.59	0.49	-10.00	10.00

Table C3

Statistics for the model fit of RFT and GEMS to individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

	RFT			GEMS			γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
1.46	.69	316.32	1.46	.69	0.02	316.32	0.06	-0.32	0.56
4.91	.19	529.64	4.91	.19	0.01	529.62	0.01	-0.11	0.15
13.61	.43	709.18	13.60	.43	0.01	709.18	0.02	-0.19	0.29
3.65	.58	477.66	3.65	.58	0.03	477.47	0.03	-0.09	0.17
6.22	.09	571.49	6.22	.09	0.00	571.49	0.00	-0.14	0.17
4.33	.45	507.69	4.33	.45	0.02	507.56	0.03	-0.09	0.17
8.89	.33	634.29	8.89	.33	0.01	634.26	0.01	-0.10	0.14
7.69	.62	608.69	7.67	.62	0.05	608.37	0.06	-0.12	0.27
4.01	1	494.16	4.01	1	1.25	494.16	0.27	-10.00	10.00
14.43	.74	719.50	14.43	.74	0.01	719.50	0.09	-0.39	0.78
4.78	.58	525.21	4.78	.58	0.04	525.12	0.05	-0.20	0.35
9.72	.28	650.08	9.72	.28	-0.02	650.04	-0.01	-0.22	0.24
8.79	.41	632.20	8.78	.41	0.02	632.15	0.02	-0.17	0.23
2.96	.62	440.65	2.96	.62	0.00	440.65	0.00	-0.09	0.09
13.15	.41	703.14	13.14	.41	0.01	703.13	0.02	-0.16	0.22
11.15	.25	674.11	11.15	.25	0.01	674.11	0.01	-0.17	0.22
9.38	.23	643.76	9.38	.23	0.00	643.76	0.01	-0.14	0.17
15.80	.80	735.54	15.80	.80	0.01	735.54	0.39	-0.57	10.00
6.47	.71	578.42	6.47	.71	0.02	578.36	0.03	-0.17	0.28
6.80	.58	587.24	6.79	.58	0.04	586.93	0.04	-0.10	0.22
3.68	.86	478.97	3.68	.86	0.06	478.85	0.09	-0.25	0.57
4.19	.71	501.81	4.19	.71	0.01	501.77	0.02	-0.14	0.18
5.44	.73	547.76	5.44	.73	-0.01	547.75	0.00	-0.16	0.19
7.71	.89	609.25	7.71	.89	0.03	609.24	0.51	-0.56	10.00
5.78	.84	558.43	5.77	.84	0.07	558.28	0.10	-0.30	0.61
5.09	1	536.04	5.09	1	0.49	536.04	0.16	-10.00	10.00
7.02	.38	592.83	7.03	.38	0.00	592.83	0.00	-0.11	0.12
7.28	.82	599.23	7.27	.82	0.10	598.91	0.14	-0.26	0.65
16.12	.86	739.02	16.11	.86	0.13	738.96	1.53	-0.97	10.00
5.97	.61	564.11	5.95	.61	0.05	563.59	0.05	-0.10	0.22
7.60	.42	606.69	7.60	.42	0.02	606.65	0.02	-0.14	0.21
4.87	.86	528.47	4.87	.86	0.10	528.19	0.16	-0.29	0.82

(continued)

Table C3

Statistics for the model fit of RFT and GEMS to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
5.23	.95	540.80	5.23	.95	0.05	540.79	1.34	-0.94	10.00
7.02	.86	592.83	7.02	.86	0.05	592.80	0.17	-0.39	0.93
4.98	.55	532.21	4.97	.55	0.03	531.98	0.03	-0.09	0.15
16.48	.42	742.95	16.48	.42	0.01	742.94	0.02	-0.21	0.30
4.32	.65	507.08	4.31	.65	0.04	506.75	0.05	-0.10	0.21
10.16	1	657.75	10.16	1	10.00	657.75	-1.68	-10.00	10.00
10.96	.97	671.17	10.96	.97	10.00	671.06	3.95	-10.00	10.00
1.18	.62	279.11	1.18	.62	0.03	279.01	0.04	-0.15	0.25
12.90	0	699.78	12.70	0	4.18	697.11	3.72	-0.71	5.37
8.94	.69	635.18	8.91	.69	2.50	634.61	1.99	-0.79	5.03
7.46	.51	603.34	7.38	.51	2.35	601.40	2.05	-0.40	4.37
8.66	.44	629.73	8.28	.44	2.37	621.87	2.12	-0.28	3.97
2.07	.83	378.20	2.03	.83	2.30	374.40	2.03	-1.17	6.28
7.04	.32	593.17	7.04	.32	0.11	593.11	0.87	-0.60	3.81
8.53	0	626.99	7.84	0	4.67	612.06	4.59	3.63	5.57
12.90	0	699.78	13.12	0	3.01	702.76	2.91	0.37	4.13
6.03	.30	566.00	6.01	.30	0.19	565.49	0.45	-0.27	2.83
0.00	0	-1054.03	0.00	0	0.00	-1054.03	0.00	0.00	0.00
27.45	.93	832.73	27.45	.93	2.17	832.70	-2.02	-10.00	7.94
3.15	.60	451.88	3.11	.60	2.35	449.35	2.08	-0.80	5.74
9.84	.59	652.06	9.83	.59	0.30	651.99	1.07	-1.39	5.00
7.23	.23	597.91	7.23	.23	0.12	597.81	0.78	-0.54	3.53
9.80	.56	651.39	9.70	.56	2.31	649.70	1.63	-1.34	5.16
7.30	.49	599.59	7.15	.49	2.30	595.94	2.03	-0.12	3.86
1.07	.18	262.44	1.06	.18	2.11	259.43	1.81	-0.02	3.63
4.98	.49	532.12	4.92	.49	2.29	530.01	1.94	-0.63	4.85
6.24	.51	572.09	6.19	.51	2.24	570.53	2.05	-0.05	3.85
5.48	.23	549.04	5.48	.23	0.05	549.02	0.64	-0.58	3.72
13.40	0	706.49	13.06	0	3.17	701.92	3.09	0.79	4.25
23.54	.72	805.69	23.49	.72	2.39	805.31	0.62	-10.00	6.64
14.24	.37	717.18	14.20	.37	2.29	716.65	1.97	-0.78	4.75
11.49	0	679.48	12.26	0	3.32	690.85	3.25	2.24	4.29

(continued)

Table C3

Statistics for the model fit of RFT and GEMS to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
16.43	.69	742.42	16.43	.69	-0.18	742.40	0.68	-2.43	5.82
11.32	.26	676.80	11.20	.26	2.24	674.98	2.03	0.06	3.84
5.99	.30	564.79	6.01	.30	2.12	565.50	1.88	-0.05	3.46
8.26	.00	621.41	8.23	.00	-0.27	620.60	0.04	-0.83	3.61
6.44	.09	577.41	6.42	.09	0.13	577.04	0.38	-0.22	2.75
10.65	.73	666.09	10.65	.73	1.95	666.06	1.50	-1.39	5.17
6.83	.40	587.77	6.84	.40	2.18	588.08	2.02	0.10	3.50
7.37	.27	601.22	7.51	.27	2.15	604.61	2.01	0.23	3.20
14.83	.86	724.38	14.80	.86	2.24	724.01	0.35	-10.00	6.63
13.88	.52	712.73	13.86	.52	2.41	712.40	1.91	-0.91	5.14
6.28	.21	573.03	6.27	.21	-0.10	572.75	0.04	-0.42	2.54
10.13	.79	657.25	10.10	.79	2.06	656.66	1.19	-2.38	5.56
8.23	.52	620.74	8.17	.52	2.14	619.46	1.85	-0.27	3.70
9.00	.64	636.47	8.93	.64	2.29	635.13	1.81	-0.69	4.68
6.04	.45	566.37	6.04	.45	0.11	566.25	0.57	-0.43	3.28
15.91	1	736.69	15.91	1	1.21	736.69	-3.71	-10.00	7.55
7.80	.70	611.36	7.65	.70	2.16	607.91	1.78	-0.79	4.84

Table C4

Statistics for the model fit of RFT and GEMS to individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
0.30	.56	21.71	0.30	.56	0.92	21.03	0.98	-0.96	3.62
0.00	1	18.00	0.00	1	-0.59	18.00	-0.67	-10.00	10.00
0.22	1	16.67	0.22	1	10.00	16.67	2.76	-10.00	10.00
0.29	.42	22.52	0.28	.42	-0.49	22.32	-0.58	-4.28	1.42
0.02	.15	19.38	0.02	.15	1.65	17.55	1.65	1.59	1.70
0.00	.01	8.00	0.00	.01	-0.12	8.00	-0.12	-0.13	-0.12
0.24	.38	20.44	0.24	.38	0.11	20.41	0.14	-0.37	0.71
0.17	1	15.95	0.17	1	-1.82	15.95	2.60	-10.00	10.00
0.78	.86	37.59	0.67	.86	10.00	36.60	5.59	-7.31	10.00
0.00	1	10.00	0.00	1	0.91	10.00	2.18	-10.00	10.00
0.03	.08	25.97	0.00	.08	0.45	24.00	0.45	0.45	0.46
0.32	1	23.24	0.32	1	-1.83	23.24	2.19	-10.00	10.00
0.07	.94	13.61	0.07	.94	-0.57	13.60	-0.80	-6.13	1.98
0.01	0	26.00	0.00	0	0.25	26.00	0.25	0.24	0.25
0.23	1	20.76	0.23	1	-0.15	20.76	2.46	-10.00	10.00
0.07	.86	14.78	0.08	.86	1.30	14.76	1.45	-0.62	4.17
0.35	1	25.85	0.35	1	-1.95	25.85	2.24	-10.00	10.00
1.29	1	44.26	1.29	1	-2.02	44.26	2.67	-10.00	10.00
0.09	.48	11.80	0.09	.48	0.12	11.80	0.14	-0.28	0.64
0.00	1	6.00	0.00	1	-0.65	6.00	0.59	-10.00	10.00
0.00	1	8.00	0.00	1	0.91	8.00	2.26	-10.00	10.00
0.26	0	26.38	0.07	0	5.31	19.14	5.31	4.95	5.68
0.01	0	32.00	0.01	0	8.09	20.00	8.09	8.03	8.16
0.04	0	25.24	0.07	0	5.31	21.14	5.30	4.90	5.69
0.04	.22	11.12	0.07	.22	-1.51	10.83	-1.51	-1.71	-1.33
0.00	.80	18.00	0.00	.80	-0.59	18.00	-0.59	-0.61	-0.56
0.05	.80	11.19	0.05	.80	-0.92	11.16	-0.93	-1.46	-0.44
0.06	1	17.15	0.06	1	-2.25	17.15	2.09	-10.00	10.00
0.08	1	15.30	0.08	1	-10.00	15.30	2.66	-10.00	10.00
0.49	1	31.50	0.49	1	0.39	31.50	2.66	-10.00	10.00
0.04	.58	5.96	0.06	.58	0.52	5.86	0.52	0.12	1.00
0.00	.67	2.00	0.01	.67	0.78	2.00	0.78	0.68	0.88

(continued)

Table C4

Statistics for the model fit of RFT and GEMS to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
1.28	1	45.75	1.28	1	-2.19	45.75	2.49	-10.00	10.00
0.19	1	14.81	0.19	1	-2.45	14.81	2.31	-10.00	10.00
0.38	.67	26.78	0.37	.67	0.84	26.43	0.91	-1.38	4.06
0.00	1	26.00	0.00	1	-0.59	26.00	2.04	-10.00	10.00
0.01	.34	20.00	0.00	.34	-0.31	20.00	-0.31	-0.32	-0.29
0.25	.66	20.50	0.18	.66	-10.00	17.60	-7.65	-10.00	-0.24
0.35	.00	25.54	0.35	.00	-0.05	25.43	-0.05	-0.25	0.12
0.35	.04	25.88	0.35	.04	0.01	25.88	0.00	-0.15	0.17
0.24	0	28.21	0.01	0	-0.49	22.00	-0.49	-0.50	-0.49
0.53	0	32.00	0.03	0	-0.67	15.29	-0.67	-0.68	-0.66
0.11	1	16.18	0.11	1	9.44	16.18	0.41	-10.00	9.97
0.02	.63	4.55	0.02	.63	0.00	4.55	0.00	-0.02	0.03
0.32	.74	22.79	0.32	.74	0.21	22.67	0.67	-0.74	10.00
0.84	.82	39.57	0.83	.82	0.34	39.52	2.03	-5.57	10.00
0.48	0	31.05	0.48	0	-0.02	31.03	-0.02	-0.24	0.19
0.24	0	26.30	0.00	0	-0.22	20.00	-0.22	-0.22	-0.22
0.22	1	16.36	0.22	1	9.98	16.36	1.22	-10.00	10.00
0.00	.67	30.00	0.00	.67	5.39	28.00	5.52	4.76	7.15
0.01	0	20.00	0.02	0	0.35	20.00	0.35	0.34	0.35
0.00	.79	10.00	0.00	.79	5.39	8.00	6.74	3.79	10.00
0.42	.46	29.99	0.42	.46	-0.01	29.99	0.00	-0.48	0.52
0.16	0	17.88	0.05	0	-0.13	16.33	-0.13	-0.14	-0.11
0.20	1	14.32	0.20	1	9.96	14.32	1.39	-10.00	10.00
0.01	0	18.00	0.01	0	-0.13	16.00	-0.13	-0.13	-0.13
0.01	.45	8.00	0.01	.45	0.00	8.00	0.00	-0.01	0.01
0.09	1	14.37	0.09	1	10.00	14.37	0.63	-10.00	10.00
0.00	.95	22.00	0.00	.95	5.39	22.00	6.72	3.78	10.00
0.10	1	9.26	0.10	1	7.69	9.26	1.09	-10.00	10.00
0.15	1	21.30	0.15	1	10.00	21.30	1.38	-10.00	10.00
0.12	1	11.17	0.12	1	6.62	11.17	1.40	-10.00	10.00
0.17	.77	15.66	0.17	.77	-0.02	15.66	-0.02	-0.35	0.31
0.01	0	20.00	0.01	0	1.71	20.00	1.71	1.67	1.76

(continued)

Table C4

Statistics for the model fit of RFT and GEMS to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
0.09	1	4.69	0.09	1	9.95	4.69	0.27	-10.00	9.99
0.05	0	11.86	0.05	0	-0.03	11.85	-0.03	-0.06	0.01
0.49	.06	31.40	0.49	.06	0.00	31.40	0.00	-0.30	0.34
0.08	.75	17.94	0.08	.75	0.01	17.94	0.01	-0.12	0.14
0.29	.22	22.20	0.29	.22	0.04	22.14	0.04	-0.19	0.33
0.00	.79	20.00	0.01	.79	0.00	20.00	0.00	-0.02	0.02
0.50	.82	32.84	0.50	.82	0.10	32.82	1.13	-1.24	10.00
0.09	.28	13.16	0.09	.28	0.00	13.16	0.00	-0.06	0.06
0.15	.21	15.38	0.16	.21	0.02	15.35	0.02	-0.06	0.10
0.30	1	21.95	0.30	1	0.61	21.95	1.18	-10.00	10.00
0.04	.50	7.85	0.04	.50	0.01	7.85	0.01	-0.03	0.04
0.26	0	31.92	0.01	0	3.83	20.00	3.83	3.62	4.08
0.09	1	9.86	0.09	1	1.64	9.86	0.39	-10.00	9.99
0.18	.52	12.05	0.19	.52	0.05	11.97	0.05	-0.15	0.23

Table C5

Statistics for the model fit of RFT and GEMS to individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
0.03	0	11.27	0.04	0	2.14	11.25	2.14	1.95	2.32
0.27	.02	32.01	0.26	.02	0.36	31.64	0.37	-0.36	1.26
0.56	.16	50.37	0.55	.16	-0.37	50.08	-0.42	-1.61	0.96
3.99	1	74.08	3.99	1	-1.87	74.08	2.16	-10.00	10.00
3.19	1	71.61	3.19	1	-1.52	71.61	2.24	-10.00	10.00
6.19	1	74.88	6.19	1	-1.06	74.88	1.35	-9.92	10.00
0.28	0	33.15	0.26	0	0.55	32.41	0.55	0.13	1.04
0.05	0	12.52	0.01	0	0.43	7.95	0.43	0.42	0.45
0.66	.18	53.13	0.66	.18	0.03	53.13	0.04	-2.18	2.17
0.42	.47	46.12	0.39	.47	-1.71	45.08	-1.90	-4.87	-0.41
1.10	0	64.35	1.08	0	0.50	64.14	0.57	-0.90	2.61
0.60	.11	55.09	0.59	.11	-0.30	54.75	-0.32	-1.20	0.83
0.38	.00	45.26	0.38	.00	-0.07	45.23	-0.07	-0.57	0.49
0.28	.04	39.97	0.28	.04	0.26	39.72	0.28	-0.14	0.74
0.29	0	39.29	0.16	0	1.32	29.80	1.32	1.05	1.61
0.37	0	40.14	0.37	0	-0.05	40.13	-0.03	-0.44	0.45
0.23	0	31.29	0.15	0	1.15	29.02	1.14	0.84	1.44
1.24	.48	62.83	1.22	.48	1.21	62.69	1.39	-10.00	10.00
2.63	.46	73.23	2.60	.46	-0.69	73.18	-0.58	-10.00	10.00
1.38	.09	69.52	1.38	.09	-0.13	69.50	-0.22	-4.98	2.43
1.67	.17	71.72	1.67	.17	-0.16	71.70	-0.37	-6.88	2.53
0.33	.05	35.20	0.31	.05	0.53	33.64	0.55	0.08	1.09
0.26	.03	30.53	0.24	.03	-0.47	26.91	-0.47	-0.71	-0.25
0.84	0	59.31	0.83	0	-0.27	59.20	-0.42	-4.47	1.25
0.29	.08	37.46	0.29	.08	0.15	37.37	0.16	-0.31	0.72
0.63	.04	51.82	0.60	.04	-0.59	51.24	-0.62	-1.66	0.37
0.20	0	36.70	0.05	0	1.73	29.50	1.73	1.60	1.86
0.70	0	56.94	0.69	0	0.27	56.45	0.30	-0.32	1.03
0.51	.12	49.44	0.49	.12	-0.31	48.66	-0.31	-0.75	0.16
0.63	0	53.86	0.63	0	0.20	53.69	0.23	-0.38	0.98
2.18	.05	68.73	2.16	.05	-0.33	68.69	-0.51	-7.13	3.18
0.47	0	45.17	0.46	0	0.49	44.19	0.50	-0.05	1.14

(continued)

Table C5

Statistics for the model fit of RFT and GEMS to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
0.98	0	61.30	0.92	0	1.02	60.24	1.02	-0.18	2.56
0.02	.14	23.40	0.04	.14	0.56	23.27	0.56	0.49	0.64
0.80	0	55.94	0.80	0	1.19	55.34	1.23	-0.03	2.67
0.69	0	52.44	0.69	0	0.01	52.44	0.02	-0.68	0.75
0.42	.04	43.84	0.41	.04	-0.20	43.40	-0.20	-0.61	0.22
1.44	.48	63.90	1.43	.48	0.73	63.85	0.90	-10.00	10.00
0.01	.65	42.00	0.02	.65	0.29	42.00	0.30	0.22	0.38
1.08	0	64.41	0.88	0	1.45	61.16	1.50	0.38	2.85
0.98	0	60.06	0.26	0	4.00	27.33	4.01	3.55	4.47
1.11	0	61.09	1.10	0	1.52	59.50	1.50	0.05	2.97
0.87	.17	60.53	0.83	.17	-0.54	59.87	-0.64	-3.40	0.58
0.32	0	35.73	0.00	0	2.09	0.00	2.09	2.09	2.09
1.08	0	62.10	0.58	0	2.35	52.27	2.37	1.41	3.33
0.41	0	45.58	0.40	0	0.74	42.10	0.75	0.30	1.26
0.23	0	37.11	0.04	0	2.09	25.25	2.09	2.03	2.15
0.51	0	50.80	0.44	0	1.18	47.08	1.19	0.61	1.79
1.74	0	64.17	1.66	0	1.40	63.54	1.54	-0.50	4.23
2.18	.66	68.27	2.18	.66	0.89	68.24	1.52	-10.00	10.00
0.15	0	45.18	0.22	0	3.69	40.69	3.70	3.34	4.07
0.29	0	42.29	0.01	0	3.36	31.73	3.36	3.34	3.37
0.31	0	37.54	0.30	0	-0.25	37.01	-0.24	-0.53	0.08
0.26	.41	35.61	0.26	.41	0.00	35.61	0.02	-0.51	0.69
0.30	0	34.15	0.21	0	2.03	24.60	2.02	1.43	2.63
0.00	.19	38.00	0.00	.19	5.42	34.00	5.42	5.41	5.42
0.19	.08	31.14	0.07	.08	1.55	30.30	1.55	1.45	1.65
0.79	0	55.78	0.35	0	2.34	45.74	2.34	1.83	2.90
3.65	.46	71.57	3.64	.46	1.23	71.51	1.47	-10.00	10.00
0.52	0	45.60	0.27	0	2.42	36.29	2.42	1.91	2.90
0.35	.14	39.40	0.35	.14	-0.21	39.16	-0.21	-0.75	0.36
2.66	1	68.21	2.66	1	9.99	68.21	2.56	-10.00	10.00
0.42	.31	44.38	0.42	.31	-0.21	44.33	-0.21	-1.25	0.99
0.44	.05	45.46	0.43	.05	-0.28	45.10	-0.28	-0.93	0.46

(continued)

Table C5

Statistics for the model fit of RFT and GEMS to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

	RFT			GEMS			γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
0.00	.05	34.00	0.00	.05	9.98	26.00	9.98	9.97	9.99
0.00	1	42.00	0.00	1	-0.59	42.00	-0.48	-10.00	10.00
0.28	0	48.08	0.01	0	8.53	34.00	8.53	8.49	8.56
0.52	0	51.43	0.46	0	4.83	49.84	4.97	2.61	8.15
2.24	0	71.76	1.61	0	-0.44	68.92	-0.45	-0.87	-0.07
0.26	0	29.48	0.00	0	-0.29	4.00	-0.29	-0.29	-0.29
0.80	0	60.98	0.79	0	-0.02	60.96	-0.02	-0.24	0.23
2.12	.34	70.60	2.12	.34	0.02	70.59	0.05	-0.61	0.95
0.43	.86	49.18	0.43	.86	-0.46	49.11	-0.52	-6.68	1.73
0.67	0	55.57	0.49	0	-0.23	48.53	-0.23	-0.34	-0.12
0.71	0	59.01	0.68	0	-0.15	58.12	-0.15	-0.42	0.13
0.92	0	62.74	0.71	0	-0.24	58.29	-0.24	-0.41	-0.07
0.46	0	46.08	0.44	0	-0.10	45.33	-0.10	-0.24	0.04
0.85	0	60.28	0.78	0	-0.20	58.70	-0.20	-0.44	0.05
0.54	0	50.55	0.54	0	-0.06	49.99	-0.07	-0.19	0.05
3.36	0	75.86	2.33	0	-0.38	74.47	-0.39	-0.91	0.10
0.78	0	57.28	0.76	0	-0.08	56.86	-0.08	-0.28	0.12
0.62	0	52.52	0.62	0	-0.06	52.32	-0.05	-0.23	0.13
2.33	0	70.47	2.03	0	-0.27	69.36	-0.27	-0.64	0.10
0.30	0	33.32	0.25	0	-0.12	27.81	-0.12	-0.18	-0.06
0.35	0	42.73	0.35	0	-0.01	42.72	-0.01	-0.08	0.07
0.65	.02	50.67	0.65	.02	-0.01	50.66	-0.01	-0.13	0.12
2.75	0	72.68	2.01	0	-0.40	70.53	-0.40	-0.76	-0.05
0.63	0	54.01	0.57	0	-0.13	51.93	-0.13	-0.26	0.01
0.78	0	54.10	0.78	0	-0.05	53.98	-0.05	-0.20	0.11
0.04	0	5.59	0.13	0	-0.11	4.58	-0.11	-0.13	-0.09
0.04	0	7.18	0.04	0	0.00	7.16	0.00	-0.01	0.00
0.31	0	33.61	0.28	0	-0.13	26.58	-0.13	-0.18	-0.08
0.31	0	37.07	0.30	0	-0.07	34.68	-0.07	-0.13	-0.01
0.33	.11	36.97	0.34	.11	0.03	36.85	0.03	-0.05	0.10
0.66	0	53.64	0.40	0	-0.28	41.86	-0.28	-0.36	-0.19
0.40	0	48.03	0.37	0	-0.11	45.78	-0.11	-0.19	-0.02

(continued)

Table C5

Statistics for the model fit of RFT and GEMS to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
1.49	0	65.25	1.42	0	-0.22	64.57	-0.21	-0.54	0.11
0.43	0	46.23	0.42	0	-0.03	45.99	-0.03	-0.11	0.05
0.99	0	61.00	0.77	0	-0.23	56.55	-0.23	-0.38	-0.09
1.15	0	63.60	1.05	0	-0.17	62.28	-0.16	-0.37	0.05
1.14	0	65.65	1.04	0	-0.14	64.42	-0.14	-0.34	0.08
0.16	0	17.81	0.17	0	-0.01	17.74	-0.01	-0.04	0.02
3.57	0	73.66	3.24	0	-0.48	73.10	-0.60	-4.08	0.72
0.62	0	52.32	0.40	0	-0.28	43.23	-0.28	-0.36	-0.20
3.88	0	72.78	3.81	0	-0.18	72.67	-0.10	-0.97	0.72
0.24	0	33.04	0.24	0	-0.06	32.46	-0.06	-0.14	0.02
1.56	0	62.60	1.52	0	-0.15	62.43	-0.15	-0.46	0.20
0.63	0	51.51	0.58	0	-0.15	50.24	-0.15	-0.28	-0.01
0.98	0	61.18	0.83	0	-0.22	58.16	-0.22	-0.39	-0.06
0.88	0	58.31	0.77	0	-0.22	55.73	-0.22	-0.41	-0.04
7.87	0	77.26	7.49	0	-0.25	77.19	0.00	-9.41	10.00
0.41	0	48.89	0.31	0	-0.29	44.86	-0.29	-0.34	-0.23
0.60	1	51.87	0.60	1	2.77	51.87	0.13	-9.29	7.01
1.62	0	64.20	1.56	0	-0.21	63.76	-0.21	-0.58	0.20
0.86	.07	59.40	0.86	.07	-0.03	59.37	-0.02	-0.29	0.29
0.50	0	47.36	0.39	0	-0.18	44.56	-0.18	-0.25	-0.11
0.18	0	29.06	0.00	0	-0.27	12.00	-0.27	-0.27	-0.27
0.94	.72	58.17	1.27	.72	-1.12	58.24	-1.69	-9.88	-0.12
0.00	.77	32.00	0.00	.77	0.01	32.00	0.01	0.01	0.01
3.06	.00	72.54	2.79	.00	-0.38	72.01	-0.40	-1.36	0.48
0.00	1	42.00	0.00	1	2.52	42.00	1.56	-10.00	10.00
3.72	0	72.61	3.04	0	-1.13	71.16	-2.01	-10.00	0.14
5.07	0	74.49	3.70	0	-1.04	73.03	-1.98	-10.00	0.13
0.31	0	37.06	0.31	0	0.00	37.05	-0.01	-0.07	0.06
5.95	0	74.76	5.48	0	-0.63	74.45	-1.24	-10.00	10.00
0.10	.56	40.84	0.01	.56	0.13	39.62	0.13	0.12	0.14
0.22	.36	42.59	0.22	.36	0.02	42.55	0.02	-0.04	0.09
0.81	0	55.96	0.45	0	-0.40	46.95	-0.40	-0.51	-0.31

(continued)

Table C5

Statistics for the model fit of RFT and GEMS to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			GEMS				γ bootstrap		
SD	*w*	-2lnL	SD	w	γ	-2lnL	Mean	95% CI	
2.43	0	66.59	2.25	0	-0.52	65.67	-0.54	-1.15	-0.02
0.36	.47	44.43	0.36	.47	0.06	44.40	0.06	-0.18	0.31
6.32	0	75.07	4.66	0	-0.88	73.77	-1.34	-9.12	-0.04
4.88	0	73.66	4.34	0	-0.55	73.06	-0.75	-4.73	0.43
1.45	0	63.14	1.43	0	-0.35	62.69	-0.35	-1.06	0.33
0.35	1	49.72	0.35	1	8.92	49.72	0.58	-10.00	10.00
0.00	1	48.00	0.00	1	5.39	48.00	1.29	-10.00	10.00

Appendix D:

RFT and GEMS Comparison

Table D1

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
80.68	84.68	84.87	89.06	77.86	83.86	84.25	90.43	.40	.60	.67	.33
34.95	38.95	39.14	43.33	28.87	34.87	35.26	41.44	.12	.88	.28	.72
80.21	84.21	84.41	88.59	79.35	85.35	85.73	91.92	.64	.36	.84	.16
-5.59	-1.59	-1.40	2.78	-5.60	0.40	0.79	6.97	.73	.27	.89	.11
-49.96	-45.96	-45.77	-41.58	-49.96	-43.96	-43.57	-37.39	.73	.27	.89	.11
72.29	76.29	76.48	80.67	71.92	77.92	78.30	84.49	.69	.31	.87	.13
-20.43	-16.43	-16.24	-12.05	-20.77	-14.77	-14.39	-8.20	.70	.30	.87	.13
97.51	101.51	101.70	105.88	92.99	98.99	99.38	105.56	.22	.78	.46	.54
109.72	113.72	113.91	118.10	109.28	115.28	115.66	121.85	.69	.31	.87	.13
93.64	97.64	97.84	102.02	93.14	99.14	99.53	105.71	.68	.32	.86	.14
77.00	81.00	81.19	85.38	76.82	82.82	83.21	89.39	.71	.29	.88	.12
75.15	79.15	79.34	83.53	74.77	80.77	81.15	87.34	.69	.31	.87	.13
76.53	80.53	80.72	84.91	76.06	82.06	82.45	88.63	.68	.32	.87	.13
19.05	23.05	23.24	27.42	17.27	23.27	23.66	29.84	.53	.47	.77	.23
-30.79	-26.79	-26.60	-22.41	-31.66	-25.66	-25.27	-19.09	.64	.36	.84	.16
90.52	94.52	94.71	98.90	90.12	96.12	96.50	102.68	.69	.31	.87	.13
-50.13	-46.13	-45.93	-41.75	-50.14	-44.14	-43.75	-37.57	.73	.27	.89	.11
52.35	56.35	56.54	60.73	52.30	58.30	58.69	64.87	.73	.27	.89	.11
115.03	119.03	119.22	123.41	84.45	90.45	90.83	97.02	0	1	0	1
3.66	7.66	7.85	12.04	1.29	7.29	7.68	13.86	.45	.55	.71	.29
62.24	66.24	66.43	70.62	62.16	68.16	68.55	74.73	.72	.28	.89	.11
37.95	41.95	42.14	46.32	37.41	43.41	43.80	49.98	.68	.32	.86	.14
-21.07	-17.07	-16.88	-12.69	-21.78	-15.78	-15.39	-9.21	.66	.34	.85	.15
20.33	24.33	24.52	28.71	19.70	25.70	26.09	32.27	.67	.33	.86	.14

Table D2

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
343.83	347.83	348.02	352.21	342.97	348.97	349.35	355.54	.64	.36	.84	.16
445.45	449.45	449.64	453.83	442.79	448.79	449.18	455.36	.42	.58	.68	.32
554.93	558.93	559.12	563.31	554.28	560.28	560.67	566.85	.66	.34	.85	.15
369.95	373.95	374.14	378.33	369.65	375.65	376.04	382.22	.70	.30	.87	.13
537.92	541.92	542.11	546.30	537.86	543.86	544.25	550.43	.72	.28	.89	.11
533.45	537.45	537.64	541.83	533.43	539.43	539.82	546.00	.73	.27	.89	.11
580.84	584.84	585.03	589.22	580.75	586.75	587.13	593.31	.72	.28	.89	.11
540.78	544.78	544.97	549.16	540.78	546.78	547.17	553.35	.73	.27	.89	.11
489.50	493.50	493.69	497.88	488.96	494.96	495.35	501.53	.68	.32	.86	.14
535.61	539.61	539.80	543.99	535.05	541.05	541.43	547.62	.67	.33	.86	.14
423.28	427.28	427.47	431.66	419.47	425.47	425.85	432.03	.29	.71	.55	.45
567.59	571.59	571.78	575.97	567.56	573.56	573.95	580.13	.73	.27	.89	.11
486.98	490.98	491.17	495.36	486.98	492.98	493.36	499.55	.73	.27	.89	.11
476.99	480.99	481.18	485.37	476.84	482.84	483.23	489.41	.72	.28	.88	.12
469.50	473.50	473.69	477.88	462.51	468.51	468.89	475.08	.08	.92	.20	.80
450.68	454.68	454.88	459.06	449.88	455.88	456.27	462.45	.65	.35	.84	.16
472.94	476.94	477.13	481.32	471.53	477.53	477.91	484.10	.57	.43	.80	.20
446.90	450.90	451.09	455.28	445.75	451.75	452.14	458.32	.60	.40	.82	.18
385.33	389.33	389.52	393.71	384.39	390.39	390.78	396.96	.63	.37	.84	.16
410.18	414.18	414.37	418.56	409.80	415.80	416.18	422.37	.69	.31	.87	.13
458.56	462.56	462.75	466.93	458.54	464.54	464.93	471.11	.73	.27	.89	.11
537.46	541.46	541.65	545.84	537.28	543.28	543.67	549.85	.71	.29	.88	.12
520.75	524.75	524.94	529.13	520.68	526.68	527.07	533.25	.72	.28	.89	.11
414.70	418.70	418.89	423.08	414.65	420.65	421.04	427.22	.73	.27	.89	.11
477.35	481.35	481.54	485.73	477.34	483.34	483.72	489.91	.73	.27	.89	.11
462.21	466.21	466.40	470.59	462.21	468.21	468.59	474.77	.73	.27	.89	.11
343.61	347.61	347.80	351.99	343.56	349.56	349.95	356.13	.73	.27	.89	.11
423.02	427.02	427.21	431.40	423.02	429.02	429.40	435.58	.73	.27	.89	.11
340.52	344.52	344.71	348.90	340.52	346.52	346.91	353.09	.73	.27	.89	.11
455.37	459.37	459.57	463.75	455.36	461.36	461.75	467.93	.73	.27	.89	.11
371.98	375.98	376.17	380.36	371.98	377.98	378.37	384.55	.73	.27	.89	.11
421.28	425.28	425.48	429.66	421.28	427.28	427.67	433.85	.73	.27	.89	.11

(continued)

Table D2

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Melrose et al. (2012) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
407.78	411.78	411.97	416.16	407.75	413.75	414.14	420.32	.73	.27	.89	.11
330.08	334.08	334.27	338.46	330.08	336.08	336.46	342.65	.73	.27	.89	.11
301.66	305.66	305.85	310.04	301.46	307.46	307.85	314.03	.71	.29	.88	.12
503.54	507.54	507.73	511.92	503.54	509.54	509.93	516.11	.73	.27	.89	.11
337.90	341.90	342.09	346.28	337.79	343.79	344.18	350.36	.72	.28	.88	.12
393.36	397.36	397.55	401.74	393.34	399.34	399.72	405.91	.73	.27	.89	.11
434.21	438.21	438.40	442.59	434.21	440.21	440.60	446.78	.73	.27	.89	.11
462.35	466.35	466.54	470.73	461.56	467.56	467.95	474.13	.65	.35	.85	.15
424.17	428.17	428.36	432.55	424.17	430.17	430.56	436.74	.73	.27	.89	.11
502.92	506.92	507.11	511.30	502.92	508.92	509.30	515.49	.73	.27	.89	.11
542.74	546.74	546.93	551.11	542.74	548.74	549.12	555.30	.73	.27	.89	.11
267.53	271.53	271.72	275.91	267.40	273.40	273.78	279.97	.72	.28	.88	.12
301.57	305.57	305.76	309.95	301.57	307.57	307.96	314.14	.73	.27	.89	.11
488.90	492.90	493.09	497.28	488.90	494.90	495.28	501.47	.73	.27	.89	.11
517.58	521.58	521.77	525.96	517.58	523.58	523.97	530.15	.73	.27	.89	.11
458.34	462.34	462.53	466.72	458.34	464.34	464.73	470.91	73	.27	.89	.11
254.16	258.16	258.35	262.54	254.14	260.14	260.52	266.71	.73	.27	.89	.11
391.94	395.94	396.13	400.32	391.94	397.94	398.33	404.51	.73	.27	.89	.11
409.18	413.18	413.37	417.56	409.17	415.17	415.56	421.74	.73	.27	.89	.11
454.59	458.59	458.78	462.97	454.59	460.59	460.98	467.16	.73	.27	.89	.11

Table D3

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
316.32	320.32	320.61	323.89	316.32	322.32	322.92	327.67	.73	.27	.87	.13
529.64	533.64	533.93	537.20	529.62	535.62	536.22	540.97	.73	.27	.87	.13
709.18	713.18	713.48	716.75	709.18	715.18	715.78	720.53	.73	.27	.87	.13
477.66	481.66	481.95	485.23	477.47	483.47	484.07	488.82	.71	.29	.86	.14
571.49	575.49	575.78	579.06	571.49	577.49	578.09	582.84	.73	.27	.87	.13
507.69	511.69	511.98	515.26	507.56	513.56	514.16	518.91	.72	.28	.86	.14
634.29	638.29	638.58	641.86	634.26	640.26	640.86	645.61	.73	.27	.87	.13
608.69	612.69	612.99	616.26	608.37	614.37	614.97	619.73	.70	.30	.85	.15
494.16	498.16	498.46	501.73	494.16	500.16	500.76	505.52	.73	.27	.87	.13
719.50	723.50	723.80	727.07	719.50	725.50	726.10	730.85	.73	.27	.87	.13
525.21	529.21	529.50	532.78	525.12	531.12	531.72	536.48	.72	.28	.86	.14
650.08	654.08	654.37	657.65	650.04	656.04	656.64	661.39	.73	.27	.87	.13
632.20	636.20	636.49	639.77	632.15	638.15	638.75	643.51	.73	.27	.87	.13
440.65	444.65	444.95	448.22	440.65	446.65	447.25	452.00	.73	.27	.87	.13
703.14	707.14	707.43	710.70	703.13	709.13	709.73	714.48	.73	.27	.87	.13
674.11	678.11	678.40	681.68	674.11	680.11	680.71	685.46	.73	.27	.87	.13
643.76	647.76	648.06	651.33	643.76	649.76	650.36	655.11	.73	.27	.87	.13
735.54	739.54	739.83	743.11	735.54	741.54	742.14	746.89	.73	.27	.87	.13
578.42	582.42	582.71	585.99	578.36	584.36	584.96	589.71	.73	.27	.87	.13
587.24	591.24	591.53	594.81	586.93	592.93	593.53	598.29	.70	.30	.85	.15
478.97	482.97	483.26	486.54	478.85	484.85	485.45	490.20	.72	.28	.86	.14
501.81	505.81	506.10	509.38	501.77	507.77	508.37	513.13	.73	.27	.87	.13
547.76	551.76	552.05	555.33	547.75	553.75	554.35	559.11	.73	.27	.87	.13
609.25	613.25	613.54	616.82	609.24	615.24	615.84	620.59	.73	.27	.87	.13
558.43	562.43	562.73	566.00	558.28	564.28	564.88	569.63	.72	.28	.86	.14
536.04	540.04	540.33	543.61	536.04	542.04	542.64	547.39	.73	.27	.87	.13
592.83	596.83	597.12	600.40	592.83	598.83	599.43	604.18	.73	.27	.87	.13
599.23	603.23	603.53	606.80	598.91	604.91	605.51	610.26	.70	.30	.85	.15
739.02	743.02	743.32	746.59	738.96	744.96	745.56	750.31	.72	.28	.87	.13
564.11	568.11	568.40	571.67	563.59	569.59	570.19	574.94	.68	.32	.84	.16
606.69	610.69	610.99	614.26	606.65	612.65	613.25	618.00	.73	.27	.87	.13
528.47	532.47	532.76	536.04	528.19	534.19	534.79	539.54	.70	.30	.85	.15

(continued)

Table D3

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
540.80	544.80	545.10	548.37	540.79	546.79	547.39	552.15	.73	.27	.87	.13
592.83	596.83	597.12	600.40	592.80	598.80	599.40	604.15	.73	.27	.87	.13
532.21	536.21	536.50	539.77	531.98	537.98	538.58	543.33	.71	.29	.86	.14
742.95	746.95	747.24	750.51	742.94	748.94	749.54	754.30	.73	.27	.87	.13
507.08	511.08	511.37	514.65	506.75	512.75	513.35	518.10	.70	.30	.85	.15
657.75	661.75	662.05	665.32	657.75	663.75	664.35	669.11	.73	.27	.87	.13
671.17	675.17	675.47	678.74	671.06	677.06	677.66	682.42	.72	.28	.86	.14
279.11	283.11	283.41	286.68	279.01	285.01	285.61	290.36	.72	.28	.86	.14
699.78	703.78	704.07	707.35	697.11	703.11	703.71	708.47	.42	.58	.64	.36
635.18	639.18	639.47	642.75	634.61	640.61	641.21	645.96	.67	.33	.83	.17
603.34	607.34	607.63	610.91	601.40	607.40	608.00	612.75	.51	.49	.72	.28
629.73	633.73	634.03	637.30	621.87	627.87	628.47	633.23	.05	.95	.12	.88
378.20	382.20	382.50	385.77	374.40	380.40	381.00	385.75	.29	.71	.50	.50
593.17	597.17	597.46	600.73	593.11	599.11	599.71	604.46	.73	.27	.87	.13
626.99	630.99	631.28	634.56	612.06	618.06	618.66	623.41	.00	1	0	1
699.78	703.78	704.07	707.35	702.76	708.76	709.36	714.11	.92	.08	.97	.03
566.00	570.00	570.29	573.57	565.49	571.49	572.09	576.84	.68	.32	.84	.16
832.73	836.73	837.02	840.30	832.70	838.70	839.30	844.05	.73	.27	.87	.13
451.88	455.88	456.17	459.45	449.35	455.35	455.95	460.70	.43	.57	.65	.35
652.06	656.06	656.36	659.63	651.99	657.99	658.59	663.34	.72	.28	.86	.14
597.91	601.91	602.20	605.48	597.81	603.81	604.41	609.16	.72	.28	.86	.14
651.39	655.39	655.68	658.96	649.70	655.70	656.30	661.05	.54	.46	.74	.26
599.59	603.59	603.88	607.16	595.94	601.94	602.54	607.29	.30	.70	.52	.48
262.44	266.44	266.73	270.01	259.43	265.43	266.03	270.78	.38	.62	.60	.40
532.12	536.12	536.42	539.69	530.01	536.01	536.61	541.36	.49	.51	.70	.30
572.09	576.09	576.38	579.66	570.53	576.53	577.13	581.89	.56	.44	.75	.25
549.04	553.04	553.33	556.61	549.02	555.02	555.62	560.38	.73	.27	.87	.13
706.49	710.49	710.78	714.06	701.92	707.92	708.52	713.27	.22	.78	.40	.60
805.69	809.69	809.98	813.26	805.31	811.31	811.91	816.66	.69	.31	.85	.15
717.18	721.18	721.47	724.75	716.65	722.65	723.25	728.00	.68	.32	.84	.16
679.48	683.48	683.77	687.04	690.85	696.85	697.45	702.21	1	0	1	0

(continued)

Table D3

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
742.42	746.42	746.71	749.99	742.40	748.40	749.00	753.75	.73	.27	.87	.13
676.80	680.80	681.09	684.37	674.98	680.98	681.58	686.33	.52	.48	.73	.27
564.79	568.79	569.08	572.36	565.50	571.50	572.10	576.85	.79	.21	.90	.10
621.41	625.41	625.71	628.98	620.60	626.60	627.20	631.95	.64	.36	.82	.18
577.41	581.41	581.70	584.98	577.04	583.04	583.64	588.39	.69	.31	.85	.15
666.09	670.09	670.39	673.66	666.06	672.06	672.66	677.41	.73	.27	.87	.13
587.77	591.77	592.06	595.34	588.08	594.08	594.68	599.43	.76	.24	.89	.11
601.22	605.22	605.51	608.79	604.61	610.61	611.21	615.96	.94	.06	.97	.03
724.38	728.38	728.67	731.95	724.01	730.01	730.61	735.36	.69	.31	.85	.15
712.73	716.73	717.03	720.30	712.40	718.40	719.00	723.75	.70	.30	.85	.15
573.03	577.03	577.32	580.59	572.75	578.75	579.35	584.11	.70	.30	.85	.15
657.25	661.25	661.54	664.82	656.66	662.66	663.26	668.01	.67	.33	.83	.17
620.74	624.74	625.04	628.31	619.46	625.46	626.06	630.81	.59	.41	.78	.22
636.47	640.47	640.77	644.04	635.13	641.13	641.73	646.48	.58	.42	.77	.23
566.37	570.37	570.66	573.94	566.25	572.25	572.85	577.60	.72	.28	.86	.14
736.69	740.69	740.98	744.26	736.69	742.69	743.29	748.04	.73	.27	.87	.13
611.36	615.36	615.65	618.93	607.91	613.91	614.51	619.26	.33	.67	.54	.46

Table D4

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
21.71	25.71	26.34	27.89	21.03	27.03	28.36	30.30	.66	.34	.77	.23
18.00	22.00	22.63	24.18	18.00	24.00	25.33	27.27	.73	.27	.82	.18
16.67	20.67	21.31	22.86	16.67	22.67	24.01	25.95	.73	.27	.82	.18
22.52	26.52	27.15	28.70	22.32	28.32	29.65	31.59	.71	.29	.81	.19
19.38	23.38	24.01	25.56	17.55	23.55	24.89	26.83	.52	.48	.65	.35
8.00	12.00	12.63	14.18	8.00	14.00	15.33	17.27	.73	.27	.82	.18
20.44	24.44	25.08	26.63	20.41	26.41	27.74	29.68	.73	.27	.82	.18
15.95	19.95	20.58	22.13	15.95	21.95	23.28	25.22	.73	.27	.82	.18
37.59	41.59	42.22	43.77	36.60	42.60	43.93	45.87	.62	.38	.74	.26
10.00	14.00	14.63	16.18	10.00	16.00	17.33	19.27	.73	.27	.82	.18
25.97	29.97	30.60	32.15	24.00	30.00	31.33	33.27	.50	.50	.64	.36
23.24	27.24	27.88	29.43	23.24	29.24	30.58	32.52	.73	.27	.82	.18
13.61	17.61	18.24	19.79	13.60	19.60	20.93	22.87	.73	.27	.82	.18
26.00	30.00	30.63	32.18	26.00	32.00	33.33	35.27	.73	.27	.82	.18
20.76	24.76	25.39	26.94	20.76	26.76	28.09	30.03	.73	.27	.82	.18
14.78	18.78	19.41	20.96	14.76	20.76	22.09	24.03	.73	.27	.82	.18
25.85	29.85	30.48	32.03	25.85	31.85	33.18	35.12	.73	.27	.82	.18
44.26	48.26	48.89	50.44	44.26	50.26	51.60	53.54	.73	.27	.82	.18
11.80	15.80	16.43	17.98	11.80	17.80	19.13	21.07	.73	.27	.82	.18
6.00	10.00	10.63	12.18	6.00	12.00	13.33	15.27	.73	.27	.82	.18
8.00	12.00	12.63	14.18	8.00	14.00	15.33	17.27	.73	.27	.82	.18
26.38	30.38	31.01	32.56	19.14	25.14	26.47	28.41	.07	.93	.11	.89
32.00	36.00	36.63	38.18	20.00	26.00	27.33	29.27	.01	.99	.01	.99
25.24	29.24	29.87	31.42	21.14	27.14	28.47	30.41	.26	.74	.38	.62
11.12	15.12	15.76	17.31	10.83	16.83	18.17	20.11	.70	.30	.80	.20
18.00	22.00	22.63	24.18	18.00	24.00	25.33	27.27	.73	.27	.82	.18
11.19	15.19	15.82	17.38	11.16	17.16	18.50	20.44	.73	.27	.82	.18
17.15	21.15	21.78	23.33	17.15	23.15	24.48	26.42	.73	.27	.82	.18
15.30	19.30	19.94	21.49	15.30	21.30	22.64	24.58	.73	.27	.82	.18
31.50	35.50	36.13	37.68	31.50	37.50	38.83	40.77	.73	.27	.82	.18
5.96	9.96	10.59	12.14	5.86	11.86	13.19	15.13	.72	.28	.82	.18
2.00	6.00	6.63	8.18	2.00	8.00	9.33	11.27	.73	.27	.82	.18

(continued)

Table D4

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
45.75	49.75	50.38	51.93	45.75	51.75	53.08	55.02	.73	.27	.82	.18
14.81	18.81	19.44	20.99	14.81	20.81	22.15	24.09	.73	.27	.82	.18
26.78	30.78	31.41	32.96	26.43	32.43	33.76	35.70	.70	.30	.80	.20
26.00	30.00	30.63	32.18	26.00	32.00	33.33	35.27	.73	.27	.82	.18
20.00	24.00	24.63	26.18	20.00	26.00	27.33	29.27	.73	.27	.82	.18
20.50	24.50	25.13	26.68	17.60	23.60	24.94	26.88	.39	.61	.52	.48
25.54	29.54	30.17	31.72	25.43	31.43	32.76	34.70	.72	.28	.82	.18
25.88	29.88	30.51	32.06	25.88	31.88	33.21	35.15	.73	.27	.82	.18
28.21	32.21	32.84	34.39	22.00	28.00	29.33	31.27	.11	.89	.17	.83
32.00	36.00	36.63	38.19	15.29	21.29	22.62	24.56	0	1	0	1
16.18	20.18	20.81	22.36	16.18	22.18	23.51	25.45	.73	.27	.82	.18
4.55	8.55	9.18	10.73	4.55	10.55	11.88	13.82	.73	.27	.82	.18
22.79	26.79	27.42	28.97	22.67	28.67	30.01	31.95	.72	.28	.82	.18
39.57	43.57	44.20	45.75	39.52	45.52	46.86	48.80	.73	.27	.82	.18
31.05	35.05	35.68	37.23	31.03	37.03	38.36	40.30	.73	.27	.82	.18
26.30	30.30	30.93	32.48	20.00	26.00	27.33	29.27	.10	.90	.17	.83
16.36	20.36	20.99	22.54	16.36	22.36	23.69	25.63	.73	.27	.82	.18
20.00	24.00	24.63	26.18	20.00	26.00	27.33	29.27	.73	.27	.82	.18
10.00	14.00	14.63	16.18	8.00	14.00	15.33	17.27	.50	.50	.63	.37
29.99	33.99	34.62	36.17	29.99	35.99	37.32	39.26	.73	.27	.82	.18
17.88	21.88	22.51	24.06	16.33	22.33	23.66	25.60	.56	.44	.68	.32
14.32	18.32	18.95	20.50	14.32	20.32	21.65	23.59	.73	.27	.82	.18
18.00	22.00	22.63	24.18	16.00	22.00	23.33	25.27	.50	.50	.63	.37
8.00	12.00	12.63	14.18	8.00	14.00	15.33	17.27	.73	.27	.82	.18
14.37	18.37	19.00	20.55	14.37	20.37	21.70	23.64	.73	.27	.82	.18
22.00	26.00	26.63	28.18	22.00	28.00	29.33	31.27	.73	.27	.82	.18
9.26	13.26	13.89	15.44	9.26	15.26	16.59	18.53	.73	.27	.82	.18
21.30	25.30	25.93	27.49	21.30	27.30	28.64	30.58	.73	.27	.82	.18
11.17	15.17	15.80	17.35	11.17	17.17	18.50	20.44	.73	.27	.82	.18
15.66	19.66	20.29	21.84	15.66	21.66	22.99	24.93	.73	.27	.82	.18
20.00	24.00	24.63	26.18	20.00	26.00	27.33	29.27	.73	.27	.82	.18

(continued)

Table D4

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
4.69	8.69	9.32	10.87	4.69	10.69	12.02	13.96	.73	.27	.82	.18
11.86	15.86	16.50	18.05	11.85	17.85	19.18	21.12	.73	.27	.82	.18
31.40	35.40	36.04	37.59	31.40	37.40	38.74	40.68	.73	.27	.82	.18
17.94	21.94	22.57	24.12	17.94	23.94	25.27	27.21	.73	.27	.82	.18
22.20	26.20	26.83	28.38	22.14	28.14	29.47	31.41	.73	.27	.82	.18
20.00	24.00	24.63	26.18	20.00	26.00	27.33	29.27	.73	.27	.82	.18
32.84	36.84	37.47	39.02	32.82	38.82	40.15	42.09	.73	.27	.82	.18
13.16	17.16	17.79	19.34	13.16	19.16	20.49	22.43	.73	.27	.82	.18
15.38	19.38	20.01	21.56	15.35	21.35	22.68	24.62	.73	.27	.82	.18
21.95	25.95	26.58	28.13	21.95	27.95	29.28	31.22	.73	.27	.82	.18
7.85	11.85	12.48	14.03	7.85	13.85	15.18	17.12	.73	.27	.82	.18
31.92	35.92	36.55	38.10	20.00	26.00	27.33	29.27	.01	.99	.01	.99
9.86	13.86	14.49	16.04	9.86	15.86	17.19	19.13	.73	.27	.82	.18
12.05	16.05	16.69	18.24	11.97	17.97	19.30	21.24	.72	.28	.82	.18

Table D5

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
11.27	15.27	15.67	18.27	11.25	17.25	18.08	21.74	.73	.27	.85	.15
32.01	36.01	36.41	39.01	31.64	37.64	38.47	42.13	.69	.31	.83	.17
50.37	54.37	54.77	57.37	50.08	56.08	56.91	60.57	.70	.30	.83	.17
74.08	78.08	78.48	81.08	74.08	80.08	80.91	84.57	.73	.27	.85	.15
71.61	75.61	76.01	78.60	71.61	77.61	78.43	82.10	.73	.27	.85	.15
74.88	78.88	79.28	81.87	74.88	80.88	81.71	85.37	.73	.27	.85	.15
33.15	37.15	37.55	40.14	32.41	38.41	39.24	42.90	.65	.35	.80	.20
12.52	16.52	16.92	19.51	7.95	13.95	14.78	18.44	.22	.78	.37	.63
53.13	57.13	57.53	60.12	53.13	59.13	59.96	63.62	.73	.27	.85	.15
46.12	50.12	50.52	53.11	45.08	51.08	51.91	55.57	.62	.38	.77	.23
64.35	68.35	68.75	71.35	64.14	70.14	70.97	74.63	.71	.29	.84	.16
55.09	59.09	59.49	62.09	54.75	60.75	61.58	65.24	.70	.30	.83	.17
45.26	49.26	49.66	52.26	45.23	51.23	52.06	55.72	.73	.27	.85	.15
39.97	43.97	44.37	46.96	39.72	45.72	46.55	50.21	.71	.29	.84	.16
39.29	43.29	43.69	46.28	29.80	35.80	36.62	40.28	.02	.98	.05	.95
40.14	44.14	44.54	47.14	40.13	46.13	46.96	50.62	.73	.27	.85	.15
31.29	35.29	35.69	38.29	29.02	35.02	35.85	39.51	.47	.53	.65	.35
62.83	66.83	67.23	69.82	62.69	68.69	69.52	73.18	.72	.28	.84	.16
73.23	77.23	77.63	80.23	73.18	79.18	80.01	83.67	.73	.27	.85	.15
69.52	73.52	73.92	76.51	69.50	75.50	76.33	79.99	.73	.27	.85	.15
71.72	75.72	76.12	78.72	71.70	77.70	78.53	82.19	.73	.27	.85	.15
35.20	39.20	39.60	42.20	33.64	39.64	40.47	44.13	.55	.45	.72	.28
30.53	34.53	34.93	37.52	26.91	32.91	33.74	37.40	.31	.69	.48	.52
59.31	63.31	63.71	66.30	59.20	65.20	66.03	69.69	.72	.28	.84	.16
37.46	41.46	41.86	44.45	37.37	43.37	44.20	47.86	.72	.28	.85	.15
51.82	55.82	56.22	58.81	51.24	57.24	58.07	61.73	.67	.33	.81	.19
36.70	40.70	41.10	43.69	29.50	35.50	36.33	39.99	.07	.93	.14	.86
56.94	60.94	61.34	63.93	56.45	62.45	63.28	66.94	.68	.32	.82	.18
49.44	53.44	53.84	56.43	48.66	54.66	55.49	59.15	.65	.35	.80	.20
53.86	57.86	58.26	60.85	53.69	59.69	60.52	64.18	.71	.29	.84	.16
68.73	72.73	73.13	75.72	68.69	74.69	75.52	79.18	.73	.27	.85	.15
45.17	49.17	49.57	52.16	44.19	50.19	51.02	54.68	.62	.38	.78	.22

(continued)

Table D5

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
61.30	65.30	65.70	68.29	60.24	66.24	67.07	70.73	.62	.38	.77	.23
23.40	27.40	27.80	30.39	23.27	29.27	30.10	33.76	.72	.28	.84	.16
55.94	59.94	60.34	62.93	55.34	61.34	62.17	65.83	.67	.33	.81	.19
52.44	56.44	56.84	59.44	52.44	58.44	59.27	62.93	.73	.27	.85	.15
43.84	47.84	48.24	50.84	43.40	49.40	50.23	53.89	.69	.31	.82	.18
63.90	67.90	68.30	70.89	63.85	69.85	70.68	74.34	.73	.27	.85	.15
46.00	50.00	50.40	52.99	36.00	42.00	42.83	46.49	.02	.98	.04	.96
64.41	68.41	68.81	71.40	61.16	67.16	67.99	71.65	.35	.65	.53	.47
60.06	64.06	64.46	67.05	27.33	33.33	34.16	37.82	0	1	0	1
61.09	65.09	65.49	68.08	59.50	65.50	66.33	69.99	.55	.45	.72	.28
60.53	64.53	64.93	67.52	59.87	65.87	66.70	70.36	.66	.34	.81	.19
35.73	39.73	40.13	42.72	0.00	6.00	6.83	10.49	0	1	0	1
62.10	66.10	66.50	69.09	52.27	58.27	59.10	62.76	.02	.98	.04	.96
45.58	49.58	49.98	52.57	42.10	48.10	48.93	52.59	.32	.68	.50	.50
37.11	41.11	41.51	44.10	25.25	31.25	32.08	35.74	.01	.99	.02	.98
50.80	54.80	55.20	57.80	47.08	53.08	53.91	57.57	.30	.70	.47	.53
64.17	68.17	68.57	71.16	63.54	69.54	70.37	74.03	.67	.33	.81	.19
45.18	49.18	49.58	52.18	40.69	46.69	47.52	51.18	.22	.78	.38	.62
42.29	46.29	46.69	49.29	31.73	37.73	38.56	42.22	.01	.99	.03	.97
37.54	41.54	41.94	44.53	37.01	43.01	43.84	47.50	.68	.32	.82	.18
35.61	39.61	40.01	42.60	35.61	41.61	42.43	46.09	.73	.27	.85	.15
34.15	38.15	38.55	41.14	24.60	30.60	31.43	35.09	.02	.98	.05	.95
38.00	42.00	42.40	44.99	34.00	40.00	40.83	44.49	.27	.73	.44	.56
31.14	35.14	35.54	38.13	30.30	36.30	37.12	40.79	.64	.36	.79	.21
55.78	59.78	60.18	62.78	45.74	51.74	52.57	56.23	.02	.98	.04	.96
71.57	75.57	75.97	78.57	71.51	77.51	78.34	82.00	.72	.28	.85	.15
45.60	49.60	50.00	52.59	36.29	42.29	43.12	46.78	.03	.97	.05	.95
39.40	43.40	43.80	46.40	39.16	45.16	45.99	49.65	.71	.29	.84	.16
68.21	72.21	72.61	75.20	68.21	74.21	75.04	78.70	.73	.27	.85	.15
44.38	48.38	48.78	51.37	44.33	50.33	51.16	54.82	.73	.27	.85	.15
45.46	49.46	49.86	52.45	45.10	51.10	51.92	55.59	.69	.31	.83	.17

(continued)

Table D5

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
34.00	38.00	38.40	40.99	26.00	32.00	32.83	36.49	.05	.95	.10	.90
47.11	51.11	51.51	54.10	30.00	36.00	36.83	40.49	0	1	0	1
48.08	52.08	52.48	55.07	34.00	40.00	40.83	44.49	0	1	.01	.99
51.43	55.43	55.83	58.42	49.84	55.84	56.67	60.33	.55	.45	.72	.28
71.76	75.76	76.16	78.75	68.92	74.92	75.74	79.41	.40	.60	.58	.42
29.48	33.48	33.88	36.47	4.00	10.00	10.83	14.49	0	1	0	1
60.98	64.98	65.38	67.97	60.96	66.96	67.79	71.45	.73	.27	.85	.15
70.60	74.60	75.00	77.59	70.59	76.59	77.42	81.08	.73	.27	.85	.15
49.18	53.18	53.58	56.17	49.11	55.11	55.94	59.60	.72	.28	.85	.15
55.57	59.57	59.97	62.57	48.53	54.53	55.36	59.02	.07	.93	.15	.85
59.01	63.01	63.41	66.00	58.12	64.12	64.94	68.61	.64	.36	.79	.21
62.74	66.74	67.14	69.74	58.29	64.29	65.12	68.78	.23	.77	.38	.62
46.08	50.08	50.48	53.08	45.33	51.33	52.16	55.82	.65	.35	.80	.20
60.28	64.28	64.68	67.27	58.70	64.70	65.52	69.19	.55	.45	.72	.28
50.55	54.55	54.95	57.54	49.99	55.99	56.82	60.48	.67	.33	.81	.19
75.86	79.86	80.26	82.85	74.47	80.47	81.30	84.96	.58	.42	.74	.26
57.28	61.28	61.68	64.28	56.86	62.86	63.69	67.35	.69	.31	.82	.18
52.52	56.52	56.92	59.52	52.32	58.32	59.14	62.81	.71	.29	.84	.16
70.47	74.47	74.87	77.46	69.36	75.36	76.19	79.85	.61	.39	.77	.23
33.32	37.32	37.72	40.31	27.81	33.81	34.64	38.30	.15	.85	.27	.73
42.73	46.73	47.13	49.72	42.72	48.72	49.55	53.21	.73	.27	.85	.15
50.67	54.67	55.07	57.66	50.66	56.66	57.49	61.15	.73	.27	.85	.15
72.68	76.68	77.08	79.67	70.53	76.53	77.36	81.02	.48	.52	.66	.34
54.01	58.01	58.41	61.01	51.93	57.93	58.76	62.42	.49	.51	.67	.33
54.10	58.10	58.50	61.09	53.98	59.98	60.81	64.47	.72	.28	.84	.16
5.59	9.59	9.99	12.59	4.58	10.58	11.41	15.07	.62	.38	.78	.22
7.18	11.18	11.58	14.17	7.16	13.16	13.99	17.65	.73	.27	.85	.15
33.61	37.61	38.01	40.61	26.58	32.58	33.41	37.07	.07	.93	.15	.85
37.07	41.07	41.47	44.06	34.68	40.68	41.51	45.17	.45	.55	.63	.37
36.97	40.97	41.37	43.97	36.85	42.85	43.68	47.34	.72	.28	.84	.16
53.64	57.64	58.04	60.63	41.86	47.86	48.69	52.35	.01	.99	.02	.98
48.03	52.03	52.43	55.02	45.78	51.78	52.60	56.26	.47	.53	.65	.35

(continued)

Table D5

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
65.25	69.25	69.65	72.24	64.57	70.57	71.40	75.06	.66	.34	.80	.20
46.23	50.23	50.63	53.22	45.99	51.99	52.81	56.47	.71	.29	.84	.16
61.00	65.00	65.40	68.00	56.55	62.55	63.38	67.04	.23	.77	.38	.62
63.60	67.60	68.00	70.59	62.28	68.28	69.11	72.77	.58	.42	.75	.25
65.65	69.65	70.05	72.64	64.42	70.42	71.25	74.91	.60	.40	.76	.24
17.81	21.81	22.21	24.80	17.74	23.74	24.57	28.23	.72	.28	.85	.15
73.66	77.66	78.06	80.65	73.10	79.10	79.93	83.59	.67	.33	.81	.19
52.32	56.32	56.72	59.31	43.23	49.23	50.06	53.72	.03	.97	.06	.94
72.78	76.78	77.18	79.77	72.67	78.67	79.49	83.15	.72	.28	.84	.16
33.04	37.04	37.44	40.03	32.46	38.46	39.29	42.95	.67	.33	.81	.19
62.60	66.60	67.00	69.59	62.43	68.43	69.26	72.92	.71	.29	.84	.16
51.51	55.51	55.91	58.51	50.24	56.24	57.07	60.73	.59	.41	.75	.25
61.18	65.18	65.58	68.18	58.16	64.16	64.99	68.65	.37	.63	.56	.44
58.31	62.31	62.71	65.31	55.73	61.73	62.56	66.22	.43	.57	.61	.39
77.26	81.26	81.66	84.25	77.19	83.19	84.02	87.68	.72	.28	.85	.15
48.89	52.89	53.29	55.88	44.86	50.86	51.69	55.35	.27	.73	.43	.57
51.87	55.87	56.27	58.86	51.87	57.87	58.70	62.36	.73	.27	.85	.15
64.20	68.20	68.60	71.20	63.76	69.76	70.58	74.25	.68	.32	.82	.18
59.40	63.40	63.80	66.39	59.37	65.37	66.19	69.86	.73	.27	.85	.15
47.36	51.36	51.76	54.36	44.56	50.56	51.39	55.05	.40	.60	.59	.41
29.06	33.06	33.46	36.06	12.00	18.00	18.83	22.49	0	1	0	1
58.17	62.17	62.57	65.16	58.24	64.24	65.07	68.73	.74	.26	.86	.14
38.00	42.00	42.40	44.99	38.16	44.16	44.99	48.65	.75	.25	.86	.14
72.54	76.54	76.94	79.53	72.01	78.01	78.84	82.50	.68	.32	.82	.18
42.00	46.00	46.40	48.99	42.00	48.00	48.83	52.49	.73	.27	.85	.15
72.61	76.61	77.01	79.61	71.16	77.16	77.98	81.65	.57	.43	.73	.27
74.49	78.49	78.89	81.48	73.03	79.03	79.86	83.52	.57	.43	.73	.27
37.06	41.06	41.46	44.05	37.05	43.05	43.88	47.54	.73	.27	.85	.15
74.76	78.76	79.16	81.76	74.45	80.45	81.28	84.94	.70	.30	.83	.17
40.84	44.84	45.24	47.83	39.62	45.62	46.45	50.11	.60	.40	.76	.24
42.59	46.59	46.99	49.58	42.55	48.55	49.38	53.04	.73	.27	.85	.15
55.96	59.96	60.36	62.95	46.95	52.95	53.78	57.44	.03	.97	.06	.94

(continued)

Table D5

Model comparison statistics comparing range-frequency theory (RFT) and the generalized exemplar model of sampling (GEMS) using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT				GEMS				AICw		BICw	
-2lnL	AIC	AICc	BIC	-2lnL	AIC	AICc	BIC	RFT	GEMS	RFT	GEMS
66.59	70.59	70.99	73.58	65.67	71.67	72.50	76.16	.63	.37	.78	.22
44.43	48.43	48.83	51.42	44.40	50.40	51.22	54.88	.73	.27	.85	.15
75.07	79.07	79.47	82.06	73.77	79.77	80.59	84.26	.59	.41	.75	.25
73.66	77.66	78.06	80.66	73.06	79.06	79.88	83.55	.67	.33	.81	.19
63.14	67.14	67.54	70.14	62.69	68.69	69.52	73.18	.68	.32	.82	.18
49.72	53.72	54.12	56.72	49.72	55.72	56.55	60.21	.73	.27	.85	.15
50.00	54.00	54.40	56.99	30.00	36.00	36.83	40.49	0	1	0	1

Appendix E:

RFT and SDbS Fit

Table E1

Best fit statistics for the RFT and SDbS models to individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
0.45	0	80.68	0.40	11.35	73.49	.91	65.31
0.32	.49	34.95	0.30	0.84	5.00	.85	28.17
0.44	.60	80.21	0.44	2.49	7.50	.67	79.27
0.23	.55	-5.59	0.25	1.15	17.00	.40	4.57
0.17	.29	-49.96	0.18	0.69	11.65	.34	-40.57
0.42	.27	72.29	0.41	1.44	15.83	.38	70.23
0.21	.35	-20.43	0.18	2.15	9.39	.52	-38.44
0.51	.62	97.51	0.49	1.57	11.80	.50	94.40
0.56	.62	109.72	0.55	0.56	30.77	.28	108.58
0.49	.31	93.64	0.49	1.09	15.04	.35	93.31
0.43	.50	77.00	0.44	1.20	5.58	.67	77.83
0.43	.21	75.15	0.42	3.73	8.24	.60	73.71
0.43	.56	76.53	0.44	3.62	7.87	.77	77.87
0.28	.45	19.05	0.26	3.92	6.03	.85	10.06
0.19	.17	-30.79	0.19	6.74	6.07	.83	-34.74
0.48	.29	90.52	0.47	1.01	13.57	.35	88.98
0.17	.20	-50.13	0.17	1.38	3.47	.84	-45.82
0.36	.38	52.35	0.36	1.44	4.99	.68	51.57
0.58	0	115.03	0.50	0.08	100.00	.13	96.44
0.25	.30	3.66	0.24	0.80	4.38	.87	-2.70
0.39	.59	62.24	0.38	2.73	5.43	.84	60.68
0.32	.44	37.95	0.34	4.82	7.86	.84	43.72
0.21	.29	-21.07	0.22	4.70	5.28	.84	-13.07
0.28	.20	20.33	0.27	1.53	4.91	.59	15.34

Table E2

Best fit statistics for the RFT and SDbS models to individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
3.27	.40	343.83	3.25	0.37	9.53	.43	342.96
7.07	.19	445.45	7.08	0.08	20.56	.50	445.59
16.20	0	554.93	16.20	0.98	0.00	.50	554.93
3.99	.19	369.95	3.98	0.61	6.97	.47	369.55
14.24	.87	537.92	14.29	0.14	14.89	.51	538.35
13.77	.13	533.45	13.75	0.89	83.68	.51	533.26
19.71	.67	580.84	19.69	0.10	21.13	.50	580.66
14.55	.42	540.78	14.56	0.90	8.91	.58	540.87
9.87	.43	489.50	9.77	0.07	34.09	.50	488.18
14.00	.22	535.61	13.99	0.08	20.94	.50	535.61
5.98	0	423.28	5.88	0.02	81.17	.54	420.35
17.83	.85	567.59	17.83	0.19	11.18	.50	567.55
9.68	.48	486.98	9.68	0.79	100.00	.50	487.01
8.98	.51	476.99	8.97	0.35	8.14	.51	476.86
8.48	0	469.50	8.48	0.52	0.00	.50	469.50
7.35	.89	450.68	7.18	0.89	11.58	.66	447.44
8.71	.86	472.94	8.61	0.39	11.70	.46	471.42
7.15	0	446.90	7.14	0.99	100.00	.53	446.73
4.48	0	385.33	4.41	0.03	69.59	.46	383.96
5.41	.48	410.18	5.38	0.64	8.34	.54	409.45
7.81	.43	458.56	7.82	0.18	9.55	.50	458.81
14.19	.24	537.46	14.19	0.08	20.91	.50	537.41
12.51	.39	520.75	12.50	0.10	17.96	.50	520.67
5.60	.18	414.70	5.58	0.06	31.49	.51	414.27
9.00	.40	477.35	8.97	0.88	38.19	.54	476.83
8.03	.45	462.21	8.03	0.13	12.93	.50	462.24
3.27	.71	343.61	3.27	0.48	7.04	.79	343.52
5.96	.62	423.02	5.96	0.78	9.09	.69	423.02
3.19	1	340.52	3.12	0.82	8.79	.88	337.62
7.62	.97	455.37	7.77	0.45	7.64	.85	455.49
4.05	.11	371.98	4.05	0.32	7.14	.52	372.00
5.89	.09	421.28	5.88	1.25	8.85	.55	421.23

(continued)

Table E2

Best fit statistics for the RFT and SDbS models to individual participant data from Melrose et al. (2012) (continued). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
5.31	.70	407.78	5.32	0.95	8.14	.80	408.02
2.95	1	330.08	2.85	1.20	28.39	.81	325.44
2.38	.11	301.66	2.31	2.45	10.23	.82	297.98
10.98	.43	503.54	10.96	0.66	31.13	.55	503.38
3.13	.14	337.90	3.06	2.06	10.03	.78	335.10
4.76	.40	393.36	4.76	1.03	8.12	.70	393.33
6.49	1	434.21	6.11	1.13	14.21	.88	426.08
8.03	0	462.35	7.93	0.03	90.44	.44	460.94
6.02	1	424.17	5.46	1.09	12.55	.90	411.44
10.92	1	502.92	9.54	0.85	30.21	.74	485.02
14.77	1	542.74	13.93	0.90	100.00	.70	535.03
1.84	.43	267.53	1.79	0.36	7.71	.68	263.95
2.38	1	301.57	2.37	0.63	14.39	.64	301.04
9.82	1	488.90	9.43	0.82	44.35	.67	483.52
12.21	1	517.58	12.19	0.71	10.99	.76	517.42
7.79	.62	458.34	7.85	0.52	7.53	.69	459.22
1.66	.53	254.16	1.63	0.68	6.52	.76	251.48
4.71	1	391.94	4.39	1.18	11.82	.92	382.53
5.37	.84	409.18	5.44	0.88	7.69	.85	410.29
7.58	1	454.59	7.14	1.17	27.84	.80	446.81

Table E3

Best fit statistics for the RFT and SDbS models to individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

	RFT				SDbS		
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
1.46	.69	316.32	1.46	0.49	35.30	.55	316.45
4.91	.19	529.64	4.92	0.00	85.47	.52	529.16
13.61	.43	709.18	13.58	0.48	14.07	.53	708.84
3.65	.58	477.66	3.58	0.96	10.37	.77	474.07
6.22	.09	571.49	6.22	0.26	5.52	.50	571.48
4.33	.45	507.69	4.26	1.35	22.19	.80	504.61
8.89	.33	634.29	8.89	0.99	9.31	.69	634.33
7.69	.62	608.69	7.31	1.24	29.96	.80	599.77
4.01	1	494.16	4.02	0.58	9.74	.78	494.44
14.43	.74	719.50	14.45	0.41	9.32	.63	719.74
4.78	.58	525.21	4.74	0.84	14.91	.70	523.76
9.72	.28	650.08	9.71	0.15	9.71	.46	649.79
8.79	.41	632.20	8.81	0.92	8.52	.70	632.64
2.96	.62	440.65	2.97	0.45	7.27	.69	441.30
13.15	.41	703.14	13.15	0.62	5.68	.76	703.12
11.15	.25	674.11	11.15	0.76	10.90	.58	674.09
9.38	.23	643.76	9.37	0.52	16.77	.49	643.58
15.80	.80	735.54	16.02	0.51	6.63	.84	735.69
6.47	.71	578.42	6.46	0.67	8.62	.75	578.08
6.80	.58	587.24	6.55	1.13	15.83	.79	580.44
3.68	.86	478.97	3.63	0.80	10.04	.82	476.65
4.19	.71	501.81	4.09	0.68	6.53	.87	500.23
5.44	.73	547.76	5.46	0.38	7.85	.68	548.59
7.71	.89	609.25	7.73	0.49	9.56	.70	609.72
5.78	.84	558.43	5.72	0.83	10.71	.80	556.73
5.09	1	536.04	4.99	0.58	12.43	.72	532.65
7.02	.38	592.83	7.03	0.37	7.59	.57	593.04
7.28	.82	599.23	6.86	0.83	24.53	.73	588.55
16.12	.86	739.02	15.86	0.66	100.00	.63	736.12
5.97	.61	564.11	5.54	1.18	16.74	.81	551.08
7.60	.42	606.69	7.59	0.74	9.03	.65	606.40
4.87	.86	528.47	4.63	0.78	32.80	.70	519.48

(continued)

Table E3

Best fit statistics for the RFT and SDbS models to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
5.23	.95	540.80	5.33	0.47	7.94	.78	544.28
7.02	.86	592.83	7.06	0.53	8.87	.73	593.76
4.98	.55	532.21	4.93	1.03	8.87	.80	530.54
16.48	.42	742.95	16.48	0.45	8.85	.56	742.90
4.32	.65	507.08	4.23	1.24	12.94	.85	503.49
10.16	1	657.75	9.86	0.31	15.18	.62	652.42
10.96	.97	671.17	10.85	0.22	13.01	.53	669.40
1.18	.62	279.11	1.18	0.82	9.53	.75	278.47
12.90	0	699.78	12.80	0.01	54.35	.77	694.89
8.94	.69	635.18	8.95	0.42	8.71	.58	635.45
7.46	.51	603.34	7.37	0.52	13.39	.49	601.18
8.66	.44	629.73	8.22	0.72	100.00	.50	620.41
2.07	.83	378.20	1.95	0.40	20.92	.48	367.12
7.04	.32	593.17	7.04	0.37	6.93	.51	593.26
8.53	0	626.99	8.53	13.66	49.63	.50	626.99
12.90	0	699.78	12.77	0.00	80.90	.51	698.41
6.03	.30	566.00	5.98	0.47	8.48	.47	564.63
0.00	0	-1054.03	0.10	8.47	34.17	.50	-243.52
27.45	.93	832.73	27.45	0.28	11.97	.51	832.68
3.15	.60	451.88	3.09	0.54	22.06	.50	448.46
9.84	.59	652.06	9.83	0.35	9.18	.52	652.03
7.23	.23	597.91	7.22	0.51	6.79	.50	597.71
9.80	.56	651.39	9.69	0.57	24.40	.50	649.37
7.30	.49	599.59	7.15	0.53	13.55	.49	595.88
1.07	.18	262.44	1.07	0.55	8.94	.45	261.00
4.98	.49	532.12	4.91	0.56	16.15	.50	529.89
6.24	.51	572.09	6.19	0.44	8.48	.54	570.55
5.48	.23	549.04	5.48	0.22	6.34	.50	549.07
13.40	0	706.49	13.16	0.00	80.51	.51	704.22
23.54	.72	805.69	23.47	0.48	18.68	.50	805.11
14.24	.37	717.18	14.18	0.53	11.09	.48	716.46
11.49	0	679.48	11.49	3.37	0.01	.50	679.48

(continued)

Table E3

Best fit statistics for the RFT and SDbS models to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
16.43	.69	742.42	16.42	0.25	7.44	.61	742.35
11.32	.26	676.80	11.21	0.70	12.90	.49	675.10
5.99	.30	564.79	5.97	0.42	7.62	.49	564.16
8.26	0	621.41	8.20	0.01	81.52	.48	620.62
6.44	.09	577.41	6.41	1.07	11.41	.50	576.81
10.65	.73	666.09	10.58	0.21	13.99	.40	665.40
6.83	.40	587.77	6.76	0.38	10.16	.45	586.11
7.37	.27	601.22	7.35	0.40	6.70	.51	600.86
14.83	.86	724.38	14.81	0.49	8.50	.69	723.62
13.88	.52	712.73	13.84	0.48	11.36	.49	712.23
6.28	.21	573.03	6.26	0.15	6.53	.50	572.53
10.13	.79	657.25	10.06	0.27	12.69	.46	655.97
8.23	.52	620.74	8.14	0.30	11.89	.42	618.74
9.00	.64	636.47	8.91	0.43	9.24	.56	635.15
6.04	.45	566.37	6.08	0.39	6.86	.57	567.30
15.91	1	736.69	15.77	0.28	20.22	.50	735.21
7.80	.70	611.36	7.53	0.33	16.93	.43	605.12

Table E4

Best fit statistics for the RFT and SDbS models to individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
0.30	.56	21.71	0.30	0.05	16.73	.51	21.38
0.00	1	18.00	0.10	0.24	18.60	.50	18.06
0.22	1	16.67	0.20	0.07	16.04	.64	15.15
0.29	.42	22.52	0.28	0.20	8.94	.45	22.39
0.02	.15	19.38	0.10	0.03	33.51	.50	19.83
0.00	.01	8.00	0.10	0.35	100.00	.48	8.75
0.24	.38	20.44	0.24	0.04	24.37	.50	20.15
0.17	1	15.95	0.16	0.02	77.48	.50	15.54
0.78	.86	37.59	0.76	0.06	15.94	.52	37.42
0.00	1	10.00	0.16	0.12	84.05	.50	9.50
0.03	.08	25.97	0.10	0.03	32.39	.50	26.19
0.32	1	23.24	0.17	0.07	19.44	.51	17.16
0.07	.94	13.61	0.10	0.13	10.35	.49	13.75
0.01	0	26.00	0.10	5.03	0.00	.50	27.20
0.23	1	20.76	0.10	0.07	17.96	.51	14.42
0.07	.86	14.78	0.10	0.06	16.49	.50	14.84
0.35	1	25.85	0.21	0.19	16.35	.50	15.31
1.29	1	44.26	1.00	0.21	89.10	.50	41.76
0.09	.48	11.80	0.10	0.05	17.31	.50	11.79
0.00	1	6.00	0.10	0.07	18.87	.50	5.27
0.00	1	8.00	0.10	0.07	16.33	.52	9.42
0.26	0	26.38	0.26	2.83	0.00	.50	26.38
0.01	0	32.00	0.17	3.43	0.00	.50	32.06
0.04	0	25.24	0.10	8.26	0.66	.50	25.57
0.04	.22	11.12	0.10	0.39	20.13	.50	11.35
0.00	.80	18.00	0.10	0.31	15.41	.48	18.27
0.05	.80	11.19	0.10	0.06	17.16	.50	11.24
0.06	1	17.15	0.10	0.23	11.64	.51	17.16
0.08	1	15.30	0.10	0.07	14.89	.50	15.30
0.49	1	31.50	0.49	0.06	17.50	.50	31.34
0.04	.58	5.96	0.10	0.05	17.96	.50	6.37
0.00	.67	2.00	0.10	0.05	18.02	.50	2.93

(continued)

Table E4

Best fit statistics for the RFT and SDbS models to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

	RFT				SDbS		
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
1.28	1	45.75	1.04	0.19	16.22	.50	44.06
0.19	1	14.81	0.10	0.23	18.35	.51	7.25
0.38	.67	26.78	0.37	0.06	16.76	.51	26.40
0.00	1	26.00	0.13	0.23	100.00	.50	21.07
0.01	.34	20.00	0.10	0.02	48.02	.51	20.33
0.25	.66	20.50	0.16	0.28	100.00	.42	18.42
0.35	0	25.54	0.32	0.32	100.00	.53	24.76
0.35	.04	25.88	0.33	0.32	99.64	.53	25.08
0.24	0	28.21	0.24	9.46	1.30	.50	28.21
0.53	0	32.00	0.10	0.04	99.99	.46	15.23
0.11	1	16.18	0.14	0.12	39.55	.50	16.12
0.02	.63	4.55	0.10	1.06	50.30	.96	6.33
0.32	.74	22.79	0.29	0.98	90.98	.94	20.63
0.84	.82	39.57	0.83	0.42	100.00	.64	39.35
0.48	0	31.05	0.48	0.40	100.00	.58	30.98
0.24	0	26.30	0.10	0.11	100.00	.35	18.01
0.22	1	16.36	0.18	0.09	37.52	.50	15.26
0.00	.67	30.00	0.10	0.11	91.16	.50	20.00
0.01	0	20.00	0.10	0.21	100.00	.50	18.00
0.00	.79	10.00	0.10	0.12	72.65	.50	8.00
0.42	.46	29.99	0.42	0.19	10.51	.52	29.95
0.16	0	17.88	0.16	7.68	0.09	.50	17.88
0.20	1	14.32	0.10	0.16	17.69	.53	11.81
0.01	0	18.00	0.13	7.90	0.49	.50	17.96
0.01	.45	8.00	0.10	0.15	11.07	.65	8.67
0.09	1	14.37	0.13	0.12	98.54	.50	7.07
0.00	.95	22.00	0.10	0.16	13.65	.72	19.43
0.10	1	9.26	0.10	0.42	12.91	.79	9.32
0.15	1	21.30	0.10	0.66	100.00	.88	15.21
0.12	1	11.17	0.10	0.50	18.67	.81	8.63
0.17	.77	15.66	0.12	0.14	12.36	.62	15.22
0.01	0	20.00	0.10	7.25	18.00	.50	20.39

(continued)

Table E4

Best fit statistics for the RFT and SDbS models to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
0.09	1	4.69	0.14	0.12	59.90	.50	4.58
0.05	0	11.86	0.10	4.24	0.00	.50	11.97
0.49	.06	31.40	0.49	0.09	11.92	.50	31.38
0.08	.75	17.94	0.10	0.17	8.68	.72	18.77
0.29	.22	22.20	0.29	0.56	10.88	.67	22.19
0.00	.79	20.00	0.18	0.15	7.01	.64	22.00
0.50	.82	32.84	0.52	0.18	9.39	.66	33.43
0.09	.28	13.16	0.12	0.07	42.51	.50	12.96
0.15	.21	15.38	0.15	0.16	6.70	.61	15.71
0.30	1	21.95	0.30	0.15	11.32	.64	22.52
0.04	.50	7.85	0.10	0.17	7.51	.69	9.63
0.26	0	31.92	0.26	5.18	0.01	.50	31.92
0.09	1	9.86	0.10	0.17	10.38	.66	10.56
0.18	.52	12.05	0.12	1.01	39.86	.94	11.82

Table E5

Best fit statistics for the RFT and SDbS models to individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

	RFT				SDbS		
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
0.03	0	11.27	0.10	13.57	6.18	.50	12.39
0.27	.02	32.01	0.27	1.89	1.30	.50	32.02
0.56	.16	50.37	0.54	0.09	100.00	.50	49.88
3.99	1	74.08	3.36	0.00	99.21	.48	63.42
3.19	1	71.61	3.01	0.02	22.18	.50	70.82
6.19	1	74.88	0.31	0.01	62.66	.91	58.32
0.28	0	33.15	0.28	7.03	10.72	.50	33.15
0.05	0	12.52	0.10	7.29	20.59	.50	13.79
0.66	.18	53.13	0.66	0.03	6.44	.50	53.13
0.42	.47	46.12	0.37	0.08	24.74	.52	44.71
1.10	0	64.35	1.10	4.58	18.07	.50	64.35
0.60	.11	55.09	0.58	0.10	29.96	.50	54.69
0.38	0	45.26	0.38	1.05	0.40	.50	45.26
0.28	.04	39.97	0.28	29.52	8.79	.50	39.99
0.29	0	39.29	0.29	6.05	11.14	.50	39.29
0.37	0	40.14	0.36	0.00	70.84	.50	39.87
0.23	0	31.29	0.23	22.89	3.86	.50	31.29
1.24	.48	62.83	1.24	0.01	16.15	.50	62.83
2.63	.46	73.23	2.58	0.09	16.76	.50	73.15
1.38	.09	69.52	1.38	0.09	100.00	.50	69.48
1.67	.17	71.72	1.67	0.11	11.05	.50	71.68
0.33	.05	35.20	0.32	0.00	52.24	.50	34.61
0.26	.03	30.53	0.26	1.02	0.29	.50	30.62
0.84	0	59.31	0.84	0.49	5.98	.50	59.31
0.29	.08	37.46	0.29	0.10	4.34	.55	37.47
0.63	.04	51.82	0.61	0.09	79.80	.50	51.46
0.20	0	36.70	0.20	21.79	2.86	.50	36.70
0.70	0	56.94	0.70	4.32	15.68	.50	56.94
0.51	.12	49.44	0.47	0.08	100.00	.43	47.72
0.63	0	53.86	0.63	2.61	8.90	.50	53.86
2.18	.05	68.73	2.17	0.09	100.00	.50	68.69
0.47	0	45.17	0.47	2.93	14.19	.50	45.17

(continued)

Table E5

Best fit statistics for the RFT and SDbS models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

	RFT				SDbS		
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
0.98	0	61.30	0.98	0.24	0.01	.50	61.30
0.02	.14	23.40	0.10	0.01	29.89	.50	23.63
0.80	0	55.94	0.80	5.37	15.23	.50	55.94
0.69	0	52.44	0.69	14.06	4.30	.50	52.44
0.42	.04	43.84	0.42	1.70	4.49	.50	43.93
1.44	.48	63.90	1.44	0.01	17.55	.50	63.88
0.01	0	46.00	0.21	1.21	2.60	.50	46.03
1.08	0	64.41	0.91	0.00	81.15	.46	62.04
0.98	0	60.06	0.98	4.78	14.00	.50	60.06
1.11	0	61.09	1.12	0.00	88.36	.51	59.69
0.87	.17	60.53	0.83	0.09	62.53	.50	59.81
0.32	0	35.73	0.32	2.37	1.08	.50	35.73
1.08	0	62.10	1.08	7.09	0.49	.50	62.10
0.41	0	45.58	0.39	0.00	78.46	.54	40.43
0.23	0	37.11	0.23	22.95	3.97	.50	37.11
0.51	0	50.80	0.51	3.20	0.34	.50	50.80
1.74	0	64.17	1.74	15.28	2.98	.50	64.17
2.18	.66	68.27	2.18	0.03	8.65	.58	68.26
0.15	0	45.18	0.15	18.72	5.83	.50	45.18
0.29	0	42.29	0.26	0.08	100.00	.44	41.10
0.31	0	37.54	0.31	0.88	2.20	.50	37.54
0.26	.41	35.61	0.33	13.88	0.18	.50	39.37
0.30	0	34.15	0.30	2.27	0.78	.50	34.15
0.00	.25	38.00	0.10	5.91	15.27	.50	39.20
0.19	.08	31.14	0.18	12.88	0.22	.50	31.28
0.79	0	55.78	0.79	15.07	3.51	.50	55.78
3.65	.46	71.57	3.65	0.01	17.62	.50	71.56
0.52	0	45.60	0.52	3.22	0.43	.50	45.60
0.35	.14	39.40	0.36	1.67	0.00	.50	40.14
2.66	1	68.21	2.65	0.02	17.32	.50	68.18
0.42	.31	44.38	0.42	0.09	8.39	.50	44.15
0.44	.05	45.46	0.43	0.09	76.11	.50	45.15

(continued)

Table E5

Best fit statistics for the RFT and SDbS models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

	RFT				SDbS		
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
0.01	0	34.00	0.10	6.19	11.26	.50	35.18
0.19	0	47.11	0.19	12.28	9.73	.50	47.11
0.28	0	48.08	0.23	0.00	70.18	.51	46.87
0.52	0	51.43	0.52	4.65	13.51	.50	51.43
2.24	0	71.76	2.24	15.35	2.20	.50	71.76
0.26	0	29.48	0.26	2.81	6.05	.50	29.48
0.80	0	60.98	0.80	0.42	6.13	.50	60.98
2.12	.34	70.60	2.14	0.08	10.13	.55	70.63
0.43	.86	49.18	0.40	0.02	69.54	.50	47.82
0.67	0	55.57	0.67	1.97	0.13	.50	55.57
0.71	0	59.01	0.67	0.00	80.30	.51	58.01
0.92	0	62.74	0.92	2.51	1.92	.50	62.74
0.46	0	46.08	0.46	5.31	14.47	.50	46.08
0.85	0	60.28	0.85	2.75	3.60	.50	60.28
0.54	0	50.55	0.54	14.24	13.77	.50	50.55
3.36	0	75.86	2.09	0.01	84.54	.50	73.94
0.78	0	57.28	0.78	2.35	0.25	.50	57.28
0.62	0	52.52	0.62	0.09	40.92	.48	52.52
2.33	0	70.47	2.25	0.00	99.33	.48	69.48
0.30	0	33.32	0.30	2.08	0.16	.50	33.32
0.35	0	42.73	0.35	1.43	0.03	.50	42.73
0.65	.02	50.67	0.65	2.73	0.26	.50	50.67
2.75	0	72.68	2.75	0.25	0.00	.50	72.68
0.63	0	54.01	0.63	3.19	0.26	.50	54.01
0.78	0	54.10	0.78	2.20	0.23	.50	54.10
0.04	0	5.59	0.10	14.21	5.76	.50	7.35
0.04	0	7.18	0.10	6.59	17.41	.50	9.26
0.31	0	33.61	0.31	2.57	8.88	.50	33.61
0.31	0	37.07	0.31	0.01	29.09	.50	35.76
0.33	.11	36.97	0.33	1.38	4.29	.50	37.43
0.66	0	53.64	0.66	36.10	16.62	.50	53.64
0.40	0	48.03	0.40	2.70	9.98	.50	48.03

(continued)

Table E5

Best fit statistics for the RFT and SDbS models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

	RFT				SDbS		
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
1.49	0	65.25	1.49	4.71	13.07	.50	65.25
0.43	0	46.23	0.43	15.17	4.07	.50	46.23
0.99	0	61.00	0.99	4.63	15.38	.50	61.00
1.15	0	63.60	1.15	1.78	0.19	.50	63.60
1.14	0	65.65	1.14	12.01	3.51	.50	65.65
0.16	0	17.81	0.16	7.28	13.31	.50	17.81
3.57	0	73.66	3.57	0.39	0.00	.50	73.66
0.62	0	52.32	0.62	2.50	0.35	.50	52.32
3.88	0	72.78	3.88	2.71	17.39	.50	72.78
0.24	0	33.04	0.24	5.25	12.04	.50	33.04
1.56	0	62.60	1.56	2.67	17.02	.50	62.60
0.63	0	51.51	0.55	0.02	36.57	.47	49.71
0.98	0	61.18	0.98	0.15	0.01	.50	61.18
0.88	0	58.31	0.88	2.57	0.12	.50	58.31
7.87	0	77.26	7.87	10.08	0.57	.50	77.26
0.41	0	48.89	0.41	2.53	11.69	.50	48.89
0.60	1	51.87	0.67	0.05	11.82	.56	52.07
1.62	0	64.20	1.62	9.26	7.98	.50	64.20
0.86	.07	59.40	0.83	0.02	30.30	.50	58.63
0.50	0	47.36	0.50	2.75	0.39	.50	47.36
0.18	0	29.06	0.18	10.64	13.43	.50	29.06
0.94	.72	58.17	1.12	0.08	100.00	.55	58.15
0.01	0	38.00	0.10	0.04	22.17	.43	32.81
3.06	0	72.54	3.06	14.40	1.75	.50	72.54
0.00	.88	42.00	0.21	4.73	2.15	.50	45.57
3.72	0	72.61	3.72	12.24	1.50	.50	72.61
5.07	0	74.49	5.07	1.96	9.98	.50	74.49
0.31	0	37.06	0.31	13.16	5.92	.50	37.06
5.95	0	74.76	5.95	3.82	5.41	.50	74.76
0.10	.56	40.84	0.19	12.12	1.40	.50	43.01
0.22	.36	42.59	0.22	0.04	7.71	.63	43.25
0.81	0	55.96	0.81	6.45	14.99	.50	55.96

(continued)

Table E5

Best fit statistics for the RFT and SDbS models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			SDbS				
SD	*w*	-2lnL	SD	*c*	*s*	*t*	-2lnL
2.43	0	66.59	2.43	10.02	20.69	.50	66.59
0.36	.47	44.43	0.36	0.04	8.24	.62	44.83
6.32	0	75.07	6.32	12.83	2.16	.50	75.07
4.88	0	73.66	4.88	9.97	0.00	.50	73.66
1.45	0	63.14	1.45	3.63	12.74	.50	63.14
0.35	1	49.72	0.42	0.04	8.35	.45	52.92
0.01	0	50.00	0.16	0.97	2.44	.50	50.07

Appendix F:

RFT and SDbS Comparison

Table F1

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

	RFT			SDbS		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
84.68	84.87	89.06	73.31	73.96	90.07	0	1	.62	.38
38.95	39.14	43.33	36.17	36.83	52.93	.20	.80	.99	.01
84.21	84.41	88.59	87.27	87.93	104.03	.82	.18	1	0
-1.59	-1.40	2.78	12.57	13.23	29.33	1	0	1	0
-45.96	-45.77	-41.58	-32.57	-31.91	-15.81	1	0	1	0
76.29	76.48	80.67	78.23	78.89	94.99	.73	.27	1	0
-16.43	-16.24	-12.05	-30.44	-29.78	-13.68	0	1	.31	.69
101.51	101.70	105.88	102.40	103.06	119.16	.61	.39	1	0
113.72	113.91	118.10	116.58	117.23	133.33	.81	.19	1	0
97.64	97.84	102.02	101.31	101.97	118.07	.86	.14	1	0
81.00	81.19	85.38	85.83	86.49	102.59	.92	.08	1	0
79.15	79.34	83.53	81.71	82.36	98.47	.78	.22	1	0
80.53	80.72	84.91	85.87	86.52	102.63	.94	.06	1	0
23.05	23.24	27.42	18.06	18.72	34.82	.08	.92	.98	.02
-26.79	-26.60	-22.41	-26.74	-26.08	-9.98	.51	.49	1	0
94.52	94.71	98.90	96.98	97.63	113.74	.77	.23	1	0
-46.13	-45.93	-41.75	-37.82	-37.16	-21.06	.98	.02	1	0
56.35	56.54	60.73	59.57	60.23	76.33	.83	.17	1	0
119.03	119.22	123.41	104.44	105.10	121.20	0	1	.25	.75
7.66	7.85	12.04	5.30	5.96	22.06	.23	.77	.99	.01
66.24	66.43	70.62	68.68	69.34	85.44	.77	.23	1	0
41.95	42.14	46.32	51.72	52.37	68.47	.99	.01	1	0
-17.07	-16.88	-12.69	-5.07	-4.42	11.69	1	0	1	0
24.33	24.52	28.71	23.34	24.00	40.10	.38	.62	1	0

Table F2.

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

	RFT			SDbS		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
347.83	348.02	352.21	350.96	351.62	359.72	.83	.17	.98	.02
449.45	449.64	453.83	453.59	454.24	462.35	.89	.11	.99	.01
558.93	559.12	563.31	562.93	563.58	571.69	.88	.12	.99	.01
373.95	374.14	378.33	377.55	378.21	386.31	.86	.14	.98	.02
541.92	542.11	546.30	546.35	547.01	555.11	.90	.10	.99	.01
537.45	537.64	541.83	541.26	541.92	550.02	.87	.13	.98	.02
584.84	585.03	589.22	588.66	589.31	597.42	.87	.13	.98	.02
544.78	544.97	549.16	548.87	549.53	557.63	.89	.11	.99	.01
493.50	493.69	497.88	496.18	496.84	504.94	.79	.21	.97	.03
539.61	539.80	543.99	543.61	544.26	552.37	.88	.12	.99	.01
427.28	427.47	431.66	428.35	429.01	437.11	.63	.37	.94	.06
571.59	571.78	575.97	575.55	576.21	584.31	.88	.12	.98	.02
490.98	491.17	495.36	495.01	495.67	503.77	.88	.12	.99	.01
480.99	481.18	485.37	484.86	485.51	493.62	.87	.13	.98	.02
473.50	473.69	477.88	477.50	478.15	486.26	.88	.12	.99	.01
454.68	454.88	459.06	455.44	456.10	464.20	.59	.41	.93	.07
476.94	477.13	481.32	479.42	480.08	488.18	.78	.22	.97	.03
450.90	451.09	455.28	454.73	455.38	463.49	.87	.13	.98	.02
389.33	389.52	393.71	391.96	392.62	400.72	.79	.21	.97	.03
414.18	414.37	418.56	417.45	418.11	426.21	.84	.16	.98	.02
462.56	462.75	466.93	466.81	467.46	475.56	.89	.11	.99	.01
541.46	541.65	545.84	545.41	546.07	554.17	.88	.12	.98	.02
524.75	524.94	529.13	528.67	529.33	537.43	.88	.12	.98	.02
418.70	418.89	423.08	422.27	422.92	431.02	.86	.14	.98	.02
481.35	481.54	485.73	484.83	485.49	493.59	.85	.15	.98	.02
466.21	466.40	470.59	470.24	470.90	479.00	.88	.12	.99	.01
347.61	347.80	351.99	351.52	352.17	360.27	.88	.12	.98	.02
427.02	427.21	431.40	431.02	431.67	439.78	.88	.12	.99	.01
344.52	344.71	348.90	345.62	346.27	354.38	.63	.37	.94	.06
459.37	459.57	463.75	463.49	464.15	472.25	.89	.11	.99	.01
375.98	376.17	380.36	380.00	380.66	388.76	.88	.12	.99	.01
425.28	425.48	429.66	429.23	429.88	437.99	.88	.12	.98	.02

(continued)

Table F2

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Melrose et al. (2012) (continued). Note: each row is the data for one participant.

	RFT			SDbS		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
411.78	411.97	416.16	416.02	416.67	424.77	.89	.11	.99	.01
334.08	334.27	338.46	333.44	334.09	342.20	.42	.58	.87	.13
305.66	305.85	310.04	305.98	306.63	314.74	.54	.46	.91	.09
507.54	507.73	511.92	511.38	512.04	520.14	.87	.13	.98	.02
341.90	342.09	346.28	343.10	343.76	351.86	.65	.35	.94	.06
397.36	397.55	401.74	401.33	401.98	410.08	.88	.12	.98	.02
438.21	438.40	442.59	434.08	434.73	442.83	.11	.89	.53	.47
466.35	466.54	470.73	468.94	469.59	477.70	.78	.22	.97	.03
428.17	428.36	432.55	419.44	420.10	428.20	.01	.99	.10	.90
506.92	507.11	511.30	493.02	493.68	501.78	0	1	.01	.99
546.74	546.93	551.11	543.03	543.69	551.79	.14	.86	.58	.42
271.53	271.72	275.91	271.95	272.60	280.71	.55	.45	.92	.08
305.57	305.76	309.95	309.04	309.70	317.80	.85	.15	.98	.02
492.90	493.09	497.28	491.52	492.17	500.28	.33	.67	.82	.18
521.58	521.77	525.96	525.42	526.08	534.18	.87	.13	.98	.02
462.34	462.53	466.72	467.22	467.87	475.98	.92	.08	.99	.01
258.16	258.35	262.54	259.48	260.14	268.24	.66	.34	.95	.05
395.94	396.13	400.32	390.53	391.19	399.29	.06	.94	.37	.63
413.18	413.37	417.56	418.29	418.95	427.05	.93	.07	.99	.01
458.59	458.78	462.97	454.81	455.46	463.56	.13	.87	.57	.43

Table F3

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

	RFT			SDbS		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
320.32	320.61	323.89	324.45	325.47	331.58	.89	.11	.98	.02
533.64	533.93	537.20	537.16	538.19	544.30	.85	.15	.97	.03
713.18	713.48	716.75	716.84	717.86	723.97	.86	.14	.97	.03
481.66	481.95	485.23	482.07	483.09	489.21	.55	.45	.88	.12
575.49	575.78	579.06	579.48	580.50	586.61	.88	.12	.98	.02
511.69	511.98	515.26	512.61	513.64	519.75	.61	.39	.90	.10
638.29	638.58	641.86	642.33	643.36	649.47	.88	.12	.98	.02
612.69	612.99	616.26	607.77	608.79	614.91	.08	.92	.34	.66
498.16	498.46	501.73	502.44	503.46	509.57	.89	.11	.98	.02
723.50	723.80	727.07	727.74	728.77	734.88	.89	.11	.98	.02
529.21	529.50	532.78	531.76	532.79	538.90	.78	.22	.96	.04
654.08	654.37	657.65	657.79	658.81	664.92	.86	.14	.97	.03
636.20	636.49	639.77	640.64	641.67	647.78	.90	.10	.98	.02
444.65	444.95	448.22	449.30	450.33	456.44	.91	.09	.98	.02
707.14	707.43	710.70	711.12	712.14	718.26	.88	.12	.98	.02
678.11	678.40	681.68	682.09	683.11	689.22	.88	.12	.98	.02
647.76	648.06	651.33	651.58	652.61	658.72	.87	.13	.98	.02
739.54	739.83	743.11	743.69	744.72	750.83	.89	.11	.98	.02
582.42	582.71	585.99	586.08	587.11	593.22	.86	.14	.97	.03
591.24	591.53	594.81	588.44	589.46	595.57	.20	.80	.59	.41
482.97	483.26	486.54	484.65	485.67	491.79	.70	.30	.93	.07
505.81	506.10	509.38	508.23	509.25	515.36	.77	.23	.95	.05
551.76	552.05	555.33	556.59	557.62	563.73	.92	.08	.99	.01
613.25	613.54	616.82	617.72	618.74	624.86	.90	.10	.98	.02
562.43	562.73	566.00	564.73	565.76	571.87	.76	.24	.95	.05
540.04	540.33	543.61	540.65	541.68	547.79	.58	.42	.89	.11
596.83	597.12	600.40	601.04	602.06	608.17	.89	.11	.98	.02
603.23	603.53	606.80	596.55	597.58	603.69	.03	.97	.17	.83
743.02	743.32	746.59	744.12	745.14	751.25	.63	.37	.91	.09
568.11	568.40	571.67	559.08	560.11	566.22	.01	.99	.06	.94
610.69	610.99	614.26	614.40	615.42	621.53	.86	.14	.97	.03
532.47	532.76	536.04	527.48	528.51	534.62	.08	.92	.33	.67

(continued)

Table F3

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
544.80	545.10	548.37	552.28	553.31	559.42	.98	.02	1	0
596.83	597.12	600.40	601.76	602.79	608.90	.92	.08	.99	.01
536.21	536.50	539.77	538.54	539.56	545.67	.76	.24	.95	.05
746.95	747.24	750.51	750.90	751.92	758.04	.88	.12	.98	.02
511.08	511.37	514.65	511.49	512.51	518.62	.55	.45	.88	.12
661.75	662.05	665.32	660.42	661.45	667.56	.34	.66	.75	.25
675.17	675.47	678.74	677.40	678.43	684.54	.75	.25	.95	.05
283.11	283.41	286.68	286.47	287.49	293.60	.84	.16	.97	.03
703.78	704.07	707.35	702.89	703.92	710.03	.39	.61	.79	.21
639.18	639.47	642.75	643.45	644.48	650.59	.89	.11	.98	.02
607.34	607.63	610.91	609.18	610.21	616.32	.72	.28	.94	.06
633.73	634.03	637.30	628.41	629.43	635.55	.07	.93	.29	.71
382.20	382.50	385.77	375.12	376.14	382.25	.03	.97	.15	.85
597.17	597.46	600.73	601.26	602.28	608.39	.89	.11	.98	.02
630.99	631.28	634.56	634.99	636.01	642.13	.88	.12	.98	.02
703.78	704.07	707.35	706.41	707.43	713.54	.79	.21	.96	.04
570.00	570.29	573.57	572.63	573.66	579.77	.79	.21	.96	.04
-1050.03	-1049.74	-1046.46	-235.52	-234.50	-228.39	1	0	1	0
836.73	837.02	840.30	840.68	841.71	847.82	.88	.12	.98	.02
455.88	456.17	459.45	456.46	457.49	463.60	.57	.43	.89	.11
656.06	656.36	659.63	660.03	661.05	667.16	.88	.12	.98	.02
601.91	602.20	605.48	605.71	606.73	612.85	.87	.13	.98	.02
655.39	655.68	658.96	657.37	658.39	664.50	.73	.27	.94	.06
603.59	603.88	607.16	603.88	604.91	611.02	.54	.46	.87	.13
266.44	266.73	270.01	269.00	270.03	276.14	.78	.22	.96	.04
536.12	536.42	539.69	537.89	538.92	545.03	.71	.29	.94	.06
576.09	576.38	579.66	578.55	579.57	585.68	.77	.23	.95	.05
553.04	553.33	556.61	557.07	558.09	564.20	.88	.12	.98	.02
710.49	710.78	714.06	712.22	713.25	719.36	.70	.30	.93	.07
809.69	809.98	813.26	813.11	814.14	820.25	.85	.15	.97	.03
721.18	721.47	724.75	724.46	725.49	731.60	.84	.16	.97	.03
683.48	683.77	687.04	687.48	688.50	694.61	.88	.12	.98	.02

(continued)

Table F3

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Wood et al (2011a) (continued). Note: each row is the data for one participant.

	RFT			SDbS		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
746.42	746.71	749.99	750.35	751.38	757.49	.88	.12	.98	.02
680.80	681.09	684.37	683.10	684.13	690.24	.76	.24	.95	.05
568.79	569.08	572.36	572.16	573.19	579.30	.84	.16	.97	.03
625.41	625.71	628.98	628.62	629.65	635.76	.83	.17	.97	.03
581.41	581.70	584.98	584.81	585.83	591.94	.85	.15	.97	.03
670.09	670.39	673.66	673.40	674.42	680.53	.84	.16	.97	.03
591.77	592.06	595.34	594.11	595.13	601.24	.76	.24	.95	.05
605.22	605.51	608.79	608.86	609.89	616.00	.86	.14	.97	.03
728.38	728.67	731.95	731.62	732.64	738.75	.83	.17	.97	.03
716.73	717.03	720.30	720.23	721.26	727.37	.85	.15	.97	.03
577.03	577.32	580.59	580.53	581.56	587.67	.85	.15	.97	.03
661.25	661.54	664.82	663.97	664.99	671.10	.80	.20	.96	.04
624.74	625.04	628.31	626.74	627.76	633.87	.73	.27	.94	.06
640.47	640.77	644.04	643.15	644.17	650.29	.79	.21	.96	.04
570.37	570.66	573.94	575.30	576.33	582.44	.92	.08	.99	.01
740.69	740.98	744.26	743.21	744.24	750.35	.78	.22	.95	.05
615.36	615.65	618.93	613.12	614.15	620.26	.25	.75	.66	.34

Table F4

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
25.71	26.34	27.89	29.38	31.74	33.75	.86	.14	.95	.05
22.00	22.63	24.18	26.06	28.41	30.42	.88	.12	.96	.04
20.67	21.31	22.86	23.15	25.50	27.51	.78	.22	.91	.09
26.52	27.15	28.70	30.39	32.74	34.75	.87	.13	.95	.05
23.38	24.01	25.56	27.83	30.18	32.20	.90	.10	.97	.03
12.00	12.63	14.18	16.75	19.10	21.12	.91	.09	.97	.03
24.44	25.08	26.63	28.15	30.50	32.51	.86	.14	.95	.05
19.95	20.58	22.13	23.54	25.90	27.91	.86	.14	.95	.05
41.59	42.22	43.77	45.42	47.78	49.79	.87	.13	.95	.05
14.00	14.63	16.18	17.50	19.86	21.87	.85	.15	.94	.06
29.97	30.60	32.15	34.19	36.54	38.56	.89	.11	.96	.04
27.24	27.88	29.43	25.16	27.52	29.53	.26	.74	.51	.49
17.61	18.24	19.79	21.75	24.10	26.11	.89	.11	.96	.04
30.00	30.63	32.18	35.20	37.55	39.57	.93	.07	.98	.02
24.76	25.39	26.94	22.42	24.77	26.78	.24	.76	.48	.52
18.78	19.41	20.96	22.84	25.20	27.21	.88	.12	.96	.04
29.85	30.48	32.03	23.31	25.66	27.67	.04	.96	.10	.90
48.26	48.89	50.44	49.76	52.11	54.12	.68	.32	.86	.14
15.80	16.43	17.98	19.79	22.14	24.15	.88	.12	.96	.04
10.00	10.63	12.18	13.27	15.63	17.64	.84	.16	.94	.06
12.00	12.63	14.18	17.42	19.78	21.79	.94	.06	.98	.02
30.38	31.01	32.56	34.38	36.73	38.74	.88	.12	.96	.04
36.00	36.63	38.18	40.06	42.41	44.42	.88	.12	.96	.04
29.24	29.87	31.42	33.57	35.92	37.93	.90	.10	.96	.04
15.12	15.76	17.31	19.35	21.70	23.71	.89	.11	.96	.04
22.00	22.63	24.18	26.27	28.62	30.63	.89	.11	.96	.04
15.19	15.82	17.38	19.24	21.59	23.60	.88	.12	.96	.04
21.15	21.78	23.33	25.16	27.52	29.53	.88	.12	.96	.04
19.30	19.94	21.49	23.30	25.65	27.66	.88	.12	.96	.04
35.50	36.13	37.68	39.34	41.69	43.71	.87	.13	.95	.05
9.96	10.59	12.14	14.37	16.72	18.73	.90	.10	.96	.04
6.00	6.63	8.18	10.93	13.28	15.29	.92	.08	.97	.03

(continued)

Table F4

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
49.75	50.38	51.93	52.06	54.42	56.43	.76	.24	.90	.10
18.81	19.44	20.99	15.25	17.60	19.61	.14	.86	.33	.67
30.78	31.41	32.96	34.40	36.76	38.77	.86	.14	.95	.05
30.00	30.63	32.18	29.07	31.42	33.43	.39	.61	.65	.35
24.00	24.63	26.18	28.33	30.69	32.70	.90	.10	.96	.04
24.50	25.13	26.68	26.42	28.77	30.78	.72	.28	.89	.11
29.54	30.17	31.72	32.76	35.11	37.12	.83	.17	.94	.06
29.88	30.51	32.06	33.08	35.44	37.45	.83	.17	.94	.06
32.21	32.84	34.39	36.21	38.56	40.58	.88	.12	.96	.04
36.00	36.63	38.19	23.23	25.58	27.59	0	1	0	1
20.18	20.81	22.36	24.12	26.48	28.49	.88	.12	.96	.04
8.55	9.18	10.73	14.33	16.68	18.69	.95	.05	.98	.02
26.79	27.42	28.97	28.63	30.99	33.00	.72	.28	.88	.12
43.57	44.20	45.75	47.35	49.70	51.72	.87	.13	.95	.05
35.05	35.68	37.23	38.98	41.33	43.34	.88	.12	.96	.04
30.30	30.93	32.48	26.01	28.36	30.37	.10	.90	.26	.74
20.36	20.99	22.54	23.26	25.61	27.62	.81	.19	.93	.07
34.00	34.63	36.18	28.00	30.35	32.36	.05	.95	.13	.87
24.00	24.63	26.18	26.00	28.36	30.37	.73	.27	.89	.11
14.00	14.63	16.18	16.00	18.35	20.36	.73	.27	.89	.11
33.99	34.62	36.17	37.95	40.30	42.31	.88	.12	.96	.04
21.88	22.51	24.06	25.88	28.23	30.24	.88	.12	.96	.04
18.32	18.95	20.50	19.81	22.17	24.18	.68	.32	.86	.14
22.00	22.63	24.18	25.96	28.32	30.33	.88	.12	.96	.04
12.00	12.63	14.18	16.67	19.03	21.04	.91	.09	.97	.03
18.37	19.00	20.55	15.07	17.42	19.43	.16	.84	.36	.64
26.00	26.63	28.18	27.43	29.78	31.79	.67	.33	.86	.14
13.26	13.89	15.44	17.32	19.67	21.68	.88	.12	.96	.04
25.30	25.93	27.49	23.21	25.57	27.58	.26	.74	.51	.49
15.17	15.80	17.35	16.63	18.99	21.00	.68	.32	.86	.14
19.66	20.29	21.84	23.22	25.58	27.59	.86	.14	.95	.05
24.00	24.63	26.18	28.39	30.75	32.76	.90	.10	.96	.04

(continued)

Table F4

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
8.69	9.32	10.87	12.58	14.93	16.94	.87	.13	.95	.05
15.86	16.50	18.05	19.97	22.32	24.33	.89	.11	.96	.04
35.40	36.04	37.59	39.38	41.73	43.75	.88	.12	.96	.04
21.94	22.57	24.12	26.77	29.13	31.14	.92	.08	.97	.03
26.20	26.83	28.38	30.19	32.55	34.56	.88	.12	.96	.04
24.00	24.63	26.18	30.00	32.36	34.37	.95	.05	.98	.02
36.84	37.47	39.02	41.43	43.78	45.79	.91	.09	.97	.03
17.16	17.79	19.34	20.96	23.32	25.33	.87	.13	.95	.05
19.38	20.01	21.56	23.71	26.06	28.07	.90	.10	.96	.04
25.95	26.58	28.13	30.52	32.87	34.88	.91	.09	.97	.03
11.85	12.48	14.03	17.63	19.99	22.00	.95	.05	.98	.02
35.92	36.55	38.10	39.92	42.27	44.28	.88	.12	.96	.04
13.86	14.49	16.04	18.56	20.91	22.93	.91	.09	.97	.03
16.05	16.69	18.24	19.82	22.17	24.18	.87	.13	.95	.05

Table F5

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
15.27	15.67	18.27	20.39	21.82	26.38	.93	.07	.98	.02
36.01	36.41	39.01	40.02	41.45	46.01	.88	.12	.97	.03
54.37	54.77	57.37	57.88	59.30	63.86	.85	.15	.96	.04
78.08	78.48	81.08	71.42	72.85	77.41	.03	.97	.14	.86
75.61	76.01	78.60	78.82	80.25	84.81	.83	.17	.96	.04
78.88	79.28	81.87	66.32	67.75	72.31	0	1	.01	.99
37.15	37.55	40.14	41.15	42.58	47.14	.88	.12	.97	.03
16.52	16.92	19.51	21.79	23.22	27.78	.93	.07	.98	.02
57.13	57.53	60.12	61.13	62.56	67.12	.88	.12	.97	.03
50.12	50.52	53.11	52.71	54.14	58.70	.79	.21	.94	.06
68.35	68.75	71.35	72.35	73.78	78.34	.88	.12	.97	.03
59.09	59.49	62.09	62.69	64.12	68.67	.86	.14	.96	.04
49.26	49.66	52.26	53.26	54.69	59.25	.88	.12	.97	.03
43.97	44.37	46.96	47.99	49.42	53.98	.88	.12	.97	.03
43.29	43.69	46.28	47.29	48.72	53.28	.88	.12	.97	.03
44.14	44.54	47.14	47.87	49.30	53.85	.87	.13	.97	.03
35.29	35.69	38.29	39.29	40.72	45.28	.88	.12	.97	.03
66.83	67.23	69.82	70.83	72.26	76.81	.88	.12	.97	.03
77.23	77.63	80.23	81.15	82.57	87.13	.88	.12	.97	.03
73.52	73.92	76.51	77.48	78.91	83.46	.88	.12	.97	.03
75.72	76.12	78.72	79.68	81.11	85.67	.88	.12	.97	.03
39.20	39.60	42.20	42.61	44.04	48.60	.85	.15	.96	.04
34.53	34.93	37.52	38.62	40.05	44.61	.89	.11	.97	.03
63.31	63.71	66.30	67.31	68.74	73.29	.88	.12	.97	.03
41.46	41.86	44.45	45.47	46.90	51.46	.88	.12	.97	.03
55.82	56.22	58.81	59.46	60.89	65.45	.86	.14	.97	.03
40.70	41.10	43.69	44.70	46.13	50.68	.88	.12	.97	.03
60.94	61.34	63.93	64.94	66.37	70.92	.88	.12	.97	.03
53.44	53.84	56.43	55.72	57.15	61.70	.76	.24	.93	.07
57.86	58.26	60.85	61.86	63.29	67.85	.88	.12	.97	.03
72.73	73.13	75.72	76.69	78.12	82.68	.88	.12	.97	.03
49.17	49.57	52.16	53.17	54.60	59.16	.88	.12	.97	.03

(continued)

Table F5

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
65.30	65.70	68.29	69.30	70.73	75.28	.88	.12	.97	.03
27.40	27.80	30.39	31.63	33.06	37.62	.89	.11	.97	.03
59.94	60.34	62.93	63.94	65.37	69.93	.88	.12	.97	.03
56.44	56.84	59.44	60.44	61.87	66.43	.88	.12	.97	.03
47.84	48.24	50.84	51.93	53.36	57.91	.89	.11	.97	.03
67.90	68.30	70.89	71.88	73.30	77.86	.88	.12	.97	.03
50.00	50.40	52.99	54.03	55.46	60.02	.88	.12	.97	.03
68.41	68.81	71.40	70.04	71.47	76.03	.69	.31	.91	.09
64.06	64.46	67.05	68.06	69.49	74.04	.88	.12	.97	.03
65.09	65.49	68.08	67.69	69.11	73.67	.79	.21	.94	.06
64.53	64.93	67.52	67.81	69.24	73.80	.84	.16	.96	.04
39.73	40.13	42.72	43.73	45.16	49.72	.88	.12	.97	.03
66.10	66.50	69.09	70.10	71.53	76.09	.88	.12	.97	.03
49.58	49.98	52.57	48.43	49.86	54.42	.36	.64	.72	.28
41.11	41.51	44.10	45.11	46.54	51.10	.88	.12	.97	.03
54.80	55.20	57.80	58.80	60.23	64.79	.88	.12	.97	.03
68.17	68.57	71.16	72.17	73.60	78.15	.88	.12	.97	.03
72.27	72.67	75.26	76.26	77.69	82.24	.88	.12	.97	.03
49.18	49.58	52.18	53.18	54.61	59.17	.88	.12	.97	.03
46.29	46.69	49.29	49.10	50.53	55.08	.80	.20	.95	.05
41.54	41.94	44.53	45.54	46.97	51.52	.88	.12	.97	.03
39.61	40.01	42.60	47.37	48.80	53.36	.98	.02	1	0
38.15	38.55	41.14	42.15	43.58	48.13	.88	.12	.97	.03
42.00	42.40	44.99	47.20	48.62	53.18	.93	.07	.98	.02
35.14	35.54	38.13	39.28	40.70	45.26	.89	.11	.97	.03
59.78	60.18	62.78	63.78	65.21	69.77	.88	.12	.97	.03
75.57	75.97	78.57	79.56	80.98	85.54	.88	.12	.97	.03
49.60	50.00	52.59	53.60	55.02	59.58	.88	.12	.97	.03
43.40	43.80	46.40	48.14	49.57	54.12	.91	.09	.98	.02
72.21	72.61	75.20	76.18	77.61	82.16	.88	.12	.97	.03
48.38	48.78	51.37	52.15	53.58	58.14	.87	.13	.97	.03
49.46	49.86	52.45	53.15	54.58	59.14	.86	.14	.97	.03

(continued)

Table F5

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

	RFT			SDbS		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
38.00	38.40	40.99	43.18	44.61	49.17	.93	.07	.98	.02
51.11	51.51	54.10	55.11	56.54	61.10	.88	.12	.97	.03
52.08	52.48	55.07	54.87	56.30	60.86	.80	.20	.95	.05
55.43	55.83	58.42	59.43	60.85	65.41	.88	.12	.97	.03
75.76	76.16	78.75	79.76	81.18	85.74	.88	.12	.97	.03
33.48	33.88	36.47	37.48	38.91	43.47	.88	.12	.97	.03
64.98	65.38	67.97	68.98	70.41	74.97	.88	.12	.97	.03
74.60	75.00	77.59	78.63	80.06	84.62	.88	.12	.97	.03
53.18	53.58	56.17	55.82	57.25	61.81	.79	.21	.94	.06
59.57	59.97	62.57	63.57	65.00	69.56	.88	.12	.97	.03
63.01	63.41	66.00	66.01	67.43	71.99	.82	.18	.95	.05
66.74	67.14	69.74	70.74	72.17	76.73	.88	.12	.97	.03
50.08	50.48	53.08	54.08	55.51	60.07	.88	.12	.97	.03
64.28	64.68	67.27	68.28	69.71	74.26	.88	.12	.97	.03
54.55	54.95	57.54	58.55	59.98	64.53	.88	.12	.97	.03
79.86	80.26	82.85	81.94	83.37	87.93	.74	.26	.93	.07
61.28	61.68	64.28	65.28	66.71	71.27	.88	.12	.97	.03
56.52	56.92	59.52	60.52	61.95	66.50	.88	.12	.97	.03
74.47	74.87	77.46	77.48	78.91	83.47	.82	.18	.95	.05
37.32	37.72	40.31	41.32	42.75	47.31	.88	.12	.97	.03
46.73	47.13	49.72	50.73	52.16	56.71	.88	.12	.97	.03
54.67	55.07	57.66	58.67	60.10	64.66	.88	.12	.97	.03
76.68	77.08	79.67	80.68	82.11	86.66	.88	.12	.97	.03
58.01	58.41	61.01	62.01	63.44	68.00	.88	.12	.97	.03
58.10	58.50	61.09	62.10	63.53	68.08	.88	.12	.97	.03
9.59	9.99	12.59	15.35	16.78	21.34	.95	.05	.99	.01
11.18	11.58	14.17	17.26	18.69	23.24	.95	.05	.99	.01
37.61	38.01	40.61	41.61	43.04	47.60	.88	.12	.97	.03
41.07	41.47	44.06	43.76	45.18	49.74	.79	.21	.94	.06
40.97	41.37	43.97	45.43	46.86	51.41	.90	.10	.98	.02
57.64	58.04	60.63	61.64	63.07	67.62	.88	.12	.97	.03
52.03	52.43	55.02	56.03	57.46	62.02	.88	.12	.97	.03

(continued)

Table F5

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
69.25	69.65	72.24	73.25	74.67	79.23	.88	.12	.97	.03
50.23	50.63	53.22	54.23	55.66	60.21	.88	.12	.97	.03
65.00	65.40	68.00	69.00	70.43	74.99	.88	.12	.97	.03
67.60	68.00	70.59	71.60	73.03	77.59	.88	.12	.97	.03
69.65	70.05	72.64	73.65	75.07	79.63	.88	.12	.97	.03
21.81	22.21	24.80	25.81	27.23	31.79	.88	.12	.97	.03
77.66	78.06	80.65	81.66	83.09	87.65	.88	.12	.97	.03
56.32	56.72	59.31	60.32	61.74	66.30	.88	.12	.97	.03
76.78	77.18	79.77	80.78	82.21	86.77	.88	.12	.97	.03
37.04	37.44	40.03	41.04	42.47	47.02	.88	.12	.97	.03
66.60	67.00	69.59	70.60	72.03	76.59	.88	.12	.97	.03
55.51	55.91	58.51	57.71	59.14	63.69	.75	.25	.93	.07
65.18	65.58	68.18	69.18	70.61	75.17	.88	.12	.97	.03
62.31	62.71	65.31	66.31	67.74	72.30	.88	.12	.97	.03
81.26	81.66	84.25	85.26	86.69	91.25	.88	.12	.97	.03
52.89	53.29	55.88	56.89	58.32	62.88	.88	.12	.97	.03
55.87	56.27	58.86	60.07	61.50	66.05	.89	.11	.97	.03
68.20	68.60	71.20	72.20	73.63	78.19	.88	.12	.97	.03
63.40	63.80	66.39	66.63	68.06	72.61	.83	.17	.96	.04
51.36	51.76	54.36	55.36	56.79	61.35	.88	.12	.97	.03
33.06	33.46	36.06	37.06	38.49	43.05	.88	.12	.97	.03
62.17	62.57	65.16	66.15	67.57	72.13	.88	.12	.97	.03
42.00	42.40	44.99	40.81	42.24	46.80	.36	.64	.71	.29
76.54	76.94	79.53	80.54	81.97	86.53	.88	.12	.97	.03
46.00	46.40	48.99	53.57	55.00	59.56	.98	.02	.99	.01
76.61	77.01	79.61	80.61	82.04	86.60	.88	.12	.97	.03
78.49	78.89	81.48	82.49	83.92	88.48	.88	.12	.97	.03
41.06	41.46	44.05	45.06	46.49	51.04	.88	.12	.97	.03
78.76	79.16	81.76	82.76	84.19	88.75	.88	.12	.97	.03
44.84	45.24	47.83	51.01	52.44	57.00	.96	.04	.99	.01
46.59	46.99	49.58	51.25	52.68	57.23	.91	.09	.98	.02
59.96	60.36	62.95	63.96	65.39	69.94	.88	.12	.97	.03

(continued)

Table F5

Penalized model fit statistics comparing the RFT and SDbS model using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

RFT			SDbS			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	RFT	SDbS	RFT	SDbS
70.59	70.99	73.58	74.59	76.02	80.58	.88	.12	.97	.03
48.43	48.83	51.42	52.83	54.25	58.81	.90	.10	.98	.02
79.07	79.47	82.06	83.07	84.50	89.06	.88	.12	.97	.03
77.66	78.06	80.66	81.66	83.09	87.65	.88	.12	.97	.03
67.14	67.54	70.14	71.14	72.57	77.13	.88	.12	.97	.03
53.72	54.12	56.72	60.92	62.34	66.90	.97	.03	.99	.01
54.00	54.40	56.99	58.07	59.50	64.06	.88	.12	.97	.03

Appendix G:

SDbS and SDbS+Range Fit

Table G1

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
0.40	11.35	73.49	.91	65.31	0.40	11.35	73.49	.91	0	65.31	.06	0	.25
0.30	0.84	5.00	.85	28.17	0.29	0.60	5.45	.79	.25	24.81	.31	0	.52
0.44	2.49	7.50	.67	79.27	0.44	2.57	7.42	.63	.27	78.67	.50	0	.73
0.25	1.15	17.00	.40	4.57	0.20	0.19	100.00	.15	.46	-24.64	.50	.26	.61
0.18	0.69	11.65	.34	-40.57	0.16	0.08	88.37	.12	.28	-55.47	.26	.09	.33
0.41	1.44	15.83	.38	70.23	0.41	0.70	31.24	.24	.18	68.71	.16	0	.38
0.18	2.15	9.39	.52	-38.44	0.18	2.15	9.35	.52	0	-38.44	.18	0	.41
0.49	1.57	11.80	.50	94.40	0.49	1.57	11.81	.50	0	94.40	.50	0	.76
0.55	0.56	30.77	.28	108.58	0.53	0.32	54.79	.20	.38	103.24	.53	0	.79
0.49	1.09	15.04	.35	93.31	0.49	0.74	27.47	.24	.23	92.12	.22	0	.46
0.44	1.20	5.58	.67	77.83	0.42	0.06	65.26	.39	.47	74.21	.37	0	.60
0.42	3.73	8.24	.60	73.71	0.42	4.51	11.94	.57	.13	73.18	.12	0	.33
0.44	3.62	7.87	.77	77.87	0.43	5.24	11.32	.73	.44	74.37	.51	.02	.69
0.26	3.92	6.03	.85	10.06	0.26	4.16	5.90	.86	.07	10.00	.29	0	.53
0.19	6.74	6.07	.83	-34.74	0.18	14.05	17.47	.94	.11	-39.50	.10	0	.20
0.47	1.01	13.57	.35	88.98	0.48	4.16	75.16	.50	.28	90.30	.21	0	.45
0.17	1.38	3.47	.84	-45.82	0.17	4.70	100.00	.51	.19	-50.23	.16	0	.23
0.36	1.44	4.99	.68	51.57	0.36	0.71	4.80	.66	.24	50.61	.23	0	.46
0.50	0.08	100.00	.13	96.44	0.50	0.08	98.00	.13	0	96.51	.09	0	.26
0.24	0.80	4.38	.87	-2.70	0.23	0.60	4.62	.83	.13	-4.12	.20	0	.34
0.38	2.73	5.43	.84	60.68	0.37	0.20	15.34	.61	.47	54.71	.43	0	.66
0.34	4.82	7.86	.84	43.72	0.31	14.47	69.76	.96	.36	32.29	.43	.16	.55
0.22	4.70	5.28	.84	-13.07	0.20	13.13	11.34	.92	.25	-24.69	.23	.04	.35
0.27	1.53	4.91	.59	15.34	0.27	0.16	14.35	.51	.17	15.59	.13	0	.27

Table G2

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
3.25	0.37	9.53	.43	342.96	3.27	2.62	0.34	.50	.40	343.83	.35	0	.59
7.08	0.08	20.56	.50	445.59	7.06	0.03	60.90	.35	.26	444.69	.16	0	.39
16.20	0.98	0.00	.50	554.93	16.13	0.01	79.36	.52	.05	554.50	.03	0	.29
3.98	0.61	6.97	.47	369.55	3.98	0.61	6.96	.47	0	369.55	.11	0	.38
14.29	0.14	14.89	.51	538.35	13.30	0.02	99.34	.49	.89	528.92	.96	.69	1
13.75	0.89	83.68	.51	533.26	13.75	0.88	99.39	.50	.01	533.26	.14	0	.47
19.69	0.10	21.13	.50	580.66	19.69	0.04	67.43	.50	0	580.65	.58	0	1
14.56	0.90	8.91	.58	540.87	14.55	0.83	68.25	.49	.32	540.70	.34	0	.79
9.77	0.07	34.09	.50	488.18	9.83	0.22	6.30	.76	.15	488.62	.36	0	.70
13.99	0.08	20.94	.50	535.61	13.98	0.01	84.17	.52	.33	535.24	.20	0	.57
5.88	0.02	81.17	.54	420.35	5.84	0.01	76.62	.56	.06	420.39	.02	0	.14
17.83	0.19	11.18	.50	567.55	17.82	0.21	10.40	.50	.33	567.53	.73	0	1
9.68	0.79	100.00	.50	487.01	9.67	0.82	100.00	.50	.31	486.78	.39	0	.95
8.97	0.35	8.14	.51	476.86	8.97	0.38	7.79	.50	.17	476.83	.40	0	.77
8.48	0.52	0.00	.50	469.50	8.37	0.01	78.94	.52	0	466.86	.02	0	.21
7.18	0.89	11.58	.66	447.44	7.15	0.78	15.22	.58	.18	447.42	.70	0	1
8.61	0.39	11.70	.46	471.42	8.62	0.34	13.73	.41	.02	471.42	.67	0	1
7.14	0.99	100.00	.53	446.73	7.13	0.01	81.00	.54	.14	446.31	.08	0	.38
4.41	0.03	69.59	.46	383.96	4.48	20.70	18.21	.50	0	385.33	.03	0	.19
5.38	0.64	8.34	.54	409.45	5.38	0.64	8.73	.51	.05	409.45	.43	0	.71
7.82	0.18	9.55	.50	458.81	7.81	0.01	72.03	.50	.50	458.54	.37	0	.73
14.19	0.08	20.91	.50	537.41	14.19	0.01	85.07	.51	.35	537.30	.20	0	.60
12.50	0.10	17.96	.50	520.67	12.49	0.07	26.37	.50	.28	520.64	.34	0	.71
5.58	0.06	31.49	.51	414.27	5.64	0.04	54.43	.55	.02	414.29	.15	0	.36
8.97	0.88	38.19	.54	476.83	8.96	0.81	100.00	.50	.18	476.73	.33	0	.68
8.03	0.13	12.93	.50	462.24	8.03	19.55	0.27	.50	.45	462.21	.38	0	.64
3.27	0.48	7.04	.79	343.52	3.23	0.24	12.84	.55	.54	341.95	.70	.47	.82
5.96	0.78	9.09	.69	423.02	5.96	0.77	19.77	.57	.45	422.78	.58	.13	.89
3.12	0.82	8.79	.88	337.62	3.02	0.26	99.79	.50	.87	333.18	.98	.71	1
7.77	0.45	7.64	.85	455.49	7.60	0.16	71.79	.50	.86	455.07	.84	.10	1
4.05	0.32	7.14	.52	372.00	4.05	0.40	0.23	.50	.11	371.98	.13	0	.44
5.88	1.25	8.85	.55	421.23	5.88	1.25	11.21	.54	.04	421.23	.11	0	.43

(continued)

Table G3

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Melrose et al. (2012) (continued). Note: each row is the data for one participant.

		SDbS					SDbS+Range					*w* bootstrap	
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
5.32	0.95	8.14	.80	408.02	5.28	1.09	11.70	.70	.50	406.93	.67	.31	.85
2.85	1.20	28.39	.81	325.44	2.84	1.01	100.00	.73	.38	325.14	.93	0	1
2.31	2.45	10.23	.82	297.98	2.30	2.77	21.40	.88	.01	297.89	.07	0	.18
10.96	0.66	31.13	.55	503.38	10.96	0.65	100.00	.54	.22	503.30	.39	0	1
3.06	2.06	10.03	.78	335.10	3.06	2.06	10.03	.78	0	335.10	.07	0	.23
4.76	1.03	8.12	.70	393.33	4.77	0.01	81.27	.54	.38	392.81	.40	.04	.64
6.11	1.13	14.21	.88	426.08	6.49	5.35	9.09	.50	1	434.21	.89	.39	1
7.93	0.03	90.44	.44	460.94	8.03	6.89	59.58	.50	0	462.35	.13	0	.61
5.46	1.09	12.55	.90	411.44	6.02	25.26	2.60	.50	1	424.17	.93	.57	1
9.54	0.85	30.21	.74	485.02	10.92	8.39	2.20	.50	1	502.92	.85	.07	1
13.93	0.90	100.00	.70	535.03	13.93	0.90	100.00	.70	0	535.03	.94	.01	1
1.79	0.36	7.71	.68	263.95	1.84	5.69	0.00	.50	.43	267.53	.41	.20	.52
2.37	0.63	14.39	.64	301.04	2.36	0.62	100.00	.57	.64	300.42	.88	0	1
9.43	0.82	44.35	.67	483.52	9.41	0.77	100.00	.64	.26	483.28	.92	0	1
12.19	0.71	10.99	.76	517.42	12.10	0.26	100.00	.50	.85	516.37	.97	.59	1
7.85	0.52	7.53	.69	459.22	7.79	0.03	74.27	.50	.72	458.28	.58	.04	.89
1.63	0.68	6.52	.76	251.48	1.62	0.55	10.65	.52	.27	250.91	.40	.02	.59
4.39	1.18	11.82	.92	382.53	4.26	1.31	48.44	.86	.66	378.65	1.	1	1
5.44	0.88	7.69	.85	410.29	5.36	0.82	14.65	.63	.70	408.88	.79	.33	1
7.14	1.17	27.84	.80	446.81	7.58	4.93	1.43	.50	1	454.59	.90	.40	1

Table G3

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
1.46	0.49	35.30	.55	316.45	1.46	0.53	100.00	.54	.44	315.79	.56	0	.96
4.92	0.00	85.47	.52	529.16	4.89	0.70	11.09	.54	0	529.17	.15	0	.41
13.58	0.48	14.07	.53	708.84	13.58	0.49	14.52	.53	.03	708.84	.35	0	.70
3.58	0.96	10.37	.77	474.07	3.56	1.21	14.50	.78	.23	473.48	.57	.24	.70
6.22	0.26	5.52	.50	571.48	6.22	3.72	0.00	.50	.09	571.49	.11	0	.42
4.26	1.35	22.19	.80	504.61	4.26	1.95	12.91	.90	.19	504.61	.44	.02	.64
8.89	0.99	9.31	.69	634.33	8.87	1.14	11.45	.64	.21	633.95	.30	0	.52
7.31	1.24	29.96	.80	599.77	7.31	1.24	29.96	.80	0	599.77	.41	0	.75
4.02	0.58	9.74	.78	494.44	3.99	0.52	100.00	.58	.75	493.17	.93	.31	1
14.45	0.41	9.32	.63	719.74	14.43	0.78	7.57	.51	.70	719.49	.68	.01	1
4.74	0.84	14.91	.70	523.76	4.74	0.87	15.29	.71	.08	523.74	.57	.04	.85
9.71	0.15	9.71	.46	649.79	9.66	0.04	94.68	.55	.40	649.50	.23	0	.61
8.81	0.92	8.52	.70	632.64	8.77	0.01	75.55	.50	.53	632.00	.36	0	.70
2.97	0.45	7.27	.69	441.30	2.96	0.49	7.66	.51	.52	440.55	.58	.05	.73
13.15	0.62	5.68	.76	703.12	13.14	0.51	7.38	.52	.27	703.09	.37	0	.64
11.15	0.76	10.90	.58	674.09	11.14	0.89	11.12	.55	.14	674.04	.23	0	.58
9.37	0.52	16.77	.49	643.58	9.37	0.52	16.77	.49	0	643.58	.21	0	.52
16.02	0.51	6.63	.84	735.69	15.82	0.01	25.43	.50	.80	735.54	.75	.06	1
6.46	0.67	8.62	.75	578.08	6.43	0.83	16.32	.67	.47	577.23	.63	0	.85
6.55	1.13	15.83	.79	580.44	6.61	1.47	11.40	.91	.05	580.38	.47	0	.73
3.63	0.80	10.04	.82	476.65	3.60	0.72	27.36	.66	.50	475.37	.78	.22	.96
4.09	0.68	6.53	.87	500.23	4.11	0.49	10.62	.60	.25	500.17	.69	.18	.81
5.46	0.38	7.85	.68	548.59	5.43	0.14	7.58	.59	.64	547.60	.71	.53	.83
7.73	0.49	9.56	.70	609.72	7.70	0.53	54.34	.55	.74	608.95	.77	0	1
5.72	0.83	10.71	.80	556.73	5.70	0.89	17.66	.74	.43	556.02	.71	.01	.95
4.99	0.58	12.43	.72	532.65	4.98	0.50	22.81	.62	.42	532.11	.85	.01	1
7.03	0.37	7.59	.57	593.04	7.02	1.47	6.57	.51	.36	592.82	.35	0	.55
6.86	0.83	24.53	.73	588.55	6.90	1.17	18.57	.85	.02	588.44	.54	0	.94
15.86	0.66	100.00	.63	736.12	15.85	0.68	100.00	.64	.08	736.09	.76	0	1
5.54	1.18	16.74	.81	551.08	5.54	1.19	16.73	.81	0	551.07	.54	0	.74
7.59	0.74	9.03	.65	606.40	7.58	0.91	9.74	.64	.19	606.30	.35	0	.65
4.63	0.78	32.80	.70	519.48	4.59	0.68	100.00	.63	.34	517.98	.69	0	.98

(continued)

Table G3

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
5.33	0.47	7.94	.78	544.28	5.22	0.61	100.00	.59	.88	540.64	.84	.11	1
7.06	0.53	8.87	.73	593.76	7.02	0.69	14.16	.62	.74	592.64	.76	.06	1
4.93	1.03	8.87	.80	530.54	4.90	1.65	15.71	.85	.28	529.34	.54	.25	.66
16.48	0.45	8.85	.56	742.90	16.48	0.49	9.42	.51	.19	742.89	.35	0	.74
4.23	1.24	12.94	.85	503.49	4.24	1.95	31.21	.90	.30	501.27	.64	.24	.77
9.86	0.31	15.18	.62	652.42	10.16	0.06	0.14	.50	1	657.75	.88	0	1
10.85	0.22	13.01	.53	669.40	10.97	0.06	62.27	.80	.65	669.39	.91	.39	1
1.18	0.82	9.53	.75	278.47	1.17	1.08	13.51	.74	.34	277.91	.52	0	.78
12.80	0.01	54.35	.77	694.89	12.60	0.01	85.57	.49	0	695.42	0	0	.05
8.95	0.42	8.71	.58	635.45	8.89	0.67	20.65	.49	.50	634.22	.53	0	.85
7.37	0.52	13.39	.49	601.18	7.37	0.54	14.03	.49	.05	601.17	.38	0	.72
8.22	0.72	100.00	.50	620.41	8.22	0.72	100.00	.50	.00	620.41	.31	0	.67
1.95	0.40	20.92	.48	367.12	1.95	0.15	92.19	.23	.01	367.12	.41	0	.97
7.04	0.37	6.93	.51	593.26	7.03	0.86	100.00	.50	.25	593.01	.26	0	.56
8.53	13.66	49.63	.50	626.99	8.53	2.46	88.48	.50	0	626.99	.05	0	.23
12.77	0.00	80.90	.51	698.41	12.80	0.00	67.57	.51	0	698.43	.01	0	.09
5.98	0.47	8.48	.47	564.63	5.99	0.45	9.09	.45	.01	564.63	.23	0	.43
0.10	8.47	34.17	.50	-243.52	0.10	29.10	7.42	.50	0	-243.52	0	0	0
27.45	0.28	11.97	.51	832.68	27.44	0.31	12.57	.50	.49	832.66	.74	0	1
3.09	0.54	22.06	.50	448.46	3.09	0.54	22.05	.50	0	448.46	.43	0	.83
9.83	0.35	9.18	.52	652.03	9.83	0.61	9.11	.49	.44	651.90	.46	0	.93
7.22	0.51	6.79	.50	597.71	7.22	0.58	6.69	.50	.04	597.71	.17	0	.43
9.69	0.57	24.40	.50	649.37	9.69	0.57	24.40	.50	0	649.37	.39	0	.88
7.15	0.53	13.55	.49	595.88	7.15	0.54	13.02	.50	0	595.88	.34	0	.67
1.07	0.55	8.94	.45	261.00	1.07	0.14	68.14	.20	0	260.98	.11	0	.44
4.91	0.56	16.15	.50	529.89	4.91	0.56	16.15	.50	0	529.89	.45	0	.77
6.19	0.44	8.48	.54	570.55	6.22	0.67	13.43	.49	.27	570.46	.44	0	.67
5.48	0.22	6.34	.50	549.07	5.47	0.90	100.00	.51	.19	548.94	.18	0	.48
13.16	0.00	80.51	.51	704.22	13.29	0.00	98.56	.43	.01	704.23	0	0	.03
23.47	0.48	18.68	.50	805.11	23.46	0.52	26.35	.50	.17	805.09	.57	0	1
14.18	0.53	11.09	.48	716.46	14.18	0.53	11.09	.48	0	716.46	.29	0	.75
11.49	3.37	0.01	.50	679.48	11.28	0.01	86.45	.50	0	676.17	0	0	.02

(continued)

Table G3

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
16.42	0.25	7.44	.61	742.35	16.43	19.96	21.84	.50	.69	742.42	.62	0	1
11.21	0.70	12.90	.49	675.10	11.21	0.70	12.90	.49	0	675.10	.23	0	.50
5.97	0.42	7.62	.49	564.16	5.97	0.43	7.50	.49	0	564.16	.23	0	.50
8.20	0.01	81.52	.48	620.62	8.23	0.01	77.08	.50	0	620.62	.02	0	.18
6.41	1.07	11.41	.50	576.81	6.41	1.07	11.41	.50	0	576.81	.08	0	.26
10.58	0.21	13.99	.40	665.40	10.61	0.25	11.18	.47	0	665.40	.58	0	.99
6.76	0.38	10.16	.45	586.11	6.80	0.33	11.88	.40	.01	586.10	.33	0	.56
7.35	0.40	6.70	.51	600.86	7.35	0.88	74.73	.51	.19	600.73	.23	0	.43
14.81	0.49	8.50	.69	723.62	14.77	0.35	13.25	.50	.11	723.61	.80	.01	1
13.84	0.48	11.36	.49	712.23	13.84	0.56	12.30	.49	.18	712.16	.38	0	.82
6.26	0.15	6.53	.50	572.53	6.28	20.49	0.26	.50	.21	573.03	.17	0	.35
10.06	0.27	12.69	.46	655.97	10.06	0.27	12.69	.46	0	655.97	.60	0	1
8.14	0.30	11.89	.42	618.74	8.14	0.34	10.15	.47	0	618.75	.36	0	.67
8.91	0.43	9.24	.56	635.15	8.93	0.47	11.62	.49	.25	635.01	.45	0	.82
6.08	0.39	6.86	.57	567.30	6.03	0.90	18.12	.50	.36	565.91	.41	0	.59
15.77	0.28	20.22	.50	735.21	15.77	0.28	20.32	.50	0	735.21	.96	.33	1
7.53	0.33	16.93	.43	605.12	7.53	0.33	16.93	.43	0	605.12	.52	0	.90

Table G4

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

		SDbS					SDbS+Range				*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
0.30	0.05	16.73	.51	21.38	0.30	0.02	60.29	.52	.03	21.48	.51	0	.96
0.10	0.24	18.60	.50	18.06	0.10	0.29	99.49	.50	.32	18.05	.91	.65	1
0.20	0.07	16.04	.64	15.15	0.22	19.39	32.36	.50	1	16.67	.84	.62	1
0.28	0.20	8.94	.45	22.39	0.26	0.30	100.00	.43	.14	21.75	.42	0	.95
0.10	0.03	33.51	.50	19.83	0.10	9.41	12.13	.50	.10	19.97	.34	.07	.45
0.10	0.35	100.00	.48	8.75	0.10	9.68	1.41	.50	.27	8.84	.46	.37	.48
0.24	0.04	24.37	.50	20.15	0.24	6.47	0.10	.50	.38	20.44	.37	.09	.59
0.16	0.02	77.48	.50	15.54	0.16	0.02	86.85	.51	.17	15.54	.78	.61	1
0.76	0.06	15.94	.52	37.42	0.76	0.07	14.14	.52	.20	37.46	.68	0	1
0.16	0.12	84.05	.50	9.50	0.16	0.13	66.86	.50	0	9.50	.90	.27	1
0.10	0.03	32.39	.50	26.19	0.10	10.61	11.47	.50	.13	26.56	.37	.13	.46
0.17	0.07	19.44	.51	17.16	0.18	0.09	49.56	.50	.57	17.31	.98	.74	1
0.10	0.13	10.35	.49	13.75	0.10	0.36	100.00	.49	.93	13.76	.67	.62	.82
0.10	5.03	0.00	.50	27.20	0.10	9.80	7.37	.50	.01	27.20	.18	.04	.33
0.10	0.07	17.96	.51	14.42	0.23	95.93	80.78	.50	1	20.76	.84	.61	1
0.10	0.06	16.49	.50	14.84	0.10	9.01	5.90	.50	.88	14.88	.61	.60	.63
0.21	0.19	16.35	.50	15.31	0.35	97.05	2.74	.50	1	25.85	.84	.16	1
1.00	0.21	89.10	.50	41.76	1.00	0.02	99.86	.49	.21	41.99	.91	0	1
0.10	0.05	17.31	.50	11.79	0.10	9.49	3.57	.50	.47	11.81	.49	.48	.53
0.10	0.07	18.87	.50	5.27	0.10	0.09	52.58	.50	.54	5.28	.84	.63	1
0.10	0.07	16.33	.52	9.42	0.10	0.07	16.36	.54	0	9.43	.86	.67	1
0.26	2.83	0.00	.50	26.38	0.10	0.31	100.00	.52	0	22.45	.96	.63	1
0.17	3.43	0.00	.50	32.06	0.17	13.76	72.19	.50	0	32.06	.13	0	.30
0.10	8.26	0.66	.50	25.57	0.15	0.01	91.17	.47	0	25.15	.15	0	.41
0.10	0.39	20.13	.50	11.35	0.10	0.39	20.14	.50	0	11.35	.40	.23	.47
0.10	0.31	15.41	.48	18.27	0.10	2.77	15.12	.50	.75	18.63	.62	.60	.66
0.10	0.06	17.16	.50	11.24	0.10	0.06	17.16	.50	0	11.24	.78	.63	.99
0.10	0.23	11.64	.51	17.16	0.10	0.26	13.56	.51	.35	17.16	.68	.60	.95
0.10	0.07	14.89	.50	15.30	0.10	0.02	85.04	.50	.79	15.31	.63	.59	.69
0.49	0.06	17.50	.50	31.34	0.49	0.06	17.51	.50	0	31.34	.87	0	1
0.10	0.05	17.96	.50	6.37	0.10	0.07	12.25	.51	.07	6.38	.49	.48	.53
0.10	0.05	18.02	.50	2.93	0.10	0.01	86.02	.51	.52	2.94	.49	.48	.61

(continued)

Table G4

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
1.04	0.19	16.22	.50	44.06	1.04	0.20	33.81	.50	.34	44.06	.91	0	1
0.10	0.23	18.35	.51	7.25	0.10	0.25	82.52	.50	.37	7.19	.99	1	1
0.37	0.06	16.76	.51	26.40	0.38	0.06	0.20	.50	.67	26.78	.61	0	1
0.13	0.23	100.00	.50	21.07	0.13	0.23	100.00	.50	0	21.07	1	1	1
0.10	0.02	48.02	.51	20.33	0.10	0.05	17.53	.50	0	20.31	.49	.48	.61
0.16	0.28	100.00	.42	18.42	0.17	0.28	100.00	.42	.01	18.42	.55	.03	.83
0.32	0.32	100.00	.53	24.76	0.32	0.32	100.00	.53	0	24.76	.29	0	.81
0.33	0.32	99.64	.53	25.08	0.33	0.32	100.00	.53	0	25.08	.26	0	.73
0.24	9.46	1.30	.50	28.21	0.10	0.04	99.85	.54	.03	20.44	.01	0	.09
0.10	0.04	99.99	.46	15.23	0.53	15.89	7.81	.50	0	32.00	.10	0	.45
0.14	0.12	39.55	.50	16.12	0.14	0.17	98.05	.50	.07	16.12	.71	.59	1
0.10	1.06	50.30	.96	6.33	0.10	0.00	82.25	.52	.52	6.86	.53	.48	.61
0.29	0.98	90.98	.94	20.63	0.29	0.49	100.00	.70	.24	21.94	.51	0	1
0.83	0.42	100.00	.64	39.35	0.82	0.40	100.00	.62	.28	39.33	.50	0	1
0.48	0.40	100.00	.58	30.98	0.48	0.40	100.00	.58	0	30.98	.14	0	.69
0.10	0.11	100.00	.35	18.01	0.24	10.15	23.01	.50	0	26.30	.18	0	.46
0.18	0.09	37.52	.50	15.26	0.19	0.00	96.87	.48	.88	12.19	.56	0	1
0.10	0.11	91.16	.50	20.00	0.10	39.63	26.34	.50	1	30.12	.67	.61	.86
0.10	0.21	100.00	.50	18.00	0.10	4.54	0.04	.50	1	20.06	.67	.61	.86
0.10	0.12	72.65	.50	8.00	0.10	0.12	100.00	.50	0	8.00	.96	.29	1
0.42	0.19	10.51	.52	29.95	0.42	96.92	5.90	.50	.46	29.99	.42	0	1
0.16	7.68	0.09	.50	17.88	0.10	0.07	30.74	.50	.01	16.27	.02	0	.15
0.10	0.16	17.69	.53	11.81	0.20	0.46	10.20	.50	1	14.32	.77	.60	1
0.13	7.90	0.49	.50	17.96	0.13	66.69	76.13	.50	0	17.96	.21	.01	.44
0.10	0.15	11.07	.65	8.67	0.10	0.01	80.77	.50	.94	8.38	.62	.60	.68
0.13	0.12	98.54	.50	7.07	0.13	0.12	100.00	.50	0	7.07	.98	1	1
0.10	0.16	13.65	.72	19.43	0.10	1.80	3.10	.50	1	24.24	.74	.66	.93
0.10	0.42	12.91	.79	9.32	0.10	0.12	100.00	.50	.92	8.82	.62	.60	.68
0.10	0.66	100.00	.88	15.21	0.15	9.79	16.96	.50	1	21.30	.71	.60	1
0.10	0.50	18.67	.81	8.63	0.10	0.12	100.00	.50	.85	8.98	.64	.60	.88
0.12	0.14	12.36	.62	15.22	0.17	7.93	64.80	.50	.77	15.66	.63	.51	.94
0.10	7.25	18.00	.50	20.39	0.10	6.29	4.39	.50	0	20.39	.48	.43	.50

(continued)

Table G4

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

		SDbS					SDbS+Range				*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
0.14	0.12	59.90	.50	4.58	0.14	0.12	100.00	.50	.01	4.58	.74	.60	1
0.10	4.24	0.00	.50	11.97	0.10	13.62	1.37	.50	.08	11.96	.48	.39	.50
0.49	0.09	11.92	.50	31.38	0.49	98.39	55.90	.50	.06	31.40	.19	0	.99
0.10	0.17	8.68	.72	18.77	0.10	0.00	0.10	.50	.76	17.97	.61	.60	.64
0.29	0.56	10.88	.67	22.19	0.29	0.70	7.03	.52	.20	22.18	.27	0	.80
0.18	0.15	7.01	.64	22.00	0.10	57.44	23.12	.50	.82	20.77	.62	.60	.68
0.52	0.18	9.39	.66	33.43	0.50	1.83	9.67	.50	.82	32.84	.70	0	1
0.12	0.07	42.51	.50	12.96	0.10	7.46	2.99	.50	.28	13.17	.48	.44	.50
0.15	0.16	6.70	.61	15.71	0.15	9.86	7.66	.50	.21	15.38	.40	.19	.49
0.30	0.15	11.32	.64	22.52	0.30	52.25	34.78	.50	1	21.95	.74	0	1
0.10	0.17	7.51	.69	9.63	0.10	0.00	85.14	.50	.39	8.78	.48	.48	.49
0.26	5.18	0.01	.50	31.92	0.20	0.01	8.27	.50	1	30.01	.76	.60	1
0.10	0.17	10.38	.66	10.56	0.10	1.33	10.92	.50	1	9.89	.61	.60	.63
0.12	1.01	39.86	.94	11.82	0.17	0.01	81.21	.52	.49	11.72	.50	.31	.65

Table G5

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

		SDbS					SDbS+Range				*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
0.10	13.57	6.18	.50	12.39	0.10	6.87	77.99	.50	0	12.39	.46	.31	.50
0.27	1.89	1.30	.50	32.02	0.27	98.32	45.92	.50	.02	32.01	.13	0	.48
0.54	0.09	100.00	.50	49.88	0.56	14.53	22.35	.50	.16	50.37	.18	0	.64
3.36	0.00	99.21	.48	63.42	3.99	2.95	33.26	.50	1	74.08	.48	0	1
3.01	0.02	22.18	.50	70.82	0.41	0.00	96.27	.50	.45	57.40	.88	.19	1
0.31	0.01	62.66	.91	58.32	5.17	0.02	97.08	.50	0	74.14	.60	0	1
0.28	7.03	10.72	.50	33.15	0.26	0.00	39.45	.34	0	31.08	.02	0	.16
0.10	7.29	20.59	.50	13.79	0.10	13.00	6.68	.50	0	13.79	.34	.18	.47
0.66	0.03	6.44	.50	53.13	0.66	9.72	2.48	.50	.18	53.13	.21	0	.86
0.37	0.08	24.74	.52	44.71	0.42	0.96	1.57	.50	.47	46.12	.44	0	.81
1.10	4.58	18.07	.50	64.35	1.10	19.08	4.81	.50	0	64.35	.13	0	.72
0.58	0.10	29.96	.50	54.69	0.60	9.80	41.04	.50	.11	55.09	.13	0	.45
0.38	1.05	0.40	.50	45.26	0.38	21.39	20.09	.50	0	45.26	.06	0	.28
0.28	29.52	8.79	.50	39.99	0.28	1.61	31.43	.50	.04	39.97	.08	0	.32
0.29	6.05	11.14	.50	39.29	0.29	60.65	24.69	.50	0	39.29	.05	0	.23
0.36	0.00	70.84	.50	39.87	0.37	2.57	7.25	.50	0	40.14	.05	0	.26
0.23	22.89	3.86	.50	31.29	0.15	0.00	44.73	.93	.11	25.43	.01	0	.07
1.24	0.01	16.15	.50	62.83	1.24	0.08	0.01	.50	.48	62.83	.42	0	1
2.58	0.09	16.76	.50	73.15	2.58	0.09	16.76	.50	0	73.15	.40	0	1
1.38	0.09	100.00	.50	69.48	1.38	12.89	6.07	.50	.09	69.52	.18	0	.83
1.67	0.11	11.05	.50	71.68	1.67	5.19	0.62	.50	.17	71.72	.23	0	.92
0.32	0.00	52.24	.50	34.61	0.33	98.81	8.04	.50	.05	35.20	.07	0	.29
0.26	1.02	0.29	.50	30.62	0.26	14.01	19.08	.50	.03	30.53	.07	0	.27
0.84	0.49	5.98	.50	59.31	0.84	0.00	37.31	.50	0	59.27	.09	0	.53
0.29	0.10	4.34	.55	37.47	0.29	51.04	34.29	.50	.08	37.46	.11	0	.38
0.61	0.09	79.80	.50	51.46	0.63	6.56	4.23	.50	.04	51.82	.11	0	.49
0.20	21.79	2.86	.50	36.70	0.20	10.77	19.31	.50	0	36.70	.12	0	.34
0.70	4.32	15.68	.50	56.94	0.70	74.64	22.19	.50	0	56.94	.06	0	.29
0.47	0.08	100.00	.43	47.72	0.51	14.67	4.87	.50	.12	49.44	.13	0	.37
0.63	2.61	8.90	.50	53.86	0.63	6.94	8.33	.50	0	53.86	.06	0	.31
2.17	0.09	100.00	.50	68.69	2.18	11.49	1.19	.50	.05	68.73	.20	0	1
0.47	2.93	14.19	.50	45.17	0.47	25.75	0.53	.50	0	45.17	.05	0	.26

(continued)

Table G5

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS					SDbS+Range						w bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
0.98	0.24	0.01	.50	61.30	0.98	29.49	37.76	.50	0	61.30	.09	0	.49
0.10	0.01	29.89	.50	23.63	0.10	4.28	2.31	.50	.15	23.72	.39	.27	.47
0.80	5.37	15.23	.50	55.94	0.80	4.73	34.89	.50	0	55.94	.09	0	.44
0.69	14.06	4.30	.50	52.44	0.69	15.59	11.60	.50	0	52.44	.06	0	.33
0.42	1.70	4.49	.50	43.93	0.42	35.44	49.26	.50	.04	43.84	.06	0	.25
1.44	0.01	17.55	.50	63.88	1.44	0.01	28.67	.50	.28	63.88	.42	0	1
0.21	1.21	2.60	.50	46.03	0.10	0.09	15.62	.50	.47	42.02	.61	.60	.65
0.91	0.00	81.15	.46	62.04	1.08	27.56	0.60	.50	0	64.41	.08	0	.45
0.98	4.78	14.00	.50	60.06	0.98	16.96	1.34	.50	0	60.06	.09	0	.44
1.12	0.00	88.36	.51	59.69	1.11	2.68	49.22	.50	0	61.09	.09	0	.45
0.83	0.09	62.53	.50	59.81	0.87	11.61	3.91	.50	.17	60.53	.17	0	.57
0.32	2.37	1.08	.50	35.73	0.19	0.00	46.95	.60	0	18.11	.01	0	.04
1.08	7.09	0.49	.50	62.10	0.58	0.00	55.63	.76	.04	52.24	0	0	0
0.39	0.00	78.46	.54	40.43	0.41	13.76	4.80	.50	0	45.58	.04	0	.19
0.23	22.95	3.97	.50	37.11	0.23	12.76	16.07	.50	0	37.11	.06	0	.24
0.51	3.20	0.34	.50	50.80	0.32	0.00	88.88	.48	0	44.92	0	0	0
1.74	15.28	2.98	.50	64.17	1.50	0.01	90.30	.61	.43	62.58	.57	0	1
2.18	0.03	8.65	.58	68.26	2.18	71.50	31.30	.50	.66	68.27	.51	0	1
0.15	18.72	5.83	.50	45.18	0.10	0.00	53.42	.50	.40	44.90	.48	.45	.49
0.26	0.08	100.00	.44	41.10	0.22	0.00	71.93	.58	0	41.69	.00	0	.03
0.31	0.88	2.20	.50	37.54	0.31	78.69	15.40	.50	0	37.54	.05	0	.23
0.33	13.88	0.18	.50	39.37	0.26	0.03	6.51	.51	.26	35.60	.40	.16	.65
0.30	2.27	0.78	.50	34.15	0.27	0.00	57.27	.26	0	29.91	.03	0	.23
0.10	5.91	15.27	.50	39.20	0.10	0.00	82.39	.50	0	38.52	.13	0	.28
0.18	12.88	0.22	.50	31.28	0.19	5.92	46.87	.50	.08	31.14	.12	0	.29
0.79	15.07	3.51	.50	55.78	0.77	0.00	95.02	.49	0	50.00	0	0	0
3.65	0.01	17.62	.50	71.56	3.65	0.63	0.00	.50	.46	71.57	.39	0	1
0.52	3.22	0.43	.50	45.60	0.52	51.07	30.13	.50	0	45.60	.06	0	.31
0.36	1.67	0.00	.50	40.14	0.35	0.11	12.77	.51	0	38.98	.17	0	.45
2.65	0.02	17.32	.50	68.18	2.65	0.02	17.32	.50	0	68.18	.61	0	1
0.42	0.09	8.39	.50	44.15	0.42	20.48	5.26	.50	.31	44.38	.29	0	.66
0.43	0.09	76.11	.50	45.15	0.44	99.81	0.01	.50	.05	45.46	.10	0	.41

(continued)

Table G5

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
0.10	6.19	11.26	.50	35.18	0.10	12.46	6.48	.50	0	35.18	.30	.15	.46
0.19	12.28	9.73	.50	47.11	0.10	0.03	14.51	.43	0	42.23	.83	.68	1
0.23	0.00	70.18	.51	46.87	0.24	0.00	84.67	.49	0	46.77	0	0	.02
0.52	4.65	13.51	.50	51.43	0.52	85.23	0.00	.50	0	51.43	.12	0	.65
2.24	15.35	2.20	.50	71.76	2.24	8.36	5.79	.50	0	71.76	.22	0	1
0.26	2.81	6.05	.50	29.48	0.26	97.74	41.49	.50	0	29.48	.12	0	.46
0.80	0.42	6.13	.50	60.98	0.80	72.40	1.61	.50	0	60.98	.13	0	.64
2.14	0.08	10.13	.55	70.63	2.13	0.00	15.67	.49	.30	70.58	.34	0	1
0.40	0.02	69.54	.50	47.82	0.41	0.02	89.12	.67	.49	48.04	.78	.11	1
0.67	1.97	0.13	.50	55.57	0.67	0.75	55.01	.50	0	55.57	.09	0	.43
0.67	0.00	80.30	.51	58.01	0.71	31.30	39.50	.50	0	59.01	.16	0	.79
0.92	2.51	1.92	.50	62.74	0.92	15.58	11.86	.50	0	62.74	.11	0	.61
0.46	5.31	14.47	.50	46.08	0.46	56.61	95.72	.50	0	46.08	.09	0	.44
0.85	2.75	3.60	.50	60.28	0.85	39.48	0.12	.50	0	60.28	.15	0	.76
0.54	14.24	13.77	.50	50.55	0.54	41.54	52.75	.50	0	50.55	.06	0	.29
2.09	0.01	84.54	.50	73.94	3.36	16.24	23.02	.50	0	75.86	.29	0	1
0.78	2.35	0.25	.50	57.28	0.78	1.61	32.20	.50	0	57.28	.12	0	.59
0.62	0.09	40.92	.48	52.52	0.62	18.46	0.17	.50	0	52.52	.10	0	.51
2.25	0.00	99.33	.48	69.48	2.03	0.00	95.71	.51	0	69.39	.03	0	.32
0.30	2.08	0.16	.50	33.32	0.30	53.95	40.55	.50	0	33.32	.06	0	.29
0.35	1.43	0.03	.50	42.73	0.35	59.65	1.22	.50	0	42.73	.05	0	.28
0.65	2.73	0.26	.50	50.67	0.65	50.55	0.26	.50	.02	50.67	.08	0	.37
2.75	0.25	0.00	.50	72.68	2.75	66.52	24.30	.50	0	72.68	.24	0	1
0.63	3.19	0.26	.50	54.01	0.63	4.48	8.21	.50	0	54.01	.10	0	.45
0.78	2.20	0.23	.50	54.10	0.77	0.02	19.56	.49	0	53.93	.04	0	.33
0.10	14.21	5.76	.50	7.35	0.10	4.71	13.87	.50	0	7.35	.17	.02	.29
0.10	6.59	17.41	.50	9.26	0.10	16.11	33.64	.50	0	9.26	.18	.03	.29
0.31	2.57	8.88	.50	33.61	0.31	34.19	55.15	.50	0	33.61	.05	0	.25
0.31	0.01	29.09	.50	35.76	0.31	45.88	32.86	.50	0	37.07	.04	0	.23
0.33	1.38	4.29	.50	37.43	0.33	0.61	5.82	.50	.11	36.97	.12	0	.34
0.66	36.10	16.62	.50	53.64	0.66	18.88	3.12	.50	0	53.64	.08	0	.38
0.40	2.70	9.98	.50	48.03	0.40	43.42	22.45	.50	0	48.03	.06	0	.29

(continued)

Table G5

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

		SDbS					SDbS+Range				*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
1.49	4.71	13.07	.50	65.25	1.49	21.50	43.82	.50	0	65.25	.17	0	.87
0.43	15.17	4.07	.50	46.23	0.43	10.28	5.35	.50	0	46.23	.05	0	.24
0.99	4.63	15.38	.50	61.00	0.65	0.01	68.93	.48	.20	53.27	.01	0	.21
1.15	1.78	0.19	.50	63.60	1.02	0.01	31.42	.50	.01	62.09	.02	0	.28
1.14	12.01	3.51	.50	65.65	1.14	16.74	34.88	.50	0	65.65	.12	0	.63
0.16	7.28	13.31	.50	17.81	0.16	57.71	0.51	.50	0	17.81	.11	0	.26
3.57	0.39	0.00	.50	73.66	3.57	0.80	51.37	.50	0	73.66	.35	0	1
0.62	2.50	0.35	.50	52.32	0.62	36.15	2.51	.50	0	52.32	.08	0	.37
3.88	2.71	17.39	.50	72.78	0.18	0.01	59.59	.85	.14	58.06	.02	0	.12
0.24	5.25	12.04	.50	33.04	0.24	68.70	15.33	.50	0	33.04	.12	0	.39
1.56	2.67	17.02	.50	62.60	1.56	10.79	53.11	.50	0	62.60	.18	0	.90
0.55	0.02	36.57	.47	49.71	0.63	36.51	18.82	.50	0	51.51	.09	0	.45
0.98	0.15	0.01	.50	61.18	0.98	11.91	22.70	.50	0	61.18	.11	0	.57
0.88	2.57	0.12	.50	58.31	0.88	30.02	45.18	.50	0	58.31	.12	0	.64
7.87	10.08	0.57	.50	77.26	7.42	0.01	46.24	.50	0	77.18	.25	0	1
0.41	2.53	11.69	.50	48.89	0.41	1.12	31.73	.50	0	48.89	.05	0	.23
0.67	0.05	11.82	.56	52.07	0.60	36.86	91.36	.50	1	51.87	.91	.50	1
1.62	9.26	7.98	.50	64.20	1.62	14.87	4.64	.50	0	64.20	.21	0	1
0.83	0.02	30.30	.50	58.63	0.86	4.32	23.12	.50	.07	59.40	.18	0	.81
0.50	2.75	0.39	.50	47.36	0.50	10.62	1.78	.50	0	47.36	.06	0	.27
0.18	10.64	13.43	.50	29.06	0.18	27.66	10.14	.50	0	29.06	.09	0	.27
1.12	0.08	100.00	.55	58.15	1.02	0.08	100.00	.55	.38	58.14	.51	0	1
0.10	0.04	22.17	.43	32.81	0.10	0.04	22.17	.43	0	32.81	.60	.45	.68
3.06	14.40	1.75	.50	72.54	3.06	19.42	0.85	.50	0	72.54	.31	0	1
0.21	4.73	2.15	.50	45.57	0.10	1.12	9.18	.50	.93	42.37	.65	.61	.78
3.72	12.24	1.50	.50	72.61	0.10	0.00	75.45	.60	.12	53.19	.00	0	0
5.07	1.96	9.98	.50	74.49	5.07	45.77	0.46	.50	0	74.49	.34	0	1
0.31	13.16	5.92	.50	37.06	0.31	78.85	25.40	.50	0	37.06	.05	0	.23
5.95	3.82	5.41	.50	74.76	5.18	0.01	100.00	.50	0	74.21	.19	0	1
0.19	12.12	1.40	.50	43.01	0.10	14.98	16.18	.50	.56	40.84	.53	.48	.61
0.22	0.04	7.71	.63	43.25	0.22	0.45	1.30	.50	.36	42.59	.37	.18	.53
0.81	6.45	14.99	.50	55.96	0.47	0.01	75.83	.50	0	47.48	.01	0	.11

(continued)

Table G5

Best fitting model parameters of the SDbS and SDbS+ Range models to individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS					SDbS+Range						*w* bootstrap		
SD	*c*	*s*	*t*	-2lnL	SD	*c*	*s*	*t*	*w*	-2lnL	Mean	95% CI	
2.43	10.02	20.69	.50	66.59	2.43	21.89	27.01	.50	0	66.59	.26	0	1
0.36	0.04	8.24	.62	44.83	0.36	0.11	0.05	.50	.47	44.43	.47	.03	.88
6.32	12.83	2.16	.50	75.07	6.32	25.96	8.62	.50	0	75.07	.32	0	1
4.88	9.97	0.00	.50	73.66	4.88	18.17	14.55	.50	0	73.66	.32	0	1
1.45	3.63	12.74	.50	63.14	1.45	2.51	6.38	.50	0	63.14	.27	0	1
0.42	0.04	8.35	.45	52.92	0.35	24.61	4.83	.50	1	49.72	.92	.63	1
0.16	0.97	2.44	.50	50.07	0.18	54.47	99.20	.50	1	48.64	.87	.67	1

Appendix H:

SDbS and SDbS+Range Comparison

Table H1

Model comparison of SDbS and SDbS + Range using individual participant data from Brown et al. (2008). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
73.31	73.96	82.07	75.31	76.31	86.26	.73	.27	.89	.11
36.17	36.83	44.93	34.81	35.81	45.76	.34	.66	.60	.40
87.27	87.93	96.03	88.67	89.67	99.61	.67	.33	.86	.14
12.57	13.23	21.33	-14.64	-13.64	-3.69	0	1	0	1
-32.57	-31.91	-23.81	-45.47	-44.47	-34.52	0	1	0	1
78.23	78.89	86.99	78.71	79.71	89.66	.56	.44	.79	.21
-30.44	-29.78	-21.68	-28.44	-27.44	-17.49	.73	.27	.89	.11
102.40	103.06	111.16	104.40	105.40	115.35	.73	.27	.89	.11
116.58	117.23	125.33	113.24	114.24	124.18	.16	.84	.36	.64
101.31	101.97	110.07	102.12	103.12	113.07	.60	.40	.82	.18
85.83	86.49	94.59	84.21	85.21	95.16	.31	.69	.57	.43
81.71	82.36	90.47	83.18	84.18	94.13	.68	.32	.86	.14
85.87	86.52	94.63	84.37	85.37	95.32	.32	.68	.59	.41
18.06	18.72	26.82	20.00	21.00	30.95	.73	.27	.89	.11
-26.74	-26.08	-17.98	-29.50	-28.50	-18.55	.20	.80	.43	.57
96.98	97.63	105.74	100.30	101.30	111.25	.84	.16	.94	.06
-37.82	-37.16	-29.06	-40.23	-39.23	-29.28	.23	.77	.47	.53
59.57	60.23	68.33	60.61	61.61	71.55	.63	.37	.83	.17
104.44	105.10	113.20	106.51	107.51	117.46	.74	.26	.89	.11
5.30	5.96	14.06	5.88	6.88	16.83	.57	.43	.80	.20
68.68	69.34	77.44	64.71	65.71	75.66	.12	.88	.29	.71
51.72	52.37	60.47	42.29	43.29	53.24	.01	.99	.03	.97
-5.07	-4.42	3.69	-14.69	-13.69	-3.74	.01	.99	.02	.98
23.34	24.00	32.10	25.59	26.59	36.54	.75	.25	.90	.10

Table H2

Model comparison of SDbS and SDbS + Range using individual participant data from Melrose et al. (2012). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
350.96	359.62	376.37	353.83	364.83	385.78	.81	.19	.99	.01
453.59	462.24	479.00	454.69	465.69	486.64	.63	.37	.98	.02
562.93	571.58	588.34	564.50	575.50	596.45	.69	.31	.98	.02
377.55	386.21	402.96	379.55	390.55	411.50	.73	.27	.99	.01
546.35	555.01	571.77	538.92	549.92	570.87	.02	.98	.39	.61
541.26	549.92	566.67	543.26	554.26	575.21	.73	.27	.99	.01
588.66	597.31	614.07	590.65	601.65	622.60	.73	.27	.99	.01
548.87	557.53	574.29	550.70	561.70	582.64	.71	.29	.98	.02
496.18	504.84	521.60	498.62	509.62	530.57	.77	.23	.99	.01
543.61	552.26	569.02	545.24	556.24	577.19	.69	.31	.98	.02
428.35	437.01	453.77	430.39	441.39	462.34	.74	.26	.99	.01
575.55	584.21	600.97	577.53	588.53	609.48	.73	.27	.99	.01
495.01	503.67	520.43	496.78	507.78	528.73	.71	.29	.98	.02
484.86	493.51	510.27	486.83	497.83	518.78	.73	.27	.99	.01
477.50	486.15	502.91	476.86	487.86	508.81	.42	.58	.95	.05
455.44	464.10	480.86	457.42	468.42	489.36	.73	.27	.99	.01
479.42	488.08	504.84	481.42	492.42	513.37	.73	.27	.99	.01
454.73	463.38	480.14	456.31	467.31	488.26	.69	.31	.98	.02
391.96	400.62	417.38	395.33	406.33	427.28	.84	.16	.99	.01
417.45	426.11	442.87	419.45	430.45	451.40	.73	.27	.99	.01
466.81	475.46	492.22	468.54	479.54	500.49	.70	.30	.98	.02
545.41	554.07	570.83	547.30	558.30	579.24	.72	.28	.99	.01
528.67	537.33	554.09	530.64	541.64	562.59	.73	.27	.99	.01
422.27	430.92	447.68	424.29	435.29	456.24	.73	.27	.99	.01
484.83	493.49	510.25	486.73	497.73	518.68	.72	.28	.99	.01
470.24	478.90	495.66	472.21	483.21	504.16	.73	.27	.99	.01
351.52	360.17	376.93	351.95	362.95	383.90	.55	.45	.97	.03
431.02	439.67	456.43	432.78	443.78	464.73	.71	.29	.98	.02
345.62	354.27	371.03	343.18	354.18	375.13	.23	.77	.89	.11
463.49	472.15	488.91	465.07	476.07	497.02	.69	.31	.98	.02
380.00	388.66	405.42	381.98	392.98	413.93	.73	.27	.99	.01
429.23	437.88	454.64	431.23	442.23	463.18	.73	.27	.99	.01

(continued)

Table H2

Model comparison of SDbS and SDbS + Range using individual participant data from Melrose et al. (2012) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
416.02	424.67	441.43	416.93	427.93	448.88	.61	.39	.98	.02
333.44	342.09	358.85	335.14	346.14	367.09	.70	.30	.98	.02
305.98	314.63	331.39	307.89	318.89	339.84	.72	.28	.99	.01
511.38	520.04	536.80	513.30	524.30	545.24	.72	.28	.99	.01
343.10	351.76	368.52	345.10	356.10	377.05	.73	.27	.99	.01
401.33	409.98	426.74	402.81	413.81	434.76	.68	.32	.98	.02
434.08	442.73	459.49	444.21	455.21	476.16	.99	.01	1	0
468.94	477.59	494.35	472.35	483.35	504.30	.85	.15	.99	.01
419.44	428.10	444.85	434.17	445.17	466.12	1	0	1	0
493.02	501.68	518.44	512.92	523.92	544.87	1	0	1	0
543.03	551.69	568.44	545.03	556.03	576.98	.73	.27	.99	.01
271.95	280.60	297.36	277.53	288.53	309.48	.94	.06	1	0
309.04	317.70	334.45	310.42	321.42	342.37	.67	.33	.98	.02
491.52	500.17	516.93	493.28	504.28	525.23	.71	.29	.98	.02
525.42	534.08	550.84	526.37	537.37	558.32	.62	.38	.98	.02
467.22	475.87	492.63	468.28	479.28	500.23	.63	.37	.98	.02
259.48	268.14	284.90	260.91	271.91	292.86	.67	.33	.98	.02
390.53	399.19	415.95	388.65	399.65	420.60	.28	.72	.91	.09
418.29	426.95	443.71	418.88	429.88	450.82	.57	.43	.97	.03
454.81	463.46	480.22	464.59	475.59	496.54	.99	.01	1	0

Table H3

Model comparison of SDbS and SDbS + Range using individual participant data from Wood et al. (2011a). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
324.45	325.47	331.58	325.79	327.37	334.71	.66	.34	.83	.17
537.16	538.19	544.30	539.17	540.75	548.09	.73	.27	.87	.13
716.84	717.86	723.97	718.84	720.42	727.76	.73	.27	.87	.13
482.07	483.09	489.21	483.48	485.05	492.40	.67	.33	.83	.17
579.48	580.50	586.61	581.49	583.07	590.41	.73	.27	.87	.13
512.61	513.64	519.75	514.61	516.19	523.53	.73	.27	.87	.13
642.33	643.36	649.47	643.95	645.53	652.87	.69	.31	.85	.15
607.77	608.79	614.91	609.77	611.35	618.69	.73	.27	.87	.13
502.44	503.46	509.57	503.17	504.75	512.09	.59	.41	.78	.22
727.74	728.77	734.88	729.49	731.07	738.42	.71	.29	.85	.15
531.76	532.79	538.90	533.74	535.32	542.66	.73	.27	.87	.13
657.79	658.81	664.92	659.50	661.08	668.42	.70	.30	.85	.15
640.64	641.67	647.78	642.00	643.58	650.92	.66	.34	.83	.17
449.30	450.33	456.44	450.55	452.13	459.47	.65	.35	.82	.18
711.12	712.14	718.26	713.09	714.67	722.02	.73	.27	.87	.13
682.09	683.11	689.22	684.04	685.62	692.96	.73	.27	.87	.13
651.58	652.61	658.72	653.58	655.16	662.50	.73	.27	.87	.13
743.69	744.72	750.83	745.54	747.12	754.46	.72	.28	.86	.14
586.08	587.11	593.22	587.23	588.81	596.15	.64	.36	.81	.19
588.44	589.46	595.57	590.38	591.96	599.30	.73	.27	.87	.13
484.65	485.67	491.79	485.37	486.95	494.29	.59	.41	.78	.22
508.23	509.25	515.36	510.17	511.75	519.09	.73	.27	.87	.13
556.59	557.62	563.73	557.60	559.18	566.52	.62	.38	.80	.20
617.72	618.74	624.86	618.95	620.53	627.88	.65	.35	.82	.18
564.73	565.76	571.87	566.02	567.60	574.94	.66	.34	.82	.18
540.65	541.68	547.79	542.11	543.69	551.04	.68	.32	.84	.16
601.04	602.06	608.17	602.82	604.40	611.74	.71	.29	.86	.14
596.55	597.58	603.69	598.44	600.02	607.36	.72	.28	.86	.14
744.12	745.14	751.25	746.09	747.67	755.01	.73	.27	.87	.13
559.08	560.11	566.22	561.07	562.65	569.99	.73	.27	.87	.13
614.40	615.42	621.53	616.30	617.88	625.22	.72	.28	.86	.14
527.48	528.51	534.62	527.98	529.56	536.90	.56	.44	.76	.24

(continued)

Table H3

Model comparison of SDbS and SDbS + Range using individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
552.28	553.31	559.42	550.64	552.22	559.56	.31	.69	.52	.48
601.76	602.79	608.90	602.64	604.22	611.56	.61	.39	.79	.21
538.54	539.56	545.67	539.34	540.92	548.26	.60	.40	.78	.22
750.90	751.92	758.04	752.89	754.46	761.81	.73	.27	.87	.13
511.49	512.51	518.62	511.27	512.85	520.19	.47	.53	.69	.31
660.42	661.45	667.56	667.75	669.33	676.67	.98	.02	.99	.01
677.40	678.43	684.54	679.39	680.97	688.31	.73	.27	.87	.13
286.47	287.49	293.60	287.91	289.49	296.83	.67	.33	.83	.17
702.89	703.92	710.03	705.42	707.00	714.34	.78	.22	.90	.10
643.45	644.48	650.59	644.22	645.80	653.14	.59	.41	.78	.22
609.18	610.21	616.32	611.17	612.75	620.09	.73	.27	.87	.13
628.41	629.43	635.55	630.41	631.99	639.33	.73	.27	.87	.13
375.12	376.14	382.25	377.12	378.69	386.04	.73	.27	.87	.13
601.26	602.28	608.39	603.01	604.59	611.93	.71	.29	.85	.15
634.99	636.01	642.13	636.99	638.57	645.91	.73	.27	.87	.13
706.41	707.43	713.54	708.43	710.01	717.35	.73	.27	.87	.13
572.63	573.66	579.77	574.63	576.21	583.55	.73	.27	.87	.13
-235.52	-234.50	-228.39	-233.52	-231.94	-224.60	.73	.27	.87	.13
840.68	841.71	847.82	842.66	844.24	851.58	.73	.27	.87	.13
456.46	457.49	463.60	458.46	460.04	467.38	.73	.27	.87	.13
660.03	661.05	667.16	661.90	663.48	670.82	.72	.28	.86	.14
605.71	606.73	612.85	607.71	609.29	616.63	.73	.27	.87	.13
657.37	658.39	664.50	659.37	660.95	668.29	.73	.27	.87	.13
603.88	604.91	611.02	605.88	607.46	614.81	.73	.27	.87	.13
269.00	270.03	276.14	270.98	272.56	279.90	.73	.27	.87	.13
537.89	538.92	545.03	539.89	541.47	548.81	.73	.27	.87	.13
578.55	579.57	585.68	580.46	582.04	589.38	.72	.28	.86	.14
557.07	558.09	564.20	558.94	560.52	567.86	.72	.28	.86	.14
712.22	713.25	719.36	714.23	715.81	723.15	.73	.27	.87	.13
813.11	814.14	820.25	815.09	816.67	824.02	.73	.27	.87	.13
724.46	725.49	731.60	726.46	728.04	735.38	.73	.27	.87	.13
687.48	688.50	694.61	686.17	687.75	695.09	.34	.66	.56	.44

(continued)

Table H3

Model comparison of SDbS and SDbS + Range using individual participant data from Wood et al. (2011a) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
750.35	751.38	757.49	752.42	754.00	761.34	.74	.26	.87	.13
683.10	684.13	690.24	685.10	686.68	694.02	.73	.27	.87	.13
572.16	573.19	579.30	574.16	575.74	583.08	.73	.27	.87	.13
628.62	629.65	635.76	630.62	632.20	639.54	.73	.27	.87	.13
584.81	585.83	591.94	586.81	588.38	595.73	.73	.27	.87	.13
673.40	674.42	680.53	675.40	676.98	684.32	.73	.27	.87	.13
594.11	595.13	601.24	596.10	597.68	605.03	.73	.27	.87	.13
608.86	609.89	616.00	610.73	612.30	619.65	.72	.28	.86	.14
731.62	732.64	738.75	733.61	735.19	742.54	.73	.27	.87	.13
720.23	721.26	727.37	722.16	723.74	731.09	.72	.28	.87	.13
580.53	581.56	587.67	583.03	584.61	591.95	.78	.22	.89	.11
663.97	664.99	671.10	665.97	667.55	674.89	.73	.27	.87	.13
626.74	627.76	633.87	628.75	630.33	637.67	.73	.27	.87	.13
643.15	644.17	650.29	645.01	646.59	653.93	.72	.28	.86	.14
575.30	576.33	582.44	575.91	577.49	584.83	.57	.43	.77	.23
743.21	744.24	750.35	745.21	746.79	754.13	.73	.27	.87	.13
613.12	614.15	620.26	615.12	616.70	624.04	.73	.27	.87	.13

Table H4

Model comparison of SDbS and SDbS + Range using individual participant data from Wood et al. (2011b). Note: each row is the data for one participant.

	SDbS			SDbS+Range		AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
29.38	39.74	52.10	31.48	45.23	60.69	.74	.26	.99	.01
26.06	36.41	48.77	28.05	41.80	57.26	.73	.27	.99	.01
23.15	33.50	45.87	26.67	40.42	55.88	.85	.15	.99	.01
30.39	40.74	53.10	31.75	45.50	60.96	.66	.34	.98	.02
27.83	38.18	50.55	29.97	43.72	59.17	.74	.26	.99	.01
16.75	27.10	39.47	18.84	32.59	48.04	.74	.26	.99	.01
28.15	38.50	50.86	30.44	44.19	59.65	.76	.24	.99	.01
23.54	33.90	46.26	25.54	39.29	54.74	.73	.27	.99	.01
45.42	55.78	68.14	47.46	61.21	76.67	.73	.27	.99	.01
17.50	27.86	40.22	19.50	33.25	48.71	.73	.27	.99	.01
34.19	44.54	56.91	36.56	50.31	65.77	.77	.23	.99	.01
25.16	35.52	47.88	27.31	41.06	56.52	.75	.25	.99	.01
21.75	32.10	44.46	23.76	37.51	52.96	.73	.27	.99	.01
35.20	45.55	57.92	37.20	50.95	66.40	.73	.27	.99	.01
22.42	32.77	45.14	30.76	44.51	59.96	.98	.02	1	0
22.84	33.20	45.56	24.88	38.63	54.09	.74	.26	.99	.01
23.31	33.66	46.03	35.85	49.60	65.05	1	0	1	0
49.76	60.11	72.47	51.99	65.74	81.19	.75	.25	.99	.01
19.79	30.14	42.51	21.81	35.56	51.02	.73	.27	.99	.01
13.27	23.63	35.99	15.28	29.03	44.49	.73	.27	.99	.01
17.42	27.78	40.14	19.43	33.18	48.63	.73	.27	.99	.01
34.38	44.73	57.09	32.45	46.20	61.65	.28	.72	.91	.09
40.06	50.41	62.78	42.06	55.81	71.26	.73	.27	.99	.01
33.57	43.92	56.28	35.15	48.90	64.36	.69	.31	.98	.02
19.35	29.70	42.06	21.35	35.10	50.55	.73	.27	.99	.01
26.27	36.62	48.99	28.63	42.38	57.83	.76	.24	.99	.01
19.24	29.59	41.96	21.24	34.99	50.45	.73	.27	.99	.01
25.16	35.52	47.88	27.16	40.91	56.36	.73	.27	.99	.01
23.30	33.65	46.02	25.31	39.06	54.52	.73	.27	.99	.01
39.34	49.69	62.06	41.34	55.09	70.55	.73	.27	.99	.01
14.37	24.72	37.08	16.38	30.13	45.59	.73	.27	.99	.01
10.93	21.28	33.64	12.94	26.69	42.14	.73	.27	.99	.01

(continued)

Table H4

Model comparison of SDbS and SDbS + Range using individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
52.06	62.42	74.78	54.06	67.81	83.26	.73	.27	.99	.01
15.25	25.60	37.97	17.19	30.94	46.40	.73	.27	.99	.01
34.40	44.76	57.12	36.78	50.53	65.98	.77	.23	.99	.01
29.07	39.42	51.78	31.07	44.82	60.27	.73	.27	.99	.01
28.33	38.69	51.05	30.31	44.06	59.52	.73	.27	.99	.01
26.42	36.77	49.14	28.42	42.17	57.62	.73	.27	.99	.01
32.76	43.11	55.48	34.76	48.51	63.97	.73	.27	.99	.01
33.08	43.44	55.80	35.08	48.83	64.29	.73	.27	.99	.01
36.21	46.56	58.93	30.44	44.19	59.65	.05	.95	.59	.41
23.23	33.58	45.95	42.00	55.75	71.21	1	0	1	0
24.12	34.48	46.84	26.12	39.87	55.33	.73	.27	.99	.01
14.33	24.68	37.05	16.86	30.61	46.07	.78	.22	.99	.01
28.63	38.99	51.35	31.94	45.69	61.14	.84	.16	.99	.01
47.35	57.70	70.07	49.33	63.08	78.54	.73	.27	.99	.01
38.98	49.33	61.70	40.98	54.73	70.19	.73	.27	.99	.01
26.01	36.36	48.73	36.30	50.05	65.50	.99	.01	1	0
23.26	33.61	45.97	22.19	35.94	51.40	.37	.63	.94	.06
28.00	38.35	50.72	40.12	53.87	69.32	1	0	1	0
26.00	36.36	48.72	30.06	43.81	59.27	.88	.12	.99	.01
16.00	26.35	38.72	18.00	31.75	47.21	.73	.27	.99	.01
37.95	48.30	60.67	39.99	53.74	69.19	.73	.27	.99	.01
25.88	36.23	48.60	26.27	40.02	55.47	.55	.45	.97	.03
19.81	30.17	42.53	24.32	38.07	53.52	.90	.10	1	.00
25.96	36.32	48.68	27.96	41.71	57.17	.73	.27	.99	.01
16.67	27.03	39.39	18.38	32.13	47.58	.70	.30	.98	.02
15.07	25.42	37.78	17.07	30.82	46.27	.73	.27	.99	.01
27.43	37.78	50.14	34.24	47.99	63.44	.97	.03	1	0
17.32	27.67	40.03	18.82	32.57	48.02	.68	.32	.98	.02
23.21	33.57	45.93	31.30	45.05	60.51	.98	.02	1	0
16.63	26.99	39.35	18.98	32.73	48.18	.76	.24	.99	.01
23.22	33.58	45.94	25.66	39.41	54.86	.77	.23	.99	.01
28.39	38.75	51.11	30.39	44.14	59.60	.73	.27	.99	.01

(continued)

Table H4

Model comparison of SDbS and SDbS + Range using individual participant data from Wood et al. (2011b) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
12.58	22.93	35.29	14.58	28.33	43.78	.73	.27	.99	.01
19.97	30.32	42.69	21.96	35.71	51.17	.73	.27	.99	.01
39.38	49.73	62.10	41.40	55.15	70.61	.73	.27	.99	.01
26.77	37.13	49.49	27.97	41.72	57.18	.65	.35	.98	.02
30.19	40.55	52.91	32.18	45.93	61.39	.73	.27	.99	.01
30.00	40.36	52.72	30.77	44.52	59.97	.59	.41	.97	.03
41.43	51.78	64.14	42.84	56.59	72.04	.67	.33	.98	.02
20.96	31.32	43.68	23.17	36.92	52.38	.75	.25	.99	.01
23.71	34.06	46.42	25.38	39.13	54.58	.70	.30	.98	.02
30.52	40.87	53.24	31.95	45.70	61.15	.67	.33	.98	.02
17.63	27.99	40.35	18.78	32.53	47.99	.64	.36	.98	.02
39.92	50.27	62.64	40.01	53.76	69.21	.51	.49	.96	.04
18.56	28.91	41.28	19.89	33.64	49.09	.66	.34	.98	.02
19.82	30.17	42.54	21.72	35.47	50.92	.72	.28	.99	.01

Table H5

Model comparison of SDbS and SDbS + Range using individual participant data from Maltby et al. (2012). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
20.39	29.82	43.81	22.39	34.62	52.10	.73	.27	.98	.02
40.02	49.45	63.43	42.01	54.24	71.72	.73	.27	.98	.02
57.88	67.30	81.29	60.37	72.59	90.08	.78	.22	.99	.01
71.42	80.85	94.83	84.08	96.31	113.79	1	0	1	0
78.82	88.25	102.24	67.40	79.63	97.11	0	1	.07	.93
66.32	75.75	89.73	84.14	96.36	113.84	1	0	1	0
41.15	50.58	64.57	41.08	53.31	70.79	.49	.51	.96	.04
21.79	31.22	45.21	23.79	36.01	53.50	.73	.27	.98	.02
61.13	70.56	84.54	63.13	75.35	92.84	.73	.27	.98	.02
52.71	62.14	76.12	56.12	68.34	85.82	.85	.15	.99	.01
72.35	81.78	95.77	74.35	86.58	104.06	.73	.27	.98	.02
62.69	72.12	86.10	65.09	77.32	94.80	.77	.23	.99	.01
53.26	62.69	76.68	55.26	67.49	84.97	.73	.27	.98	.02
47.99	57.42	71.41	49.97	62.19	79.67	.73	.27	.98	.02
47.29	56.72	70.71	49.29	61.51	79.00	.73	.27	.98	.02
47.87	57.30	71.28	50.14	62.37	79.85	.76	.24	.99	.01
39.29	48.72	62.71	35.43	47.65	65.13	.13	.87	.77	.23
70.83	80.26	94.24	72.83	85.05	102.53	.73	.27	.98	.02
81.15	90.57	104.56	83.15	95.37	112.85	.73	.27	.98	.02
77.48	86.91	100.89	79.52	91.74	109.22	.73	.27	.98	.02
79.68	89.11	103.10	81.72	93.94	111.43	.74	.26	.98	.02
42.61	52.04	66.03	45.20	57.43	74.91	.78	.22	.99	.01
38.62	48.05	62.04	40.53	52.75	70.24	.72	.28	.98	.02
67.31	76.74	90.72	69.27	81.49	98.97	.73	.27	.98	.02
45.47	54.90	68.88	47.46	59.68	77.17	.73	.27	.98	.02
59.46	68.89	82.88	61.82	74.04	91.52	.76	.24	.99	.01
44.70	54.13	68.11	46.70	58.92	76.40	.73	.27	.98	.02
64.94	74.37	88.35	66.94	79.16	96.64	.73	.27	.98	.02
55.72	65.15	79.13	59.44	71.66	89.14	.87	.13	.99	.01
61.86	71.29	85.28	63.86	76.08	93.57	.73	.27	.98	.02
76.69	86.12	100.11	78.73	90.95	108.44	.74	.26	.98	.02
53.17	62.60	76.59	55.17	67.39	84.88	.73	.27	.98	.02

(continued)

Table H5

Model comparison of SDbS and SDbS + Range using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
69.30	78.73	92.71	71.30	83.52	101.00	.73	.27	.98	.02
31.63	41.06	55.05	33.72	45.94	63.43	.74	.26	.99	.01
63.94	73.37	87.36	65.94	78.16	95.65	.73	.27	.98	.02
60.44	69.87	83.86	62.44	74.67	92.15	.73	.27	.98	.02
51.93	61.36	75.34	53.84	66.07	83.55	.72	.28	.98	.02
71.88	81.30	95.29	73.88	86.10	103.59	.73	.27	.98	.02
54.03	63.46	77.44	52.02	64.24	81.72	.27	.73	.89	.11
70.04	79.47	93.45	74.41	86.63	104.11	.90	.10	1	0
68.06	77.49	91.47	70.06	82.28	99.76	.73	.27	.98	.02
67.69	77.11	91.10	71.09	83.31	100.80	.85	.15	.99	.01
67.81	77.24	91.22	70.53	82.75	100.23	.80	.20	.99	.01
43.73	53.16	67.14	28.11	40.33	57.82	0	1	.01	.99
70.10	79.53	93.52	62.24	74.46	91.95	.02	.98	.31	.69
48.43	57.86	71.85	55.58	67.80	85.28	.97	.03	1	0
45.11	54.54	68.52	47.11	59.33	76.81	.73	.27	.98	.02
58.80	68.23	82.22	54.92	67.14	84.63	.13	.87	.77	.23
72.17	81.60	95.58	72.58	84.80	102.28	.55	.45	.97	.03
76.26	85.69	99.67	78.27	90.49	107.98	.73	.27	.98	.02
53.18	62.61	76.60	54.90	67.12	84.60	.70	.30	.98	.02
49.10	58.53	72.51	51.69	63.91	81.40	.79	.21	.99	.01
45.54	54.97	68.95	47.54	59.76	77.24	.73	.27	.98	.02
47.37	56.80	70.79	45.60	57.82	75.31	.29	.71	.91	.09
42.15	51.58	65.56	39.91	52.13	69.61	.25	.75	.88	.12
47.20	56.62	70.61	48.52	60.74	78.22	.66	.34	.98	.02
39.28	48.70	62.69	41.14	53.36	70.84	.72	.28	.98	.02
63.78	73.21	87.20	60.00	72.22	89.71	.13	.87	.78	.22
79.56	88.98	102.97	81.57	93.80	111.28	.73	.27	.98	.02
53.60	63.02	77.01	55.60	67.82	85.30	.73	.27	.98	.02
48.14	57.57	71.55	48.98	61.20	78.69	.60	.40	.97	.03
76.18	85.61	99.59	78.18	90.40	107.88	.73	.27	.98	.02
52.15	61.58	75.57	54.38	66.60	84.09	.75	.25	.99	.01
53.15	62.58	76.57	55.46	67.68	85.16	.76	.24	.99	.01

(continued)

Table H5

Model comparison of SDbS and SDbS + Range using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
43.18	52.61	66.60	45.18	57.40	74.89	.73	.27	.98	.02
55.11	64.54	78.53	52.23	64.45	81.93	.19	.81	.85	.15
54.87	64.30	78.29	56.77	69.00	86.48	.72	.28	.98	.02
59.43	68.85	82.84	61.43	73.65	91.13	.73	.27	.98	.02
79.76	89.18	103.17	81.76	93.98	111.46	.73	.27	.98	.02
37.48	46.91	60.89	39.48	51.70	69.18	.73	.27	.98	.02
68.98	78.41	92.39	70.98	83.20	100.68	.73	.27	.98	.02
78.63	88.06	102.05	80.58	92.80	110.29	.73	.27	.98	.02
55.82	65.25	79.24	58.04	70.26	87.74	.75	.25	.99	.01
63.57	73.00	86.99	65.57	77.79	95.28	.73	.27	.98	.02
66.01	75.43	89.42	69.01	81.23	98.71	.82	.18	.99	.01
70.74	80.17	94.16	72.74	84.96	102.45	.73	.27	.98	.02
54.08	63.51	77.50	56.08	68.31	85.79	.73	.27	.98	.02
68.28	77.71	91.69	70.28	82.50	99.98	.73	.27	.98	.02
58.55	67.98	81.96	60.55	72.77	90.25	.73	.27	.98	.02
81.94	91.37	105.36	85.86	98.08	115.57	.88	.12	.99	.01
65.28	74.71	88.70	67.28	79.51	96.99	.73	.27	.98	.02
60.52	69.95	83.93	62.52	74.75	92.23	.73	.27	.98	.02
77.48	86.91	100.89	79.39	91.62	109.10	.72	.28	.98	.02
41.32	50.75	64.74	43.32	55.54	73.03	.73	.27	.98	.02
50.73	60.16	74.14	52.73	64.95	82.43	.73	.27	.98	.02
58.67	68.10	82.08	60.67	72.89	90.37	.73	.27	.98	.02
80.68	90.11	104.09	82.68	94.90	112.38	.73	.27	.98	.02
62.01	71.44	85.43	64.01	76.23	93.72	.73	.27	.98	.02
62.10	71.53	85.51	63.93	76.15	93.63	.71	.29	.98	.02
15.35	24.78	38.77	17.35	29.57	47.06	.73	.27	.98	.02
17.26	26.69	40.67	19.26	31.48	48.96	.73	.27	.98	.02
41.61	51.04	65.03	43.61	55.84	73.32	.73	.27	.98	.02
43.76	53.18	67.17	47.07	59.29	76.77	.84	.16	.99	.01
45.43	54.86	68.84	46.97	59.20	76.68	.68	.32	.98	.02
61.64	71.07	85.05	63.64	75.86	93.34	.73	.27	.98	.02
56.03	65.46	79.45	58.03	70.25	87.74	.73	.27	.98	.02

(continued)

Table H5

Model comparison of SDbS and SDbS + Range using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
73.25	82.67	96.66	75.25	87.47	104.95	.73	.27	.98	.02
54.23	63.66	77.64	56.23	68.45	85.93	.73	.27	.98	.02
69.00	78.43	92.42	63.27	75.49	92.97	.05	.95	.57	.43
71.60	81.03	95.01	72.09	84.31	101.79	.56	.44	.97	.03
73.65	83.07	97.06	75.65	87.87	105.35	.73	.27	.98	.02
25.81	35.23	49.22	27.81	40.03	57.51	.73	.27	.98	.02
81.66	91.09	105.07	83.66	95.88	113.36	.73	.27	.98	.02
60.32	69.74	83.73	62.32	74.54	92.02	.73	.27	.98	.02
80.78	90.21	104.19	68.06	80.28	97.76	0	1	.04	.96
41.04	50.47	64.45	43.04	55.26	72.74	.73	.27	.98	.02
70.60	80.03	94.01	72.60	84.82	102.30	.73	.27	.98	.02
57.71	67.14	81.12	61.51	73.74	91.22	.87	.13	.99	.01
69.18	78.61	92.60	71.18	83.41	100.89	.73	.27	.98	.02
66.31	75.74	89.73	68.31	80.53	98.02	.73	.27	.98	.02
85.26	94.69	108.68	87.18	99.40	116.89	.72	.28	.98	.02
56.89	66.32	80.30	58.89	71.11	88.59	.73	.27	.98	.02
60.07	69.50	83.48	61.87	74.09	91.57	.71	.29	.98	.02
72.20	81.63	95.62	74.20	86.43	103.91	.73	.27	.98	.02
66.63	76.06	90.04	69.40	81.62	99.10	.80	.20	.99	.01
55.36	64.79	78.78	57.36	69.59	87.07	.73	.27	.98	.02
37.06	46.49	60.48	39.06	51.29	68.77	.73	.27	.98	.02
66.15	75.57	89.56	68.14	80.36	97.85	.73	.27	.98	.02
40.81	50.24	64.23	42.81	55.03	72.52	.73	.27	.98	.02
80.54	89.97	103.95	82.54	94.76	112.24	.73	.27	.98	.02
53.57	63.00	76.98	52.37	64.59	82.08	.35	.65	.93	.07
80.61	90.04	104.03	63.19	75.41	92.89	0	1	0	1
82.49	91.92	105.91	84.49	96.71	114.20	.73	.27	.98	.02
45.06	54.49	68.47	47.06	59.28	76.76	.73	.27	.98	.02
82.76	92.19	106.18	84.21	96.43	113.91	.67	.33	.98	.02
51.01	60.44	74.43	50.84	63.06	80.54	.48	.52	.96	.04
51.25	60.68	74.66	52.59	64.81	82.29	.66	.34	.98	.02
63.96	73.39	87.37	57.48	69.70	87.18	.04	.96	.48	.52

(continued)

Table H5

Model comparison of SDbS and SDbS + Range using individual participant data from Maltby et al. (2012) (continued). Note: each row is the data for one participant.

SDbS			SDbS+Range			AICw		BICw	
AIC	AICc	BIC	AIC	AICc	BIC	SDbS	+Range	SDbS	+Range
74.59	84.02	98.01	76.59	88.81	106.30	.73	.27	.98	.02
52.83	62.25	76.24	54.43	66.65	84.14	.69	.31	.98	.02
83.07	92.50	106.48	85.07	97.29	114.77	.73	.27	.98	.02
81.66	91.09	105.08	83.66	95.89	113.37	.73	.27	.98	.02
71.14	80.57	94.56	73.14	85.37	102.85	.73	.27	.98	.02
60.92	70.34	84.33	59.72	71.95	89.43	.36	.64	.93	.07
58.07	67.50	81.49	58.64	70.86	88.34	.57	.43	.97	.03

Appendix I:

SIMPLE Variations Fit

Table I1

Parameter fits and -2lnL for individual participant data from Experiment 4. Note: each row is the data for one participant. Experimental conditions are: none (N), same (S), increasing (I), decreasing (D), middle high (MH).

	Time-based Forgetting				Output Interference					Both					
Pay	*c*	*s*	*t*	-2lnL	*c*	*s*	*t*	*o*	-2lnL	*c*	*s*	*t*	*o*	*w*	-2lnL
N	13.44	6.32	.81	57.69	42.31	6.24	.74	.01	60.24	8.24	12.68	.87	1	.82	54.34
N	0.75	100.00	.11	88.64	84.39	99.98	1	.55	85.39	0.20	43.55	.18	1	.02	68.28
N	10.38	5.17	.69	58.99	0.14	53.47	.14	.93	56.23	0.62	25.82	.20	.92	.63	56.60
N	15.61	5.11	.86	56.49	6.99	4.94	.86	.36	58.05	10.08	6.86	.83	1	.90	54.88
N	8.91	5.36	.75	68.96	0.31	24.85	.25	1	54.51	0.31	24.84	.25	1	0	54.51
N	1.03	100.00	.11	69.94	100.00	4.15	.87	.34	88.51	3.60	33.63	.91	.97	.03	78.02
N	13.62	10.18	.61	67.44	3.67	10.44	.61	.66	58.96	15.37	9.98	.68	.75	.85	58.29
N	0.00	4.18	.71	70.09	3.59	5.95	.70	.86	61.88	47.64	6.19	.82	.64	.03	61.59
N	14.72	6.81	.74	59.46	4.21	7.20	.79	.70	57.25	18.29	6.70	.77	.83	.96	56.27
N	19.59	4.19	.76	52.64	1.24	9.92	.35	.70	52.38	15.35	6.63	.48	.74	1	51.84
N	16.03	4.93	.75	71.79	98.80	89.40	.99	.64	66.10	0.77	18.76	.28	.93	.35	60.51
N	8.95	7.43	.56	66.36	100.00	6.17	.78	.28	61.51	10.85	6.25	.69	.56	.55	62.02
N	8.23	5.88	.68	64.01	100.00	3.94	1	.12	65.64	0.63	13.89	.37	.96	.09	54.85
N	0.71	97.57	.11	72.21	0.26	29.74	.22	1	57.87	1.09	18.51	.26	0	.70	73.41
N	0.00	4.37	.82	69.27	1.05	8.84	.46	.12	66.25	1.05	8.84	.46	.12	0	66.25
N	0.79	100.00	.11	87.59	47.32	11.30	.36	.01	89.95	100.00	9.68	.41	.04	1	89.85
S	15.69	5.89	.75	65.41	62.61	5.65	.92	.46	65.92	0.39	28.27	.22	.89	0	58.55
S	13.75	5.80	.67	66.03	98.43	91.82	.99	.62	53.02	88.70	43.83	.97	.61	.02	53.07
S	0.00	5.65	.63	82.54	100.00	11.08	1	.57	80.17	0.00	6.25	.58	.77	.91	82.54
S	14.00	7.76	.61	69.39	0.53	25.83	.24	.87	67.59	7.27	10.65	.56	.98	.91	65.58
S	24.71	5.15	.65	65.88	4.39	6.44	.62	.87	58.57	5.27	6.57	.63	.88	.17	58.70
S	0.82	100.00	.11	75.07	100.00	100.00	1	.55	81.28	0.52	16.84	.36	1	0	61.38
S	10.76	2.87	1	76.18	0.16	43.94	.15	0	66.26	8.35	5.98	.51	.63	1	75.97
S	0.99	100.00	.11	68.64	0.35	30.70	.22	.96	58.86	0.88	33.80	.21	1	.74	58.23
S	1.12	99.97	.11	66.75	0.34	32.58	.19	.97	59.15	1.10	36.70	.18	1	.82	57.88
S	9.22	6.54	.59	82.17	0.10	100.00	.12	1	56.19	0.27	41.82	.18	1	.17	57.62
S	26.13	6.23	.72	55.04	64.53	5.07	.73	.45	57.54	20.88	7.09	.72	1	.98	54.63
S	15.83	5.52	.72	81.56	0.11	99.35	.11	.96	60.99	10.16	79.11	.98	.88	0	70.30
S	1.17	100.00	.12	66.26	0.29	47.47	.19	.94	56.83	0.32	48.19	.18	.94	.15	56.78
S	9.13	7.10	.66	81.57	0.23	47.10	.18	1	54.69	0.22	47.21	.18	1	0	54.69
S	0.83	100.00	.11	79.77	100.00	4.75	1	.22	95.00	0.31	34.37	.22	1	.11	64.83
I	16.76	7.13	.68	71.54	100.00	5.50	.83	.35	82.10	3.89	24.27	.29	.98	.94	64.07
I	15.17	7.50	.66	52.11	0.19	79.30	.12	.82	52.26	2.53	53.96	.15	.85	1	50.17
I	43.30	17.83	.86	68.88	12.35	10.95	.38	.47	74.08	58.94	10.97	.38	.51	.96	73.75
I	14.04	5.42	.68	76.22	98.05	18.08	.25	.06	76.48	33.61	12.80	.32	.32	1	76.73
I	28.13	6.02	.72	64.72	100.00	6.32	.74	.54	55.46	99.96	6.63	.74	.63	.66	55.50
I	24.82	12.91	.79	52.57	18.73	10.95	.70	.51	54.25	55.78	99.98	.98	.89	.95	43.88
I	1.82	28.98	.21	71.22	100.00	5.79	1	.30	72.52	1.80	29.59	.21	.94	1	70.87
I	11.06	7.81	.54	67.82	0.15	78.00	.12	.92	60.15	0.13	88.92	.12	.92	.01	60.13
I	10.72	10.50	.52	67.18	100.00	4.17	1	.11	106.86	7.01	13.65	.52	1	.94	62.31
I	24.41	8.59	.78	78.76	100.00	6.84	.68	.32	71.36	100.00	6.84	.68	.32	0	71.36
I	24.13	7.52	.76	69.13	100.00	7.24	.74	.28	68.58	8.67	85.19	.96	.97	.52	60.15
I	12.97	5.43	.62	59.92	0.22	35.24	.16	.95	56.24	0.23	34.42	.17	.95	.01	56.24
I	29.15	10.87	.84	66.81	4.97	22.30	.90	.87	71.24	23.12	43.05	.95	.98	.91	64.13
I	42.93	24.35	.93	85.71	100.00	100.00	.99	.57	114.57	42.94	24.36	.93	1	1	85.71
I	44.69	6.10	.98	62.45	100.00	6.28	.93	.43	56.74	100.00	6.28	.93	.43	0	56.74
D	3.43	3.87	1	103.98	7.51	8.98	.61	.68	81.41	100.00	7.98	.62	.36	0	81.38
D	12.34	5.45	.82	97.75	100.00	10.44	.94	.53	84.90	0.13	97.69	.13	1	0	70.45
D	0.44	87.87	.11	83.19	11.51	10.57	.47	.13	67.48	3.49	58.09	.93	.97	0	64.71
D	5.20	6.20	.81	99.12	100.00	6.68	.79	.36	93.60	3.94	9.35	.55	0	1	99.40
D	0.76	100.00	.10	91.73	0.33	77.37	.11	.77	70.52	2.13	33.09	.17	.72	.63	71.44
D	4.97	4.70	.91	98.63	9.75	37.63	.23	.41	107.71	0.00	5.42	.77	.54	1	99.15
D	4.10	4.31	1	74.83	0.23	38.29	.18	.66	61.69	0.27	32.34	.20	.65	.03	61.76
D	0.98	100.00	.11	65.82	0.17	65.00	.15	.96	53.92	0.21	73.96	.14	.97	.34	53.91
D	4.54	5.91	.80	100.35	0.45	64.99	.15	.51	92.98	61.89	97.91	.12	0	.18	92.71
D	2.00	20.29	.32	99.31	100.00	10.28	.68	.40	44.50	100.00	10.28	.68	.40	0	44.50
D	10.00	5.68	.63	82.15	2.97	12.72	.78	.93	68.80	3.17	14.53	.81	.94	0	68.81
D	5.00	6.53	.52	88.82	0.00	3.87	.78	.72	89.63	5.39	2.98	1	0	1	89.69
D	2.94	30.56	.25	107.26	100.00	6.92	.69	.28	98.59	100.00	6.92	.69	.28	0	98.59
D	1.00	100.00	.11	63.64	100.00	5.32	.68	.01	69.91	0.29	88.32	.12	.96	.68	57.36
D	14.67	5.67	.71	65.92	0.27	37.68	.17	.87	60.11	0.46	33.20	.18	.87	.37	60.14
MH	0.39	26.27	.19	76.77	5.76	12.49	1	1.00	77.33	1.34	3.33	1	.50	1	77.09
MH	0.47	18.61	.24	67.05	0.00	3.99	.86	.60	67.59	0.77	16.44	.26	.79	1	67.02
MH	0.92	42.10	.15	88.12	10.11	100.00	.25	.46	78.17	99.76	99.58	.98	.51	0	77.05

(continued)

Table I1

Parameter fits and -2lnL for individual participant data from Experiment 4 (continued). Note: each row is the data for one participant. Experimental conditions are: none (N), same (S), increasing (I), decreasing (D), middle high (MH).

	Time-based Forgetting				Output Interference					Both					
Pay	*c*	*s*	*t*	-2lnL	*c*	*s*	*t*	*o*	-2lnL	*c*	*s*	*t*	*o*	*w*	-2lnL
MH	100.00	2.82	1	46.95	21.48	2.45	1	.09	46.83	1.03	22.76	.17	.83	1	45.87
MH	13.51	14.66	.47	106.67	5.36	100.00	.99	.96	89.67	100.00	13.53	.34	.11	1	89.23
MH	0.54	28.98	.16	70.01	4.56	7.40	.76	.86	63.86	8.07	7.74	.77	.85	.45	63.83
MH	4.83	9.28	.45	74.03	0.00	3.53	1	.74	79.16	0.00	6.28	.59	.82	.94	79.16
MH	0.87	20.91	.21	82.25	100.00	4.20	.86	.39	81.58	1.83	20.88	.22	.44	.01	81.90
MH	3.21	17.96	.25	73.07	1.39	11.38	.50	.94	68.78	0.85	12.43	.41	.97	0	67.93
MH	10.02	17.52	.35	116.64	100.00	100.00	1	.63	85.61	1.19	23.32	.20	.65	1	95.88
MH	18.40	5.23	.71	59.70	4.54	61.47	.96	.97	57.86	5.61	8.35	.54	.97	.84	57.17
MH	0.00	4.54	.79	68.34	7.60	6.01	.90	.71	59.47	100.00	5.31	.95	.39	0	59.50
MH	16.47	3.56	1	56.74	0.68	5.55	.63	.15	57.40	7.58	7.41	.49	0	1	57.25
MH	9.36	18.76	.37	68.82	0.34	14.65	.29	0	99.63	0.34	14.82	.29	0	0	99.63

Table I2

Parameter fits and -2lnL for individual participant data from Experiment 5. Note: each row is the data for one participant.

	Time-based Forgetting				Output Interference					Both					
Schedule	*c*	*s*	*t*	-2lnL	*c*	*s*	*t*	*o*	-2lnL	*c*	*s*	*t*	*o*	*w*	-2lnL
N	15.34	4.70	1	75.91	0.00	3.98	1	.47	81.06	0.29	25.77	.25	1	0	63.63
N	12.69	4.53	.89	73.55	0.54	15.68	.34	.93	63.40	0.54	15.74	.34	.93	0	63.40
N	11.47	9.36	.55	62.07	5.09	100.00	.97	.94	69.31	3.62	24.68	.27	.91	.94	56.55
N	5.05	4.22	1	90.10	3.11	7.00	.74	.61	64.44	3.11	7.00	.74	.61	0	64.44
N	4.57	3.58	1	67.26	100.00	4.09	1	.18	66.48	0.44	12.22	.37	1.	.16	62.72
N	29.59	4.17	.99	63.61	100.00	5.17	.75	.03	63.36	100.00	4.57	.85	.47	1	63.18
N	3.64	3.99	1	86.32	1.32	22.14	.61	.99	76.93	21.66	4.06	1	.26	1	85.36
N	13.37	7.26	.73	54.67	0.96	12.69	.46	.81	68.58	9.35	9.57	.75	1	.92	52.46
N	29.55	4.08	.94	63.04	5.31	50.71	.98	.97	63.32	0.15	67.14	.12	.89	.03	63.29
N	0.77	49.13	.14	81.28	100.00	4.89	1	.47	80.59	1.43	21.22	.22	.96	1	81.73
S	1.51	100.00	.13	67.39	0.13	100.00	.13	.97	51.83	100.00	7.99	.89	.55	1	74.26
S	11.28	6.63	.65	74.46	0.11	99.95	.11	.93	60.42	0.11	99.99	.11	.93	.01	60.43
S	11.60	8.17	.62	67.43	0.22	66.34	.15	.89	69.03	11.60	8.16	.62	1	1	67.43
S	7.72	4.39	1	74.08	1.03	11.24	.53	.85	66.54	1.03	11.23	.53	.85	0	66.54
S	8.31	6.56	.66	60.18	0.69	12.25	.41	.88	58.66	3.77	8.78	.57	1	.85	57.64
S	16.77	6.99	.68	56.96	100.00	5.61	.76	.29	58.08	15.65	7.16	.68	.98	.98	56.86
S	12.80	6.89	.74	67.10	3.28	20.54	.93	.93	80.87	0.96	100.00	.12	.90	.96	60.61
S	0.35	33.53	.18	59.62	42.95	4.48	1	.04	67.67	5.57	13.33	.37	.46	.98	66.00
S	9.23	8.09	.55	65.26	100.00	5.03	.95	.45	76.05	0.52	29.07	.23	.95	.33	54.06
S	0.86	73.64	.13	86.73	97.08	66.16	1	.49	98.86	2.99	9.81	.50	.97	.92	87.18
I	7.82	10.25	.53	78.91	0.30	40.48	.21	.95	62.46	48.95	29.93	.22	.02	.78	80.05
I	43.52	14.62	.95	68.45	100.00	7.04	.68	.15	53.84	100.00	7.04	.68	.15	0	53.84
I	1.41	100.00	.12	60.81	0.15	100.00	.12	.95	56.35	1.61	45.65	.18	.99	.91	55.62
I	19.95	10.04	.72	65.75	0.77	36.18	.23	.76	81.60	17.76	12.08	.75	1	.98	65.39
I	10.55	10.71	.54	67.83	0.49	28.97	.25	.88	84.16	6.95	15.59	.62	1	.90	59.40
I	4.22	20.71	.29	60.54	100.00	11.37	.43	.07	70.16	1.82	42.27	.19	1	.96	60.20
I	18.82	6.91	.83	57.47	100.00	6.32	.77	.07	60.92	15.71	8.04	.83	1	.97	57.08
I	11.54	6.48	.63	70.45	0.72	15.31	.35	.93	63.13	0.76	15.17	.35	.93	.05	63.13
I	23.01	7.36	.90	59.56	100.00	6.18	.93	.21	58.32	39.81	6.28	.89	.73	1	56.63
I	33.77	14.27	.91	52.89	0.82	28.08	.29	.80	97.14	33.77	14.27	.91	1	1	52.89
D	3.88	4.22	1	79.86	0.14	92.71	.12	.76	51.73	0.17	72.06	.13	.76	0	51.82
D	0.76	57.11	.14	59.87	1.30	4.57	.98	.11	59.91	1.09	10.81	.45	.29	.71	59.14
D	27.28	9.33	1	103.47	0.18	100.00	.13	.90	75.01	0.18	100.00	.13	.90	0	75.01
D	1.49	26.74	.23	104.01	100.00	8.17	.81	.43	91.23	100.00	8.17	.81	.43	0	91.23
D	3.47	13.57	.40	106.08	100.00	7.23	.75	.32	100.43	0.35	33.18	.20	0	.03	97.58
D	0.71	100.00	.11	75.04	100.00	5.08	1	.10	64.82	0.12	99.22	.13	1	.09	51.25
D	14.07	5.49	.78	82.06	100.00	5.24	.80	.12	88.17	94.02	17.86	.28	.04	1	84.61
D	0.69	100.00	.12	84.22	100.00	5.81	1	.19	64.89	100.00	5.81	1	.19	0	64.89
D	0.87	99.84	.10	74.21	0.17	54.51	.14	1	55.36	0.17	57.98	.14	1	.06	55.39
D	0.00	4.98	.71	84.22	0.46	10.94	.40	1	81.40	9.82	3.72	.94	.18	1	84.20
MH	0.27	67.77	.12	89.20	0.21	3.90	.97	.04	91.09	0.34	25.43	.21	.16	0	93.31
MH	10.15	7.61	.59	61.08	0.47	23.36	.25	.86	56.23	1.00	23.50	.25	.88	.62	56.09
MH	0.30	78.55	.13	92.81	100.00	100.00	1	.42	116.72	1.62	8.82	.59	.74	1	93.80
MH	0.32	45.08	.14	82.59	0.05	39.73	.15	.87	82.49	0.15	40.59	.15	.91	.81	82.49
MH	19.17	5.61	.83	69.74	0.29	36.56	.19	.91	65.15	8.18	11.64	.80	1	.83	65.24
MH	0.86	74.30	.13	73.58	9.16	32.35	.21	.03	49.56	75.49	32.82	.21	0	0	49.52
MH	0.96	77.25	.12	68.22	100.00	6.56	.63	.01	71.29	2.23	28.04	.20	0	1	69.01
MH	0.32	98.79	.11	89.67	100.00	7.06	1	.41	70.64	100.00	7.06	1	.41	0	70.64
MH	0.41	49.07	.14	70.77	39.30	4.09	1	0	81.72	0.10	22.93	.23	1	.38	70.17
MH	6.26	4.06	1	65.97	1.79	6.43	.72	.57	56.46	2.39	5.67	.78	0	0	56.52

Table I3

Parameter fits and -2lnL for individual participant data from Experiment 6. Note: each row is the data for one participant.

Time-based Forgetting				Output Interference					Both					
c	*s*	*t*	-2lnL	*c*	*s*	*t*	*o*	-2lnL	*c*	*s*	*t*	*o*	*w*	-2lnL
0.56	99.87	.10	154.12	100.00	4.69	.85	.13	133.45	100.00	5.68	.70	.16	.95	133.71
0.82	99.54	.11	113.05	100.00	4.01	1	.01	126.15	0.61	85.72	.12	.96	.92	111.32
0.87	98.10	.11	121.25	4.95	4.76	1	.67	125.43	2.36	35.57	.19	.90	1	118.56
0.72	100.00	.11	150.72	0.14	65.66	.14	.96	137.83	0.13	78.17	.13	.96	.05	137.80
8.85	5.00	.76	147.35	0.13	59.52	.14	.96	134.98	0.12	69.80	.13	.96	.00	134.97
0.70	97.16	.11	140.86	87.22	3.82	1	.11	145.10	1.14	15.08	.32	.91	.70	139.70
10.83	8.18	.57	120.57	0.58	21.53	.29	.90	113.20	1.98	18.64	.32	.93	.76	112.02
17.68	4.24	1	106.50	0.12	75.54	.13	.94	99.87	0.00	5.20	.73	.40	1	110.02
13.37	7.09	.67	119.95	0.99	13.30	.38	.75	127.51	24.36	7.12	.62	.62	1	116.68
0.76	93.68	.11	89.49	7.21	5.32	1	.61	84.67	39.73	5.09	1	.59	.90	83.81
0.78	100.00	.11	132.05	0.34	27.79	.24	.87	124.57	0.47	26.64	.25	.87	.30	124.43
0.42	15.02	.29	132.36	0.00	3.63	.98	.28	132.59	0.00	5.78	.64	.54	.98	132.59
14.70	4.47	.84	133.32	0.67	11.20	.40	.92	129.96	0.88	10.73	.41	.92	.22	129.95
0.89	100.00	.11	110.34	0.36	18.78	.26	0	115.63	0.41	84.84	.12	.98	.81	107.04
19.90	5.02	1	99.53	0.06	17.09	.29	.14	105.34	4.96	10.59	.46	.88	1	98.40
20.90	4.18	1	130.43	2.66	3.74	.99	.25	132.86	3.82	19.25	.26	.75	1	127.78
1.00	100.00	.12	114.18	1.03	12.59	.52	.94	106.76	2.14	13.56	.46	.94	.67	105.48
6.66	4.96	.87	115.98	0.39	17.85	.32	.86	110.14	0.39	17.98	.32	.86	.01	110.14
7.98	3.87	1	126.71	0.17	37.09	.18	.99	121.07	0.00	5.80	.67	.32	.97	126.68

Appendix J:

SIMPLE Variations Comparison

Table J1

Model comparison of SIMPLE variations using individual participant data from Experiment 4 Note: each row is the data for one participant. Experimental conditions are: none (N), same (S), increasing (I), decreasing (D), middle high (MH).

	Time			Input interference			Both			AICw			BICw		
Schedule	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC	Time	OI	Both	Time	OI	Both
N	63.69	65.88	65.82	68.24	72.24	71.07	64.34	71.01	67.88	.55	.06	.40	.70	.05	.25
N	94.64	96.83	96.77	93.39	97.39	96.22	78.28	84.95	81.82	0	0	1	0	0	1
N	64.99	67.17	67.11	64.23	68.23	67.07	66.60	73.27	70.14	.34	.50	.15	.45	.46	.10
N	62.49	64.67	64.61	66.05	70.05	68.88	64.88	71.55	68.42	.68	.11	.21	.79	.09	.12
N	74.96	77.15	77.09	62.51	66.51	65.34	64.51	71.18	68.05	0	.73	.27	0	.79	.20
N	75.94	78.13	78.07	96.51	100.51	99.34	88.02	94.68	91.56	1	0	0	1	0	0
N	73.44	75.62	75.56	66.96	70.96	69.79	68.29	74.96	71.83	.03	.64	.33	.04	.71	.25
N	76.09	78.28	78.22	69.88	73.88	72.72	71.59	78.26	75.13	.03	.68	.29	.05	.73	.22
N	65.46	67.64	67.59	65.25	69.25	68.08	66.27	72.94	69.81	.36	.40	.24	.47	.37	.16
N	58.64	60.82	60.77	60.38	64.38	63.22	61.84	68.51	65.38	.62	.26	.12	.72	.21	.07
N	77.79	79.97	79.92	74.10	78.10	76.93	70.51	77.17	74.05	.02	.14	.84	.04	.18	.78
N	72.36	74.54	74.48	69.51	73.51	72.35	72.02	78.69	75.56	.16	.66	.19	.22	.65	.13
N	70.01	72.20	72.14	73.64	77.64	76.47	64.85	71.52	68.39	.07	.01	.92	.13	.02	.85
N	78.21	80.39	80.33	65.87	69.87	68.70	83.41	90.08	86.95	.00	1	0	0	1	0
N	75.27	77.45	77.39	74.25	78.25	77.08	76.25	82.92	79.79	.31	.51	.19	.41	.47	.12
N	93.59	95.77	95.71	97.95	101.95	100.78	99.85	106.52	103.39	.86	.10	.04	.91	.07	.02
S	71.41	73.59	73.54	73.92	77.92	76.75	68.55	75.22	72.09	.18	.05	.76	.31	.06	.63
S	72.03	74.21	74.15	61.02	65.02	63.85	63.07	69.74	66.61	0	.73	.26	0	.80	20
S	88.54	90.72	90.66	88.17	92.17	91.00	92.54	99.21	96.08	.43	.51	.06	.52	.44	.03
S	75.39	77.58	77.52	75.59	79.59	78.43	75.58	82.25	79.12	.36	.32	.32	.48	.30	.22
S	71.88	74.07	74.01	66.57	70.57	69.41	68.70	75.36	72.24	.05	.71	.24	.07	.74	.18
S	81.07	83.25	83.19	89.28	93.28	92.11	71.38	78.05	74.92	.01	0	.99	.02	0	.98
S	82.18	84.36	84.30	74.26	78.26	77.09	85.97	92.64	89.52	.02	.98	0	.03	.97	0
S	74.64	76.82	76.76	66.86	70.86	69.70	68.23	74.90	71.77	.01	.66	.33	.02	.72	.26
S	72.75	74.93	74.87	67.15	71.15	69.98	67.88	74.55	71.42	.03	.57	.40	.06	.64	.31
S	88.17	90.35	90.29	64.19	68.19	67.02	67.62	74.29	71.16	0	.85	.15	0	.89	.11
S	61.04	63.22	63.16	65.54	69.54	68.37	64.63	71.30	68.17	.79	.08	.13	.87	.06	.07
S	87.56	89.74	89.68	68.99	72.99	71.82	80.30	86.97	83.84	0	1	.00	.00	1	0
S	72.26	74.45	74.39	64.83	68.83	67.66	66.78	73.45	70.32	.02	.71	.27	.03	.77	.20
S	87.57	89.75	89.69	62.69	66.69	65.52	64.69	71.36	68.23	0	.73	.27	0	.79	.21
S	85.77	87.96	87.90	103.00	107.00	105.83	74.83	81.50	78.37	0	0	1	.01	0	.99
I	77.54	79.72	79.66	90.10	94.10	92.93	74.07	80.74	77.61	.15	0	.85	.26	0	.74
I	58.11	60.29	60.24	60.26	64.26	63.09	60.17	66.84	63.71	.59	.20	.21	.71	.17	.12
I	74.88	77.06	77.00	82.08	86.08	84.92	83.75	90.41	87.29	.96	.03	.01	.98	.02	.01
I	82.22	84.40	84.34	84.48	88.48	87.31	86.73	93.39	90.27	.70	.23	.07	.78	.18	.04
I	70.72	72.90	72.84	63.46	67.46	66.30	65.50	72.17	69.04	.02	.72	.26	.03	.77	.20
I	58.57	60.76	60.70	62.25	66.25	65.08	53.88	60.55	57.42	.09	.01	.90	.16	.02	.82
I	77.22	79.40	79.34	80.52	84.52	83.35	80.87	87.54	84.41	.74	.14	.12	.82	.11	.07
I	73.82	76.00	75.94	68.15	72.15	70.99	70.13	76.80	73.67	.04	.70	.26	.06	.74	.19
I	73.18	75.36	75.30	114.86	118.86	117.69	72.31	78.98	75.85	.39	0	.61	.57	0	.43
I	84.76	86.95	86.89	79.36	83.36	82.19	81.36	88.02	84.90	.05	.70	.26	.07	.74	.19
I	75.13	77.32	77.26	76.58	80.58	79.41	70.15	76.82	73.69	.07	.04	.89	.14	.05	.82
I	65.92	68.11	68.05	64.24	68.24	67.07	66.24	72.91	69.79	.24	.56	.20	.33	.53	.14
I	72.81	74.99	74.93	79.24	83.24	82.07	74.13	80.79	77.67	.64	.03	.33	.78	.02	.20
I	91.71	93.90	93.84	122.57	126.57	125.40	95.71	102.38	99.25	.88	0	.12	.94	0	.06
I	68.45	70.63	70.57	64.74	68.74	67.57	66.74	73.41	70.28	.10	.66	.24	.15	.67	.17
D	109.98	112.16	112.10	89.41	93.41	92.25	91.38	98.04	94.92	0	.73	.27	0	.79	.21
D	103.75	105.93	105.88	92.90	96.90	95.73	80.45	87.12	83.99	0	0	1	0	0	1
D	89.19	91.38	91.32	75.48	79.48	78.31	74.71	81.38	78.25	0	.40	.60	0	.49	.51
D	105.12	107.30	107.25	101.60	105.60	104.44	109.40	116.07	112.94	.14	.84	.02	.19	.79	.01
D	97.73	99.91	99.85	78.52	82.52	81.35	81.44	88.11	84.98	0	.81	.19	0	.86	.14
D	104.63	106.81	106.75	115.71	119.71	118.54	109.15	115.82	112.69	.90	0	.09	.95	0	.05
D	80.83	83.01	82.95	69.69	73.69	72.52	71.76	78.43	75.30	0	.74	.26	0	.80	.20
D	71.82	74.00	73.94	61.92	65.92	64.76	63.91	70.58	67.45	.01	.73	.27	.01	.79	.20
D	106.35	108.53	108.47	100.98	104.98	103.81	102.71	109.38	106.25	.05	.67	.28	.07	.72	.21
D	105.31	107.50	107.44	52.50	56.50	55.33	54.50	61.16	58.04	0	.73	.27	0	.79	.21
D	88.15	90.33	90.28	76.80	80.80	79.64	78.81	85.48	82.35	0	.73	.27	0	.79	.20
D	94.82	97.00	96.94	97.63	101.63	100.46	99.69	106.36	103.23	.75	.18	.07	.82	.14	.04
D	113.26	115.44	115.38	106.59	110.59	109.42	108.59	115.25	112.13	.03	.71	.26	.04	.76	.20
D	69.64	71.82	71.77	77.91	81.91	80.74	67.36	74.02	70.90	.24	0	.76	.39	0	.60
D	71.92	74.10	74.04	68.11	72.11	70.94	70.14	76.81	73.68	.10	.66	.24	.14	.68	.17
MH	82.77	84.95	84.89	85.33	89.33	88.16	87.09	93.76	90.63	.72	.20	.08	.80	.16	.05
MH	73.05	75.23	75.18	75.59	79.59	78.42	77.02	83.69	80.56	.70	.20	.10	.79	.16	.05
MH	94.12	96.30	96.24	86.17	90.17	89.00	87.05	93.71	90.59	.01	.60	.39	.02	.68	.31

(continued)

Table J1

Model comparison of SIMPLE variations using individual participant data from Experiment 4 (continued). Note: each row is the data for one participant. Experimental conditions are: none (N), same (S), increasing (I), decreasing (D), middle high (MH).

	Time			Input interference			Both			AICw			BICw		
Schedule	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC	Time	OI	Both	Time	OI	Both
MH	52.95	55.13	55.08	54.83	58.83	57.66	55.87	62.54	59.41	.62	.24	.14	.72	.20	.08
MH	112.67	114.85	114.80	97.67	101.67	100.50	99.23	105.89	102.77	0	.69	.31	0	.76	.24
MH	76.01	78.19	78.14	71.86	75.86	74.69	73.83	80.50	77.37	.08	.67	.25	.12	.69	.18
MH	80.03	82.22	82.16	87.16	91.16	89.99	89.16	95.82	92.70	.96	.03	.01	.98	.02	.01
MH	88.25	90.43	90.37	89.58	93.58	92.41	91.90	98.56	95.44	.60	.31	.10	.69	.25	.06
MH	79.07	81.26	81.20	76.78	80.78	79.61	77.93	84.59	81.47	.17	.53	.30	.24	.54	.21
MH	122.64	124.82	124.76	93.61	97.61	96.44	105.88	112.55	109.42	0	1	0	.00	1	0
MH	65.70	67.88	67.83	65.86	69.86	68.69	67.17	73.83	70.71	.42	.38	.20	.53	.34	.13
MH	74.34	76.52	76.47	67.47	71.47	70.30	69.50	76.17	73.04	.02	.72	.26	.04	.77	.20
MH	62.74	64.92	64.86	65.40	69.40	68.23	67.25	73.92	70.79	.73	.19	.08	.81	.15	.04
MH	74.82	77.00	76.94	107.63	111.63	110.47	109.63	116.30	113.17	1	0	0	1	0	0

Table J2

Model comparison of SIMPLE variations using individual participant data from Experiment 5. Note: each row is the data for one participant. Experimental conditions are: none (N), same (S), increasing (I), decreasing (D), middle high (MH).

	Time			Input interference			Both			AICw			BICw		
Schedule	AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC	Time	OI	Both	Time	OI	Both
N	81.91	84.09	84.03	89.06	93.06	91.89	73.63	80.29	77.17	.02	0	.98	.03	0	.97
N	79.55	81.73	81.67	71.40	75.40	74.23	73.40	80.07	76.94	.01	.72	.27	.02	.78	.20
N	68.07	70.25	70.19	77.31	81.31	80.14	66.55	73.21	70.09	.32	0	.68	.49	.00	.51
N	96.10	98.28	98.22	72.44	76.44	75.28	74.44	81.11	77.98	0	.73	.27	0	.79	.21
N	73.26	75.44	75.39	74.48	78.48	77.32	72.72	79.39	76.26	.35	.19	.46	.49	.19	.32
N	69.61	71.79	71.73	71.36	75.36	74.19	73.18	79.85	76.72	.63	.26	.11	.73	.21	.06
N	92.32	94.50	94.44	84.93	88.93	87.77	95.36	102.02	98.90	.02	.97	.01	.03	.96	0
N	60.67	62.85	62.79	76.58	80.58	79.41	62.46	69.12	66.00	.71	0	.29	.83	0	.17
N	69.04	71.22	71.16	71.32	75.32	74.15	73.29	79.96	76.83	.69	.22	.08	.78	.18	.05
N	87.28	89.46	89.41	88.59	92.59	91.42	91.73	98.40	95.27	.61	.32	.07	.70	.26	.04
S	73.39	75.57	75.51	59.83	63.83	62.66	84.26	90.93	87.80	0	1	0	0	1	0
S	80.46	82.65	82.59	68.42	72.42	71.26	70.43	77.09	73.97	0	.73	.27	0	.79	.20
S	73.43	75.62	75.56	77.03	81.03	79.87	77.43	84.10	80.97	.77	.13	.10	.85	.10	.06
S	80.08	82.26	82.20	74.54	78.54	77.37	76.54	83.20	80.08	.04	.70	.26	.07	.74	.19
S	66.18	68.36	68.30	66.66	70.66	69.49	67.64	74.31	71.18	.44	.35	.21	.56	.31	.13
S	62.96	65.15	65.09	66.08	70.08	68.92	66.86	73.53	70.40	.74	.16	.11	.82	.12	.06
S	73.10	75.29	75.23	88.87	92.87	91.70	70.61	77.28	74.15	.22	0	.78	.37	0	.63
S	65.62	67.80	67.74	75.67	79.67	78.50	76.00	82.66	79.54	.99	.01	.01	.99	0	0
S	71.26	73.44	73.39	84.05	88.05	86.88	64.06	70.73	67.60	.03	0	.97	.05	0	.95
S	92.73	94.91	94.85	106.86	110.86	109.69	97.18	103.85	100.72	.90	0	.10	.95	0	.05
I	84.91	87.09	87.04	70.46	74.46	73.30	90.05	96.72	93.59	0	1	0	0	0	0
I	74.45	76.63	76.57	61.84	65.84	64.67	63.84	70.50	67.38	0	.73	.27	0	.79	.20
I	66.81	68.99	68.93	64.35	68.35	67.19	65.62	72.29	69.16	.16	.55	.29	.23	.56	.21
I	71.75	73.93	73.88	89.60	93.60	92.44	75.39	82.06	78.93	.86	0	.14	.93	0	.07
I	73.83	76.02	75.96	92.16	96.16	94.99	69.40	76.07	72.94	.10	0	.90	.18	0	.82
I	66.54	68.72	68.67	78.16	82.16	80.99	70.20	76.86	73.74	.86	0	.14	.92	0	.07
I	63.47	65.65	65.59	68.92	72.92	71.75	67.08	73.75	70.62	.81	.05	.13	.89	.04	.07
I	76.45	78.64	78.58	71.13	75.13	73.97	73.13	79.80	76.68	.05	.70	.26	.07	.74	.19
I	65.56	67.74	67.68	66.32	70.32	69.15	66.63	73.29	70.17	.44	.30	.26	.57	.27	.16
I	58.89	61.07	61.02	105.14	109.14	107.97	62.89	69.56	66.43	.88	0	.12	.94	0	.06
D	85.86	88.04	87.99	59.73	63.73	62.57	61.82	68.49	65.36	.00	.74	.26	0	.80	.20
D	65.87	68.05	67.99	67.91	71.91	70.74	69.14	75.81	72.68	.64	.23	.13	.74	.19	.07
D	109.47	111.65	111.59	83.01	87.01	85.84	85.01	91.68	88.55	0	.73	.27	0	.79	.21
D	110.01	112.19	112.13	99.23	103.23	102.06	101.23	107.90	104.77	0	.73	.27	.01	.79	.20
D	112.08	114.26	114.20	108.43	112.43	111.26	107.58	114.24	111.12	.06	.37	.57	.10	.43	.47
D	81.04	83.22	83.17	72.82	76.82	75.66	61.25	67.91	64.79	0	0	1	0	0	1
D	88.06	90.24	90.18	96.17	100.17	99.00	94.61	101.28	98.15	.95	.02	.04	.97	.01	.02
D	90.22	92.40	92.34	72.89	76.89	75.72	74.89	81.56	78.43	0	.73	.27	0	.79	.21
D	80.21	82.39	82.34	63.36	67.36	66.19	65.39	72.06	68.93	0	.73	.27	0	.80	.20
D	90.22	92.40	92.34	89.40	93.40	92.23	94.20	100.87	97.74	.38	.57	.05	.47	.50	.03
MH	95.20	97.38	97.32	99.09	103.09	101.92	103.31	109.98	106.85	.86	.12	.01	.90	.09	.01
MH	67.08	69.26	69.20	64.23	68.23	67.07	66.09	72.75	69.63	.15	.61	.24	.21	.62	.17
MH	98.81	100.99	100.94	124.72	128.72	127.55	103.80	110.46	107.34	.92	0	.08	.96	0	.04
MH	88.59	90.77	90.71	90.49	94.49	93.32	92.49	99.16	96.03	.65	.25	.09	.75	.20	.05
MH	75.74	77.92	77.86	73.15	77.15	75.98	75.24	81.90	78.78	.17	.62	.22	.24	.61	.15
MH	79.58	81.76	81.70	57.56	61.56	60.39	59.52	66.19	63.06	0	.73	.27	.00	.79	.21
MH	74.22	76.40	76.34	79.29	83.29	82.12	79.01	85.67	82.55	.85	.07	.08	.91	.05	.04
MH	95.67	97.85	97.79	78.64	82.64	81.47	80.64	87.30	84.18	0	.73	.27	0	.79	.21
MH	76.77	78.95	78.89	89.72	93.72	92.55	80.17	86.84	83.71	.84	0	.15	.92	0	.08
MH	71.97	74.15	74.10	64.46	68.46	67.29	66.52	73.19	70.06	.02	.72	.26	.03	.78	.19

Table J3

Model comparison of SIMPLE variations using individual participant data from Experiment 6. Note: each row is the data for one participant.

Time			Input interference			Both			AICw			BICw		
AIC	AICc	BIC	AIC	AICc	BIC	AIC	AICc	BIC	Time	OI	Both	Time	OI	Both
160.12	162.30	162.24	141.45	145.45	144.28	143.71	150.38	147.25	0	.76	.24	0	.82	.18
119.05	121.23	121.17	134.15	138.15	136.98	121.32	127.99	124.86	.76	0	.24	.86	0	.14
127.25	129.44	129.38	133.43	137.43	136.26	128.56	135.22	132.10	.64	.03	.33	.78	.02	.20
156.72	158.91	158.85	145.83	149.83	148.66	147.80	154.47	151.34	0	.73	.27	0	.79	.21
153.35	155.54	155.48	142.98	146.98	145.82	144.97	151.63	148.51	0	.73	.27	.01	.79	.21
146.86	149.04	148.98	153.10	157.10	155.93	149.70	156.36	153.24	.78	.03	.19	.87	.03	.10
126.57	128.75	128.69	121.20	125.20	124.04	122.02	128.69	125.56	.04	.58	.38	.06	.64	.30
112.50	114.69	114.63	107.87	111.87	110.71	120.02	126.69	123.56	.09	.91	.00	.12	.88	0
125.95	128.13	128.07	135.51	139.51	138.34	126.68	133.35	130.22	.59	0	.41	.74	0	.25
95.49	97.67	97.61	92.67	96.67	95.50	93.81	100.48	97.35	.13	.55	.31	.20	.57	.23
138.05	140.23	140.17	132.57	136.57	135.40	134.43	141.10	137.98	.04	.69	.27	.07	.73	.20
138.36	140.54	140.48	140.59	144.59	143.43	142.59	149.26	146.13	.69	.23	.08	.78	.18	.05
139.32	141.50	141.44	137.96	141.96	140.79	139.95	146.62	143.49	.27	.53	.20	.36	.50	.13
116.34	118.52	118.46	123.63	127.63	126.46	117.04	123.70	120.58	.58	.02	.41	.73	.01	.25
105.53	107.71	107.65	113.34	117.34	116.17	108.40	115.07	111.94	.80	.02	.19	.88	.01	.10
136.43	138.61	138.55	140.86	144.86	143.70	137.78	144.45	141.32	.62	.07	.31	.75	.06	.19
120.18	122.36	122.30	114.76	118.76	117.59	115.48	122.14	119.02	.04	.57	.40	.06	.63	.31
121.98	124.16	124.11	118.14	122.14	120.98	120.14	126.81	123.68	.10	.66	.24	.14	.68	.18
132.71	134.89	134.83	129.07	133.07	131.90	136.68	143.35	140.22	.14	.84	.02	.19	.80	.01