Psychological Processes behind Phenomena in Multialternative Choices and Public Opinions

by

Neo Poon

Doctor of Philosophy

University of Warwick, Warwick Business School (Behavioural Science Group)

June 2022

Table of Contents

List of Tables and Illustrations

Introduction

Figure

Paper 1

Tables

Figures

Paper 2

Table

Figures

Paper 3

Tables

Figures

Acknowledgements

I place on record, my sincere gratitude to my supervisors, Professor Andrea Isoni, Dr Timothy L. Mullett, and Dr Ashley Luckman for their advice and support, without which this thesis would not have been possible.

Additionally, I wish to extend my gratitude to Warwick Business School, which provided the scholarship for my PhD study.

Declaration

This thesis consists of three papers. The results of another study are presented in the Appendix. The thesis has not been submitted for a degree at another university.

For the first two papers, I developed the initial concepts and the study designs, with contributions from my supervisors Andrea Isoni, Timothy L. Mullett, and Ashley Luckman. I completed all data collections except for Experiment 2 of the first paper, whose data was collected by two MSc students, Alice Pearce and Anastasia Anoprieva. They also had inputs in the design of that experiment.

The third paper involved the use of data sets provided by two external organisations, Engage Britain and The Cognitive Company. I developed the concept of studying the specific topics of the paper with these data sets, with contributions from my supervisors.

For the study presented in the Appendix, the concepts was co-developed by Andrea Isoni, Ashley Luckman, and me. This study involved a MSc student Martina Maglicic, who had inputs in the study design and collected the data.

I performed all data analyses described in this thesis.

Introduction

Introduction

In decision research, it has been accepted that investigating the cognitive processes underlying human behaviours is as important as studying choices and judgements themselves (Johnson et al., 2008; Schulte-Mecklenbeck et al., 2017). In this thesis, I present my findings on the psychological processes of well-known phenomena in two subfields of behavioural science, including multialternative choices and public opinions. For multialternative choices, with two process-tracing methods, I examined how processes at different cognitive levels contribute to producing the attraction effect (Huber et al., 1982; Simonson, 1989), which resulted in the first two papers in my thesis. For public opinions, by analysing the behaviours of individuals on online platforms over time, I investigated how psychological processes related to echo chambers (Sunstein, 2001, 2007) and confrontation (Karlsen et al., 2017) influenced online engagement and discussed their implications for political polarisation. The third paper in my thesis was developed based on these results.

Below, I provide an introduction to the terminologies of these topics and a literature review on the recent developments. At the end of each main section, I also highlight the contributions of my papers.

Multialternative Choices

In the literature, multialternative multiattribute choices are defined as decisions that involve at least three options, whose features are defined over at least two dimensions (Cohen et al., 2017; Turner et al., 2018). These choices are referred to simply as *multialternative choices* in this thesis, as the two terms are used interchangeably in psychological research (Trueblood et al., 2014; Usher & McClelland, 2004).

Context Effects

The study of multialternative choices has a strong focus on *context effects*. In the simplest form of multialternative choices (i.e., three options and two attributes), context effects are the observations that the relative choice share between two options depend on the attribute values of the third option. In other words, the positions of three options in a two-dimensional attribute space predictably influence the probability of the options being chosen. Three major context effects are commonly investigated in cognitive science and consumer research, namely

the attraction (Huber et al., 1982), compromise (Simonson, 1989), and similarity effects (Tversky, 1972).

Consider options A and B which vary on attributes 1 and 2 (e.g., two smartphones which have different levels of storage capacity and durability). Option A is strong on attribute 1 but weak on attribute 2, and vice versa for option B. When option D_A which is similar and inferior to A on both attributes is included in the choice set, the classic findings show that option A will be promoted and gain choice share. This is known as the *attraction effect*. Conventionally, option A is termed the *target*, B the *competitor*, and D_A the attraction *decoy*. When option D_E which makes A the middle option on both attributes is included, option A will be promoted over B. This is labelled the *compromise effect* and option D_E is termed the extreme decoy. Finally, when option D_S which is similar to but not dominated by B on both attributes is included, option A will again be promoted over B. This is called the *similarity effect* and option D_S is termed the similarity decoy. Figure 1 illustrates the attribute values of the options that can lead to the three context effects.

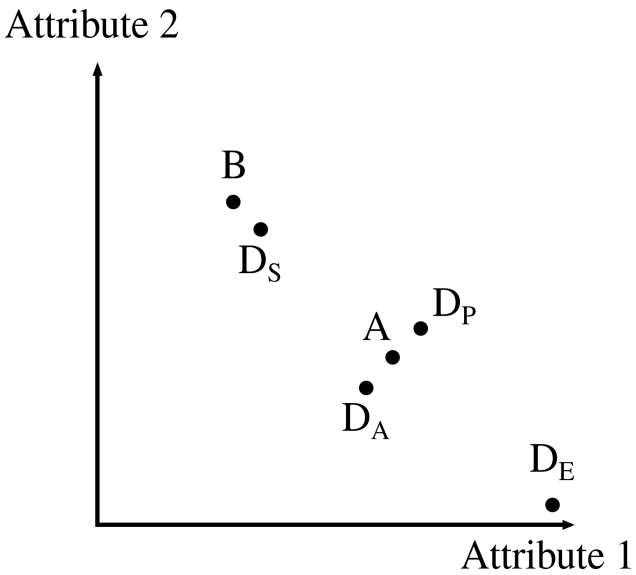


Figure 1

The positions of the options on the attribute space that are expected to create the three major context effects, as well as the phantom attraction decoy effect. Option A is the target and is expected to be promoted when any of the decoy options is included in the choice set. Option B is the competitor.

Context effects have drawn considerable attention in economics, psychology, and marketing. Initially, the context effects were documented to demonstrate that the assumptions of neoclassic economic models could be violated (Huber et al., 2014). Such models are justified with the rational theory of choice (Luce, 1959) and assume that the valuations of options are consistent in different contexts. Formally, this is related to the principles of independence of irrelevant alternatives (IIA) and regularity (Luce, 1977). The first principle suggests that, if option A is chosen more often than B in a binary choice, option A must also be chosen more often than B when additional options are added. The second principle holds that the absolute choice shares of A and B in a binary choice must not increase when additional options are added. All three major context effects can violate the principle of IIA, as they show that an option can be promoted by adding a decoy option and hence the relative choice share between options A and B can vary across contexts. The attraction effect can further violate the principle of regularity (Busemeyer et al., 2019), since the decoy option is asymmetrically dominated by the target (e.g., option A) and is not expected to be chosen, therefore option A can obtain an absolute choice share above its level in a binary choice between A and B when it is promoted. Over the past four decades, context effects have gained popularity among behaviour change and marketing practitioners, as the ability to promote an option in a choice environment is clearly desirable. Among the three major context effects, the attraction effect is the most discussed and studied in applied behavioural science, perhaps because it is practically straightforward to create an attraction decoy (Frederick et al., 2014). The following literature review, as well as the first two papers in this thesis, primarily focus on the attraction effect.

Phantom Decoy Effects

Phantom decoy effects are relatively new additions to the family of context effects. They are similar to the three major context effects and can promote one of the options, except the decoy is included in the choice set only initially. In other words, in the case of three options and two attributes, the relative choice shares between two options depends on the attribute values of a third option, which is introduced to individuals but removed at some point before a choice is made.

The first type of phantom decoy effects documented in the literature is a variant of the

attraction effect (Highhouse, 1996; Pettibone & Wedell, 2000). Consider options A and B in Figure 1. When an option D_P which is similar and superior to A on both attributes is shown to decision makers, but removed immediately prior to choices (e.g., D_P as a product is no longer available for purchase), findings show that it will promote option A. Here, option D_P is the phantom attraction decoy. The second type of phantom decoy effects is a variant of the compromise effect (Usher et al., 2008). Consider options A, B, and D_E in Figure 1. Individuals who chose one of the two extreme options (i.e., either option B or D_E) tended to switch to the target option A when their initial option was made unavailable. Here, the extreme option which was initially chosen can be described as the phantom extreme decoy. Besides the implications for marketing and applied behavioural science, the two phantom decoy effects contribute to the literature by demonstrating individuals' preference in terms of rankings. Both phantom decoy effects show that, when a favourable option is no longer available, individuals prefer the option that has a higher value on the strong attribute of the phantom decoy over the other option, which suggests that the initial choice set can draw decision makers' attention on a particular attribute.

Naturalistic Choices

A heated debate in recent years concerns whether context effects can be observed in naturalistic decision tasks, which present the options and their attributes in non-numeric formats. Trueblood et al. (2013) focused on perceptual judgement tasks, in which individuals were shown three rectangles in each trial and were asked to choose the one with the largest area. The researchers argued that height and width of rectangular shapes were perceived separately at the beginning of the task and were only integrated when an area estimate needed to be made, therefore they represented two attributes of an item. Importantly, two of the three rectangles had the same areas, but one of them was much wider than it was long, while the other was much longer than it was wide, hence they represented the two non-decoy options in multialternative choice tasks. The third rectangle acted as the decoy and was designed to promote one of the two other rectangles. For example, for the attraction effect, the decoy had the same orientation as the target and was either shorter, narrower, or both. Trueblood et al. (2013) found all three major context effects in their experiments. Trueblood & Pettibone

(2017) further found the phantom attraction decoy effect, such that a rectangle was more likely to be perceived as larger than its competitor when another clearly larger rectangle of the same orientation was included but not possible to be chosen. These findings showed that context effects could be found in low-level judgement tasks, instead of just value-based choices.

On the other hand, studies using naturalistic preferential choice tasks, which mainly focused on the attraction effect, were not able to find the expected influence of contexts on decision outcomes. Frederick et al. (2014) investigated the attraction effect with a series of experiments involving consumer choices, in which at least one attribute of the options was presented either perceptually or implicitly. For instance, participants were asked to choose between three flats based on their window views and floor spaces, the former of which were illustrated by pictures. Another example was a ternary choice between three movies, in which one of the movie options was widely considered as a bad sequel to another option, and participants were asked to make a decision based on movie posters and descriptions. Frederick et al. (2014) did not find the attraction effect in any of the naturalistic choice tasks. Their results were supported by Yang & Lynn (2014), who found the attraction effect in only two choice tasks among more than 30 ternary choices using non-numeric stimuli. A recent addition to the discussion was the study conducted by Trendl et al. (2021), who focused on movie choices and did not find the attraction effect either. Specifically, their design took individual preferences into consideration by first asking participants to rate the movies, then instructed them to choose from two movies with the same rating and another one that was more poorly rated. Taken together, these findings suggested that the attraction effect in preferential tasks is limited to environments in which numeric stimuli were used. Hence, the studies presented in the first two papers of this thesis adopted the traditional experimental setups and showed the attribute values of the options to participants with numeric matrices.

Computational Models

Over the past two decades, a focus in the study of multialternative choices has been the development of computational models to explain context effects. These models typically make assumptions on information processing, including how individuals attend to the options and their attribute values during a decision task, how they aggregate information, and under what

conditions they will terminate the decision processes and make a decision (Noguchi & Stewart, 2018; Turner et al., 2018).

Below is a verbal summary of well-known models and how they explain the attraction effect. The simplest form of multialternative choice (i.e., three options and two attributes) is assumed.

Multialternative Decision Field Theory (MDFT)

The multialternative decision field theory (Roe et al., 2001) is one of the earliest computational models for multialternative choices. This model assumes that, at every step in the decision process, either of the two attributes is attended to by the decision maker. The absolute value of an option on an attribute is first transformed into a relative value by subtracting the average of the attribute values of the two other options from it. The relative attribute values of the options are accumulated as evidence over time, depending on which attribute is being attended to. Finally, a distance-dependent lateral inhibition is assumed such that closer options inhibit each other more strongly. A decision is made when the evidence accumulated by an option exceeds a relative threshold. The attraction effect is directly explained by lateral inhibition: as the decoy is close to the target in the attribute space but not to the competitor, it inhibits the target more strongly. However, this inhibition is mathematically negative, as it is computed from the relative value of the decoy which is negative over time, and thus promotes the target.

Leaky Competing Accumulator (LCA)

The mathematical implementation of the leaky competing accumulator model (Usher & McClelland, 2004) usually requires the assumption that only one of the two attributes is attended to at a given time (Turner et al., 2018). The relative value of an option on an attribute is computed by summing the pairwise differences between it and the other two options on that attribute. These pairwise differences are transformed since loss aversion is assumed, such that negative pairwise differences in attribute values are weighted more heavily than positive ones. The relative attribute values are accumulated as evidence for the options, depending on which attribute is being attended to. The accumulation is assumed to be leaky, therefore reduced levels of relative attribute values are added as evidence at every step. In addition, lateral

inhibition is assumed. However, contrary to the MDFT, this inhibition is assumed to be uniform (Busemeyer et al., 2019) and simply implies that the options compete with each other for preference, whose strength is mathematically determined by a single parameter. The model assumes that individuals will continue the decision process until it is terminated externally, for instance, due to time limit of the task. The attraction effect is explained by loss aversion: as losses are weighted more heavily, choices depend strongly on which options having fewer comparisons that result in large negative values. Since the competitor is distant from both the target and the decoy, it has negative pairwise differences of large values when compared to them on their strong attribute and is penalised over time.

Multiattribute Linear Ballistic Accumulator (MLBA)

The multiattribute linear ballistic accumulator model (Trueblood et al., 2014) does not rely on the assumption that only one attribute is attended to at every step of the decision process. This model uses drift rates to capture the rates at which evidence is accumulated for the options. The drift rate of an option is computed as the pairwise differences between that option and the other two options on both attributes. Comparisons involving different options and different attributes are assumed to have different weights, such that the related pairwise differences are transformed by the weights. Drift rates are fixed and do not change over time, hence the model has the advantage of being mathematically tractable. Finally, it is assumed that a decision is made when the evidence accumulated by an option exceeds an absolute threshold. The attraction effect is explained by assuming heavier weights for options that are closer to each other. Thus, pairwise differences between the target and the decoy are weighted more heavily.

Associations and Accumulation Model (AAM)

The associations and accumulation model (Bhatia, 2013) similarly assumes that only one attribute is attended to at a given time. A key feature of this model is that it makes explicit assumption on the attention weights of the attributes. Specifically, the probability of an attribute being attended to by decision makers is proportional to the sum of the attribute values each option has on that particular attribute. Put simply, individuals are assumed to sample an attribute more frequently when the options have higher total values related to it. Attribute

values are transformed, commonly with a power rule, before being accumulated as evidence and no pairwise difference is computed. Similar to the LCA model, a uniform lateral inhibition is assumed. Finally, an absolute decision threshold is assumed. The attraction effect is explained by the definitions of the attention weights: the strong attribute of the target and the decoy is more likely to be attended to due to its higher total value.

Multialternative Decision by Sampling (MDbS)

The multialternative decision by sampling model (Noguchi & Stewart, 2018) assumes that decisions are made with pairwise ordinal comparisons, therefore two options are compared on one attribute at every step. Mathematically, the probability of an option being evaluated on an attribute is proportional to how similar it is to other options on that attribute. Similarity in attribute values is computed with a relative distance function. For instance, for option A on attribute 1, its distance to option B is a function of their difference in attribute values divided by the value of option A itself. A longer distance means a lower level of similarity. The distance between two options on an attribute is also used to compute the probability of which an option will be preferred over another option on that attribute. For example, if option A has a higher value than option B on attribute 1, the probability of option A being preferred is a function of the distance between the two options. If option A has a lower value, the probability of option A being favoured is zero. Putting the above together, the rate at which evidence is accumulated for an option is driven by the frequency of that option being evaluated on an attribute, as well as the frequency of that option being preferred in pairwise comparisons, summed over both attributes. Finally, the decision threshold is assumed to be relative. The attraction effect is explained by the similarity between the target and the decoy, as well as their dominance relationship. Specifically, as the target is similar to the decoy, comparisons between them occur more frequently. Furthermore, as the target is superior to the decoy on both attributes, while the competitor is preferred over the decoy only on one attribute, the target is associated with a higher number of favourable comparisons.

Summary

The above models share many similarities despite they assume different cognitive mechanisms. First, except the MLBA model, all models either assume or mathematically

require that one attribute is being attended to at a time. The more recent models, namely the AAM and the MDbS model, make explicitly assumptions on how attention is driven by the attribute values of the options and thus how comparisons are made. Second, all models assume that the attribute values of the options are transformed and accumulated as evidence in favour of the options. Most of the models compute the rate at which evidence is accumulated based on the differences in attribute values between the options, except the AAM which transforms attribute values using a power rule. A key difference between these models is their stopping rules, which range from the assumption of an internal threshold to external termination. Finally, all of the above models differ in how they explain the attraction effect.

Process Tracing

Process-tracing methods are of great importance in the study of multialternative choices and context effects, since computational models heavily rely on assumptions of how decision makers sample information and make comparisons between options, as described in the above sections.

The work of Noguchi & Stewart (2014) is the first study which adopted eye-tracking methods to examine attentional processes underlying the three major types of context effects. Specifically, it investigated whether transitions in eye-movement patterns made by individuals in such environments are primarily within-attribute or within-option. If transitions are mostly within-attribute, it suggests that multialternative choices are made mainly by evaluating options on one attribute at a time. This would provide support for most of the above computational models. If transitions are mostly within-option, it implies that individuals make decisions mainly by first integrating attribute values for the options, then comparing the options based on their integrated values. This is often assumed in economic models for binary choices, such as different variants of Prospect Theory (Stott, 2006; Tversky & Kahneman, 1992), but is not a common assumption in multialternative choices. Finally, other eye movement patterns are also possible. For instance, it is possible that individuals deliberately select a pair of options and evaluate them on a selected attribute, then move onto another pair on another attribute. This comparison pattern is assumed by variants of the Decision by Sampling model (Noguchi & Stewart, 2018; Stewart et al., 2006). Results of this study

showed that within-attribute transitions were more frequent than within-option ones. Furthermore, when participants attended to a pair of options more often, the probability of the third option being chosen decreased, which could not be explained by models that only assume within-attribute comparisons. Together, these findings suggested that multialternative choices involve the repeated comparisons between pairs of options on selected attributes.

Other eye-tracking studies on context effects also exist, with a focus on the attraction effect. Król & Król (2019) showed that the presence of the decoy option resulted in participants paying more attention to the target, and this effect was stronger when the dominance which the target had on the decoy was larger. In other words, more attention was paid to the target when it was less similar and more superior to the decoy. This was consistent with the notion that the attraction effect requires the dominance relationship between the target and the decoy to be salient to the decision makers (Huber et al., 2014). Król & Król (2019) argued that the difficulties in finding the attraction effect with naturalistic choice tasks were due to the similarity between the target and the decoy being emphasised by perceptual stimuli, while the dominance relationship was suppressed. Finally, also with eye-tracking methods, Marini et al. (2020) investigated the attentional and comparison patterns that produce the attraction effect. Findings showed that decision makers paid more attention to the target option and the attribute on which the target dominated the decoy, and made more target-decoy comparisons than competitor-decoy ones.

Contributions

While the cognitive mechanisms behind multialternative choices have been extensively investigated, many of the existing studies and models focus on low-level psychological processes, such as attention. The roles of high-level cognitive processes, including reasoning, in context effects are less explored except a few early studies (Simonson, 1989). A decision framework of particular interest is Query Theory (Johnson et al., 2007), which focuses how reasons are generated by individuals during a preferential choice task and how contexts can affect this process. Although Query Theory is not a computational model, it also relies on the assumptions that evidence needs to be gathered and informs decision outcomes, thus provides the opportunities to bridge the gaps between different research streams in value-based choices.

Query Theory

Unlike the computational models mentioned above, Query Theory does not explicitly make assumptions on how decision makers momentarily shift their attention between the attribute values of the options. However, Query Theory does assume that individuals need to sequentially process each aspect of the decision task. For instance, if the task involves a choice between two options A and B, it is assumed that decision makers will decompose the task into the two main *queries*, including 'Why should I choose option A?' and 'Why should I choose option B?' Individuals then resolve these queries by generating reasons in favour of the two options, perhaps repeatedly. Query Theory is usually regarded as a memory-based theory (Weber & Johnson, 2006) and suggests that reasons are generated by retrieving information from memory. However, the studies in this thesis show that this reason-based deliberation process can also be used to explain choices in tasks where there is no particular need to rely on memory, such as consumer choices in which the attribute values of the options are numerically defined and constantly shown to decision makers (see the Appendix of this thesis). In such tasks, each reason is naturally related to an attribute and typically involve a comparison between the options on that attribute. This mechanism shares similarities with many computational models, although these comparisons between attribute values are not strictly required by Query Theory, as reasons can also simply reflect a decision maker's general considerations of the options or the task. Additionally, Query Theory assumes that earlier queries have heavier weights and suppress the processing of later queries, hence decision outcomes are determined by which queries are executed earlier and also the number of reasons supporting each option.

One further assumption of Query Theory is that decision contexts influence the order of which queries are resolved. For example, for a choice between a default option and a non-default one, it is assumed that queries regarding why the default should be chosen would be executed earlier, and thus lead to the default effect (Johnson & Goldstein, 2003). Results demonstrated that reasons supporting the default option were indeed generated earlier by participants than those supporting the non-default option (Dinner et al., 2011). The number of reasons in favour of the default option was also greater. With similar mechanisms, Query

Theory has been used to explain a range of behavioural phenomena, including asymmetric discounting in intertemporal choices (Weber et al., 2007), observed changes in preferences due to attribute framing (Hardisty et al., 2010), and the incumbency advantage in political choices (Spälti et al., 2017).

The first paper in this thesis investigated whether Query Theory could be extended from binary to multialternative choices, specifically as an explanation of the attraction effect. Results showed that the reason generation process described by Query Theory could predict the attraction effect in the simplest form of multialternative choices, as well as when more complicated choice sets with additional attributes were presented to decision makers. This speaks to the generality of Query Theory, as well as the robustness of the attraction effect with numeric stimuli. The second paper in this thesis followed up on these findings and aimed to bridge the gaps between Query Theory and attention-based accounts of the attraction effect. Using both a mouse-tracking method and a reason listing protocol, results showed that reasons and attentional processes predicted the attraction effect individually. Furthermore, the effects of attentional process on choices were partially mediated by reason generation, which provided a major step into understanding how psychological mechanisms of different cognitive levels interact with each other and produce a classic phenomenon in multialternative choices.

Public Opinions

The dynamics of public opinions, especially political polarisation, is a key research area in social science (Baldassarri & Bearman, 2007; Fiorina & Abrams, 2008). The psychological processes behind phenomena in public opinions have long been investigated, with many studies focused on how people selectively process information (Garrett, 2009; Graf & Aday, 2008; Iyengar & Hahn, 2009; Lawrence et al., 2010) and how that results in the emergence of echo chambers (Barberá et al., 2015; Del Vicario et al., 2016, 2017; Williams et al., 2015).

Selective Attention and Confirmation Bias

In psychology, it is widely accepted that individuals have limited mental capabilities in processing information (Baddeley, 2012; Stanovich & West, 2000). The amount of information on the Internet has been increasing drastically over the years (Hilbert & López,

2011; Van den Bosch et al., 2016) and even a small proportion of it can easily lead to cognitive overload (Bawden & Robinson, 2009), therefore it is a common assumption that individuals selectively pay attention to certain information in online environments (Hills, 2019). In particular, information that is consistent with one's attitude and beliefs is more likely to be consumed (Chaffee & Miyo, 1983; Sears & Freedman, 1967).

While this selective attention to attitude-consistent information can be described as *confirmation bias*, which is a well-studied phenomenon (Nickerson, 1998), researchers have become increasingly more interested in its emergence on online platforms. The work of Graf & Aday (2008) is one of the first studies which examined selective attention without the use of self-reported metrics. Instead, they instructed participants to browse a website whose design resembled an online magazine and included a number of attentional measures, such as the order in which participants read the articles on the pages and the related durations, as well as the pages they read. Participants, who were recruited in the USA, were also asked for their views on various political issues and presidential choices. Results demonstrated that individuals generally turned to articles that were consistent with their beliefs first, and paid more attention to them in terms of both durations and numbers of pages. In a similar study, Garrett (2009) used a mocked news aggregation service and recorded whether participants expressed interest in reading articles shown to them. Findings showed that individuals were more likely to read articles which were aligned with their views on polarised social issues. Finally, Lawrence et al. (2010) found that individuals preferred to read blog posts from sites which were positioned on the same end of the political spectrum as themselves, while Iyengar & Hahn (2009) found similar results and demonstrated that individuals were more likely to read news articles from belief-consistent sources, even when the headlines were identical.

Selective attention is frequently attributed as the primary factor behind the formation of *echo chambers*, which are isolated online communities where individuals interact with others holding similar sociopolitical views and reinforce their beliefs (Sunstein, 2001, 2007). In recent years, echo chambers are usually studied by analysing contents and networks of interpersonal connections on social media. Generally, researchers were able to find echo chambers on Twitter (Barberá et al., 2015; Conover et al., 2011; Himelboim et al., 2013;

Williams et al., 2015) and Facebook (Del Vicario et al., 2015, 2017) for political issues. The existence of echo chambers on online platforms is, in turn, commonly considered as a main driver of political polarisation in offline environments.

Confrontation and Disconfirmation Bias

In recent years, researchers have become increasingly more interested in how individuals behave when they are exposed to opposing views, despite their preferences for selectively attending to information consistent with their beliefs (Bail et al., 2018; Karlsen et al., 2017). A reason is that avoiding attitude-inconsistent information entirely on online platforms can be difficult. For instance, on average, 24% to 35% of stories that an individual is exposed to on Facebook are contradictory to their sociopolitical stances (Bakshy et al., 2015). Additionally, on Twitter, it is known that users frequently address others with different political stances in their posts (Conover et al., 2011; Williams et al., 2015).

A small number of recent studies have demonstrated that political polarisation can occur even when individuals consume heterogeneous sources of information. Bail et al. (2018) measured the political attitudes of participants and instructed them to follow a Twitter bot which, unknowingly to them, was designed to circulate posts from attitude-inconsistent sources. After a month, political attitudes were measured again. Results showed that individuals either had the same sociopolitical attitudes, or held views which were on the same end of the political spectrum as before but even more extreme. With self-reported measures, Karlsen et al. (2017) found that over 70% of their respondents said that they occasionally discussed politics with individuals holding different political views, but only 6% of the respondents claimed to have changed their opinions after online debates.

Psychologically, the above results are explained by people's tendencies to refute contradictory opinions and confront others with opposing beliefs, which was termed by Taber & Lodge (2006) as the *disconfirmation bias*. Their experiments showed that participants considered attitude-consistent statements as more convincing than attitude-inconsistent ones for the same issue, even though the duration for which they read attitude-inconsistent statements was longer. With a though-listing protocol, Taber & Lodge (2006) found that individuals submitted more comments for attitude-inconsistent statements than

attitude-consistent ones. They contended that counterarguing opposing opinions is a critical part of how individuals process information and update their beliefs: the evaluation of new information is distorted by motivations and evidence contradictory to prior attitudes is discounted. Therefore, sociopolitical polarisation is a result of both confirmation and disconfirmation biases. Karlsen et al. (2017) further argued individuals are motivated to actively seek out others holding opposing views and challenge them in online debates, but mostly for the purpose of reinforcing one's existing beliefs. While evidence is limited, this view is supported by the work of Stromer-Galley & Muhlberger (2009), who found that participants were more willing to engage in future online discussions as the number of arguments challenging their views in previous discussion sections increased, but the latter did not predict whether participants were willing to change their own opinions.

Contributions

With two unique data sets, the third paper in this thesis examined the psychological factors related to public engagement. Specifically, the goals are twofold: first, I investigated the notion of echo chambers with a new methodology. Many studies on selective attention and the confirmation bias were either experimental and did not involve interactions with other individuals (Garrett, 2009; Graf & Aday, 2008; Iyengar & Hahn, 2009), or primarily focused on connectivities between users on social media networks (Barberá et al., 2015; Conover et al., 2011; Del Vicario et al., 2017; Himelboim et al., 2013; Williams et al., 2015). The data sets used in my study, on the other hand, captured the environments and behaviours of individuals on online platforms at every step. In particular, the platforms instructed participants to provide comments on social issues and allowed them to express whether they agreed or disagreed with others' opinions. The data sets recorded the comments saw by each user and the time points at which they voted or submitted their own comments. Although it is not a process-tracing method in a traditional sense (Schulte-Mecklenbeck et al., 2017), it allowed the analysis of how individuals responded to different information over time, hence provided valuable insights into the cognitive processes behind online deliberations. Second, I studied the confrontation theory and the disconfirmation bias, and whether individuals' willingness to engage in online discussions were influenced by exposure to opposing viewpoints. As stated

in the previous section, evidence related to confrontation in non-experimental settings is limited, thus this study offered an opportunity to understand how individuals counterargue attitude-inconsistent opinions. Results supported both the notion of echo chambers and the theory of confrontation, showing that individuals were generally more willing to engage in online discussions when they agreed with a higher proportion of comments on the platforms, but their engagement could also be triggered by the exposure to opposing views.

A reason-based account of the attraction effect and its dimensional extensions

Abstract

We provide novel support for Query Theory, a reason-based decision framework, extending it to multialternative choices by applying it to a standard and a novel version of the attraction effect. In Experiment 1 (N = 261), we show that Query Theory can explain the classic attraction effect： reasons supporting the target option are generated earlier and in greater quantity than those supporting the competitor. In Experiment 2 (N = 719), we develop a novel extension of the attraction effect to more complex four-attribute items. We find that the magnitude of the effect is unaffected by this dimensional extension, regardless of whether two or four attributes discriminate between target and competitor. As in Experiment 1, the reason structure is as predicted by Query Theory. Our findings speak to the robustness of the attraction effect beyond the standard setup and to the close relationship between reason generation and multialternative decision making.

A reason-based account of the attraction effect and its dimensional extensions

Most of our daily choices involve a multiplicity of alternatives often defined over a variety of different attributes. A large body of experimental evidence suggests that such multialternative, multiattribute decisions may be systematically affected by contextual features (Huber et al., 1982; Simonson, 1989; Tversky, 1972).

A classic context effect of this kind is the *attraction effect* (Huber & Puto, 1983), also known as the 'asymmetric dominance effect' or the 'decoy effect'. This effect shows that one option in a choice set – often referred to as the *target* – can gain in choice share over its *competitor* when an asymmetrically dominated *decoy* option – i.e., an option that is unambiguously worse than the target but not the competitor – is added to the set. The attraction effect is usually demonstrated when the target, the competitor, and the decoy are defined over two common attributes associated with numeric values (Bhatia & Stewart, 2018; Frederick et al., 2014). For example, consider a consumer choosing between two smartphones that differ in terms of storage capacity and malfunctioning rate. Smartphone A has a storage capacity of 16 gigabytes (GB) and a 3% probability of malfunctioning in the first two years. Smartphone B has a 32 GB storage capacity and a 5% malfunctioning rate. The attraction effect arises when the relative choice share between phones A and B is influenced by the presence of a third, asymmetrically dominated smartphone. So, adding smartphone D_A with a 12 GB storage capacity and a 3.5% malfunctioning rate is expected to promote phone A. Conversely, adding smartphone D_B with a 28 GB storage capacity and a 5.5% malfunctioning rate is expected to promote B.

The attraction effect is of particular interest in behavioural science for three major reasons. First, it violates Luce's axiom of rational choice theory and the related principle of independence of irrelevant alternatives (Luce, 1959, 1977). The axiom states that the probability of choosing one option over another should not be affected by adding or removing irrelevant options. The violation is obvious in the case of the attraction effect, as adding an irrelevant, transparently dominated option that no one is expected to choose increases the probability of people choosing the dominating option. Second, due to its potential to influence choices in a predictable direction, the effect is often viewed as a marketing tool, and hence is

one of the most widely studied phenomena in consumer behaviour (Frederick et al., 2014). Third, as the attraction effect cannot be explained by models which compute the subjective value of each option based only on its own attribute values (Turner et al., 2018; Tversky, 1972), it has led to the development of a number of process models that aim to explain multialternative choices with cognitive mechanisms such as attentional weights or loss aversion (Bhatia, 2013; Trueblood et al., 2014; Usher & McClelland, 2004).

While the attraction effect has been studied extensively (see Marini et al. (2020) for a summary of empirical studies, and Noguchi & Stewart (2018) for a review of formal models), limited attention has been devoted to understanding the high-level deliberation processes behind the effect. Here, we focus on reason-based preference construction, as individuals often produce reasons during decision making and aim to make choices they consider justifiable (Simonson, 1989). Such reasons are likely to be based on comparisons between the available options. Since the choice set is a central aspect of the decision-making context, it may exert a significant effect on which reasons are generated, and ultimately on what is chosen. Hence, reason-based decision making can provide valuable insights into the psychological mechanisms underlying multialternative choices, and shed light on the roots of context effects like the attraction effect.

The idea that reason generation plays a central role in the process of preference construction is one of the cornerstones of Query Theory (Johnson et al., 2007; Weber et al., 2007). Query Theory proposes that preference construction occurs as individuals decompose a decision task into queries that are executed sequentially. Each query involves the evaluation of possible decision outcomes and is resolved by distinct reasons generated by decision makers. For instance, for a consumer considering a choice between smartphones, examples of queries include ‘why should I buy smartphone A?’ and ‘why should I buy smartphone B?’ Examples of reasons associated with these queries are ‘I would choose this smartphone because it has enough storage for my photos’ or ‘I don’t replace my phone very often, so a low malfunctioning rate is important.’

Based on the notion of output interference, Query Theory suggests that the processing of one piece of information can suppress the processing of other information. Specifically,

output interference implies that earlier queries have greater weights than later queries, and that the execution of each query inhibits information processing and can terminate deliberation. This leads to the general prediction that reasons favouring the chosen option are generated earlier and in greater quantity than reasons favouring other options.

In the Query Theory framework, the context of a decision task may be expected to influence query order, leading to the earlier processing of salient and accessible information (Spälti et al., 2017). Although queries cannot be directly observed, the resulting reasons can be traced in experiments. Studies are typically conducted using a methodology known as *aspects listing*, which requires participants to list the reasons that come to their mind as they are evaluating the available options. This methodology allows researchers to test the general prediction of Query Theory by investigating whether the quantity and positions of the generated reasons favour the chosen option, and whether different contexts systematically affect reason generation. It has been found that, for instance, designating one of the options as the default (Dinner et al., 2011) or framing attributes differently (Hardisty et al., 2010) affects query order and, consequently, decision outcomes. Because both default settings and attribute framing are contextual features, this suggests that context effects like the attraction effect could also arise as a result of the order and quantity of reasons generated as decision makers deliberate.

While Query Theory has been used to explain a range of phenomena observed in binary decision tasks (Dinner et al., 2011; Hardisty et al., 2010; Johnson et al., 2007; Spälti et al., 2017; Weber et al., 2007), it is largely unknown whether it is applicable to multialternative choices with three or more options. In two preregistered experiments based on the aspect listing methodology, we investigated whether reason-based deliberation as characterised by Query Theory is able to explain the attraction effect. Using stimuli adapted from previous research (Noguchi & Stewart, 2014; Zhou et al., 1996), in Experiment 1 we considered whether the general prediction of Query Theory – that the quantity and order of reasons predict choice – applies to standard three-option choice sets, made of a target, a competitor and a decoy, defined over two common attributes. Extending Query Theory to multialternative choices involved generalising the definition of the *Content* and *Order* scores used in Query

Theory to allow each reason to favour multiple options, since reasons focusing on certain attributes can be supportive of more than one option if these are similar. In Experiment 2, we extended the standard attraction effect setup by doubling the number of attributes used to describe the three options. We varied whether or not these additional two attributes also discriminated between the target and the competitor. In addition to increasing the scope for reason-based deliberation, this extension allowed us to investigate the attraction effect in more complex settings and to contribute to recent debates about its robustness and generality (Frederick et al., 2014; Trendl et al., 2021; Yang & Lynn, 2014).

Experiment 1

Method

Participants were asked to make a choice between three different smartphones, and list the reasons they considered during their deliberation.

The first experiment used a 2 Target (A vs B) × 2 Sequence (Choice First vs Reason First) design and involved five main stages.[1]

- Choice: Participants were asked to make one ternary choice. As shown in Table 1, the attribute values of two options (i.e., A and B) remained the same across Target conditions, while the attribute values of the third decoy option varied to designate either option as the target (i.e., the decoy option was set to either D_A or D_B). The three options were presented in a 3 (options) × 2 (attribute) matrix in which the positions of the options and the attributes were randomised across participants.
- Aspect Listing: Participants were asked to list, one at a time, the reasons that came to their minds as they were considering the options, with a minimum of one reason. This aspect listing procedure was either manually terminated when a participant decided that they had already given all the reasons that allowed them to make their choice, or automatically terminated when they had input 10 reasons.
- Reason Coding: Participants were asked to indicate, for each of their reasons, which option(s) it supported. Participants could state that a given reason supported a single option or multiple options.

[1] See Supplementary Materials for screenshots of Experiment 1.

- Reason Weighting: Participants were asked to indicate how important each reason was in reaching their decision. The positions of the reasons shown on screen were randomly shuffled so that they did not necessarily reflect the order in which they were entered.

The standard methodology places aspect listing before choice (Johnson et al., 2007; Weber et al., 2007). We included a Sequence manipulation varying whether aspect listing took place before or after the choice stage. This manipulation was included to verify that prior reason listing has no systematic effects on choice (Adjerid et al., 2016). In both cases, reason coding and reason weighting took place after the completion of both the aspect listing and choice stages.

Table 1

Choice sets used in Experiment 1.

Smartphone	Storage Capacity	Malfunctioning Rate
A	16 GB	3%
B	32 GB	5%
D_A	12 GB	3.5%
D_B	28 GB	5.5%

Participants were recruited from Prolific. We planned to recruit 280 participants as one of our previous studies showed that 70 participants per cell was sufficient, and we aimed to recruit only English speakers since we required participants to provide reasons. We excluded mobile users due to the difference in input methods and its potential impact on aspect listing. To check that participants were paying attention, at the end of the experiment they were asked to identify which type of product their choice was about.

The experiment was approved internally by the doctoral programme office at Warwick Business School.

Predictions

All predictions were preregistered.[2]

[2] The preregistration can be found on OSF Registries: https://osf.io/b9pz8

For the attraction effect, we predicted that the decoy option would not be chosen, since it was transparently dominated by the target. Additionally, we predicted that the choice share of an option would be higher when it was the target than when it was not. From Query Theory, we predicted that reasons supporting the chosen option would be generated in greater quantity and earlier. Finally, conditional on finding the attraction effect, we predicted that reasons supporting an option would be generated in greater quantity and earlier when it was the target than when it was not.

Indices

To analyse the quantity and positions of reasons generated by the participants (hereafter called *reason structure*), we used two types of indices known as *Content* and *Order* scores.

A Content score represents the extent to which reasons supporting one of the options are generated in greater quantity by a participant. Given the three options (i.e., A, B, and the decoy), we define the Content score for option A as:

$$Content_A = \frac{n_A}{N}$$

where $Content_A$ is the proportion of reasons supporting A generated by a participant (n_A is the number of reasons supporting A of that participant and N is their total number of reasons). $Content_B$ and $Content_{Decoy}$ are defined analogously. Further, content scores can be defined with respect to the target manipulation, hence $Content_{Target}$ and $Content_{Competitor}$ represent the proportion of reasons favouring the target and competitor respectively.

Similarly, an order score represents the extent to which reasons supporting an option are generated earlier by a participant. Given the three options used in the experiments, the Order score for option A is defined as:[3]

$$Order_A = \begin{cases} 1 - (\frac{MR_A}{N+1}), & n_A > 0 \\ 1 - (\frac{N+1}{N+1}) = 0, & n_A = 0 \end{cases}$$

where MR_A is the median rank of reasons supporting A of that participant (the first

[3] This normalisation, which departs slightly from the definition used in the preregistration, has the advantage to ease the interpretation of our regressions with no material effect on the results.

reason having rank 1, the second 2, etc.), n_A and N are defined as above. Thus, $Order_A$ is the average position of reasons supporting A generated by a participant, normalised so that higher scores indicate that reasons supporting A are generated earlier. $Order_B$ and $Order_{Decoy}$ are defined analogously. Similarly to Content scores, Order scores can also be defined with respect to the target manipulation, resulting in $Order_{Target}$ and $Order_{Competitor}$.

Given the definition of Content and Order scores, when a participant generated more (respectively fewer) reasons supporting option A than those supporting option B, the difference between the Content scores of options A and B (i.e., $Content_A - Content_B$, hereafter *Content Difference*) will be positive (negative). Similarly, when reasons supporting option A were generated earlier (respectively later) than those supporting option B, the difference between the Order scores of options A and B (i.e., $Order_A - Order_B$, hereafter *Order Difference*) will be positive (negative).

Results

283 UK residents whose first language is English were recruited (53% female, 47% male, mean age = 37). 20 participants were flagged as mobile users and excluded from the analysis, while 2 other participants were excluded as they explicitly expressed confusion about the experiment during the reason listing procedure.[4] No participants failed the attention check. Hence, the data of 261 participants entered the analysis stage.

Preregistered Analysis

Very few participants chose the decoy (no more than 6% in any condition), which was as expected since the decoy was dominated by the target on both attributes. As specified in the preregistration, those who chose the decoy were removed from the analysis, allowing us to treat choice as a binary outcome. Our analysis will be based on the remaining 253 participants.

The Attraction Effect. Figure 1 shows the proportion of participants choosing the target as a function of whether it was A or B, and of whether reasons were elicited before or after choices were made. It is clear that the attraction effect was found in both the Reason First and Choice First conditions: while participants were slightly more likely to choose the target

[4] This exclusion criterion was not preregistered. Including these two participants does not alter any of our conclusions.

when it was B than A, overall they choose the target more frequently than the competitor (65.6% of the time in both the Choice First and Reason First conditions, as shown by the black diamonds in Figure 1).

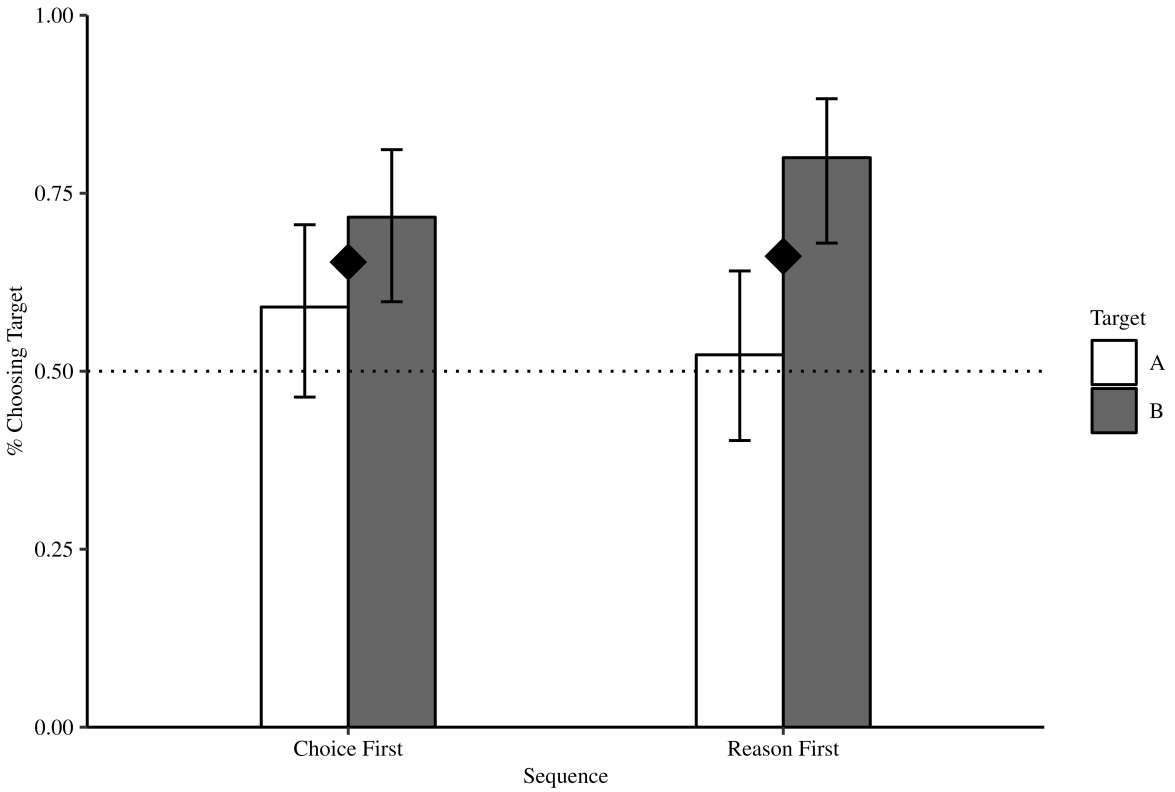


Figure 1
Proportion of participants choosing the target option by conditions in Experiment 1 (error bars are confidence intervals of a logistic regression model, black diamonds are the mean proportion in each Sequence condition).

A logistic regression shows that, in line with the attraction effect, the target was significantly more likely to be chosen than the competitor (b = 0.68, z = 4.95, p < .001, 95% CI [0.41, 0.95]). Participants were also more likely to choose the target when it was option B than when it was A (b = -0.45, z = -3.32, p < .001, 95% CI [-0.73, -0.19]), suggesting B may have generally been considered a more attractive smartphone. The sequence manipulation had no effect, that is, there is no evidence that the aspect listing procedure itself impacted choice (b = -0.02, z = -0.14, p = .885, 95% CI [-0.29, 0.25]).

Analysis of Reason Structure. Overall, participants provided 2.1 reasons on average. No participants reached the limit of ten reasons. Figure 2 shows the mean number of reasons per condition.

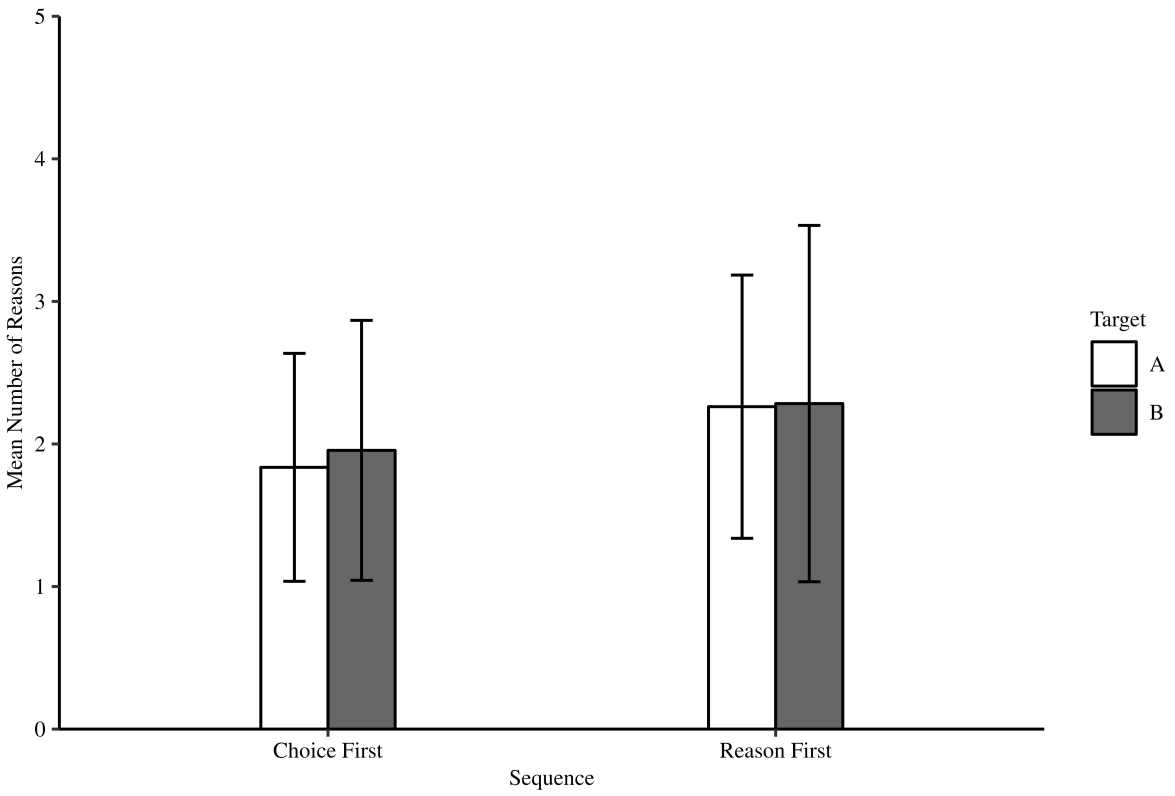


Figure 2
Mean number of reasons by conditions in Experiment 1 (error bars are standard deviations).

Reason Structure and Choices. As predicted by Query Theory, option A was more likely to be chosen by participants who generated more reasons in its favour (as captured by our Content Difference variable), and those who generated reasons in its favour earlier (as captured by our Order Difference variable). Both effects were statistically significant, as shown by the regression results below.

Quantity of Reasons: A logistic regression shows that Content Difference is a significant predictor of choosing A (b = 4.92, z = 6.19, p < .001, 95% CI [3.68, 7.12]). The sequence manipulation had no effect on choices (b = -0.20, z = -0.54, p = .589, 95% CI [-1.02, 0.52]), and did not interact with Content Difference (b = 0.74, z = 0.93, p = .350, 95% CI [-0.65, 2.92]).[5]

Positions of Reasons: A logistic regression shows that Order Difference is a significant predictor of choosing A (b = 6.60, z = 8.48, p < .001, 95% CI [5.25, 8.36]). The sequence manipulation had no effect (b = -0.35, z = -1.14, p = .255, 95% CI [-1.03, 0.21]), nor did it

[5] See Model 1 of Table 1 in the Supplementary Materials for the regression table.

interact with Order Difference (b = 1.00, z = 1.28, p = .199, 95% CI [-0.43, 2.73]).[6]

To our knowledge this is the first evidence that Query Theory can be used to predict multialternative choices.

Reason Structure and The Attraction Effect. Thus far, we have shown that we have successfully replicated the attraction effect, and that the patterns of reasons generated by participants are consistent with the general predictions of Query Theory that participants generate reasons in favour of the chosen option earlier and in greater quantity. To directly test whether the reason-based mechanism embedded in Query Theory is a possible mechanism behind the attraction effect, we look at the likelihood that the target was chosen as a function of how frequently and how early reasons were generated in its favour, by focusing on the $Content_{Target}$ and $Order_{Target}$ scores.

Figure 3 shows a very clear pattern: the target was virtually always chosen when either more reasons were generated supporting the target (left panel), or reasons supporting the target were generated earlier (right panel). As shown in the below analyses, the effects of both Content and Order are statistically significant in the direction predicted by Query Theory.

Quantity of Reasons: A logistic regression shows that $Content_{Target}$ is a significant predictor of the likelihood of choosing the target (b = 7.45, z = 7.64, p < .001, 95% CI [5.80, 9.72]), and that this effect did not differ depending on whether the target was A or B (b = -0.45, z = -0.49, p = .625, 95% CI [-2.44, 1.32]). The proportion of reasons supporting the decoy ($Content_{Decoy}$) did not affect choice (b = 0.39, z = 0.36, p = .715, 95% CI [-1.79, 2.52]), nor did it interact with the Target variable (b = -1.20, z = -1.11, p = .266, 95% CI [-3.33, 0.98]). Neither the Target (b = -0.14, z = -0.29, p = .776, 95% CI [-1.08, 0.88]) nor the Sequence (b = 0.29, z = 0.92, p = .357, 95% CI [-0.30, 0.94]) manipulation had a significant effect on choices.[7]

Positions of Reasons: A logistic regression shows that $Order_{Target}$ is a significant predictor of the likelihood of choosing the target (b = 11.01, z = 8.81, p < .001, 95% CI [8.78, 13.74]), and that this effect did not differ depending on whether the target was A or B (b =

[6] See Model 2 of Table 1 in the Supplementary Materials for the regression table.

[7] See Model 1 of Table 2 in the Supplementary Materials for the regression table.

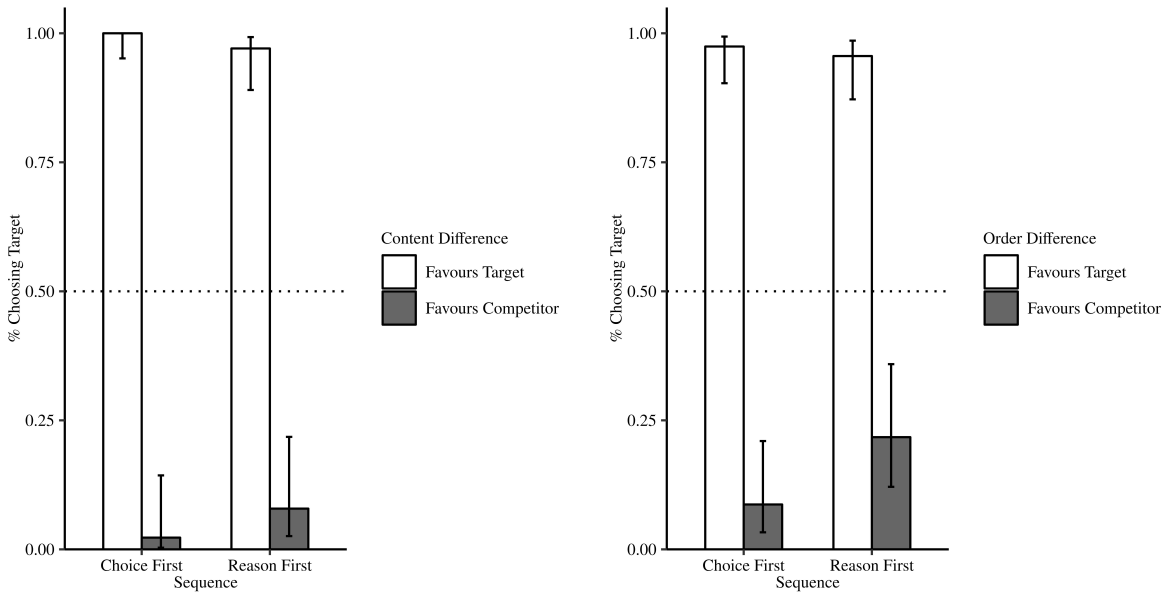


Figure 3

(Left) Proportion of participants choosing the target option by Sequence and Content Difference in Experiment 1. (Right) Proportion of participants choosing the target option by Sequence and Order Difference in Experiment 1. Error Bars are confidence intervals of logistic regression models.

-1.57, z = -1.31, p = .189, 95% CI [-4.08, 0.73]). $Order_{Decoy}$ did not influence choice (b = -0.40, z = -0.32, p = .750, 95% CI [-2.76, 2.18]), nor did it interact with which option was the target (b = -0.69, z = -0.55, p = .584, 95% CI [-3.21, 1.74]). The target manipulation did not affect choices (b = 0.04, z = 0.10, p = .924, 95% CI [-0.87, 1.05]), nor did Sequence (b = 0.42, z = 1.64, p = .100, 95% CI [-0.06, 0.94]).[8]

These results show that Query Theory can explain the attraction effect: when individuals deliberate their choice between a target and a competitor, the presence of a decoy leads participants to generate more reasons in support of the Target earlier in the deliberation process. That is, a theoretically irrelevant change in the decision context shapes peoples' reasoning, and in so doing affects their choices.

Exploratory Analysis

The Effect of Target on Reasons. In addition to the above preregistered analysis, we tested whether, independently of which option was chosen, participants generated more reasons supporting the target, and did so earlier in their deliberation. We found evidence for

[8] See Model 2 of Table 2 in the Supplementary Materials for the regression table.

both effects, supporting Query Theory's notion of output interference.

Quantity of Reasons: A linear regression model on Content Difference shows that reasons supporting option A were generated in greater quantity when option A was the target (b = 0.24, t = 4.44, p < .001, 95% CI [0.13, 0.34]).

Positions of Reasons: A linear regression model on Order Difference shows that reasons supporting option A were generated earlier when option A was the target (b = 0.11, t = 3.89, p < .001, 95% CI [0.05, 0.16]).

Weights of Reasons. As a further test of the notion of output interference, we investigated whether early reasons were rated as more important than later ones. As the majority of participants generated either two or three reasons, we ran two separate linear regression models. The first model shows that, for participants who provided exactly two reasons (41.50% of participants), their second reason was rated as significantly less important than their first reason (b = -43.70, t = -6.70, p < .001, 95% CI [-56.56, -30.83]). For those who provided exactly three reasons (22.13%), the second model shows that, later reasons were rated as less important than earlier ones (b = -7.92, t = -2.06, p = .041, 95% CI [-15.51, -0.33]).

Experiment 2

In Experiment 1, we found the attraction effect, as expected. In line with Query Theory, the quantity and order of distinct reasons predicted multialternative choices and explained the attraction effect, such that reasons supporting the chosen option as well as the target were generated earlier and in greater number. In Experiment 2, we extended the attraction effect paradigm beyond two attributes. For Query Theory, this allows us to test the model's predictions about aspect listing in a more information-rich environment. For the attraction effect, this provides a fundamental test, examining whether it can generalise beyond simple two-attribute items, to items with additional attributes that either provide additional discriminating information, or are identical for all options.

Method

The second experiment used a 2 Target (A vs B) × 3 Complexity (Standard vs Non-discriminating vs Discriminating) design.[9] The Sequence manipulation was not included

[9] See Supplementary Materials for screenshots of Experiment 2.

as Experiment 1 found no evidence that the order of aspect listing and choice has an effect on either reasons or preference. Therefore, aspect listing took place before the choice stage in all conditions.

The Complexity manipulation varied the number of attributes and their ability to discriminate between the target and the competitor. The choice sets used in Experiment 2 are shown in Table 2. In the *Standard* condition, the choice sets were the same as in Experiment 1 and each option had two attributes. In the *Non-discriminating* condition, each option had four attributes, as a result of the addition of two more attributes to the Standard choice sets. However, because the values of these additional attributes were exactly the same for all three options, the target dominated the decoy on only the two original attributes, and the two additional attributes did not provide useful information to discriminate between target and competitor. Finally, in the *Discriminating* condition, also with four attributes, the two additional attributes each favoured one of the main options. In this condition, the additional attributes provided useful information to discriminate between the options, and the target dominated the decoy on all four attributes.

Table 2

Choice sets used in Experiment 2.

Conditions	Smartphone	Storage Capacity	Malfunc. Rate	Battery Life	Camera Resolution
Standard	A	16 GB	3%	-	-
	B	32 GB	5%	-	-
	D_A	12 GB	3.5%	-	-
	D_B	28 GB	5.5%	-	-
Non-discrim.	A	16 GB	3%	12 Hours	18 MP
	B	32 GB	5%	12 Hours	18 MP
	D_A	12 GB	3.5%	12 Hours	18 MP
	D_B	28 GB	5.5%	12 Hours	18 MP
Discriminating	A	16 GB	3%	16 Hours	12 MP

Conditions	Smartphone	Storage Capacity	Malfunc. Rate	Battery Life	Camera Resolution
	B	32 GB	5%	8 Hours	24 MP
	D_A	12 GB	3.5%	15 Hours	10 MP
	D_B	28 GB	5.5%	7 Hours	22 MP

The experiment was implemented on Qualtrics and participants were recruited from Prolific. We planned to recruit 720 participants as it would maximise our sample size given our budget constraints. Only English speakers were allowed since reason listing was involved. Only users of desktop devices were allowed. Participants answered the same attention check question as in Experiment 1, and provided some basic demographic information (gender, age, country of residence, and first language).

The experiment was approved internally by the ethics board at the University of Warwick.

Predictions and Research Questions

Experiment 2 had a number of preregistered research questions.[10], including the same set of predictions as Experiment 1.

For the attraction effect, we predicted that the choice share of an option would be higher when it was the target than when it was not. Further, we were interested in whether the size of the effect depended on the complexity of the stimuli.

From Query Theory, we predicted that reasons supporting the chosen option would be generated earlier and in greater quantity. Conditional on finding the attraction effect, we expected that reasons supporting the target would be generated in greater quantity and earlier. Finally, we were interested in whether our Complexity manipulation affected the quantity and order of reasons.

[10] The preregistration can be found on OSF Registries: https://osf.io/kmn8b

Results

723 UK participants whose first language is English were recruited (65% female, 35% male, mean age = 36). Four participants were excluded as they failed the attention check, leaving 719 for the analysis.

Preregistered Analysis

Very few participants chose the decoy (no more than 4% in any condition). As in Experiment 1, participants who chose the decoy were excluded from the analysis, leaving 708 participants.

The Attraction Effect. Figure 4 clearly shows an attraction effect in all conditions: in line with the results of Experiment 1, and regardless of the number of attributes, participants chose the target more frequently than the competitor. While adding two discriminating attributes to the smartphones did change the relative attractiveness of A and B, with A becoming more attractive as shown by the increase in the target A bar and decrease in target B, this did not affect the size of the attraction effect. Averaging across the two targets in all conditions produces a similar sized preference for the target over its competitor (63.87% in the Standard condition, 65.38% in Non-discriminating, 65.68% in Discriminating, as shown by the black diamonds in Figure 4).

These results are statistically supported by a logistic regression model: the target was significantly more likely to be chosen than the competitor (b = 0.67, z = 4.60, p < .001, 95% CI [0.39, 0.96]). As shown in Figure 4, the likelihood of the target being chosen did not differ depending on whether the choice options included two non-discriminating attributes (b = -0.03, z = -0.16, p = .871, 95% CI [-0.44, 0.37]) or two discriminating attributes (b = -0.02, z = -0.08, p = .939, 95% CI [-0.41, 0.38]). As in Experiment 1, with only two attributes, participants were more likely to choose the target when it was option B (b = -0.61, z = -4.17, p < .001, 95% CI [-0.90, -0.33]). This effect was not affected by the inclusion of two non-discriminating attributes (as shown by the non-significant interaction term: b = -0.09, z = -0.42, p = .675, 95% CI [-0.49, 0.32]), but was weaker with two discriminating attributes (b = 0.67, z = 3.36, p < .001, 95% CI [0.28, 1.07]).[11]

[11] See Table 3 in the Supplementary Materials for the regression table of this model.

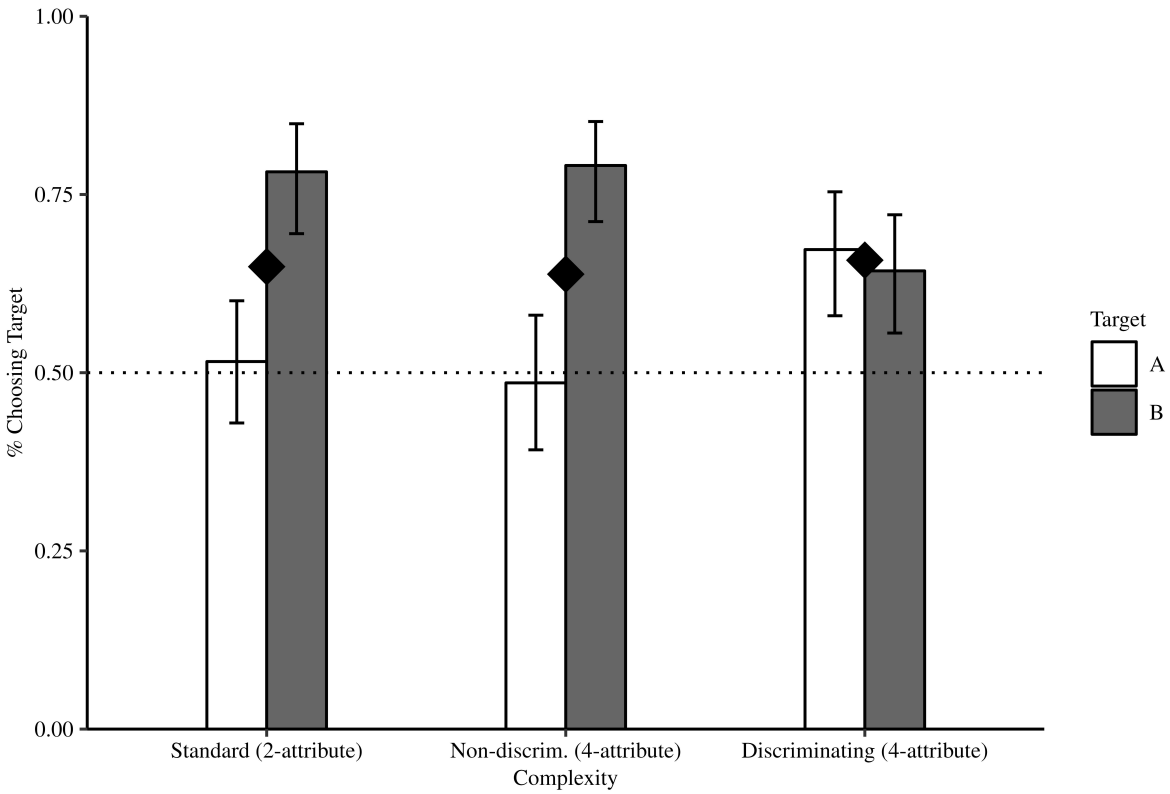


Figure 4
Proportion of participants choosing the target option by conditions in Experiment 2 (error bars are confidence intervals of a logistic regression model, black diamonds are the mean proportion in each Complexity condition).

Analysis of Reason Structure. Participants provided an average of 2.9 reasons. Fewer than 1% of all participants reached the limit of ten reasons. The breakdown by condition is shown in Figure 5. We expected that adding more attributes would increase the number of reasons people generated. A Poisson regression model shows that, adding two discriminating attributes increased the number of reasons provided by participants (b = 0.24, z = 4.44, p < .001, 95% CI [0.13, 0.34]), while adding two non-discriminating attributes had no effect (b = 0.06, z = 1.01, p = .311, 95% CI [-0.05, 0.17]). Whether the target was A or B did not affect the number of reasons generated by participants in any of the conditions.

Reason Structure and Choices. As in Experiment 1, Content Difference and Order Difference were used to study the reasons generated by participants.

Quantity of Reasons: Consistent with Query Theory and replicating the findings of Experiment 1, a logistic regression shows that, when there were only two attributes, participants were more likely to choose option A when they generated more reasons

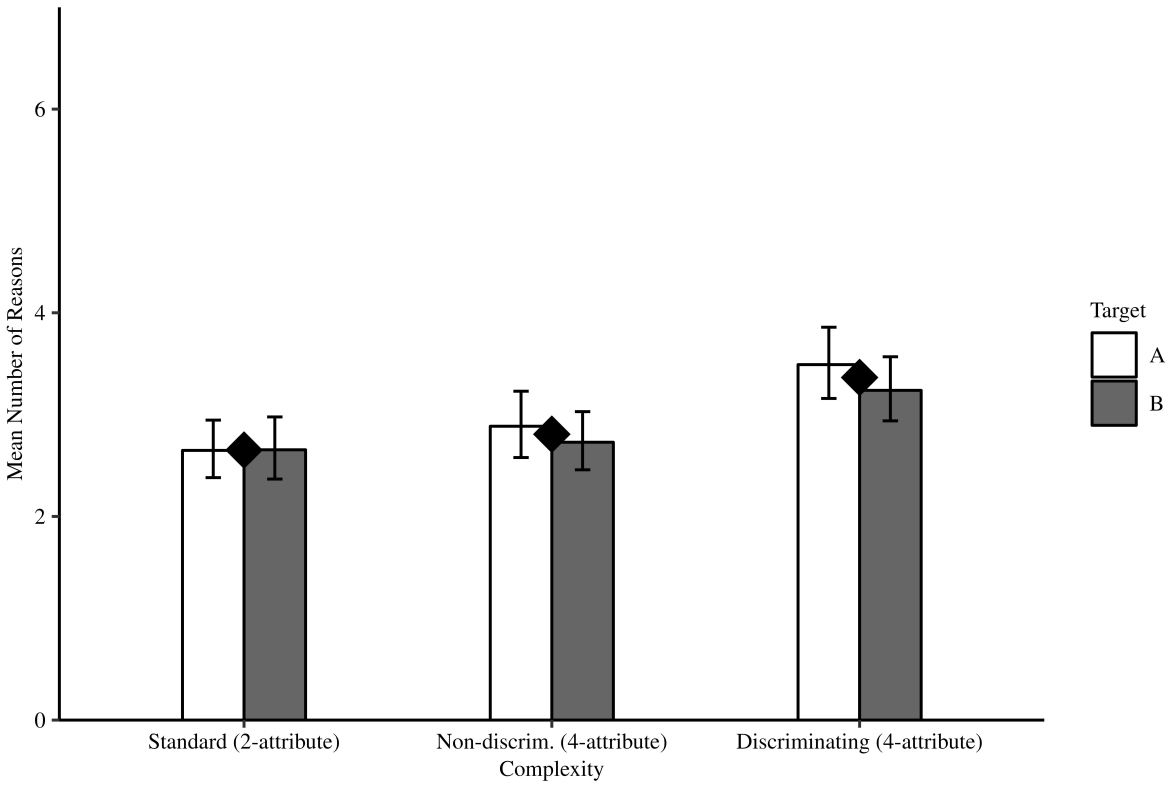


Figure 5
Mean number of reasons by conditions in Experiment 2 (error bars are confidence intervals from a Poisson regression model, black diamonds are the mean number of reasons in each Complexity condition).

supporting it than B (b = 5.69, z = 6.02, p < .001, 95% CI [4.15, 7.94]). The size of the effect was not significantly different when two non-discriminating attributes (b = 0.14, z = 0.11, p = .916, 95% CI [-2.63, 2.94]) or two discriminating attributes were added (b = -0.74, z = -0.63, p = .530, 95% CI [-3.30, 1.51]). The complexity level of the choice set was included as a control variable (for full results see Model 1 in Table 3).

Positions of Reasons: An analogous logistic regression shows that, again replicating the findings of Experiment 1, when there were two attributes participants were more likely to choose option A when they generated reasons supporting A earlier than reasons supporting B (b = 5.06, z = 9.19, p < .001, 95% CI [4.05, 6.22]). This effect was stronger when two non-discriminating attributes were added (b = 2.13, z = 2.00, p = .046, 95% CI [0.14, 4.39]), but not with two additional discriminating attributes (b = 0.54, z = 0.64, p = .522, 95% CI [-1.10, 2.24]). The complexity level of the choice set was again included as a control variable (Model 2 in Table 3).

Table 3
The effects of the proportions and positions of reasons supporting options A and B on choice, controlling for the complexity manipulation, in Experiment 2.

	Dependent variable:	
	Choice (A vs B)	
	Model 1	Model 2
	(1)	(2)
Content Difference	5.69***	
	(0.94)	
Order Difference		5.06***
		(0.55)
Complexity (Non-Discriminating vs Standard)	0.53	0.06
	(0.39)	(0.31)
Complexity (Discriminating vs Standard)	1.12***	0.38
	(0.36)	(0.28)
Content Difference × Complexity (Non-Discriminating)	0.14	
	(1.36)	
Content Difference × Complexity (Discriminating)	−0.74	
	(1.18)	
Order Difference × Complexity (Non-Discriminating)		2.13**
		(1.07)
Order Difference × Complexity (Discriminating)		0.54
		(0.84)
Intercept	−0.81***	−0.33
	(0.28)	(0.20)
Observations	708	708
Log Likelihood	−146.50	−232.52
Akaike Inf. Crit.	305.00	477.04
Note:	*p<0.1; **p<0.05; ***p<0.01	

Reason Structure and the Attraction Effect. As in Experiment 1, we investigated whether Query Theory can explain the attraction effect. Figure 6 clearly shows that the target was chosen by the vast majority of participants who generated more reasons supporting the target (left panel) or generated those reasons earlier (right panel). As shown in the following analyses, both results were significant and supported by statistical models.

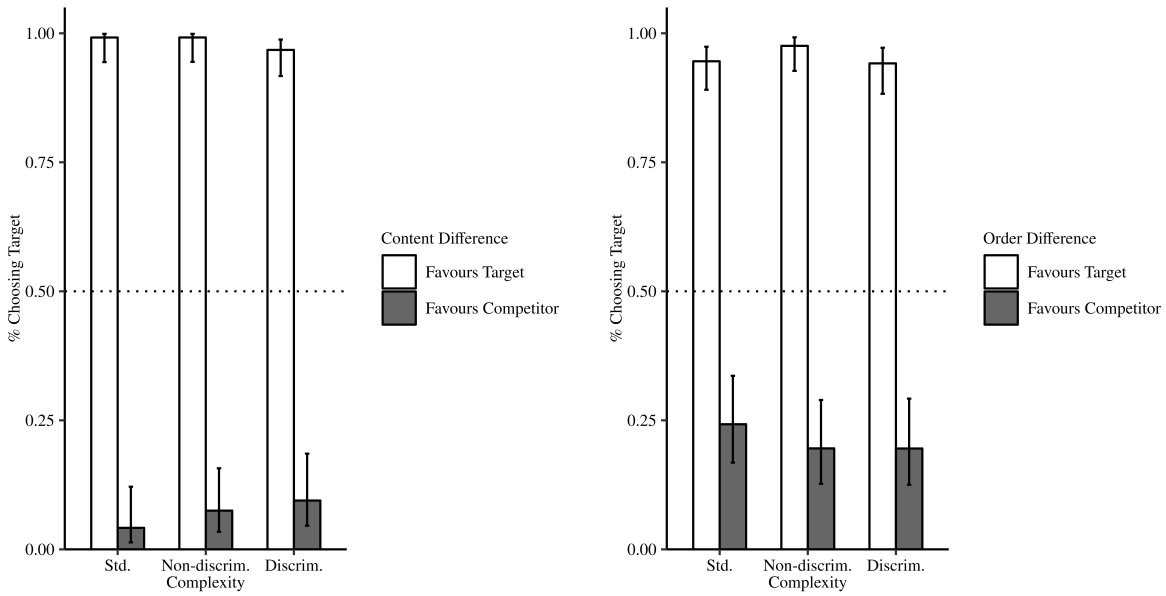


Figure 6
(Left) Proportion of participants choosing the target option by Complexity and Content Difference in Experiment 2. (Right) Proportion of participants choosing the target option by Complexity and Order Difference in Experiment 2. Error bars are confidence intervals of logistic regression models.

Quantity of Reasons: A logistic regression model shows that participants who generated more reasons supporting the target were more likely to choose it ($Content_{Target}$: b = 7.80, z = 13.06, p < .001, 95% CI [6.71, 9.06]), and this effect was the same regardless of whether the target was option A or B (b = -0.38, z = -0.64, p = .520, 95% CI [-1.59, 0.78]). The number of reasons supporting the decoy did not impact choice ($Content_{Decoy}$: b = -0.86, z = -1.42, p = .156, 95% CI [-2.03, 0.35]), nor interact with whether the target was A or B (b = 0.48, z = 0.81, p = .420, 95% CI [-0.69, 1.66]). The target manipulation and the level of complexity were both included as control variables (Model 1 in Table 4).

Positions of Reasons: Another logistic regression model similarly shows that the positions of reasons supporting the target could predict choices. That is, when participants generated reasons supporting the target earlier, they were more likely to choose it ($Order_{Target}$:

b = 11.14, z = 12.43, p < .001, 95% CI [9.51, 13.05]). This effect did not interact with whether the target was option A or B (b = -1.20, z = -1.34, p = .182, 95% CI [-3.06, 0.53]). The position of reasons supporting the decoy did not influence choice ($Order_{Decoy}$: b = -0.41, z = -0.73, p = .467, 95% CI [-1.51, 0.72]), and there was no interaction with Target (b = 0.94, z = 1.67, p = .095, 95% CI [-0.16, 2.05]). Similar to the previous model, the target manipulation and the level of complexity were included as control variables (Model 2 in Table 4).

The results of Experiment 2, corroborate our earlier findings that Query Theory can explain the attraction effect. By doing so in an environment characterized by greater complexity, they also show that the attraction effect is not confined to options defined on just two attributes, but extends to cases in which there are more attributes.

Complexity and Reason Structure. We further investigated whether the level of complexity had an impact on the proportion and position of reasons, that is, whether the number of attributes influenced the reasoning process as described by Query Theory.

Quantity of Reasons: A linear regression model shows that, when controlling for the target manipulation (b = 0.16, z = 5.53, p < .001, 95% CI [0.10, 0.21]), the relative share of reasons supporting A versus B did not change with the addition of two non-discriminating attributes (b = -0.09, t = -1.34, p = .179, 95% CI [-0.23, 0.04]) or two discriminating attributes (b = 0.14, t = 1.94, p = .053, 95% CI [-0.0018, 0.27]). Another linear regression model shows that, again controlling for the target manipulation (b = -0.07, z = -4.86, p < .001, 95% CI [-0.10, -0.04]), the relative share of reasons favouring the target did not change with two additional non-discriminating attributes (b = 0.03, t = 0.72, p = .472, 95% CI [-0.05, 0.10]) or discriminating attributes (b = -0.02, t = -0.46, p = .647, 95% CI [-0.09, 0.06]).

Positions of Reasons: A linear regression model shows that, when the target manipulation is controlled for (b = 0.06, t = 4.05, p < .001, 95% CI [0.03, 0.09]), the positions of reasons between the two options were not significantly affected by the addition of two non-discriminating attributes (b = -0.04, t = -0.98, p = .328, 95% CI [-0.11, 0.04]). However, participants did generate reasons in favour of A earlier when there were two additional discriminating attributes (b = 0.10, t = 2.76, p = .006, 95% CI [0.03, 0.17]), which is consistent with Query Theory given that the discriminating case showed a stronger preference

Table 4
The effects of the proportions and positions of reasons supporting the target option and the decoy on choice, controlling for the target and complexity manipulations, in Experiment 2.

	Dependent variable:	
	Choice (Target vs Competitor)	
	Model 1	Model 2
	(1)	(2)
$Content_{Target}$	7.80***	
	(0.60)	
$Content_{Decoy}$	−0.86	
	(0.60)	
$Order_{Target}$		11.14***
		(0.90)
$Order_{Decoy}$		−0.41
		(0.57)
Target (A vs B)	−0.16	−0.02
	(0.32)	(0.40)
Complexity (Non-Discriminating vs Standard)	−0.36	−0.09
	(0.37)	(0.31)
Complexity (Discriminating vs Standard)	0.09	−0.24
	(0.34)	(0.30)
$Content_{Target}$ × Target	−0.38	
	(0.60)	
$Content_{Decoy}$ × Target	0.48	
	(0.60)	
$Order_{Target}$ × Target		−1.20
		(0.90)
$Order_{Decoy}$ × Target		0.94*
		(0.56)
Intercept	−3.26***	−3.39***
	(0.36)	(0.42)
Observations	708	708
Log Likelihood	−167.31	−222.48
Akaike Inf. Crit.	350.62	460.96
Note:	*p<0.1; **p<0.05; ***p<0.01	

for A than the standard. Finally, another linear regression model shows that, controlling for the target manipulation (b = -0.03, z = -3.62, p < .001, 95% CI [-0.05, -0.01]), the average positions of reasons supporting the target did not differ when there were two non-discriminating attribute (b = -0.0017, t = -0.08, p = .934, 95% CI [-0.04, 0.04]) or two discriminating attributes (b = 0.02, t = 0.94, p = .349, 95% CI [-0.02, 0.06]).

The above results show that the reason generation pattern was not affected by the complexity levels of the choice set in notable ways.

Exploratory Analysis

Weights of Reasons. As in Experiment 1, we tested the effect of the positions of reasons on their subjective importance with three separate linear regression models. Since the majority of participants (90.68%) provided four reasons or fewer, running a separate regression for each unique number of reasons allows us to identify major patterns. Results show that later reasons were rated as less important than earlier ones for participants who provided exactly two reasons (29.94% of participants; b = -11.86, t = -6.08, p < .001, 95% CI [-15.70, -8.02]), exactly three reasons (32.20%; b = -9.43, t = -9.41, p < .001, 95% CI [-11.39, -7.46]), and exactly four reasons (17.09%; b = -7.08, t = -7.49, p < .001, 95% CI [-8.93, -5.22]). This is consistent with Experiment 1, which shows that earlier reasons were rated as more important and provides additional support for Query Theory.

Discussion

This paper is the first application of Query Theory to multialternative choices. By focusing on the well-known attraction effect, we have provided evidence that context effects can be understood through the lens of high-level deliberation processes. In line with the general prediction of Query Theory, reasons in support of the chosen option were generated earlier and in greater quantity than reasons supporting the other options. When one option was promoted as target by the presence of an asymmetrically dominated decoy, that option was both chosen more frequently than its competitor and supported by more reasons, which were generated by participants earlier during their deliberation. Crucially, our investigation has provided further controlled evidence that collecting data on high-level deliberation does not distort participants' choices, paving the way for the investigation of reason-based decision

making in more complex multialternative settings.

We have also extended the attraction effect to more complex settings than those in which it is commonly observed. The presence of an asymmetrically dominated decoy had the usual effect even when we doubled the amount of information about the available options by adding two non-discriminating or, more interestingly, discriminating attributes. Notably, the magnitude of the effect was not affected by this increased complexity. This finding assumes particular significance in relation to recent unsuccessful attempts to observe the effect with stimuli other than simple options defined over two numerical attributes (Frederick et al., 2014; Trendl et al., 2021; Yang & Lynn, 2014). While our studies have retained the numerical nature of the attributes, finding that the attraction effect persists in the face of doubling the complexity of the stimuli suggests a degree of robustness that warrants further investigation. There may be naturalistic settings (e.g., price comparison websites for financial products) in which most attributes are numerical and in which the presence of decoy options may exert systematic influences on people's choices.

Finding that the type of reason-based decision making assumed by Query Theory can be extended to multialternative settings and is compatible with the attraction effect opens up obvious avenues for future research, for instance, the exploration of whether the same mechanisms can explain other context effects such as the compromise and the similarity effects. More fundamentally, it raises questions about the relationship between high- and low-level psychological processes. Recent contributions (Bhatia, 2013; Noguchi & Stewart, 2018; Roe et al., 2001; Trueblood et al., 2014; Usher & McClelland, 2004) have explained context effects using accumulator models of decision-making processes at a lower level than the explicit reasons considered by Query Theory. An important question is whether explicit reasons are a direct reflection of evidence accumulation or whether high- and low-level processes are to some degree complementary. Aspect listing alone cannot answer this question, since any process-tracing method can only examine a limited number of psychological processes, typically at the same cognitive level (Schulte-Mecklenbeck et al., 2017). A key challenge for future research will be to tackle this issue by combining aspect listing with attentional tracing methods.

Supplemental Materials (Tables)

Supplemental Materials (Tables)

Table 1
The effects of the proportions and positions of reasons supporting options A and B on choice, controlling for the sequence manipulation, in Experiment 1.

	Dependent variable:	
	Choice (A vs B)	
	Model 1	Model 2
	(1)	(2)
Content Difference	4.92***	
	(0.79)	
Order Difference		6.60***
		(0.78)
Sequence (Choice First vs Reason First)	−0.20	−0.35
	(0.38)	(0.30)
Content Difference × Sequence	0.74	
	(0.79)	
Order Difference × Sequence		1.00
		(0.78)
Intercept	−0.37	−0.59*
	(0.38)	(0.30)
Observations	253	253
Log Likelihood	−37.32	−58.82
Akaike Inf. Crit.	82.65	125.63
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 2
The effects of the proportions and positions of reasons supporting the target option and the decoy on choice, controlling for the target and sequence manipulations, in Experiment 1.

	Dependent variable:	
	Choice (Target vs Competitor)	
	Model 1	Model 2
	(1)	(2)
$Content_{Target}$	7.45***	
	(0.98)	
$Content_{Decoy}$	0.39	
	(1.08)	
$Order_{Target}$		11.01***
		(1.25)
$Order_{Decoy}$		−0.40
		(1.26)
Target (A vs B)	−0.14	0.04
	(0.48)	(0.47)
Sequence (Choice First vs Reason First)	0.29	0.42
	(0.31)	(0.25)
$Content_{Target}$ × Target	−0.45	
	(0.92)	
$Content_{Decoy}$ × Target	−1.20	
	(1.08)	
$Order_{Target}$ × Target		−1.57
		(1.19)
$Order_{Decoy}$ × Target		−0.69
		(1.26)
Intercept	−2.99***	−2.79***
	(0.48)	(0.48)
Observations	253	253
Log Likelihood	−46.15	−64.93
Akaike Inf. Crit.	106.30	143.87
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 3
The effects of the target and complexity manipulations on choice in Experiment 2.

	Dependent variable:
	Choice (Target vs Competitor)
	Model 1
Target (A vs B)	−0.61***
	(0.15)
Complexity (Non-Discriminating vs Standard)	−0.03
	(0.21)
Complexity (Discriminating vs Standard)	−0.02
	(0.20)
Complexity (Non-Discriminating vs Standard) × Target	−0.09
	(0.21)
Complexity (Discriminating vs Standard) × Target	0.67***
	(0.20)
Intercept	0.67***
	(0.15)
Observations	708
Log Likelihood	−436.95
Akaike Inf. Crit.	885.90
Note:	*p<0.1; **p<0.05; ***p<0.01

Supplemental Materials (Screenshots)

Supplemental Materials (Screenshots)

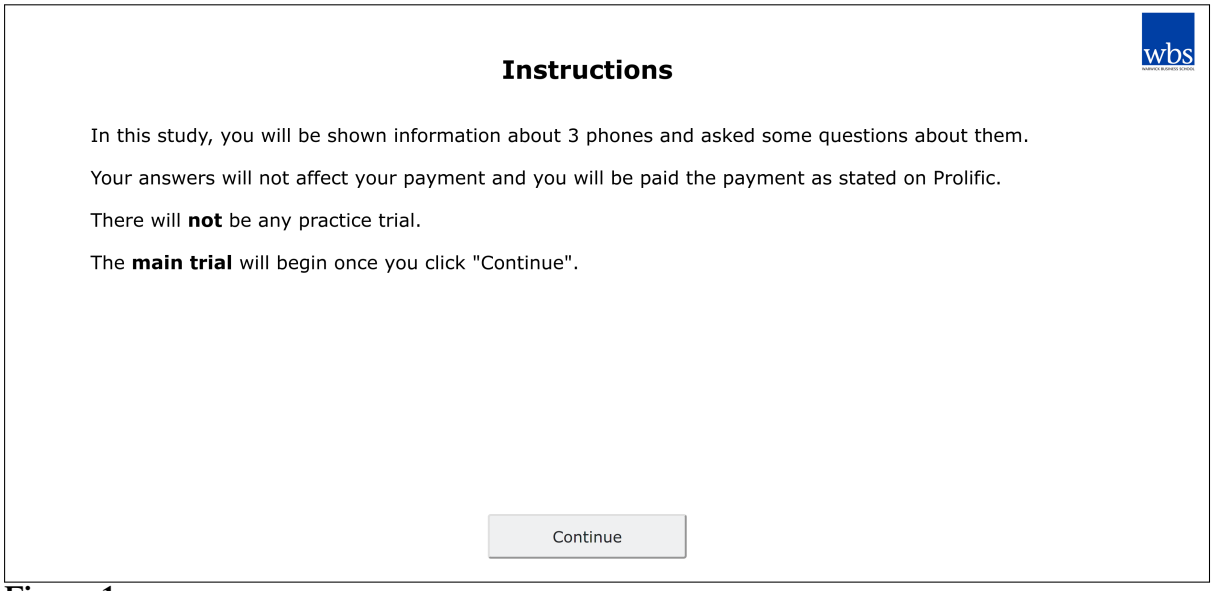


Figure 1

Instruction Screen of Experiment 1

	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Imagine that you were considering buying one of the above phones.

We are interested in the reasons which come to your mind as you are considering the options. Please list **all the reasons** that occur to you, one at a time. Type your first reason (max 130 characters) into the text field below, trying to be as clear and detailed as possible.

Click the "Submit Reason" button to submit your first reason.

Submit Reason

Figure 2

First Aspect Listing Screen of Experiment 1

	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Please continue to enter your reasons into the text field below, **one at a time** and trying to be as clear and detailed as possible.

Click the "Submit Reason" button to submit each reason.

When you have listed all the reasons which mattered for your choice, you can move on to the next stage by clicking the "No More Reasons" button.

When you click "No More Reasons", texts inputted in the box will **not** be recorded.

No More Reasons

Submit Reason

Figure 3

Second Aspect Listing Screen of Experiment 1

	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Which one would you choose?

- A
- B
- C

Continue

Figure 4

Choice Screen of Experiment 1

	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Below is a list of the reasons you provided when you were considering the options above.

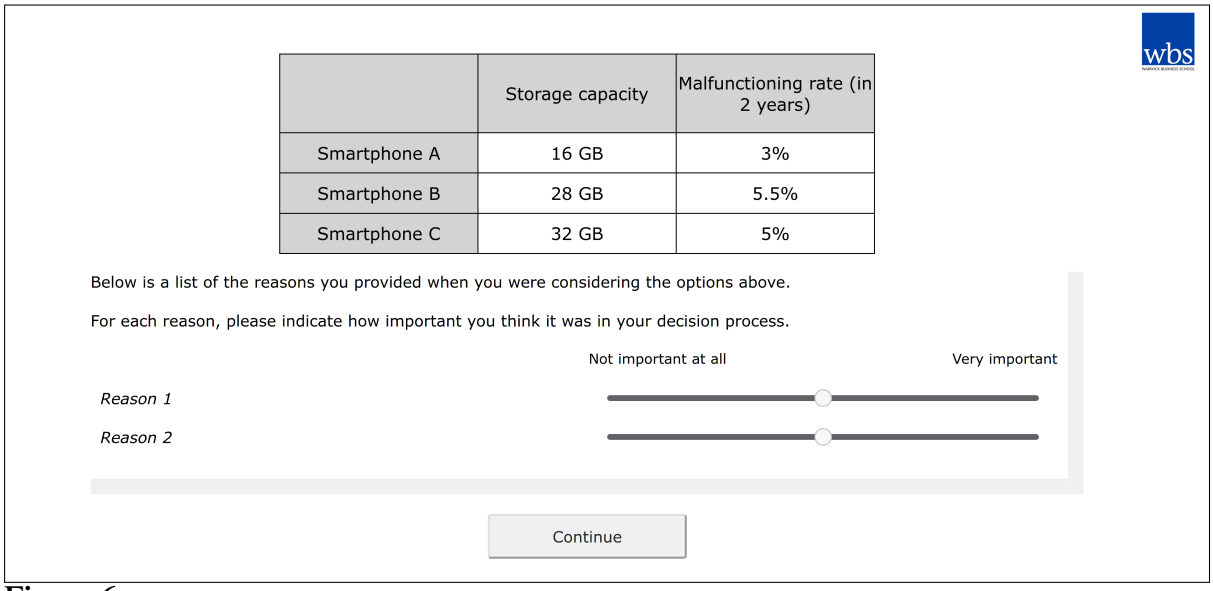
For each reason, please indicate which option(s) it supports. This can be done by ticking the corresponding box(es).

	Support A	Support B	Support C
Reason 1	☐	☐	☐
Reason 2	☐	☐	☐

Continue

Figure 5

Reason Coding Screen of Experiment 1



	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Below is a list of the reasons you provided when you were considering the options above.

For each reason, please indicate how important you think it was in your decision process.

Not important at all — Very important

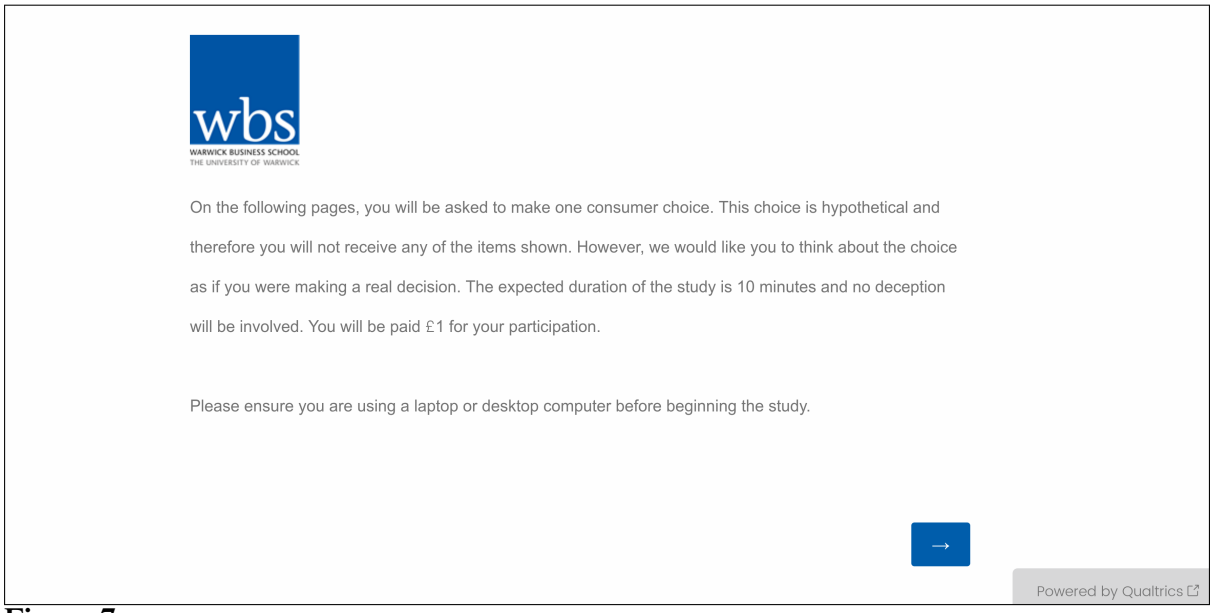
Reason 1

Reason 2

Continue

Figure 6

Reason Weighting Screen of Experiment 1



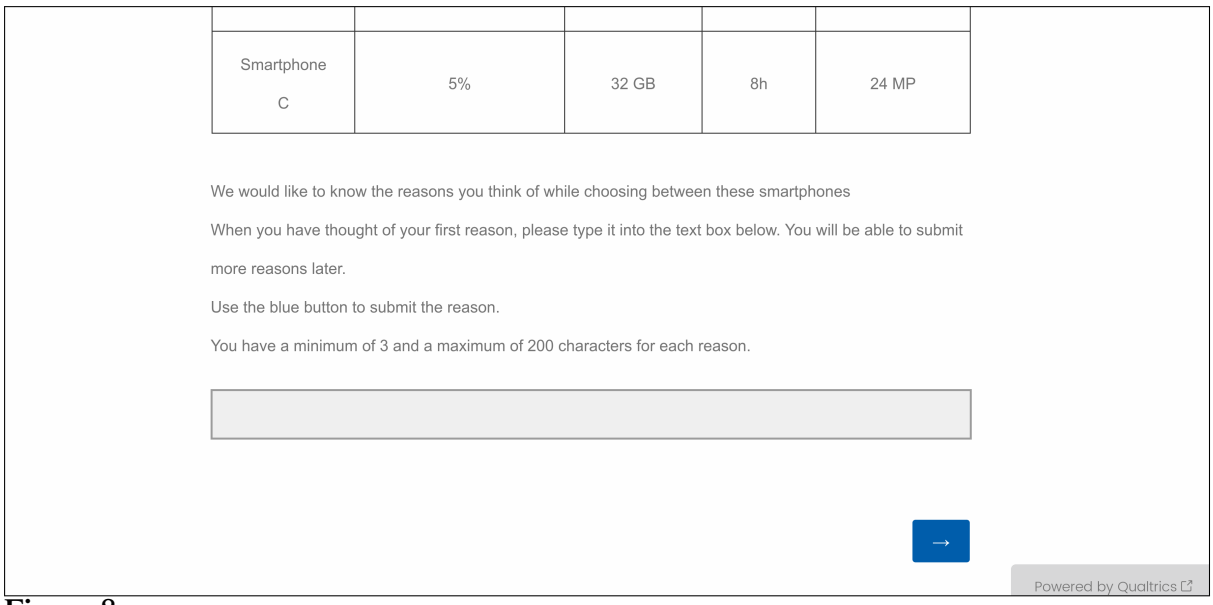
On the following pages, you will be asked to make one consumer choice. This choice is hypothetical and therefore you will not receive any of the items shown. However, we would like you to think about the choice as if you were making a real decision. The expected duration of the study is 10 minutes and no deception will be involved. You will be paid £1 for your participation.

Please ensure you are using a laptop or desktop computer before beginning the study.

Powered by Qualtrics

Figure 7

Instruction Screen of Experiment 2



Smartphone C	5%	32 GB	8h	24 MP

We would like to know the reasons you think of while choosing between these smartphones

When you have thought of your first reason, please type it into the text box below. You will be able to submit more reasons later.

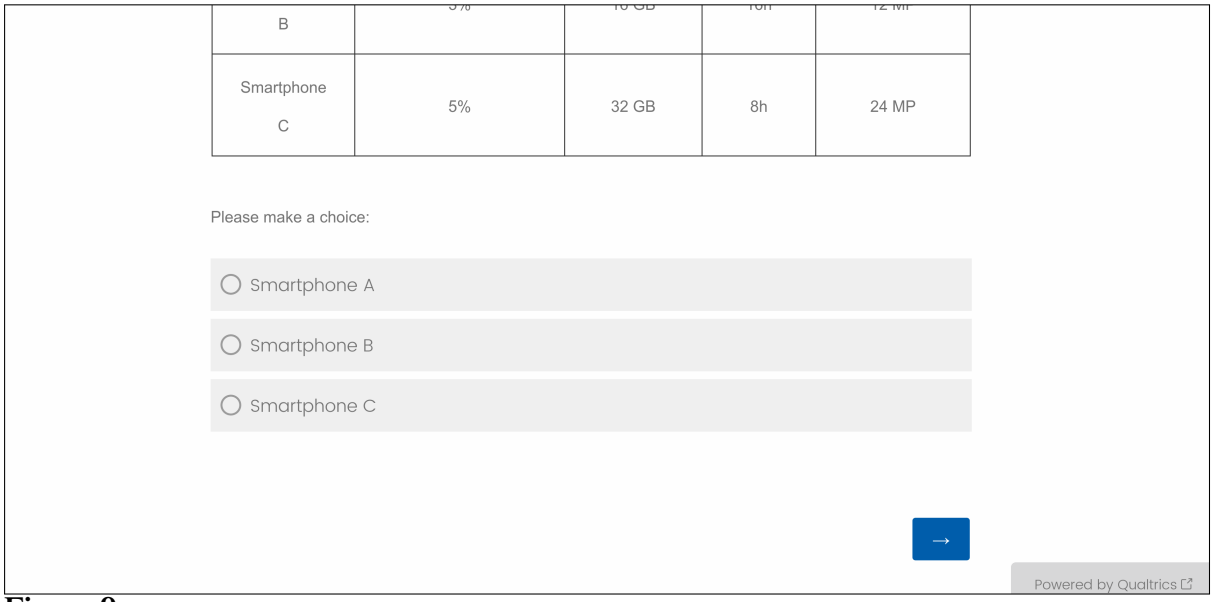
Use the blue button to submit the reason.

You have a minimum of 3 and a maximum of 200 characters for each reason.

Powered by Qualtrics

Figure 8

Aspect Listing Screen of Experiment 2



B	[illegible]	[illegible]	[illegible]	[illegible]
Smartphone C	5%	32 GB	8h	24 MP

Please make a choice:

- Smartphone A
- Smartphone B
- Smartphone C

Powered by Qualtrics

Figure 9

Choice Screen of Experiment 2

Smartphone B	3%	16 GB	16h	12 MP
Smartphone C	5%	32 GB	8h	24 MP

Below is a list of the reasons you provided when you were considering the options above. For each reason, please indicate which option(s) it supports. This can be done by ticking the corresponding box(es).

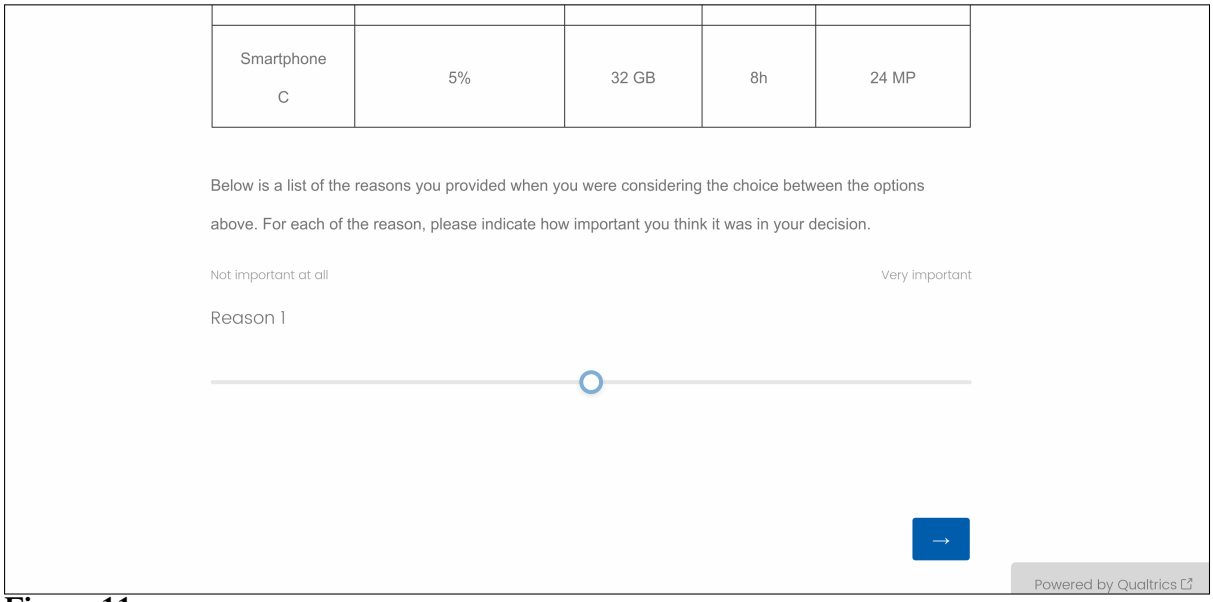
	Smartphone A	Smartphone B	Smartphone C
Reason 1	☐	☐	☐

→

Powered by Qualtrics

Figure 10

Reason Coding Screen of Experiment 2



Smartphone C	5%	32 GB	8h	24 MP

Below is a list of the reasons you provided when you were considering the choice between the options above. For each of the reason, please indicate how important you think it was in your decision.

Not important at all — Very important

Reason 1

→

Powered by Qualtrics

Figure 11

Reason Weighting Screen of Experiment 2

The interplay of multiple psychological processes underlying the attraction effect

Abstract

This paper is the first to examine the attraction effect with multiple process-tracing methods, namely mouse tracking and reason listing (N = 512). From a theoretical viewpoint, the present work allows us to study how mental processes of both high and low cognitive levels integrate to produce the effect, as well as testing the assumptions of existing cognitive models. Methodologically, the novel data obtained in this experiment provides triangulation on process-tracing methods and improves validity. After replicating the attraction effect, we first found that the quantity and position of reasons, as well as sampling patterns, could predict choice independently. We further found that reasoning partially mediated the effects of attentional processes on choice, while the frequency of mouse clicks also predicted the types of reasons generated by participants. These results demonstrate the intertwined roles of attentional processes and reasoning in the attraction effect, and provide the next step towards a deeper understanding of human cognition behind multialternative choices.

The interplay of multiple psychological processes underlying the attraction effect

Consider two options with different values on two attributes. Research has shown that, when a third option which is dominated by one of the two existing options on both attributes is added, the choice share of the dominating option can be increased (Huber et al., 1982; Simonson, 1989). For instance, imagine two smartphones, A (16 gigabytes storage capacity and a 3% probability of malfunctioning in the first two years) and B (32 GB storage and 5% malfunction rate). Adding smartphone D_A (12 GB storage and 3.5% malfunctioning rate) which is asymmetrically dominated by A but not B is expected to promote A. Here A is conventionally termed the *target*, B the *competitor*, and D_A the *decoy*. Conversely, adding decoy D_B (28 GB storage and 5.5% malfunctioning rate) is expected to promote B as the target.

This is formally known as the *attraction effect*, also called the asymmetric dominance effect or the decoy effect. The attraction effect is one of the most studied phenomena in behavioural science and has drawn considerable attention in many disciplines. For example, the effect was investigated and found in high-level decision tasks, including consumer (Heath & Chatterjee, 1995; Huber & Puto, 1983), political (Herne, 1997), risky (Cheng et al., 2012), job candidate (Highhouse, 1996; Slaughter et al., 1999), and legal choices (Kelman et al., 1996), as well as in low-level cognitive tasks such as perceptual decision making (Trueblood et al., 2013) and memory recall (Maylor & Roberts, 2007). The prevalence of the attraction effect in real-world applications has also been evaluated and debated (Frederick et al., 2014; Huber et al., 2014; Trendl et al., 2021; Yang & Lynn, 2014). Additionally, a variety of mathematical models have been proposed to explain the effect through different cognitive mechanisms (Bhatia, 2013; Noguchi & Stewart, 2018; Roe et al., 2001; Trueblood et al., 2014; Usher & McClelland, 2004). In sum, the attraction effect is of great interest to both researchers and practitioners in the field of decision making.

However, research which directly investigates the cognitive mechanisms underlying the attraction effect with process-tracing methods remains uncommon. With eye tracking, a handful of studies examined the roles of attentional and comparison patterns in the attraction effect (Król & Król, 2019; Marini et al., 2020; Noguchi & Stewart, 2014), that is, how

individuals attend to the options and their attribute values during the decision tasks, as well as how they compare the options. Noguchi & Stewart (2014) found that within-attribute transitions (i.e., shifting between options on the same attribute dimension) were more frequent than within-option ones (i.e., shifting between attribute values of the same option). Further, their results supported the notion that multialternative choices are made by comparing pairs of options on a chosen attribute, one pair at a time. Król & Król (2019) found that a decoy option could drive attention to the target, and that this effect was stronger when the attribute difference between the target and the decoy was larger. Furthermore, Marini et al. (2020) found that decision makers attended more to the target and, in tasks where the target and the decoy had the same values on some attributes, attended more to the attributes which allowed discrimination between the target and the decoy. Finally, Marini et al. (2020) also found that transitions between the target and the decoy were more common than other types of transitions.

A recent study focused on deliberation and the roles of distinct reasons in the attraction effect. With a protocol known as *aspect listing*, in which participants were instructed to provide reasons as they chose among three options, this first paper of this thesis showed that the reasoning process characterised by Query Theory (Johnson et al., 2007; Weber et al., 2007) can explain multialternative choices. The theory suggests that decision making involves the resolution of *queries* by sequentially generating reasons in favour of the options, while early reasons can suppress the generation of other reasons. Furthermore, context can change the deliberation process such that evidence in favour of a particular option is accumulated earlier, thus influence the number of reasons supporting each option and in turn choices. Results showed that the quantity and order of reasons generated by participants predicted choice, as such reasons supporting the chosen option were provided earlier and in greater amount than those supporting other options. Additionally, results showed that reasons were rated as more important when they were generated earlier in deliberation. According to this account, the attraction effect arises when the dominance relationship between the target and the decoy influences the order of decision queries and a higher amount of evidence supporting the target is gathered through reasoning.

This paper is the first to study the decision process of the attraction effect with both

attentional tracking and reason listing methods. While it is commonly accepted that process data is essential to understand cognitive mechanisms behind behavioural phenomena and test competing explanations (Johnson et al., 2008), it is less acknowledged that a single process-tracing method can only capture a limited number of mental processes, often at the same cognitive level (Schulte-Mecklenbeck et al., 2017). For instance, eye-tracking and mouse-tracking methods target attentional and information search processes, but cannot directly capture goals or attitude during decision making and rely on the assumptions that information being attended to is processed by high-level components in the cognitive system (Orquin & Holmqvist, 2018). Analysis of fixation time data particularly assumes that there is no substantial lag between attention and further processing of information. On the other hand, while aspect listing explicitly examines decision strategies and the importance of attributes or reasons, this protocol does not necessarily capture information acquisition or the frequency of comparisons, especially during the initial sampling phase. This highlights the advantages of using multiple process-tracing methods: it provides triangulation and improves validity of the methods (Holmqvist et al., 2011).

From a theoretical perspective, the goal of this study was to investigate different explanations of the psychological processes behind the attraction effect. Adopting the methodology of mouse tracking and aspect listing, we first focused on whether attentional patterns and reasons generated through deliberation could predict choices independently. Furthermore, we studied how information sampling and decision queries affected each other, and how they integrated and together impacted decision outcomes. Specifically, we examined whether reasoning mediated the effect of attentional patterns on choice, as well as whether mouse movement could predict the types of reasons when mouse-tracking data was divided into discrete stages for each distinct reason.

Method

This study involved one ternary choice task in which participants were asked to choose between three different smartphones, one of which was a decoy option. The attribute values of the smartphones are shown in Table 1. Smartphones A and B were shown to all participants, while either decoy D_A or D_B was shown to each of them (i.e., a target manipulation with two

conditions, *Target A* vs *Target B*).

Table 1. Choice sets used in this study.

Smartphone	Storage capacity	Malfunctioning rate
A	16 GB	3%
B	32 GB	5%
D_A	12 GB	3.5%
D_B	28 GB	5.5%

The choice set was presented to participants as a table in the top part of the screen, with the positions of the options and attributes randomised across participants. Stages before the choice was made, as detailed below, included mouse-tracking mechanisms. That is, all attribute values in the choice set table were initially hidden behind grey cells and an attribute value would be revealed when a participant clicked on its cell. Clicking on a cell hid the previously revealed cell, as such participants could only sample one attribute value at a time. Clicking anywhere outside of the choice set table also hid the revealed cell. The order in which each participant clicked on the cells, as well as the duration for which each cell was revealed, were recorded. Mouse-tracking mechanisms were not included after the choice was made, such that all attribute values in the choice set table were visible to the participants in those stages.

The main stages of the experiment are as follows:[1]

- Practice: Participants were given the opportunity to familiarise themselves with the mouse-tracking interface. They were presented with a table that had the same format and functionality as that in the main stages, but without the actual attribute values. Clicking a cell to reveal it would show placeholder text, e.g., “Feature 1 of phone A”. To ensure participant were familiar with the task, they were required to click at least three cells before they could progress.
- Aspect Listing: Participants were presented with the information about the choice options using the mouse-tracking mechanisms described above. Participants were

[1] See Supplemental Materials for screenshots of the Experiment.

instructed to input into a text box, one at a time, the reasons that came to their minds as they were considering the options and submit them. Each time they submitted a reason, the text box was cleared and the screen refreshed so they could continue to explore the attribute values of the options, provide another reason, or terminate the aspect listing procedure when they indicated that they had already input all the reasons which allowed them to make their decision. Participants were explicitly told to provide at least one reason, while the procedure would be terminated automatically when they had input 10 reasons.

- Choice: Participants were asked to choose one smartphone. Mouse-tracking mechanisms were included, with the same choice set table used in the aspect listing stage being shown as a final opportunity to explore the attributes before choices.
- Reason Coding: Participants were presented with the reasons which they had provided and told to code each reason indicating which option(s) it supported. Participants could select a single option or multiple options for each reason. The positions of the reasons were randomised across participants and did not reflect the order in which they were provided. Mouse-tracking mechanisms were not included in this stage and participants could view all attribute values at once.
- Reason Weighting: Participants were again presented with the reasons which they had input and were asked to rate the importance of each reason. The importance scales ranged from -100 to 100, but the exact values were not visible to participants. The positions of the reasons were the same as in the reason coding stage. Similar to the previous stage, mouse-tracking mechanisms were not included.

Participants were recruited from Prolific. Mobile users were not allowed into the experiment due to differences in both input and pointer methods and their potential impacts on process tracing. Finally, two attention check questions were included at the end of the experiment. The first was a multiple-choice question where participants were asked to select the type of product their choice was about, while the second explicitly asked whether they have paid full attention during the entire experiment and stated that they would be paid the participation fee regardless of their response.

The experiment was approved internally by the doctoral programme office at Warwick Business School.

Predictions

All predictions were preregistered.[2]

For the attraction effect, we predicted that the decoy option would not be chosen since it was asymmetrically dominated on both attributes. Further, for the two other options, we predicted that the choice share of an option would be higher when it was the target than when it was not.

For the roles of reasons in decision making, we based our predictions on Query Theory: we predicted that reasons supporting the chosen option would be generated earlier and in greater quantity. Additionally, conditional on finding the attraction effect, we predicted that reasons supporting an option would be generated earlier and in greater quantity when it was the target than when it was not.

Finally, for mouse-tracking data, we based our predictions on previous findings in eye-tracking studies, due to the similarities of the two methods. As previous findings showed that information search appeared to be more systematic with mouse tracking than eye tracking (Lohse & Johnson, 1996), we believed that patterns found in eye-tracking studies could be reproduced with mouse-tracking methods. We predicted that the chosen option would be attended to more frequently than other options (Shimojo et al., 2003; Simion & Shimojo, 2007). We further predicted that within-attributes transitions would be more common than within-option ones (Noguchi & Stewart, 2014), while transitions involving the chosen option would also be more frequent than transitions involving other options.

Indices

Reason Structure

We created two types of indices, termed *Content* and *Order* scores, to conduct analysis on the quantity and positions of reasons generated by participants.

The Content score of an option is the proportion of reasons supporting it among all

[2] The preregistration can be found on OSF Registries: https://osf.io/h2473

reasons generated by a participant, with a higher value showing that relatively more reasons supporting that option were provided. For instance, consider the three smartphones in our experiment (i.e., A, B, and either of the decoy options), the Content score for smartphone A of a participant is defined as:

$$Content_A = \frac{n_A}{N}$$

where n_A is the number of reasons supporting A provided by that participant and N is the total number of reasons. $Content_B$ and $Content_{Decoy}$ are defined analogously. Further, this index can be defined per the Target conditions, with $Content_{Target}$ and $Content_{Competitor}$ representing the proportion of reasons supporting the target and competitor respectively.

Additionally, the Order score of an option is the normalised average position of reasons supporting it among all reasons generated by a participant, with a higher value indicating that reasons supporting that option were provided relatively earlier. The Order score for smartphone A of a participant is defined as:[3]

$$Order_A = \begin{cases} 1 - (\frac{MR_A}{N+1}), & n_A > 0 \\ 1 - (\frac{N+1}{N+1}) = 0, & n_A = 0 \end{cases}$$

where MR_A is the median rank of reasons supporting A provided by that participant. $Order_B$ and $Order_{Decoy}$ are defined analogously. Further, $Order_{Target}$ and $Order_{Competitor}$ represent the normalised average position of reasons supporting the target and the competitor.

Finally, we define *Content Difference* (i.e., $Content_A - Content_B$) and *Order Difference* (i.e., $Order_A - Order_B$). When Content Difference is positive, it means that a participant generated more reasons supporting smartphone A than B. Similarly, when Order Difference is positive, it shows that reasons supporting smartphone A were provided earlier than those supporting B.

[3] This definition, which departs slightly from the one used in the preregistration, eases the interpretation of our results with no material effect, as it simply reverses the sign of the metric such that a higher value represents reasons being generated earlier.

Mouse-tracking Data

Two types of indices, *Frequency* and *Duration* scores, were used to analyse the frequency and duration of which each option was attended to by participants.

A Frequency score is the proportion of clicks which a participant performed to reveal the attribute values of an option among all clicks on the cells in the choice set table. The Frequency score for smartphone A of a participant is labelled $Frequency_A$. $Frequency_B$ and $Frequency_{Decoy}$ are termed analogously. We further defined *Frequency Difference* which is $Frequency_A - Frequency_B$ to analyse whether participants attended to smartphone A more often than B.

A Duration score, on the other hand, represents the duration for which the attribute values of an option were attended to among the total duration of all clicks. $Duration_A$, $Duration_B$, and $Duration_{Decoy}$ were created for each participant. *Duration Difference* which is $Duration_A - Duration_B$ was also used in analysis.

Finally, we created Transition scores, each of which is the proportion of transitions between a pair of options (e.g., sampling the attribute value of an option right after sampling that of another option) among all transitions carried out by a participant. The Transition score for smartphones A and B is labelled as $Transition_{A\text{-}B}$, regardless of the directions of the transitions. $Transition_{A\text{-}Decoy}$ and $Transition_{B\text{-}Decoy}$ are defined analogously. To analyse whether transitions involving smartphone A were more frequent than those involving B, we created *Transition Difference* which is defined as $Transition_{A\text{-}Decoy} - Transition_{B\text{-}Decoy}$.

Results

527 UK residents whose first language is English were recruited into the experiment (58% female, mean age = 34). 15 participants failed either of the attention check questions and their data were excluded. 512 participants remained.

Very few participants chose the decoy options (no more than 6% in either condition). Per the preregistration, the 21 participants who chose the decoy were excluded and choice was treated as binary hereafter. Hence, the data of 491 participants entered the main data analysis stage.

Preregistered Analysis

The Attraction Effect

Figure 1 shows the attraction effect and illustrates that smartphone A was chosen more often when the target was A than when it was B. This is statistically supported by a logistic regression model, which showed that participants were more likely to choose the target than the competitor on average (b = 0.46, z = 4.55, p < .001, 95% CI [0.27, 0.67]). Further, consistent with Figure 1, results also showed that participants were more likely to choose the target when the target was smartphone B than when it was A (b = -0.85, z = -8.37, p < .001, 95% CI [-1.06, -0.66])

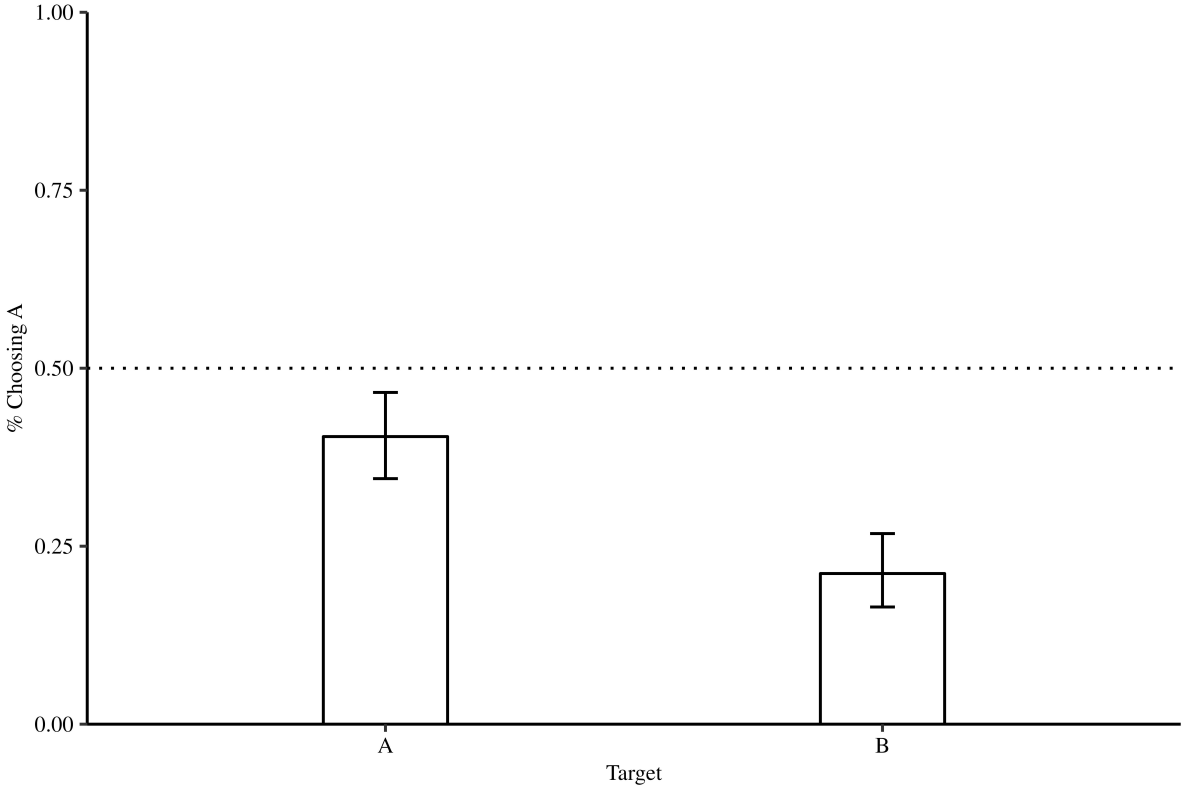


Figure 1
Proportion of participants choosing option A by the target manipulation (error bars are confidence intervals of a logistic regression model).

Analysis of Reason Structure

As shown in Figure 2, participants generated 2.56 reasons on average. A Poisson regression model showed the number of reasons submitted by participants did not depend on which option was the target (b = 0.03, z = 1.07, p = .283, 95% CI [-0.03, 0.09]).

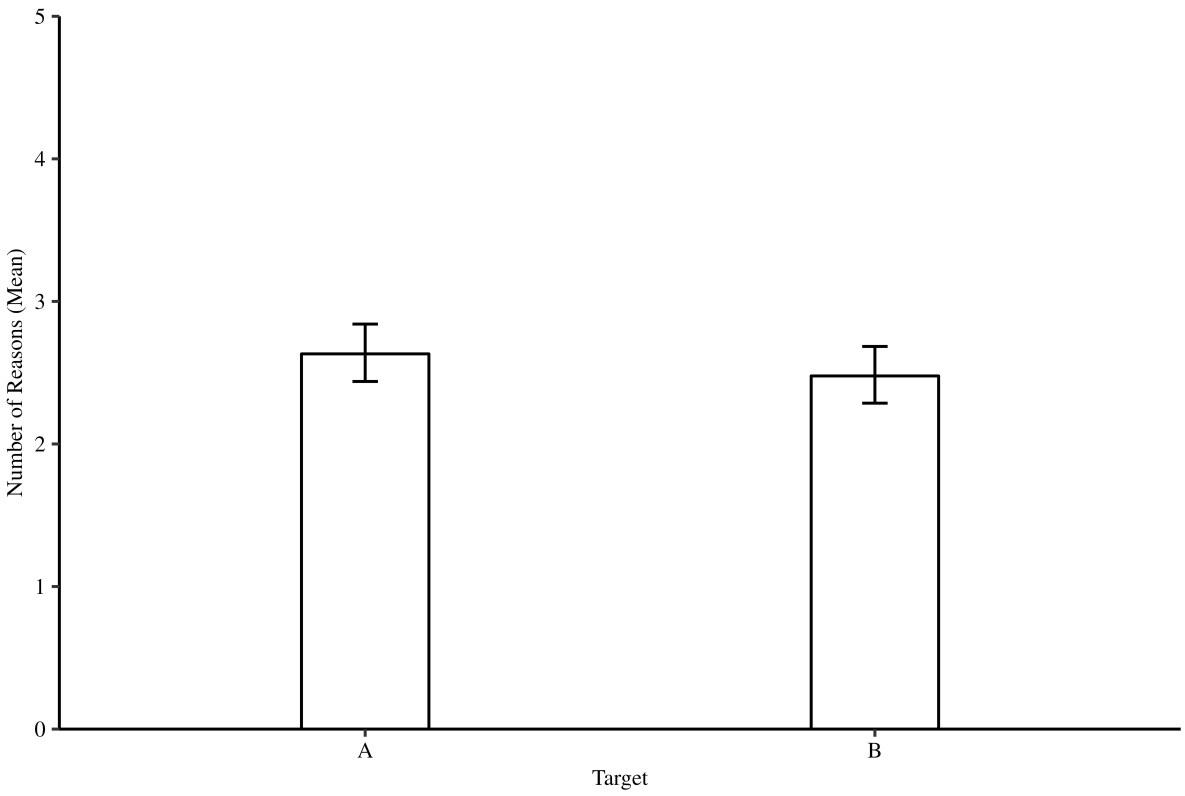


Figure 2
Mean number of reasons by the target manipulation (error bars are confidence intervals of a Poisson regression model).

Reason Structure and Choices.

Quantity of Reasons. Consistent with our predictions, a logistic regression model showed that participants were more likely to choose option A when they generated more reasons supporting smartphone A than B (as captured by our Content Difference variable: b = 6.23, z = 6.61, p < .001, 95% CI [4.69, 8.41]). This effect was weaker when the target was smartphone A, as shown by the significant interaction term (b = -2.01, z = -2.13, p = .033, 95% CI [-4.18, -0.45]). Participants were also more likely to choose option A than B when the target was option A (b = 0.68, z = 3.23, p = .001, 95% CI [0.29, 1.12]).

Positions of Reasons. Similarly, as predicted, participants were more likely to choose smartphone A when reasons supporting A were generated earlier (as captured by Order Difference: b = 4.70, z = 10.20, p < .001, 95% CI [3.86, 5.68]). Again, this effect was weaker when the target was option A (b = -1.36, z = -2.94, p = .003, 95% CI [-2.33, -0.50]). Finally, in line with previous results, participants were more likely to choose option A when the target was A (b = 0.78, z = 4.74, p < .001, 95% CI [0.48, 1.13]).

These results support Query Theory and suggest that both the quantity and positions of reasons could predict multialternative choices.

Reason Structure and the Attraction Effect. Figure 3 shows that participants who chose the target generated more reasons to support the target than the competitor (left panel), while reasons in favour of the target were provided earlier (right panel). Both effects were supported by results of statistical models, as shown below.

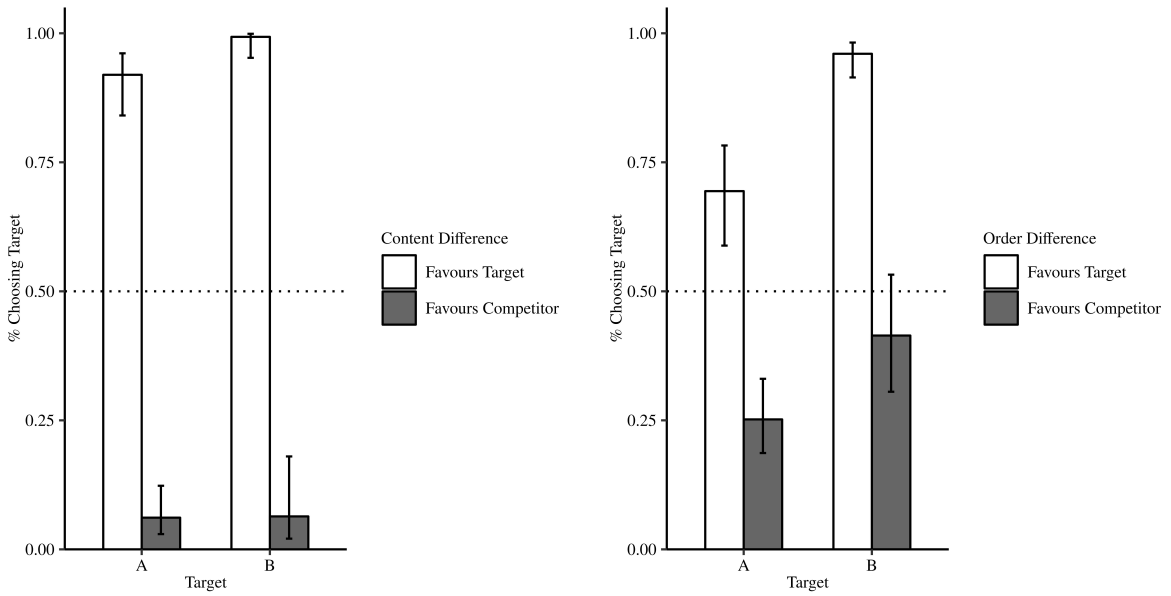


Figure 3
(Left) Proportion of participants choosing the target option by Target and Content Difference. (Right) Proportion of participants choosing the target option by Target and Order Difference. Error bars are confidence intervals of logistic regression models.

Quantity of Reasons. Results of a logistic regression model supported our predictions and suggested that participants were more likely to choose the target option when they generated more reasons supporting it (captured by our $Content_{Target}$ metric: b = 6.59, z = 10.59, p < .001, 95% CI [5.47, 7.94]). This effect did not differ depending on whether the target was smartphone A or B (b = -1.05, z = -1.69, p = .091, 95% CI [-2.38, 0.10]). The proportion of reasons supporting the decoy did not predict choice ($Content_{Decoy}$: b = -0.55, z = -0.81, p = .420, 95% CI [-1.88, 0.82]), and it did not interact with the target manipulation either (b = -0.69, z = -1.01, p = .314, 95% CI [-2.07, 0.63]). Finally, the target manipulation itself was not significant (b = -0.34, z = -1.00, p = .319, 95% CI [-0.98, 0.37]).

Positions of Reasons. As predicted, a logistic regression model showed that participants were more likely to choose the target when they generated reasons supporting it earlier (captured by $Order_{Target}$: b = 7.52, z = 10.01, p < .001, 95% CI [6.14, 9.11]). This effect was weaker when the target was option A (b = -2.54, z = -3.38, p < .001, 95% CI [-4.13, -1.14]), which was in line with the right panel of Figure 3. The position of reasons supporting the decoy had no effect on choice ($Order_{Decoy}$: b = -0.47, z = -0.75, p = .452, 95% CI [-1.64, 0.81]), and had no interaction with the target manipulation (b = -0.80, z = -1.30, p = .194, 95% CI [-2.08, 0.36]). Nor did the target manipulation have impact on choice (b = 0.20, z = 0.61, p = .544, 95% CI [-0.40, 0.89]).

In sum, results supported Query Theory as an explanation of the attraction effect, which was in line with previous research (e.g., the first paper of this thesis). That is, when the target manipulation was controlled for, the quantity and order of reasons could predict whether participants chose the target.

Analysis of Mouse-tracking Data

On average, participants clicked on 24.61 cells and spent 43.21 seconds in the stages with mouse-tracking elements.

Frequency and Choices. For mouse-tracking data, we first investigated whether participants attended to the chosen option more often than its competitor, as captured by the frequency metrics. Supporting our predictions, Figure 4 shows that participants were more likely to choose smartphone A when they attended to A more often than B. This is statistically supported by a logistic regression model (as captured by Frequency Difference: b = 10.00, z = 9.02, p < .001, 95% CI [7.92, 12.27]). This effect did not interact with the target manipulation (b = 0.07, z = 0.06, p = .953, 95% CI [-2.14, 2.24]). Finally, participants were more likely to choose option A than B when the target was A (b = 0.34, z = 2.89, p = .004, 95% CI [0.11, 0.58]).

Durations and Choices. We further tested whether participants attended to the chosen option for longer durations than its competitor. Supporting our previous results, a logistic regression model shows that participants were more likely to choose smartphone A when attended to A for longer durations than B (Duration Difference: b = 2.53, z = 6.81, p <

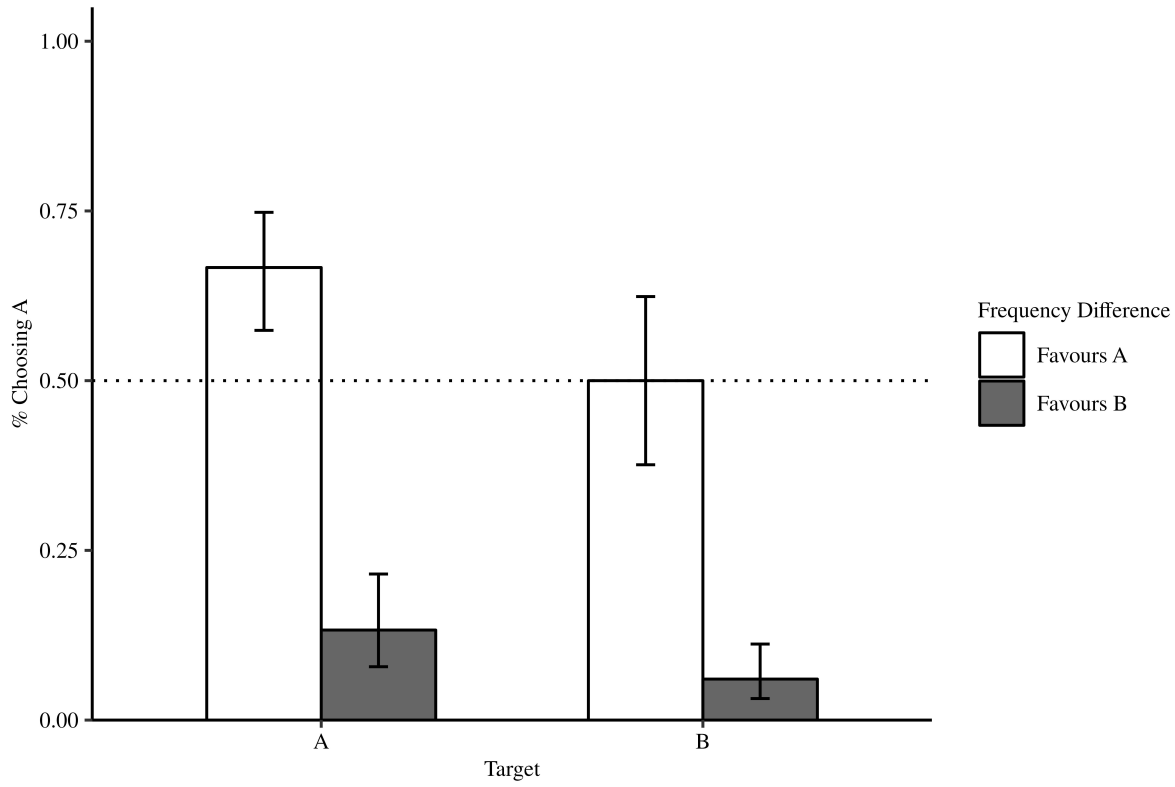


Figure 4
Proportion of participants choosing the target option by Target and Frequency Difference (Error bars are confidence intervals of a logistic regression model).

.001, 95% CI [1.82, 3.28]). This effect did not differ depending on whether the target was smartphone A or B (b = 0.48, z = 1.29, p = .199, 95% CI [-0.25, 1.21]). Again, participants were more likely to choose option A than B when the target was A (b = 0.44, z = 4.05, p < .001, 95% CI [0.23, 0.66]).

Types of Transitions. Figure 5 shows that within-attribute transitions occurred more frequently than within-option ones, and that did not differ between whether the target was smartphone A or B. This supported our predictions on the types of transitions. Statistically, these results were supported by a linear regression model whose dependent variable was the difference between the proportions of within-attribute and within-option transitions. The difference between the proportions of transitions significantly differed from zero (b = 0.23, t = 12.85, p < .001, 95% CI [0.19, 0.26]), while it did not differ between the Target conditions (b = -0.02, t = -1.13, p = .258, 95% CI [-0.05, 0.01]).

Transitions and Choices. We tested whether participants carried out more transitions involving their chosen option than its competitor with a logistic regression model, which was

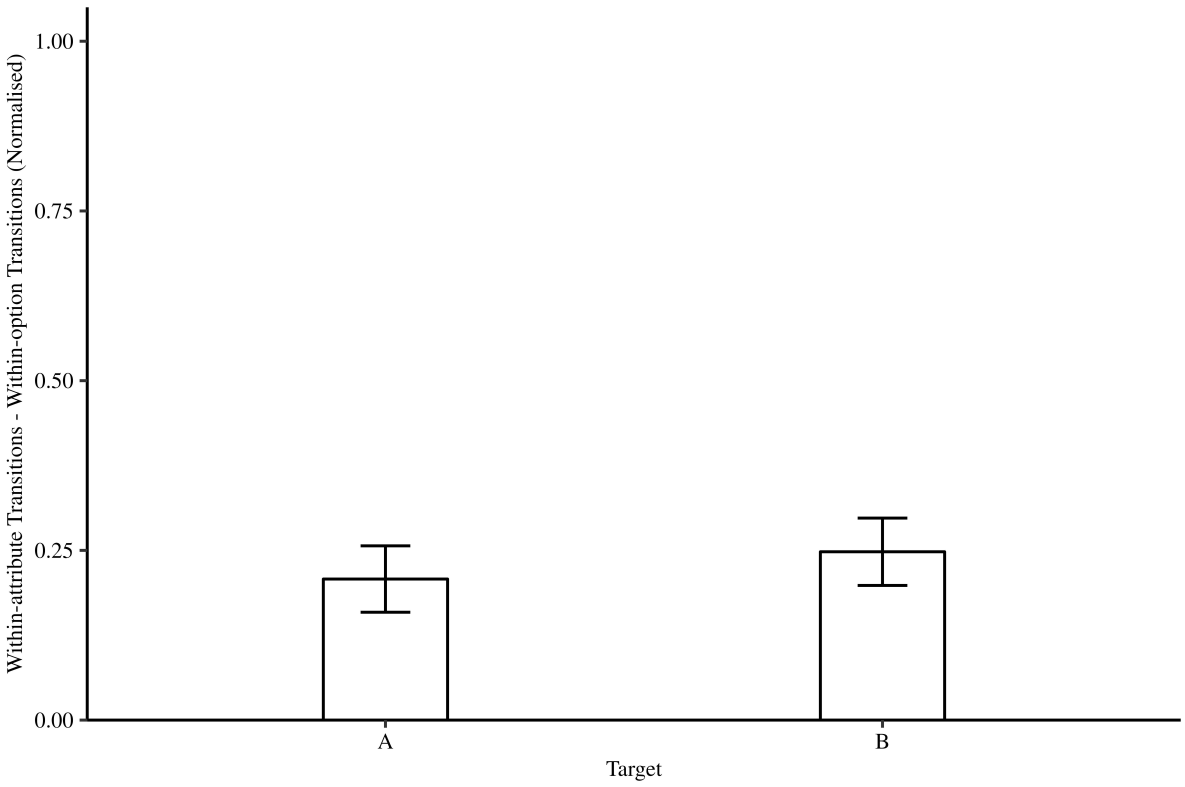


Figure 5
Difference in frequencies between types of transitions by Target (error bars are confidence intervals of a linear regression model).

captured by our Transition Difference metric. As predicted, participants were more likely to choose smartphone A when more transitions involving A than B occurred (as shown by Transition Difference: b = 1.89, z = 3.16, p = .002, 95% CI [0.73, 3.08]). This effect did not differ between Target conditions (b = 0.09, z = 0.16, p = .875, 95% CI [-1.08, 1.26]). Finally, participants were more likely to choose option A when the target was A (b = 0.42, z = 4.04, p < .001, 95% CI [0.22, 0.63]).

Exploratory Analysis

Reason Structure, Mouse-tracking Data, and Choices

Mediation analysis was used to investigate the effects of reasons and attention on choices. Results reported above already showed that the quantity and positions of reasons could separately predict choices, so could the frequency of mouse clicks. The next steps in mediation analysis were to (1) regress the reason structure metrics (i.e., separately for Content Difference and Order Difference) on Frequency Difference, (2) regress choices on both the

reason structure metrics and Frequency Difference in the same models, and finally (3) calculate the indirect effects.

It is important to point out that the dependent variable in this mediation analysis is choice and hence is binary. There is no consensus on how to calculate indirect effect for dichotomous outcome variables, but it has been suggested that the classic approach proposed by Baron & Kenny (1986) for continuous outcome variables is applicable and actually results in low bias (Rijnhart et al., 2019). Hence, Sobel tests (Sobel, 1982) were used to calculate indirect effects.

Quantity of Reasons, Frequency of Mouse Clicks, and Choices. For step (1), a linear regression model showed that participants generated more reasons supporting smartphone A when they attended to the attributes of A more frequently (b = 2.25, z = 11.18, p < .001, 95% CI [1.86, 2.65]). This effect did not interact with the target manipulation (b = 0.24, z = 1.21, p = .225, 95% CI [-0.15, 0.64]). Target itself was included as a control variable (b = 0.06, z = 2.18, p = .029, 95% CI [0.0065, 0.12]).

For step (2), a logistic regression model showed that participants were more likely to choose smartphone A when they generated more reasons supporting A than B (as capture by Content Difference: b = 5.92, z = 6.16, p < .001, 95% CI [4.37, 8.18]), and when they attended to the attribute values of A more frequently than those of B (Frequency Difference: b = 7.46, z = 3.94, p < .001, 95% CI [3.92, 11.42]). The effect of the quantity of reasons was weaker when the target was A (b = -1.96, z = -2.04, p = .041, 95% CI [-4.21, -0.39]), but the effect of frequency did not differ between the Target conditions (b = -0.75, z = -0.40, p = .691, 95% CI [-4.67, 2.85]). Finally, the target manipulation was controlled for (b = 0.56, z = 2.49, p = .013, 95% CI [0.14, 1.04]).

Frequency was significant in the logistic regression model, which showed that the quantity of reasons did not fully mediate the effect of the frequency of mouse clicks on choices. However, the indirect effect can be calculated to test whether partial mediation exists, which is step (3). A Sobel test showed that the indirect effect was significant and the quantity of reasons did partially mediate the effect of the frequency of mouse clicks on choices (ab = 13.32, z = 5.41, p < .001).

Positions of Reasons, Frequency of Mouse Clicks, and Choices. For step (1), a linear regression model showed that participants generated reasons supporting smartphone A earlier when they attended to the attributes of A more frequently (b = 1.09, z = 9.76, p < .001, 95% CI [0.87, 1.31]). This effect did not differ between the Target conditions (b = 0.04, z = 0.34, p = .737, 95% CI [-0.18, 0.26]). The target manipulation was controlled for (b = 0.0012, z = 0.07, p = .941, 95% CI [-0.03, 0.03]).

For step (2), a logistic regression model showed participants were more likely to choose smartphone A when they generated reasons supporting A earlier than those supporting B (as captured by Order Difference: b = 4.43, z = 8.39, p < .001, 95% CI [3.48, 5.57]), while the frequency of mouse clicks also had a significant effect (Frequency Difference: b = 8.95, z = 6.37, p < .001, 95% CI [6.31, 11.85]). The effect of the position of reasons was weaker when the target was A (b = -1.66, z = -3.14, p = .002, 95% CI [-2.79, -0.70]), while the effect of frequency did not differ (b = -0.63, z = -0.45, p = .653, 95% CI [-3.48, 2.07]). The target manipulation was also included as a control variable (b = 0.69, z = 3.83, p < .001, 95% CI [0.36, 1.07]).

For step (3), similar to the previous analysis, a Sobel test was used to test the indirect effect and showed that the positions of reasons partially mediated the effect of the frequency of mouse clicks on choices (ab = 4.83, z = 6.39, p < .001).

Mouse-tracking Data and Types of Reasons

We were interested in whether the frequency of mouse clicks could predict the types of reasons generated by participants, hence we split the mouse-tracking data by reasons. That is, we calculated the frequency indices (e.g., $Frequency_A$ and $Frequency_B$) by aspect listing screens, each of which represented the proportion of clicks a participant performed to reveal the attribute values of an option before submitting a particular reason. As before, Frequency Difference is defined as $Frequency_A - Frequency_B$ to capture whether a participant attended to smartphone A more frequently than B in each aspect listing screen.

We ran two random-intercept logistic regression models for this analysis. For the first model, the dependent variable was whether a reason was coded in favour of smartphone A. Results indicated that a reason was more likely to be coded as supportive of smartphone A

when participants attended to A more frequently than B in the aspect listing screen where they submitted that reason (as capture by Frequency Difference: $b = 1.64$, $z = 6.60$, $p < .001$, 95% CI [1.15, 2.13]). This effect did not differ between Target conditions ($b = 0.35$, $z = 1.44$, $p = .150$, 95% CI [-0.13, 0.83]). Results also showed that reasons were more likely to be coded in favour of smartphone A when the target was A ($b = 0.27$, $z = 3.24$, $p = .001$, 95% CI [0.10, 0.43]), which is consistent with Query Theory.

In our experiment, participants could code a reason in favour of multiple options (i.e., a reason could be coded as supportive of both smartphones A and B, which happened for 9.34% of all reasons), therefore we further ran a second model whose dependent variable is whether a reason was coded in favour of B. Similar to the previous model, results showed that a reason was more likely to be coded as supportive of B when participants attended to B more often than A in the associated aspect listing screen ($b = -1.86$, $z = -7.04$, $p < .001$, 95% CI [-2.38, -1.34]). This effect did not interact with the target manipulation ($b = -0.39$, $z = -1.51$, $p = .131$, 95% CI [-0.90, 0.12]). Further, reasons were more likely to be coded as supportive of smartphone B when the target was B ($b = -0.30$, $z = -3.56$, $p < .001$, 95% CI [-0.46, -0.13]).

Discussion

This work is the first to study the attraction effect with aspect listing and mouse tracking. The former investigated high-level deliberation processes behind multialternative choices and focused on the roles of distinct reasons. As we predicted, reasons supporting the chosen options were generated earlier and in a greater amount. Importantly, replicating the attraction effect, the option which was chosen more often by participants in our experiment was the target shown to them. Reasons in favour of this target option were generated earlier and in greater quantity, which suggested that Query Theory could explain the attraction effect. In other words, the presence of a decoy option influences the order in which mental queries are executed during decision making, which leads to the observed changes in choice shares between conditions. Combined with the previous findings of the first paper of this thesis, this is the third experiment which showed that the deliberation processes described by Query Theory can be extended beyond binary choices and explain multialternative choices.

The above results alone, however, do not tell us how individuals sample information

and attend to the attribute values during decision making. We investigated this by incorporating mouse tracking into our experimental design and results supported our predictions. The frequency of mouse clicks showed that the chosen option was attended to more often, while the durations for which information behind cells was revealed also showed that the chosen option was attended to for a longer period of time. We further investigated transitions between options, and results showed that within-attribute transitions were more frequent than within-option ones. Finally, we found that transitions between the chosen option and the decoy occurred more often than those between the competitor and the decoy. On the one hand, these results are consistent with previous findings in mouse-tracking experiments (Król & Król, 2019; Marini et al., 2020; Noguchi & Stewart, 2014) and hence provided validation to the use of process-tracing methodologies in studying multialternative choices. On the other hand, our results supported the assumptions of various cognitive models (Bhatia, 2013; Noguchi & Stewart, 2018), which are based on attentional processes and comparison patterns. This further highlights the importance of developing cognitive models which can be tested with process data (Johnson et al., 2008). Taken together, both empirical results and computational models have demonstrated that the attraction effect occurs as the dominance relationship in the choice set leads to a higher level of attention being paid to the target and more comparisons between the target and the decoy.

We additionally examined the interplay of multiple cognitive processes behind the attraction effect, and the novel findings of our study provided insights into the intertwined roles of attentional patterns and reasoning. We first showed that the quantity of reasons partially mediated the effect of the frequency of mouse clicks on choices, and so did the positions of reasons. We further found that, when dividing data into discrete stages by screens in the reason listing procedure, a reason was more likely to be supportive of an option when participants attended to its attribute values more often as they were generating that reason. These results support the notion that individuals aggregate information they sample throughout a decision task into reasons, which are further integrated into decision outcomes. This opens up opportunities into studying other context effects and behavioural phenomena with our methodology, as well as laying the foundation in developing new cognitive models.

Supplemental Materials (Screenshots)

Supplemental Materials (Screenshots)

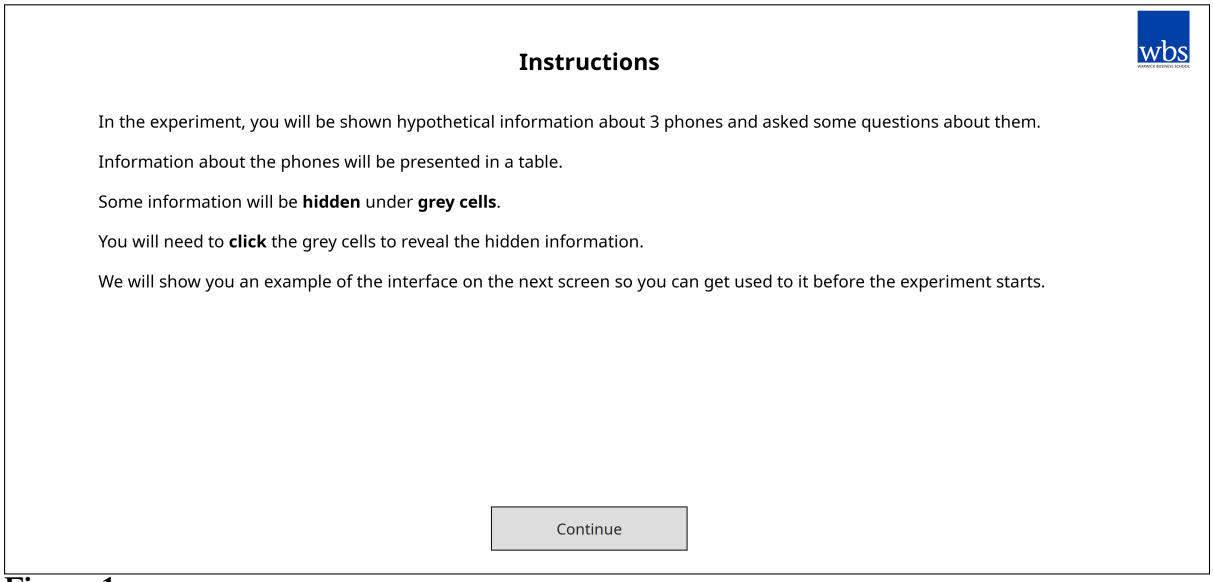


Figure 1

Instruction Screen of the Experiment

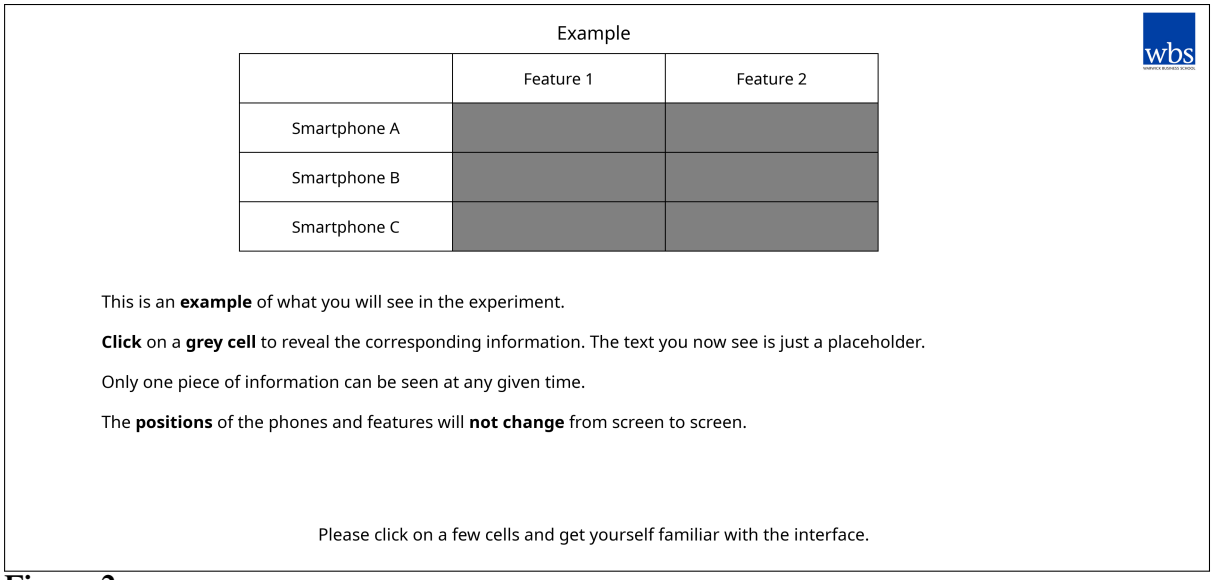


Figure 2

Practice Screen of the Experiment

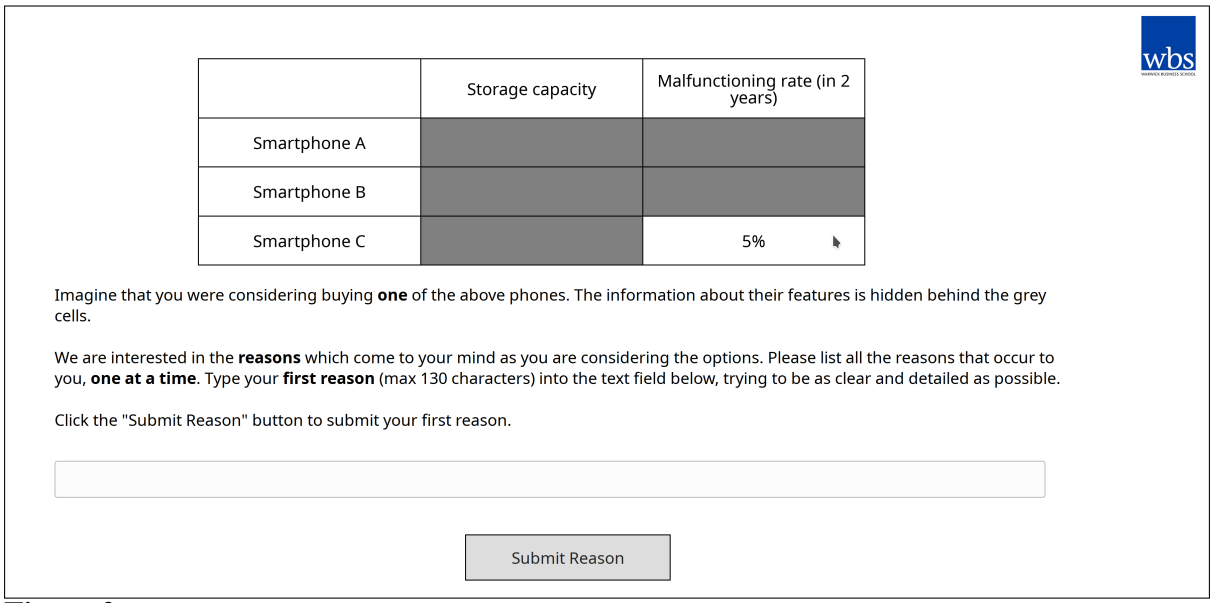
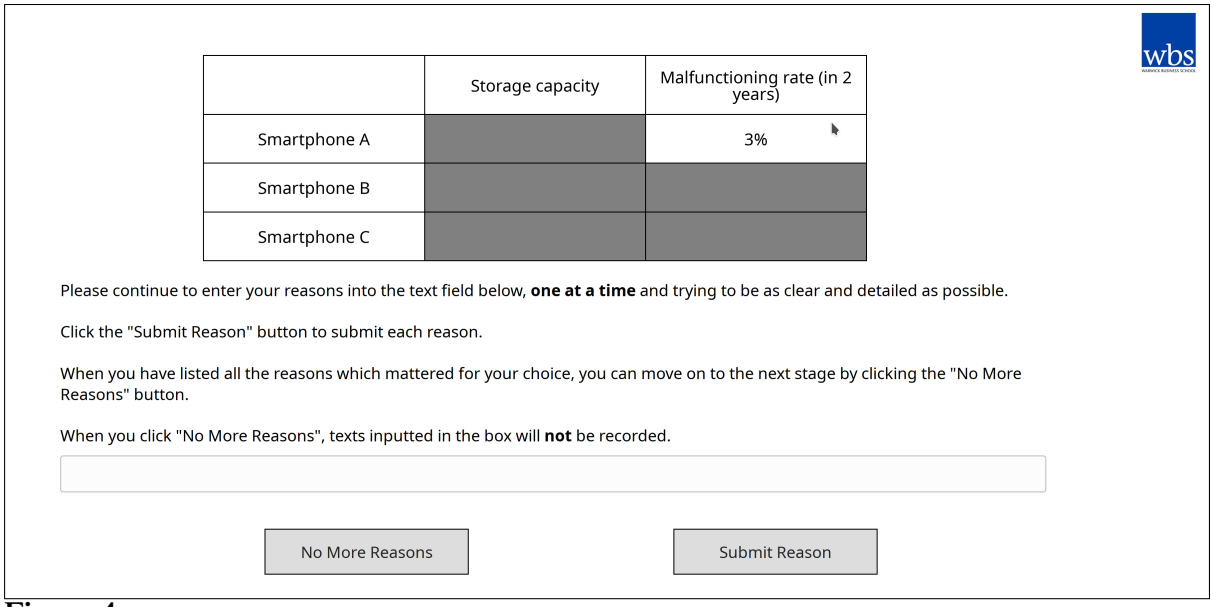


Figure 3

First Aspect Listing Screen of the Experiment



	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A		3%
Smartphone B		
Smartphone C		

Please continue to enter your reasons into the text field below, **one at a time** and trying to be as clear and detailed as possible.

Click the "Submit Reason" button to submit each reason.

When you have listed all the reasons which mattered for your choice, you can move on to the next stage by clicking the "No More Reasons" button.

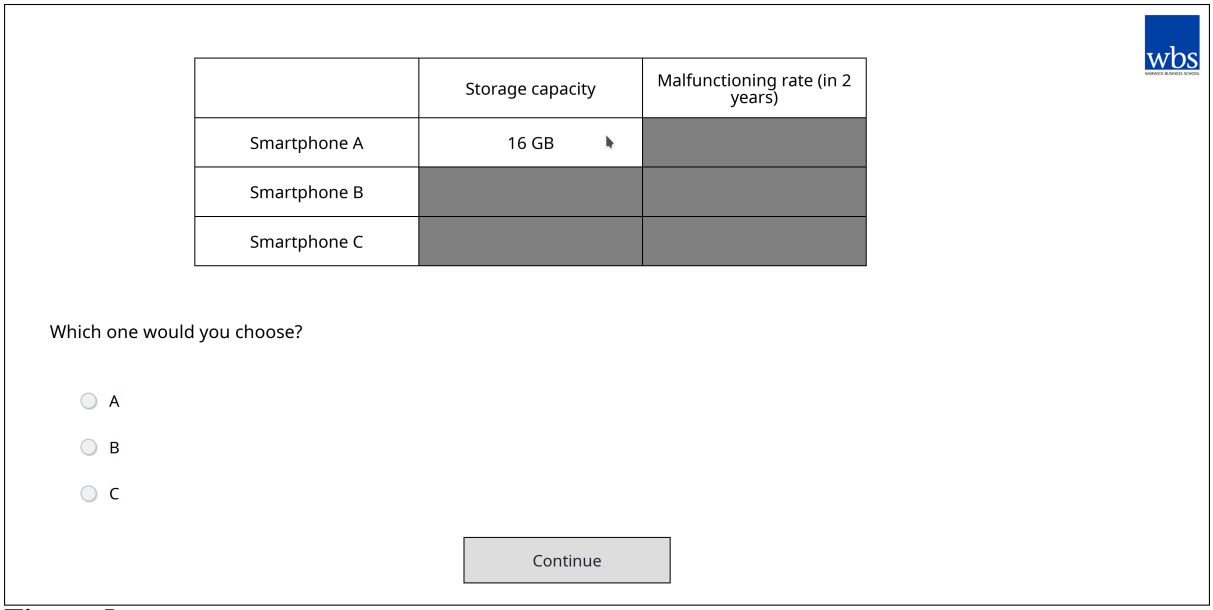
When you click "No More Reasons", texts inputted in the box will **not** be recorded.

No More Reasons

Submit Reason

Figure 4

Second Aspect Listing Screen of the Experiment



	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	
Smartphone B		
Smartphone C		

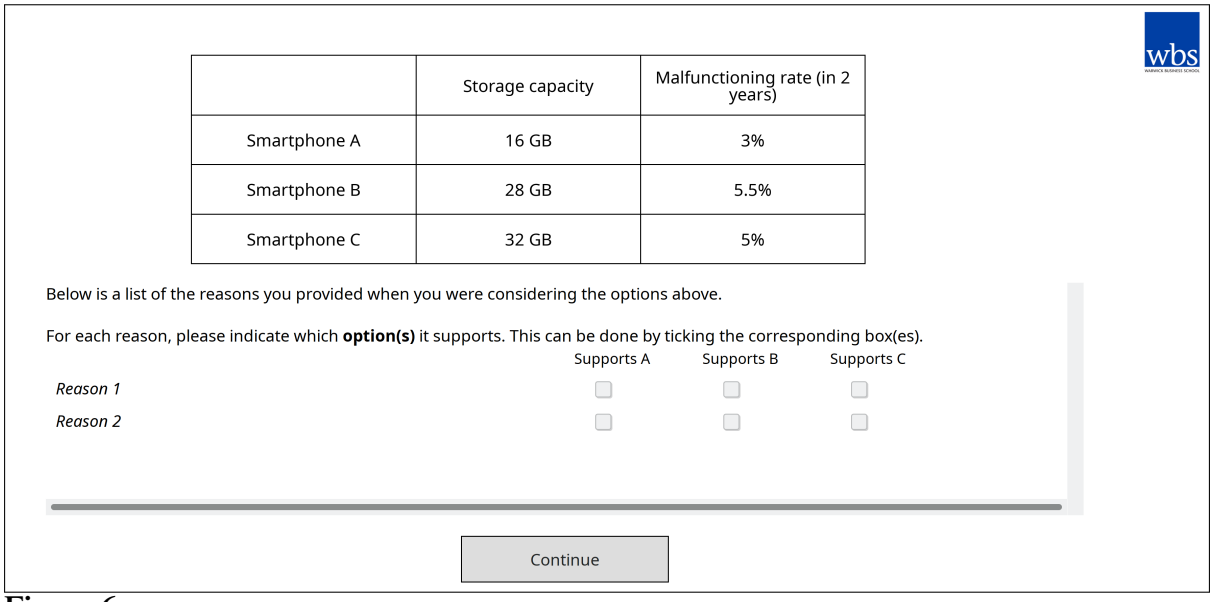
Which one would you choose?

- A
- B
- C

Continue

Figure 5

Choice Screen of the Experiment



	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Below is a list of the reasons you provided when you were considering the options above.

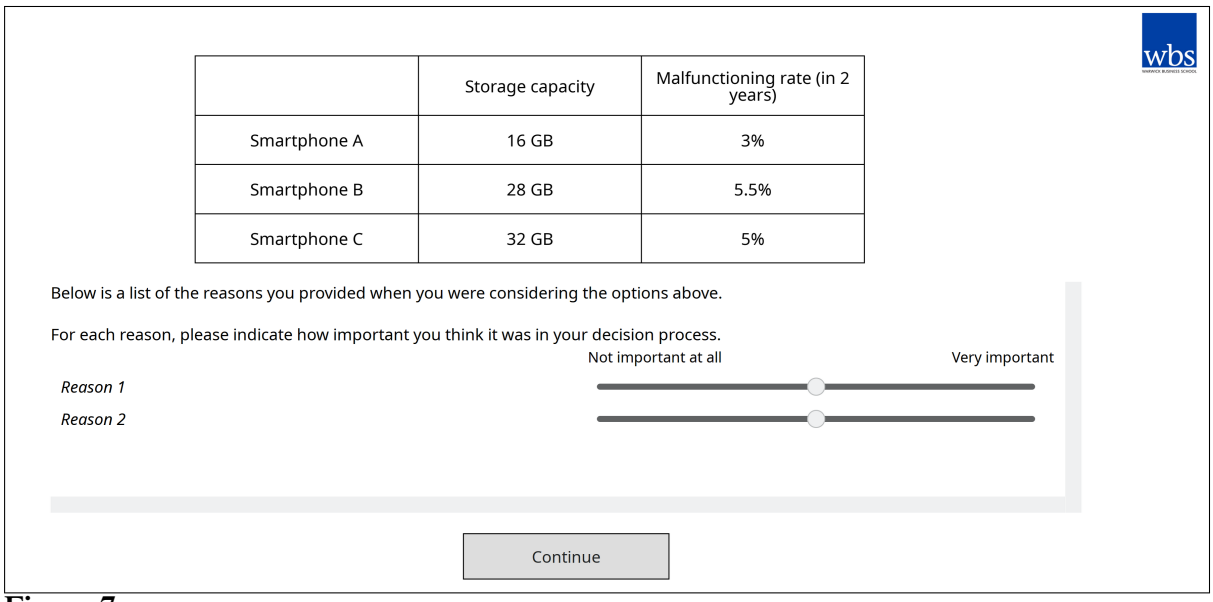
For each reason, please indicate which **option(s)** it supports. This can be done by ticking the corresponding box(es).

	Supports A	Supports B	Supports C
Reason 1	☐	☐	☐
Reason 2	☐	☐	☐

Continue

Figure 6

Reason Coding Screen of the Experiment



	Storage capacity	Malfunctioning rate (in 2 years)
Smartphone A	16 GB	3%
Smartphone B	28 GB	5.5%
Smartphone C	32 GB	5%

Below is a list of the reasons you provided when you were considering the options above.

For each reason, please indicate how important you think it was in your decision process.

	Not important at all	Very important
Reason 1		
Reason 2		

Continue

Figure 7

Reason Weighting Screen of the Experiment

Echo chambers and confrontation: Dynamics of public opinions in sociopolitically heterogeneous environments

Abstract

The echo chamber theory suggests that individuals prefer to consume attitude-consistent information and interact with like-minded others. While evidence has demonstrated the existence of echo chambers on the Internet, recent studies showed that individuals are often not in completely homogeneous environments when they engage with others and polarisation does arise despite encounters with opposing views. This paper suggests that the complex behaviours behind public opinions are driven by both echo chambers and confrontation, and provides evidence to illustrate their roles with novel data sets collected from online deliberation platforms. Study 1 (N = 5,185) showed that British participants were more willing to engage in discussions when they agreed with proportionally more comments submitted by other participants. This relationship, however, was not linear： results of regression models suggested that their willingness to engage reached a peak when they also disagreed with a small number of comments. Multilevel regression models further showed that participants were more likely to produce their own comments when they had just seen a comment with which they disagreed. With Pakistani participants, Study 2 (N = 17,620) found similar results using multilevel regression models. These findings supported the theory that individuals' willingness to engage with others on online platforms is driven by not only a subjective perception of an homogeneous environment, but also a behavioural response to contradictory views.

Echo chambers and confrontation: Dynamics of public opinions in sociopolitically heterogeneous environments

The echo chamber theory (Sunstein, 2001, 2007) is one of the leading views on how individuals process information and interact with others on the Internet. It suggests that people prefer to consume information that is consistent with their existing beliefs and maintain contacts with those who are ideologically similar. This reinforcement of ideas can lead to isolated and homogeneous online communities that consequently create sociopolitical polarisation in the physical world.

Recent studies have generally demonstrated the existence of echo chambers by analysing data collected from social media. Multiple groups of researchers (Barberá et al., 2015; Conover et al., 2011; Himelboim et al., 2013; Williams et al., 2015) found similar patterns and their results showed that Twitter users could be classified into highly homogeneous communities with limited interconnectivity, although some mixed results were also found. Conover et al. (2011) found that the network of retweeting (i.e., the acts of sharing other users' posts) could be clustered into two isolated communities, which represented individuals on the two sides of the left-right political spectrum respectively. However, the network of mentioning (i.e., the acts of directly addressing other users in one's posts) was structured differently and better represented as one heterogeneous cluster with frequent engagements between different political groups. Focusing on following (i.e., the acts of subscribing to other users' posts), Himelboim et al. (2013) also found that the network underlying each of their selected, popular political topics was best structured into only a small number of segregated clusters. Similarly, Williams et al. (2015) found that the majority of Twitter users could be categorised into activists or sceptics of climate change based on the contents of their posts, while the networks of following and retweeting showed that the majority of users only interacted with others sharing their existing views on the topic. However, activists and sceptics did interact more frequently in the network of mentioning, frequently with comments of negative sentiment. Additionally, Barberá et al. (2015) found that the retweeting of political contents mostly happened among individuals on the same side of the political spectrum, but non-political retweets were more likely to cross the ideological

line. One further piece of evidence was provided by Del Vicario et al. (2017), who instead used data from Facebook and found that the network of users engaging with posts related to Brexit could be divided into two segregated communities.

From a psychological perspective, the emergence of echo chambers is commonly explained by the selective attention theory (Chaffee & Miyo, 1983; Sears & Freedman, 1967), which is related to confirmation bias (Nickerson, 1998) and suggests that individuals pay more attention to information aligned with their worldviews. This theory is supported by a number of studies (Garrett, 2009; Graf & Aday, 2008; Iyengar & Hahn, 2009) which used different experimental methods to test how participants selected news articles. The findings of Graf & Aday (2008) showed that, when browsing freely on the website of a mocked online magazine, participants spent more time reading attitude-consistent articles, read them earlier, and read more pages in them. Similarly, with the use of a platform that resembled a news aggregation service, Garrett (2009) found that participants were more likely to mark attitude-consistent articles as of interest. By randomly assigning identical articles to news providers with different political alignments, Iyengar & Hahn (2009) also found that participants were more likely to read articles provided by attitude-consistent sources. These accounts of selective attention were further supported by Lawrence et al. (2010), who found that the majority of blog readers only browsed sites whose contents were consistent with their political stances.

While the selective attention theory can explain many phenomena in the dynamics of public opinions, it does not provide the full picture. On the one hand, the above studies using Twitter showed that individuals preferred to see and circulate content generated by other users with similar ideological views, which could be explained by selective attention and was consistent with the notion of echo chambers. On the other hand, however, the more heterogeneous network of mentioning illustrated that people did actively attempt to converse with or respond to others holding opposing opinions (Conover et al., 2011; Williams et al., 2015). Additionally, avoiding information contradictory to one's beliefs is more difficult on some online platforms than others. For instance, Bakshy et al. (2015) found that individuals on either side of the political spectrum had around 20% of Facebook friends from the other end, potentially because Facebook friendships mainly reflected relationships in daily life

instead of connections between people with similar topical interests, and as a result between 24% to 35% of news stories which Facebook users were exposed to originated from attitude-inconsistent media sources. This did not prevent homogeneous communities from being formed on Facebook for sociopolitical topics (Del Vicario et al., 2015, 2017). Furthermore, Bail et al. (2018) found that, when individuals were instructed to follow a Twitter bot which retweeted posts from accounts with opposing ideologies for a month, their political attitudes either remained the same or became more extreme. This demonstrated that political polarisation could occur even when a part of the online environment was intentionally constructed to be ideologically heterogeneous, which could not be characterised by the conventional idea of echo chambers.

Karlsen et al. (2017) proposed that confrontation and the generation of counterarguments are important components of the dynamics of public opinions, but their roles in political polarisation are often ignored. Karlsen et al. (2017) found that the vast majority of their respondents claimed to at least occasionally engage with others who held different basic values on digital platforms, although extremely few respondents reportedly changed their opinions on sociopolitical issues after debates. Furthermore, Stromer-Galley & Muhlberger (2009) found that individuals were reportedly more willing to participate in future conversations when the number of opposing statements during previous online discussions was higher, but the latter had no effect on the extent to which participants felt the need to reevaluate their viewpoints. The apparent lack of relationship between interacting with ideologically different others and changing one's own attitudes can be explained by the *disconfirmation bias*, which suggests that individuals tend to reject or devalue information challenging their beliefs. Taber & Lodge (2006) found that, when presented with statements supporting and opposing two political issues respectively, participants rated attitude-consistent statements as more convincing than attitude-inconsistent ones. Moreover, participants spent more time reading attitude-inconsistent statements and this effect was amplified when participants had stronger prior attitudes. Taber & Lodge (2006) proposed that the extra time was spent on producing counterarguments, which they tested with a thought listing protocol. Results showed that participants provided more comments for attitude-inconsistent statements

than attitude-consistent ones, while the effect was stronger for participants with a higher level of political knowledge. These results were replicated in a similar study conducted by Taber et al. (2009). Taken together, there has been evidence showing that individuals do participate in discussions with those holding opposing views, but people also tend to systemically refute or discount opinions contradicting their existing attitudes, possibly to convince themselves that their existing viewpoints are correct.

The notions of echo chamber and confrontation are not incompatible: we propose that they have intertwined roles in shaping public opinions and engagement. Specifically, we hypothesise that individuals are more willing to spend time on online discussion platforms when they agree with the opinions of many like-minded others, but exposure to opinions with which they disagree can also promote their engagement, as they feel the need to refute others' views and defend the sides which they have taken on social issues. Both types of encounters reinforce one's initial beliefs, and hence can lead to polarisation despite the individuals having experienced heterogeneous opinions. With two novel data sets that recorded how individuals responded to different opinions at a multitude of time points on deliberation platforms, in a manner similar to how individuals interact with each other on social media, this paper investigated the complex roles of the psychological drivers behind public opinions and examined evidence that supports an interplay between echo chambers and confrontation.

Study 1

Method

The data was collected by the organisation Engage Britain, which recruited participants to a platform named Polis to gather data on public opinions in the UK.

Upon entering Polis, participants were shown the question "What are the challenges that need to be tackled to make Britain a better place to live?" Participants were instructed to provide comments on the issue and were given a 140-character limit for each comment. They were also shown the comments previously submitted by other participants,[1] one at a time.

[1] Ten comments were submitted to Polis by Engage Britain before any participant was recruited, in order to ensure that the comment pool was not empty when participants entered the platform. These comments were constructed to reflect key issues in the UK based on the results of prior focus groups conducted by the organisation.

Comments were displayed to participants semi-randomly.[2] Comments were also moderated by staff members of Engage Britain, such that those submitted repeatedly by participants, unintelligible, or explicitly expressing hateful views were marked and not displayed to other participants. Participants were told to vote on each comment to express whether they agreed or disagreed with it, while they could also skip the comments on which they decided not to vote. Participants could leave Polis at any point. The interface of Polis is shown in Figure 1.

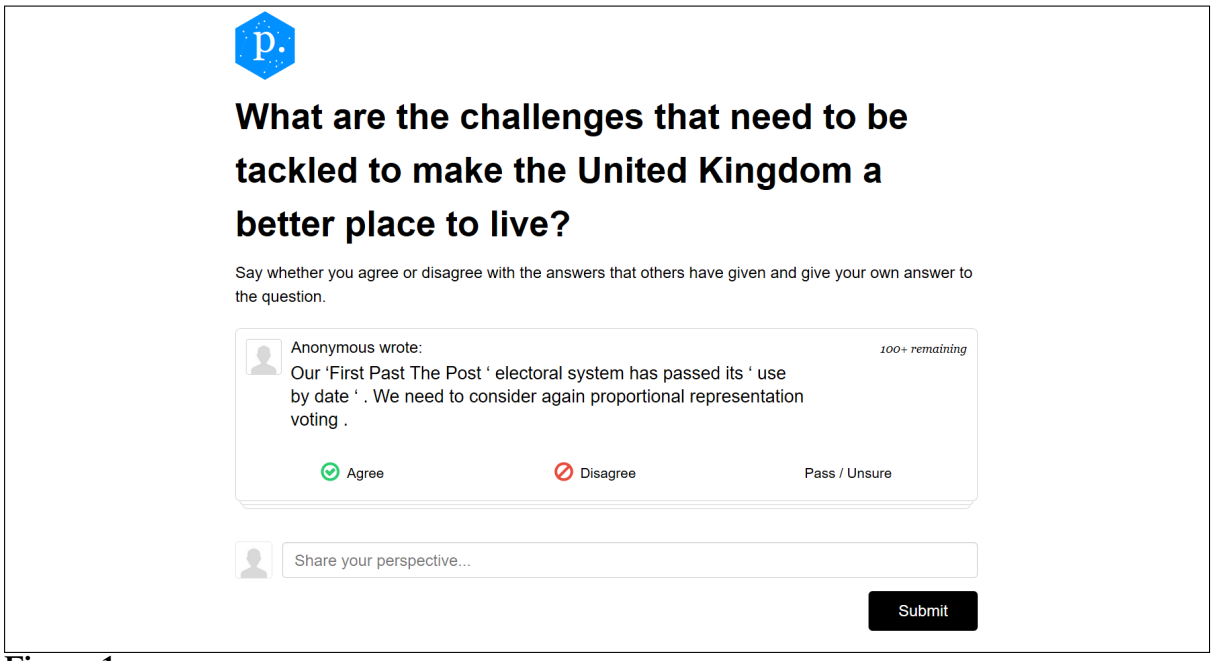


Figure 1

The interface of Polis in Study 1.

Participants were recruited from multiple sources over roughly a month. A small number of participants (n = 541) was first recruited by an external polling company. This is a *paid sample* as these participants were given a small fee for their participation. The amount of payment did not depend on the durations for which they stayed on Polis. All participants in this paid sample completed their responses within the first two days. Afterwards, recruitment from other sources (termed the *volunteer samples*) started, including advertisements on Facebook (n = 2,267), calls for participation on the online platforms of the newspaper The Times (n = 522) and other media partners (n = 121), Tweets and LinkedIn posts from Engage

[2] The display of comments was largely random, except new comments were prioritised while comments frequently skipped by participants were suppressed.

Britain (n = 115), and direct emails (n = 97). The referring sources of a number of participants (n = 251) could not be identified. The majority of participants from the volunteer samples (96.41%) entered Polis and completed their responses within two weeks. At that point, a second batch of paid sample (n = 1,271) was recruited. In total, there were 5,185 participants.

The study was approved internally by the doctoral programme office at Warwick Business School.

Prediction & Analysis

Given the hypotheses described in the introduction, we predicted that participants on Polis would reach the maximum level of engagement when they agreed with the majority of the comments —that is, when they subjectively perceived the environment as more homogeneous —but also when they disagreed with a small number of comments which led to confrontation.

To investigate the engagement of participants, we first examined the data on an *aggregated* level and focused on two dependent variables separately, namely the total number of comments provided by participants and the total number of votes cast by them. Our independent variable was the proportion of agree votes cast by the participants. As the dependent variables were count data, we ran Poisson regression models instead of simple linear regressions. Furthermore, our models included the quadratic term for the proportion of agree votes. We specifically predicted both a positive main effect and a negative quadratic term: the positive main effect would demonstrate that participants were more willing to engage when the environment subjectively resembled echo chambers, while the negative quadratic term would mean a non-linear relationship and that the mentioned effect reduced as the proportion of agree votes reached a certain level.

To more directly test how participants responded to a comment with which they disagreed, we then focused on all *events* involving the participants *over time*, including every comment provided by them and every vote they cast. We focused on the likelihood of an event being the submission of a new comment, in contrast to being a vote, which we used as dependent variable.[3] The independent variable we used were the proportion of agree votes of

[3] An event was coded as 1 when it was a new comment and 0 when it was a vote.

the participant up until the point of that event, as well as the type of the last event (i.e., whether it was a disagree vote).[4] We used random-intercept logistic regression models since the dependent variable was binary and the data involved repeated measures for each participant. We predicted a negative effect for the proportion of agree votes up until the point of an event, which would show that participants were more likely to leave their own comments when they started to disagree with the comments which they saw, as well as a positive effect of the type of the last event (since the last event was coded positively when it was a disagree vote).

All terms of the models were reported below.

Results

Summary Statistics

Participants. The vast majority of participants (92.66%) cast at least one vote. The median number of votes cast by them was 34. As stated above, there were two batches of paid samples, in addition to the volunteer samples recruited via social media and news partners. The median number of votes was 33.5 for the first batch of paid samples, 33 for the second batch, and 35 for the volunteer samples. A further breakdown of the volunteer samples, however, showed that the number of votes cast by the volunteer samples varied greatly between the sources from which participants were recruited. The median number of votes were 42 for participants recruited from Facebook, 33 for those recruited from the Times, and 20 for other sources (e.g., Twitter, LinkedIn, and direct email). All our statistical models presented below controlled for the potential difference between the paid sample and the volunteer samples.[5]

43.72% of all participants submitted at least one comment. This differed between samples, as 59.77% of the paid samples submitted at least one comment and only 35.22% of the volunteer samples did so (34.94% among those recruited from Facebook), which might reflect that the willingness to engage with Polis depended on payment. However, it is

[4] The last event was coded as 1 when it was a disagree vote and -1 otherwise.

[5] Since participants recruited from Facebook constitute 43.70% of all participants and cast more votes, we had also repeated our statistical models and controlled for the potential difference between the Facebook sample and the remaining samples. These alternative models are reported in the Supplementary Materials and do not change the conclusions of the paper.

important to note that, for participants who did submit a comment, the median number of comments provided by them was one in all samples.

Many participants in the paid samples provided their demographic information, hence allow us to investigate whether these samples represented the British population. The majority of participants within the paid samples (n = 1,518) reported their ages (median = 50) and genders, which showed that the samples were 51.03% female and 48.98% male. Furthermore, 1,509 participants stated their education level, 56.86% of whom had university degrees or further education. Finally, 1,134 participants stated their preferred political parties in the UK. Among them, 41.27% would vote for the Conservative Party, 35.10% the Labour Party, 9.61% the Liberal Democrats, and 4.59% the Scottish National Party. These demonstrated that the paid samples had similar gender split as the British population, but were older than the average British person (median age = 40.5). The paid samples were also more educated than the general British population, 42% of which hold university degrees (Office for National Statistics, 2017). Political alignments were comparable to the results of the 2019 general election, with a higher percentage of Labour and Scottish National Party supporters.

On average, for participants who would vote for these major parties, those who favoured the Labour Party agreed with 77.10% of all comments they saw, 74.60% for those who favoured the Liberal Democrats, 72.10% for the Scottish National Party, and 67.67% for the Conservative Party.

Events. In total, 422,031 events were recorded, which included the submissions of comments (n = 7,449) and votes (n = 414,582). As comments were moderated, only 1,085 comments were displayed to and voted by participants,[6] although all comments were recorded in the data set. Among the moderated comments, the median number of votes cast to them was 244.

The topics of the 150 most voted comments were coded manually and are shown in Table 1. Topics were coded by Engage Britain. In the table, prevalence refers to the proportion of comments covering a topic (however, it does not sum up to 1 as a comment can cover

[6] As described above, Engage Britain moderated out comments that covered the same topic and with similar meanings.

multiple topics). The mean proportion of agree votes of a topic is computed by averaging the proportions of agree votes of all comments covering that topic. The standard deviations of the proportions of agree votes of each topic are also reported.

Healthcare was the topic with the highest average proportion of agree votes (78.74%). Comments related to healthcare were mostly supportive of the National Health Service (NHS) and advocated for better funding or better payment for healthcare staffs. Brexit was the topic with the lowest mean proportion of agree votes (56.00%), along with a low standard deviation. In other words, the ratio of agree votes to disagree votes for Brexit was closer to 1 than any other topic, even when the range of the proportions of agree votes was taken into consideration, and this suggested that Brexit is a polarised topic among the participants.

Table 1

Topics among the top 150 comments in Study 1.

Topic	Prevalence	Mean Proportion of Agree Votes	SD of Proportions of Agree Votes
COVID-19	0.28	0.76	0.15
Healthcare	0.17	0.79	0.11
Immigration	0.15	0.63	0.11
Crime	0.13	0.70	0.15
Education	0.11	0.73	0.15
Brexit	0.11	0.56	0.08
Poverty	0.11	0.77	0.10
Racism	0.09	0.64	0.14

Inferential Statistics

To exclude participants who quickly left Polis from our analyses, we removed all participants who cast fewer than six votes, which was the lower (25%) quantile of the number of votes cast by participants.[7] This criterion was chosen such that a large amount of events

[7] We repeated the statistical models below for all participants. These alternative models are reported in the Supplementary Materials and do not change the conclusions of the paper.

would remain: indeed, 98.86% of all events remained after this exclusion. Comments generated by excluded participants were not removed from the data set per se, although they were subject to moderation just as the rest of the comments. Votes that were cast on these comments were not removed either.

We further examined the participants who were excluded and computed the proportion of participants removed in each sample. 21.10% of all participants among the first paid sample was excluded, 21.19% among the second paid sample, and 26.33% among the volunteer samples. This is due to, as stated above, participants recruited from Twitter, LinkedIn, and direct email casting fewer votes. Among the excluded participants, the median number of comments was one in all samples which is the same as the participants retained in the analyses.

Analysing the Aggregated Data of Participants. We first examined whether there was a relationship between the number of comments submitted by participants and the proportion of agree votes cast by them. As shown in Figure 2, the mean number of comments submitted by participants reached its peak for those who agreed with 60-70% of the comments they saw, and decreased for participants who agreed with proportionally more comments. This is consistent with the results of a Poisson regression model, which showed that participants tended to submit more comments when they agreed with a higher proportion of comments provided by other participants (b = 7.96, z = 12.93, p < .001, 95% CI [6.77, 9.18]). However, this relationship was weaker as the proportion of agree votes increased, which was shown by the significant quadratic term (b = -6.39, z = -13.89, p < .001, 95% CI [-7.30, -5.50]). The model also controlled for whether participants were recruited from the paid samples and showed that the paid samples submitted more comments (b = 0.08, z = 6.63, p < .001, 95% CI [0.06, 0.11]). Finally, this model fitted significantly better than the model without the quadratic term ($X^2(1)$ = 237.47, p < .001).

We further investigated whether there was a relationship between the number of votes cast by participants and their proportion of agree votes. Figure 3 showed that the relationship between them was positive, although the number of votes cast by participants dropped when the proportion of agree votes cast to others' comments exceeded 90%. Results of a Poisson regression model showed that participants tended to vote on more comments when they agreed

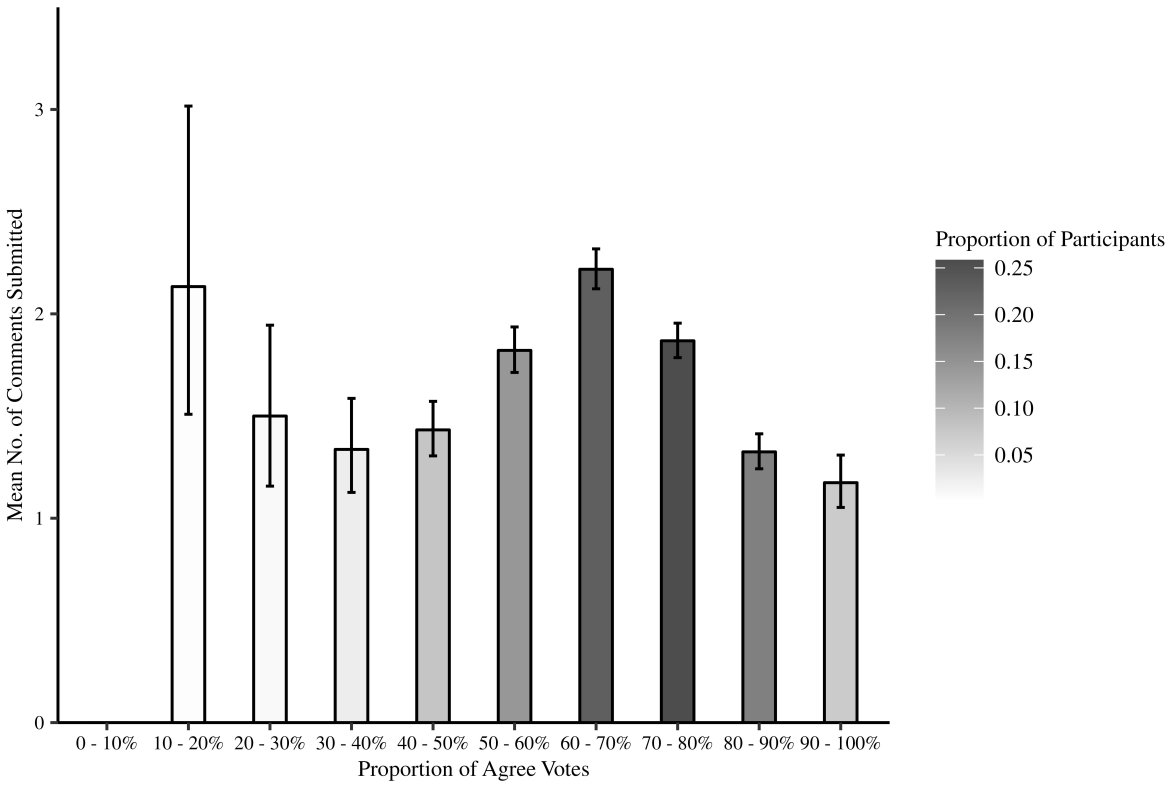


Figure 2

Mean number of comments submitted by participants, as a function of the proportions of agree votes cast by them in Study 1 (error bars are confidence intervals of a Poisson regression model).

with a higher proportion of comments (b = 5.18, z = 58.22, p < .001, 95% CI [5.01, 5.36]). The drop in the number of votes cast by participants was captured by the significant quadratic term (b = -2.74, z = -42.95, p < .001, 95% CI [-2.86, -2.61]). Similarly, whether participants were recruited from the paid samples was controlled for and the paid samples cast fewer votes than the remaining samples (b = -0.12, z = -62.37, p < .001, 95% CI [-0.12, -0.11]). This model also fitted significantly better than the model without the quadratic term ($X^2(1)$ = 2017.61, p < .001).

Analysing Events. As reported above, only a small proportion of events were the submission of new comments (1.77%). The frequency of an event being the submission of a comment was higher when a participant's last event was a disagree vote (1.81%) than when it was an agree vote (1.53%).

Statistically, we tested whether an event was more or less likely to be the submission of

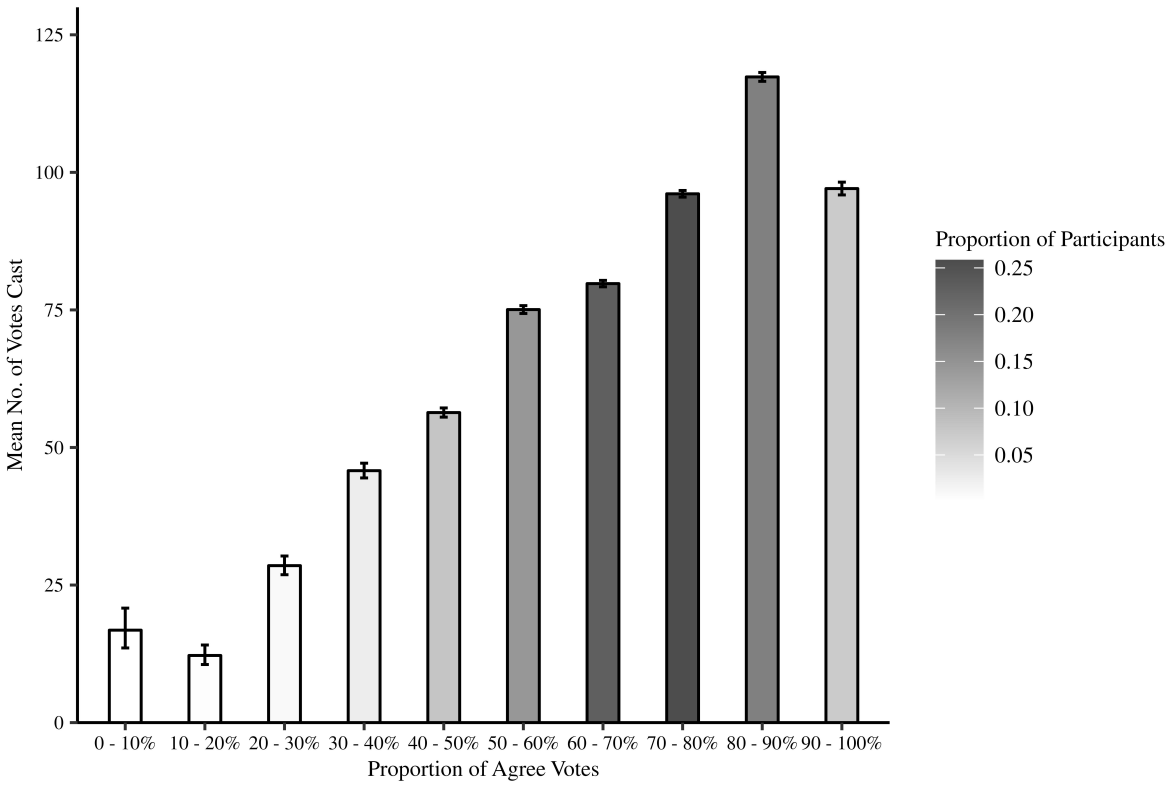


Figure 3

Mean number of votes cast by participants, as a function of the proportions of agree votes cast by them in Study 1 (error bars are confidence intervals of a Poisson regression model).

a new comment (1) when a participant agreed with proportionally more comments up until the point of the event and (2) when the last event of a participant was a disagree vote. Results of a random-intercept logistic regression model showed that a participant was more likely to submit a comment when the proportion of agree votes cast by them up until that point decreases (b = -0.35, z = -2.98, p = .003, 95% CI [-0.58, -0.12]), as well as when they disagreed with the last comment they saw (b = 0.19, z = 11.66, p < .001, 95% CI [0.16, 0.22]). Finally, whether participants were recruited from the paid samples was again controlled for. In line with previous results, an event was more likely to be a new comment when the participant was part of the paid samples (b = 0.74, z = 9.15, p < .001, 95% CI [0.58, 0.90]).

Summary

On the aggregated level, our results showed that participants submitted more comments and cast more votes when they agreed with proportionally more comments

submitted by other participants. These findings suggested that participants were more willing to engage on Polis when the environment was subjectively more attitude-consistent, which was in line with the notion of echo chambers. However, our models also showed that these relationships were not linear and that the level of engagement was influenced by more than just a subjective perception of a homogeneous environment. The relationship between the number of comments submitted by participants and the proportion of their agree votes was weaker as the latter increased. Another model showed that the relationship between the number of votes cast by participants and the proportion of their agree votes also became weaker when they agreed with the vast majority of the comments they saw.

To directly test the effect of a comment with which a participant disagreed, we further investigated the events of the participants over time. Our results showed a participant was more likely to submit a comment when their proportion of agree votes up until that point was lower, and when their last event was a disagree vote. These findings suggested that participants were more willing to actively engage on Polis at the time points which they disagreed with the comments they saw, and were consistent with the idea of the confrontation theory that people tend to counterargue with opinions which are attitude-inconsistent.

Study 2

Method

The data was collected by The Cognition Company in collaboration with United Nations Development Programme (UNDP), which recruited participants to Polis to understand public opinions in Pakistan.[8] The study was conducted entirely in English.

The question shown to participants on Polis was "What challenges need to be tackled by the Government to make Pakistan a better place to live in, both during COVID and in post-COVID times?" The interface and functionalities of Polis were otherwise identical to those in Study 1. Comments were moderated by staff members at The Cognition Company.

The data was collected over roughly three months.[9] In contrast to Study 1, all

[8] Data collection also took place in Bhutan and Timor-Leste, but the samples were small due to difficulties in recruiting participants and were therefore excluded from the analysis.

[9] Before any participant was recruited, 23 comments were submitted to Polis.

participants (*N* = 17,620) were recruited by UNDP via social media platforms, phone messages, or direct emails, and were not paid. The means through which each participant was recruited into the study was not recorded in the data set, neither was demographic information.

The study was approved internally by the doctoral programme office at Warwick Business School.

Results

Summary Statistics

Participants. The vast majority of participants (91.11%) cast at least one vote. The median number of votes was 8. The difference in the number of votes cast when compared to Study 1 was likely due to participants not being paid in Study 2. Also, while English is an official language in Pakistan, it is widely used as a second language (McArthur, 2002) which might have affected participants' engagement on Polis. In addition, 31.78% of the participants submitted at least one comment, with the median number of comments being one.

Events. 388,995 events were recorded, including 8,832 comments and 380,163 votes. After moderation, 157 comments were displayed to the participant[10] and the median number of votes cast to these comments was 2,237.

The topics of all moderated comments were manually coded and shown in Table 2. Topics were coded by The Cognition Company. On average, the comments related to any topic were agreed by more than 88% of all participants. Except for gender and governance, the standard deviations of proportions of agree votes were very low for all topics, which suggested that these topics were not polarised among participants.

Table 2

Topics among all moderated comments in Study 2.

Topic	Prevalence	Mean Proportion of Agree Votes	SD of Proportions of Agree Votes
Education	0.25	0.93	0.07
Welfare	0.17	0.94	0.05

[10] Similar to Study 1, comments that covered the same topic and with similar meanings were moderated out.

Topic	Prevalence	Mean Proportion of Agree Votes	SD of Proportions of Agree Votes
Employment	0.16	0.95	0.03
Law & Order	0.16	0.96	0.02
Gender	0.12	0.88	0.15
Governance	0.10	0.90	0.10
Technology	0.10	0.94	0.03
Corruption	0.09	0.95	0.02
Healthcare	0.07	0.95	0.02
Economy	0.06	0.95	0.01
COVID-19	0.05	0.94	0.01
Environment	0.05	0.96	0.02
Tax	0.03	0.94	0.01

Inferential Statistics

The same exclusion criterion from Study 1 were used. That is, participants whose numbers of votes were in the lower (25%) quantile were removed.[11] After this exclusion, 98.87% of all events remained in the samples.

As shown above in Table 2, participants in Study 2 agreed with most comments in most topics. When dividing participants into groups based on the proportions of agree votes cast by them (Figure 4), it was shown that 62.50% of participants agreed with over 90% of all comments they saw on Polis and 14.40% of participants agreed with 80-90% of all comments. The remaining participants constitute only 23.10% of the whole sample. Therefore, analyses on the aggregated level were not repeated for Study 2, since the models use the proportions of agree votes as independent variable to explain differences in the numbers of comments and votes of the participants, and the lack of variability in the proportions of agree votes greatly reduced the validity of such models. The following analyses focused on the events

[11] Similar to Study 1, we repeated the statistical models below for all participants. These alternative models are reported in the Supplementary Materials and do not change the conclusions of the paper.

experienced by participants over time and the same random-intercept logistic regression models in Study 1 were used.

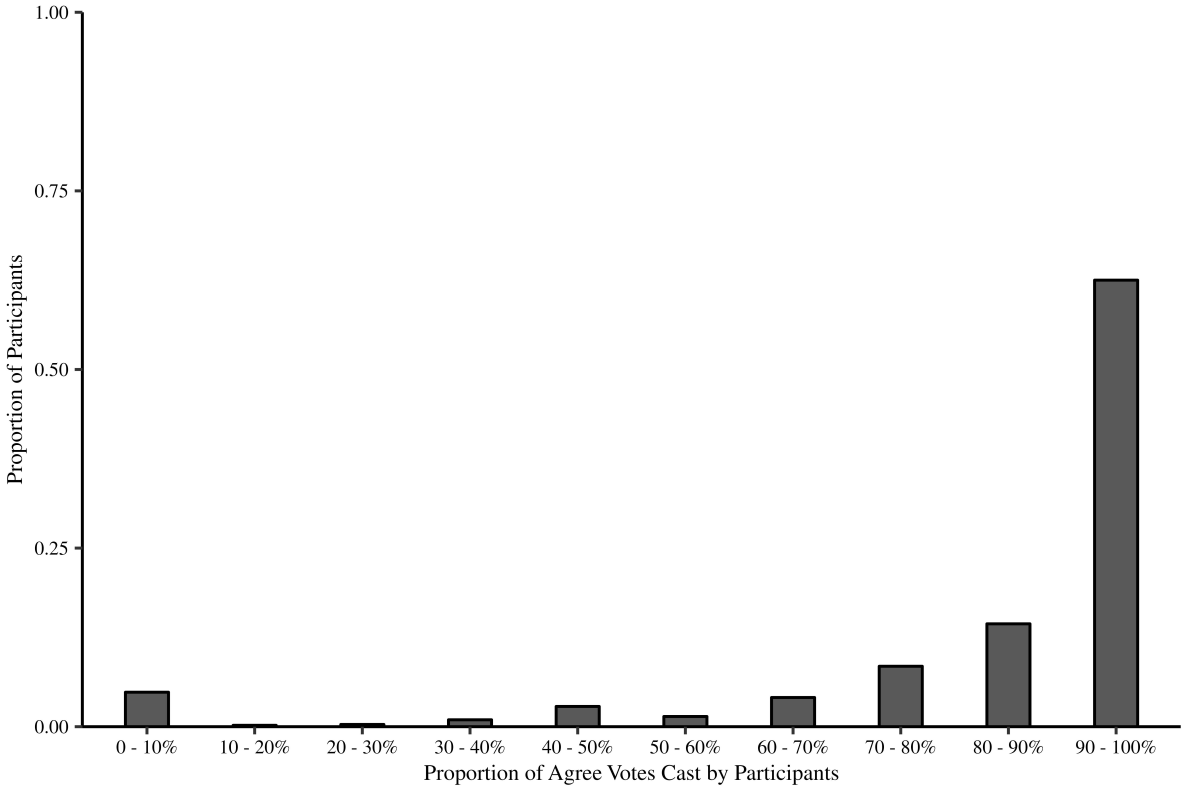


Figure 4

Proportions of participants grouped by the proportions of agree votes cast by them in Study 2.

Analysing Events. Similar to Study 1, only a small proportion of events were the submission of new comments (2.14%). Again, the frequency of an event being the submission of a comment was higher when a participant's last event was a disagree vote (2.00%) than when it was an agree vote (1.76%).

We performed the same analysis as in Study 1 and tested whether an event was more or less likely to be the submission of a new comment (1) when a participant agreed with proportionally more comments up until the point of the event and (2) when the last event of a participant was a disagree vote. Results of a random-intercept logistic regression model showed that a participant was more likely to submit a comment when they disagreed with the last comment they saw (b = 0.08, z = 3.02, p = .002, 95% CI [0.03, 0.13]), but the proportion of agree votes up until that point had no effect (b = 0.15, z = 1.54, p = .125, 95% CI [-0.04, 0.35]).

Discussion

Political polarisation is commonly attributed to echo chambers, which refer to the observations that individuals prefer to consume attitude-consistent information (Garrett, 2009; Graf & Aday, 2008; Iyengar & Hahn, 2009) and keep contacts with like-minded others (Barberá et al., 2015; Del Vicario et al., 2017; Himelboim et al., 2013). However, it has been suggested that polarisation could also be a result of individuals' tendencies to challenge those who are ideologically different and to counterargue opposing opinions. As it is unlikely that an individual's interactions with other users on a mainstream platform would be entirely homogeneous (Bakshy et al., 2015; Conover et al., 2011; Williams et al., 2015), it is important to understand the complex roles of echo chambers and confrontation in public opinions, especially how individuals' tendency to express their views can be influenced.

The two studies in this paper directly investigated how individuals deliberated on online platforms and examined their willingness to engage using the submission of comments and votes as proxies. The first study focused on participants from the UK, while data collection in the second study took place in Pakistan. With the British sample, we found that individuals submitted more comments and cast more votes on other participants' comments when they perceived the environment as more consistent to their beliefs, as captured by the proportion of agree votes they cast. Furthermore, the relationships were not linear and were weaker as the proportion of agree votes increased, especially when the proportion of agree votes was extremely high. An interpretation is that individuals' willingness to engage reached its peak when they were generally exposed to a homogeneous environment which was aligned with their sociopolitical views, but also when they encountered a small number of opposing opinions, which provided the initial support for an interplay between echo chambers and a behavioural response to contradictory opinions.

To better understand the impacts of attitude-inconsistent information on online engagement platforms and test our predictions, we directly studied individuals' behaviours over time. In the first study with Bristish samples, we found that participants were more likely to submit their own comments when they had just seen a comment with which they disagreed and when their proportion of agree votes up until that point was lower. These results were in

line with the notion of the confrontation theory, which suggested that individuals tend to refute statements which are contradictory to their worldviews. In the second study with Pakistani samples, we also found that participants were more likely to submit comments when they disagreed with the last comments they saw. However, the proportion of agree votes did not have any effect. These results could be explained by the differences between the two samples: the topics submitted and voted on by participants in the first study were much more polarised than those in the second study. Popular topics in the first, British study included Brexit, immigration, and racial tensions, which are known to divide individuals on online platforms (Del Vicario et al., 2017; Himelboim et al., 2013). On the contrary, popular topics in the second study included education, employment, and technological infrastructure, on which Pakistani participants had reached consensus. The difference in topics provided by the participants could be a result of the different development levels between the two countries, as well as cultural factors. The weaker effect in the model could then reflect the overall lower number of disagree votes and the lack of variability in the proportion of agree votes over time in the second study. In other words, the environment in the second study could be considered as much more homogeneous than that in the first study, and the encounter of opposing arguments was less common and thus played a lesser role in participants' willingness to engage in discussions.

Overall, our findings provide support for the intertwined roles of echo chambers and confrontation in public opinions. Taken together with the results of previous studies (Bail et al., 2018; Karlsen et al., 2017; Stromer-Galley & Muhlberger, 2009), it is clear that engagement on online platforms is determined by multiple psychological and social factors. Our results also have implications for political and communication science: although the exposure to heterogenous environments and thus opposing opinions can promote engagement, the underlying driver of this behaviour might increase political polarisation due to the reinforcement of existing beliefs through counterarguments. Although our studies only involved one online platform and future research is needed to assess the external validity of our results, the setups of our studies shared many similarities to popular social media platforms, and therefore improve our general understanding of how individuals interact with

others during online debates.

Naturally, political polarisation is a complex topic and can be investigated with a multitude of methodologies. For instance, natural language processing (NLP) can be used to categorise the topics of the comments submitted by participants and compute their sentiment values. These methods can be used to verify the results presented in this paper, such as to examine whether participants did express their views on the same topic when they provided a comment right after seeing another comment with which they disagreed, and thus whether they did counterargue contradictory viewpoints. Sentiment analysis would also can allow us to study whether such comments have negative emotional values, hence provides another analysis on the notion of confrontation. However, the data sets in this paper have several limitations which restricted the use of NLP techniques. The most notable ones are that participants had a limited word count for each comment and most participants provided only a small number of comments over the course of their stays on Polis. These factors pose challenges to accurately identifying the topics or calculating sentiment values of the comments, while the individual differences between how participants compose their comments might also bias the results.

The willingness to engage with other individuals on an online platform can be analysed with other techniques, such as computational models originated from cognitive psychology. The multilevel models presented in this paper treated events as binary, that is, whether an event was a new comment submitted by a participant. This can be represented as an accumulation model, such that whether a participant agreed with a comment left by other participants either increases or decreases their tendency to be involved in the discussion and express their own opinions, and participants can be assumed to be willing to do so once this tendency exceeds certain cognitive threshold. With parameter estimation techniques include maximum likelihood models or their Bayesian counterparts, this will allow researchers to capture how individuals' willingness to engage change over time and also the weight of each comment with which they disagreed.

The implications of echo chambers and confrontation can be further studied with agent-based simulations. The findings in this paper suggest that individuals have the preference to spend time in an environment which is consistent with their views and tend to

counterargue any content which are contradictory, therefore it is a possibility that political polarisation is very difficult to avoid once individuals were engaged to freely express their opinions on such platforms. In other words, individuals reinforce their beliefs by consuming information generated by like-minded others and also do so via refuting other types of information, despite that the sources of information are heterogeneous. This can be investigated by simulating how users interact with others on social media or in real life, and whether the emergence of isolated networks and the elimination of groups which hold moderate opinions are inevitable given certain assumptions on human information processing.

Supplementary Materials (Alternative Models)

Supplementary Materials (Alternative Models)

Study 1

Controlling for the Facebook Sample

In Study 1, all statistical models reported in the paper controlled for the potential differences between the paid sample and the volunteer samples. As the Facebook sample cast much more votes than the rest of the participants, we performed alternative analyses and controlled for whether participants were recruited from Facebook. The models were otherwise the same as those reported in the paper and only included the participants that passed our exclusion criterion (i.e., those who cast six or more votes). Including this control variable did not change the conclusions of the paper.

Analysing the Aggregated Data of Participants

When investigating the relationship between the number of comments submitted by participants and the proportion of agree votes cast by them, our alternative Poisson regression model showed that participants submitted more comments when they agreed with more comments submitted by others (b = 7.71, z = 12.64, p < .001, 95% CI [6.53, 8.92]). The quadratic term was significant, showing that this relationship was weaker as the proportion of agree votes increased (b = -6.15, z = -13.52, p < .001, 95% CI [-7.06, -5.27]). Finally, participants recruited from Facebook tended to submit fewer comments (b = -0.06, z = -4.93, p < .001, 95% CI [-0.085, -0.037]).

We then examined the relationship between the number of votes cast by participants and their proportion of agree votes. Our alternative Poisson regression model showed that participants cast more votes when they agreed with more comments (b = 5.04, z = 57.04, p < .001, 95% CI [4.87, 5.22]). Again, the quadratic term was significant and showed that this effect was weaker as the proportion of agree votes increased (b = -2.69, z = -42.62, p < .001, 95% CI [-2.82, -2.57]). Furthermore, participants recruited from Facebook cast more votes (b = 0.19, z = 109.20, p < .001, 95% CI [0.18, 0.19]).

Analysing Events

When we investigated the events of the participants over time, our alternative random-intercept logistic regression model showed that a participant was more likely to submit a new comment when the proportion of agree votes cast by them up until that point decreases (b = -0.30, z = -2.57, p < .001, 95% CI [-0.53, -0.07]) and when their last event was a disagree vote (b = 0.19, z = 11.69, p < .001, 95% CI [0.16, 0.22]). Additionally, an event was less likely to be a new comment when the participant was part of the Facebook samples (b = -0.82, z = -10.26, p < .001, 95% CI [-0.98, -0.66]).

Including All Participants

We further repeated the statistical models in Study 1 with all participants, that is, regardless of whether they cast six or more votes. Consistent with the models reported in the paper, these models controlled for whether participants were recruited from the paid samples. These models did not qualitatively change our conclusions.

Analysing the Aggregated Data of Participants

For the relationship between the number of comments submitted by participants and the proportion of agree votes cast by them, our alternative Poisson regression model showed that participants submitted more comments when they agreed with more comments submitted by others (b = 8.99, z = 22.05, p < .001, 95% CI [8.21, 9.81]). The quadratic term was also significant, showing that this relationship was weaker as the proportion of agree votes increased (b = -7.37, z = -24.42, p < .001, 95% CI [-7.98, -6.79]). Finally, participants from the paid samples submitted more comments (b = 0.12, z = 9.85, p < .001, 95% CI [0.10, 0.14]).

As for the relationship between the number of votes cast by participants and their proportion of agree votes, our alternative Poisson regression model showed that participants cast more votes when they agreed with more comments (b = 15.27, z = 170.90, p < .001, 95% CI [15.09, 15.44]). The quadratic term was again significant and showed that this effect was weaker as the proportion of agree votes increased (b = -10.57, z = -171.80, p < .001, 95% CI [-10.69, -10.45]). Furthermore, participants recruited from the paid sample cast fewer votes (b = -0.058, z

= -31.60, $p < .001$, 95% CI [-0.062, -0.055]).

Analysing Events

We then investigated the events of the participants over time. Our alternative random-intercept logistic regression model showed that a participant was more likely to submit a new comment when the proportion of agree votes cast by them up until that point decreases (b = -0.23, z = -2.09, p =0.04, 95% CI [-0.46, -0.014]) and when their last event was a disagree vote (b = 0.19, z = 11.99, $p < .001$, 95% CI [0.16, 0.22]). Additionally, an event was more likely to be a new comment when the participant was recruited from the paid samples (b = 0.84, z = 10.46, $p <$ .001, 95% CI [0.69, 1.00]).

Study 2

Including All Participants

In Study 2, we also repeated the statistical models with all participants regardless of whether they passed the exclusion criterion. This did not change our conclusions in Study 2.

Analysing Events

When investigating the events of the participants over time, our alternative random-intercept logistic regression model showed that a participant was more likely to submit a new comment when their last event was a disagree vote (b = 0.08, z = 3.11, p = .002, 95% CI [0.03, 0.13]). However, whether the proportion of agree votes up until that point did not have a significant effect (b = 0.17, z = 0.08, p = .10, 95% CI [-0.026, 0.36]).

Conclusion

Conclusion

The first two papers in this thesis investigated the cognitive mechanisms behind the attraction effect, while the third paper examined the psychological factors that influenced the dynamics of public opinions.

The attraction effect is a classic phenomenon in decision research and there has been a wide range of studies which attempted to explain its emergence. As summarised in the introduction of this thesis, computational models explain the attraction effect by making assumptions on different psychological processes during multialternative choice tasks, while the more recent models explicitly assume how the attribute values of the options can affect how decision makers shift their attention between the attributes and evaluate the options (Bhatia, 2013; Noguchi & Stewart, 2018). The latter is closely linked to the use of process-tracing methods in studying context effects (Król & Król, 2019; Marini et al., 2020; Noguchi & Stewart, 2014), since the results of process-tracing studies can inform the development of new models, while the predictions of models can be, in turn, tested by process data (Johnson et al., 2008).

The first two papers in this thesis focused on a high-level reason-based decision framework known as Query Theory (Johnson et al., 2007). While Query Theory is not a computational model, it similarly makes assumptions on how contexts can influence information search and how evidence is sequentially gathered by decision makers to support or oppose the options. These assumptions are commonly tested by a reason listing protocol, which is a type of process-tracing method (Schulte-Mecklenbeck et al., 2017). Query Theory can explain many phenomena in binary choices (Dinner et al., 2011; Hardisty et al., 2010; Spälti et al., 2017; Weber et al., 2007), but to my knowledge, the first paper in this thesis is the first piece of research which demonstrated that the theory can be extended from binary to multialternative choices. Our results showed that the deliberation mechanisms proposed by Query Theory can explain the attraction effect: the existence of a decoy option can lead to decision makers evaluating the target option earlier, as reflected by the observations that reasons supporting the target were generated earlier during the decision task. Due to output interference, such reasons were also generated in a larger amount. Furthermore, motivated by

the recent discussions on the prevalence of the attraction effect in multialternative choices beyond its simplest forms, we repeated our study with choice sets of different levels of complexity. Results showed that the attraction effect could also be found when we doubled the number of attributes, while the reasoning mechanisms characterised by Query Theory could explain decision outcomes regardless of the amount of information in the choice sets. These results illustrated the generality of the attraction effect when stimuli are presented numerically, as well as that of Query Theory as a decision framework.

The second paper is a natural follow-up to the above results and studies the links between reasoning and attention in multialternative choices. By combining reason-listing and mouse-tracking methods, we investigated how decision makers sample information during different stages of their reason generation processes. Results of analyses performed on the reasoning data replicated the findings of the first paper, showing that reasons supporting the target were generated earlier and in a greater amount. Results of the mouse-tracking data showed that decision makers paid more attention to the chosen option than its competitor, and transitions occurred more frequently between the chosen option and the decoy than between the competitor and the decoy, both of which supported previous findings (Marini et al., 2020; Noguchi & Stewart, 2014; Shimojo et al., 2003). We then analysed reasoning and mouse-tracking data together. Results showed that the two reason structure metrics (i.e., quantity and positions of reasons) partially mediated the effects of information search on choices. Additionally, when we examined how individuals sample information before they generated each reasons, results showed that mouse clicks predict the types of reasons. That is, a reason was more likely to support an option when a participant paid more attention to its attributes in that particular stage of the reason generation process. These findings support that notion that information sampled throughout decision making is aggregated into reasons, which are subsequently accumulated into evidences that drive choices. However, as we only found partial mediations, the relationship between attentional processes and reasoning can be more complicated and is worth further investigations.

An avenue for future research is to explain the two other main types of context effects, namely the compromise (Simonson, 1989) and similarity effects (Tversky, 1972), with Query

Theory. Unlike the attraction effect, it is expected that some participants will choose the decoy option in choice tasks set up to demonstrate these two effects, as the decoy is not asymmetrically dominated. This poses a challenge to the analyses, as how to generalise the Query Theory metrics to capture the positions and quantity of reasons supporting all three options (i.e., instead of treating the decoy option mostly as an irrelevant alternative) will be a key consideration. Another possibility for future research is to computationally implement Query Theory, taking both reasoning and attentional processes into account, but this is beyond the scope of this thesis. Finally, while previous studies have demonstrated that the attraction effect is mostly limited to numeric stimuli (Frederick et al., 2014; Trendl et al., 2021; Yang & Lynn, 2014), it is worth investigating how the strength of the effect will change when the complexity level of the choice sets systematically increases. The results from the second paper provided evidence that the attraction effect can be observed when the number of attributes is doubled, but the effect of increasing the number of options is less explored.

Similar to multialternative choices, public opinions is a popular topic in behavioural science and has drawn considerable attention in recent years, perhaps because the studying of communications between individuals has been made easier by the prevalence of social media platforms. The dynamics of public opinions is frequently characterised by echo chambers (Sunstein, 2001, 2007), which are in turn commonly explained by the theory of selective attention (Garrett, 2009; Graf & Aday, 2008; Iyengar & Hahn, 2009). Recent studies, however, highlighted that individuals tend to challenge opposing viewpoints and provide counterargument (Taber et al., 2009; Taber & Lodge, 2006), and suggested that people do communicate with others with different sociopolitical stances in the form of confrontation (Karlsen et al., 2017).

The main goal of the third paper of this thesis was to understand how individuals engage in online discussions from the perspectives of both echo chambers and confrontations, as it is possible that the formation and reinforcement of beliefs are results of multiple factors which together lead to attitude polarisation. With two data sets provided by external organisations, we investigated how participants expressed opinions and voted on others' comments across a multitude of time points. This method is different from previous studies,

which mostly relied on experiments or network analyses, and has the advantage of allowing us to understand how individuals responded to the changing environments and how they behaved when exposed to attitude-inconsistent information. Results of Poisson regression models showed that participants were more willing to engage in expressing their views when they agreed with proportionally more comments provided by others. In other words, participants were more likely to engage in discussions when the environment was perceived as homogeneously consistent with their own beliefs. However, the non-linear relationships, as reported in the third paper, suggested that a small number of comments with which participants disagreed also prompted them to express their own opinions. Multilevel models further demonstrated that participants were more likely to provide their comments when they just saw a comment with which they disagreed. In sum, the results showed supports for both the notion of echo chambers and the theory of confrontation, as individuals preferred an environment in which other people shared similar sociopolitical views, but the existence of opposing views could also promote individuals to express their own opinions.

Although the results of the third paper captured the behaviours of individuals on the specific online platforms from which the data was collected, it has important implications for political polarisation in general: our results supported the views of Bail et al. (2018) that being exposed to opposing opinions does not necessarily reduce polarisation, but instead can backfire and potentially reinforce one's existing beliefs. A number of studies have suggested that individuals on the two ends of the left-right political spectrum hold different fundamental values (Graham et al., 2009; Jost et al., 2007), therefore it is possible that political stances play a role in the counterarguing behaviours and how people react to others' opinions (Bail et al., 2018). However, as demographic information was recorded only for a small number of participants and only in the first study of the third paper, it is difficult to draw conclusions on this and further research is needed. Finally, as the comments provided by participants were relatively short, it is difficult to perform analyses involving natural language processing techniques. Future research can focus on whether the sentiment values of comments provided by individuals change when they are exposed to attitude-inconsistent opinions, and how that differs between individuals holding different sociopolitical views.

Bibliography

Bibliography

Adjerid, I., Samat, S., & Acquisti, A. (2016). A Query-Theory Perspective of Privacy Decision Making. *The Journal of Legal Studies*, *45*(S2), S97–S121.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132.

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, *72*(5), 784–811.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, *26*(10), 1531–1542.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, *35*(2), 180–191.

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, *120*(3), 522–543.

Bhatia, S., & Stewart, N. (2018). Naturalistic multiattribute choice. *Cognition*, *179*, 71–88.

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in Cognitive Sciences*, *23*(3), 251–263.

Chaffee, S. H., & Miyo, Y. (1983). Selective exposure and the reinforcement hypothesis: An intergenerational panel study of the 1980 presidential campaign. *Communication*

Research, *10*(1), 3–36.

Cheng, Y.-H., Chuang, S.-C., Huang, M. C.-J., & Hsieh, W.-C. (2012). More than two choices: The influence of context on the framing effect. *Current Psychology*, *31*(3), 325–334.

Cohen, A. L., Kang, N., & Leise, T. L. (2017). Multi-attribute, multi-alternative models of choice: Choice, reaction time, and process tracing. *Cognitive Psychology*, *98*, 45–72.

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2015). *Echo chambers in the age of misinformation*. https://arxiv.org/abs/1509.00189

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559.

Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, *50*, 6–16.

Dinner, I., Johnson, E. J., Goldstein, D. G., & Liu, K. (2011). Partitioning default effects: Why people choose not to choose. *Journal of Experimental Psychology: Applied*, *17*(4), 332–341.

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*, 563–588.

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. *Journal of Marketing Research*, *51*(4), 487–507.

Garrett, R. K. (2009). Echo chambers online? Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication*, *14*(2), 265–285.

Graf, J., & Aday, S. (2008). Selective attention to online political information. *Journal of Broadcasting & Electronic Media*, *52*(1), 86–100.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5),

1029–1046.

Hardisty, D. J., Johnson, E. J., & Weber, E. U. (2010). A dirty word or a dirty world? Attribute framing, political affiliation, and query theory. *Psychological Science*, *21*(1), 86–92.

Heath, T. B., & Chatterjee, S. (1995). Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence. *Journal of Consumer Research*, *22*(3), 268–284.

Herne, K. (1997). Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy*, *13*(3), 575–589.

Highhouse, S. (1996). Context-dependent selection: The effects of decoy and phantom job candidates. *Organizational Behavior and Human Decision Processes*, *65*(1), 68–76.

Hilbert, M., & López, P. (2011). The world's technological capacity to store, communicate, and compute information. *Science*, *332*(6025), 60–65.

Hills, T. T. (2019). The dark side of information proliferation. *Perspectives on Psychological Science*, *14*(3), 323–330.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication*, *18*(2), 154–174.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*(1), 90–98.

Huber, J., Payne, J. W., & Puto, C. P. (2014). Let's be honest about the attraction effect. *Journal of Marketing Research*, *51*(4), 520–525.

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, *10*(1), 31–44.

Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, *59*(1), 19–39.

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, *302*(5649), 1338–1339.

Johnson, E. J., 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. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, *115*(1), 263–273.

Jost, J. T., Napier, J. L., Thorisdottir, H., Gosling, S. D., Palfai, T. P., & Ostafin, B. (2007). Are needs to manage uncertainty and threat associated with political conservatism or ideological extremity? *Personality and Social Psychology Bulletin*, *33*(7), 989–1007.

Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication*, *32*(3), 257–273.

Kelman, M., Rottenstreich, Y., & Tversky, A. (1996). Context-dependence in legal decision making. *The Journal of Legal Studies*, *25*(2), 287–318.

Król, M., & Król, M. (2019). Inferiority, not similarity of the decoy to target, is what drives the transfer of attention underlying the attraction effect: Evidence from an eye-tracking study with real choices. *Journal of Neuroscience, Psychology, and Economics*, *12*(2), 88–104.

Lawrence, E., Sides, J., & Farrell, H. (2010). Self-segregation or deliberation? Blog readership, participation, and polarization in American politics. *Perspectives on Politics*, *8*(1), 141–157.

Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes*, *68*(1), 28–43.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, *15*(3), 215–233.

Marini, M., Ansani, A., & Paglieri, F. (2020). Attraction comes from many sources:

Attentional and comparative processes in decoy effects. *Judgment and Decision Making*, *15*(5), 704–726.

Maylor, E. A., & Roberts, M. A. (2007). Similarity and attraction effects in episodic memory judgments. *Cognition*, *105*(3), 715–723.

McArthur, T. (2002). *The Oxford guide to world English*. Oxford University Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Noguchi, T., & Stewart, N. (2014). In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition*, *132*(1), 44–56.

Noguchi, T., & Stewart, N. (2018). Multialternative decision by sampling: A model of decision making constrained by process data. *Psychological Review*, *125*(4), 512–544.

Office for National Statistics. (2017). *Graduates in the UK labour market*. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/graduatesintheuklabourmarket/2017

Orquin, J. L., & Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behavior Research Methods*, *50*(4), 1645–1656.

Pettibone, J. C., & Wedell, D. H. (2000). Examining models of nondominated decoy effects across judgment and choice. *Organizational Behavior and Human Decision Processes*, *81*(2), 300–328.

Rijnhart, J. J., Twisk, J. W., Eekhout, I., & Heymans, M. W. (2019). Comparison of logistic-regression based methods for simple mediation analysis with a dichotomous outcome variable. *BMC Medical Research Methodology*, *19*(1), 1–10.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*(2), 370–392.

Schulte-Mecklenbeck, M., Johnson, J. G., Böckenholt, U., Goldstein, D. G., Russo, J. E., Sullivan, N. J., & Willemsen, M. C. (2017). Process-tracing methods in decision making: On growing up in the 70s. *Current Directions in Psychological Science*,

26(5), 442–450.

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31*(2), 194–213.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322.

Simion, C., & Shimojo, S. (2007). Interrupting the cascade: Orienting contributes to decision making even in the absence of visual stimulation. *Perception & Psychophysics*, *69*(4), 591–595.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158–174.

Slaughter, J. E., Sinar, E. F., & Highhouse, S. (1999). Decoy effects and attribute-level inferences. *Journal of Applied Psychology*, *84*(5), 823–828.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhart (Ed.), *Sociological methodology 1982* (Vol. 13, pp. 290–312). Jossey-Bass.

Spälti, A. K., Brandt, M. J., & Zeelenberg, M. (2017). Memory retrieval processes help explain the incumbency advantage. *Judgment and Decision Making*, *12*(2), 173–182.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(5), 645–665.

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

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, *32*(2), 101–130.

Stromer-Galley, J., & Muhlberger, P. (2009). Agreement and disagreement in group deliberation: Effects on deliberation satisfaction, future engagement, and decision legitimacy. *Political Communication*, *26*(2), 173–192.

Sunstein, C. (2001). *Republic.com*. Princeton University Press.

Sunstein, C. (2007). *Republic.com 2.0*. Princeton University Press.

Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments.

Political Behavior, *31*(2), 137–155.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769.

Trendl, A., Stewart, N., & Mullett, T. L. (2021). A zero attraction effect in naturalistic choice. *Decision*, *8*(1), 55–68.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*(2), 179–205.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science*, *24*(6), 901–908.

Trueblood, J. S., & Pettibone, J. C. (2017). The phantom decoy effect in perceptual decision making. *Journal of Behavioral Decision Making*, *30*(2), 157–167.

Turner, B. M., Schley, D. R., Muller, C., & Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review*, *125*(3), 329–362.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Usher, M., Elhalal, A., & McClelland, J. L. (2008). The neurodynamics of choice, value-based decisions, and preference reversal. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 277–300). Oxford University.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, *111*(3), 757–769.

Van den Bosch, A., Bogers, T., & De Kunder, M. (2016). Estimating search engine index size variability: A 9-year longitudinal study. *Scientometrics*, *107*(2), 839–856.

Weber, E. U., & Johnson, E. J. (2006). Constructing preferences from memories. In S.

Lichtenstein & P. Slovic (Eds.), *The construction of preference* (pp. 397–410). Cambridge University Press.

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science*, *18*(6), 516–523.

Williams, H. T., McMurray, J. R., Kurz, T., & Lambert, F. H. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, *32*, 126–138.

Yang, S., & Lynn, M. (2014). More evidence challenging the robustness and usefulness of the attraction effect. *Journal of Marketing Research*, *51*(4), 508–513.

Zhou, L., Kim, C., & Laroche, M. (1996). Decision Processes of the Attraction Effect: A Theoretical Analysis and Some Preliminary Evidence. In K. P. Corfman & John G Lynch Jr (Eds.), *ACR North American Advances* (Vol. 23, pp. 218–224). Association for Consumer Research.

Appendix

Appendix

We conducted a partial replication of the work of Dinner et al. (2011), who explained the default effect (Johnson & Goldstein, 2003) with Query Theory. The main purpose of this study is to test whether Query Theory is an entirely memory-based framework, that is, whether the order and quantity of reasons can also predict choices when the stimuli were constantly presented to decision makers.[1]

Design

The experiment was mostly based on the design of Dinner et al. (2011), with minor changes in the attribute values (e.g., currency) in order to localise the study for British participants. A major change is the addition of a *decision environment* manipulation, in which the choice set table was either presented to the participants throughout aspect listing and choice (i.e., stimuli-based), or shown to participants only initially before they were asked to provide their reasons and make their decisions (i.e., memory-based).

Results

The Default Effect

The default effect was replicated: results of a logistic regression model showed that the default option was chosen more frequently than the non-default one (b = 0.33, z = 2.97, p = .003, 95% CI [0.11, 0.54]). This effect did not depend on which option was set up as the default (b = -0.09, z = -0.86, p = .392, 95% CI [-0.31, 0.12]), and did not depend on whether the decision environment was stimuli-based or memory-based (b = -0.07, z = -0.64, p = .521, 95% CI [-0.29, 0.14]). Their interaction did not have any effect on choice either (b = -0.05, z = -0.42, p = .674, 95% CI [-0.26, 0.17]).

Analysis of Reason Structure

The definitions of content and order scores are identical to those in the study of Dinner et al. (2011). The content score of a participant is defined as follows:

$$Content = \frac{n_{Default} - n_{Non\text{-}Default}}{N}$$

[1] The preregistration can be found on OSF Registries: https://osf.io/7u3yq

where $n_{Default}$ is the number of reasons supporting the default option provided by the participant, $n_{Non\text{-}Default}$ is the number of reasons supporting the non-default option, while N is the total number of reasons. A Content score of 1 represents that all reasons submitted by the participant were in favour of the default, while -1 means that all reasons were in favour of the non-default option.

The Order score of a participant is defined as follows:

$$Order = \frac{2 \times (MR_{Non\text{-}Default} - MR_{Default})}{N}$$

where $MR_{Default}$ is the median rank of reasons supporting the default option and $MR_{Non\text{-}Default}$ is that of reasons supporting the non-default option. A Order score of 1 represents that all reasons in favour of the default provided by the participant were generated earlier than those in favour of the non-default option, and vice versa for a score of -1.

Quantity of Reasons

Results supported the hypothesis that the quantity of reasons could predict choices, as the results of a logistic regression model showed that participants were more likely to choose the default when they generated more reasons supporting it (as captured by the Content score: b = 3.02, z = 9.83, p < .001, 95% CI [2.46, 3.67]). Whether participants chose the default option did not depend on which option was set up as the default (b = 0.33, z = 1.93, p = .053, 95% CI [0.0025, 0.67]), or whether the decision environment was stimuli-based or memory-based (b = -0.07, z = -0.41, p = .678, 95% CI [-0.38, 0.25]). The interaction between the default manipulation and Content score was not significant (b = 0.08, z = 0.26, p = .792, 95% CI [-0.53, 0.69]), nor was the interaction between decision environment and Content score (b = 0.16, z = 0.57, p = .571, 95% CI [-0.40, 0.75]).

Order of Reasons

Similarly, results supported the hypothesis that the order of reasons could predict choices. Results of a logistic regression model showed that participants were more likely to choose the default when they generated reasons supporting it earlier (as captured by the Order score: b = 1.81, z = 10.07, p < .001, 95% CI [1.47, 2.18]). Whether participants chose the default option was not associated with which option was set up as the default (b = 0.02, z =

0.11, p = .914, 95% CI [-0.26, 0.30]), or whether the decision environment was stimuli-based or memory-based (b = 0.08, z = 0.59, p = .554, 95% CI [-0.19, 0.37]). The interaction between the default manipulation and Order score did not have an effect on choices (b = 0.05, z = 0.28, p = .780, 95% CI [-0.30, 0.40]), nor did the interaction between decision environment and Order score (b = 0.14, z = 0.77, p = .443, 95% CI [-0.21, 0.50]).

Conclusion

The key findings of this study is the lack of either main effects or interactions for the decision environment in all of the above models. These results demonstrated that Query Theory can be applied to explain binary choices in a stimuli-based environment, that is, when the choice set was shown to individuals during the entirety of the decision process. This serves as a backbone of the first two papers in this thesis which investigated whether Query Theory can be extended from binary to multialternative choices, as the latter are commonly studied with stimuli-based tasks.