

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/83196>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

Perceptual Comparison of Source-Code Plagiarism within Students from UK, China, and South Cyprus Higher Education Institutions

GEORGINA COSMA, College of Science and Technology, Nottingham Trent University, UK

MIKE JOY, Department of Computer Science, University of Warwick, UK

JANE SINCLAIR, Department of Computer Science, University of Warwick, UK

MARGARITA ANDREOU, Department of Business Computing, PA College, Cyprus

DONGYONG ZHANG, College of Information and Management Science, Henan Agricultural University, Zhengzhou, PR China

BEVERLEY COOK, College of Science and Technology, Nottingham Trent University, UK

RUSSELL BOYATT, Department of Computer Science, University of Warwick, UK

Perspectives of students on what constitutes source-code plagiarism may differ based on their educational background. Surveys have been conducted with home students undertaking computing and joint computing subject degrees at Higher Education Institutions throughout the UK, China and South Cyprus, and a total of 984 responses have been statistically analysed to determine the common areas of understanding and misunderstanding among students on various topics related to source-code plagiarism. The study identifies those topics which are well understood, and those topics which are not properly understood across the different groups of students, and is the first study which specifically discusses Cypriot student perceptions on source-code plagiarism. This study provides useful information to educators (teaching home and international students) who wish to better inform their students on the issues of plagiarism and source-code plagiarism. Finally, the survey results revealed that although students who were informed about plagiarism better understood what actions constitute plagiarism, some topics were still unclear among students regardless of the students' educational background and whether they had been previously informed about plagiarism.

Categories and Subject Descriptors: C.2.2 [Computer-Communication Networks]: Network Protocols

General Terms: Plagiarism, academic integrity, China, Cyprus, UK

Additional Key Words and Phrases: Plagiarism, academic integrity, China, Cyprus, UK

ACM Reference Format:

Georgina Cosma, Mike Joy, Jane Sinclair, Margarita Andreou, Dongyong Zhang, Beverley Cook and Russell Boyatt. Perceptual Comparison of Source-Code Plagiarism within Students from UK, China, and South Cyprus Higher Education Institutions *ACM Trans. Comput. Educ.* 9, 4, Article 39 (March 2014), 16 pages.
DOI: <http://dx.doi.org/10.1145/0000000.0000000>

1. INTRODUCTION

Plagiarism is the intentional or unintentional use of someone else's work without providing appropriate acknowledgement of the sources. Students from different national and cultural backgrounds may have different attitudes and perspectives on what constitutes plagiarism [Foltýnek and Čech 2012]. Hence, plagiarism may be a problem when students choose to study away from their home country.

1.1. International students — is there a problem?

Much research has been conducted on plagiarism in Higher Education Institutions (HEI) on the topics of plagiarism detection, investigation and prevention [Bradley 2011], [Twomey et al. 2009]. Research also exists on identifying student perspectives on plagiarism, but these are surveys which focus English speaking HEIs [Chuda et al. 2012], [Joy et al. 2011]. In such studies, the “home” university has typically been English speaking (predominantly UK or US) and the “international” students referred to are generally from countries where English is not the first language. These studies have often reported a higher rate of plagiarism amongst international students compared to local students [Park 2003]. Some studies of international students in English

speaking universities, such as those of Bamford and Sergiou in the UK [Bamford and Sergiou 2005], and Bista [Bista 2011] and Qi [Qi 2008] in the US, have reported instances of plagiarism accounting for up to 50% of the international students in the cohorts studied. However, other studies have found no correlation between incidence of plagiarism and students' nationality, with Chinese students who took their undergraduate degree in a Chinese university discovered to be less likely to plagiarise than those whose first degree was in the UK [Barrett and Malcolm 2006]. The situation is complex, and Montgomery [Montgomery 2010] warns against simplistic attribution of plagiarism to different learning backgrounds. She suggests instead that academic integrity is a contextual educational value which needs to be embedded throughout the teaching and learning process.

1.2. Factors contributing to misunderstanding

In some cases, plagiarism occurs through lack of understanding [Nadelson 2007]. Even within a country with a "robust" approach to plagiarism education and detection, policy details and penalties may vary between institutions [Tennant et al. 2007], and lack of awareness of expected citation behaviour is consistently cited as a major reason for text-based plagiarism [Bamford and Sergiou 2005], [Marshall and Garry 2005]. Few institutions have specific guidance on the issue of source-code plagiarism and there is evidence that aspects of this are commonly misunderstood by students both in the UK [Joy et al. 2011] and in China [Zhang et al. 2013].

Yakovchuk [Yakovchuk 2008] undertook a study of non-native university students in the UK, and identified a number of reasons for plagiarising given by those students. However, the only reason which was specific to international students was language difficulty. This is reinforced by Timm [Timm et al. 2008] who highlights the relatively low level of English ability required of international students in some UK universities, and notes that such students' vocabulary is insufficient for effective paraphrasing. Since the majority of comparative studies have been conducted in a single institution using a single language (mainly English) the issue of language difficulties has often become linked in the literature with plagiarism, although the effect of this on source-code plagiarism has not been explored.

Similarly, the effects on source-code plagiarism of other factors known to affect text plagiarism are as yet unclear. For example, in many Asian countries (including China), Cyprus and Greece, learning and assessment are textbook oriented and students are encouraged to memorise works of others. As a result, when these students enter UK Higher Education, they experience difficulties in being critical and stating their own opinions [Hayes and Introna 2005].

Ireland and English [Ireland and English 2011] raise the issue of individual learner differences based on previous learning styles adopted from their country of origin. Mattisson [Mattisson 2010], Montgomery [Montgomery 2010], Hayes and Introna [Hayes and Introna 2005] and Pennycook [Pennycook 1996] share the same opinion that the writing style of international students influences their understanding of plagiarism and argue that it is very important for academics to familiarise themselves with the learning styles of their students prior to accusing them of plagiarism.

Hayes and Introna [Hayes and Introna 2005] investigated the understanding of overseas postgraduate students studying on a Masters course at a UK HEI. The study gathered information using focus groups, and questionnaires from five groups of students, Asian, Chinese, Greek, British and rest of the world. They found that memorisation is a form of learning enforced in China and Greece. In addition to the educational background of overseas students, "linguistic competence", and "the ability to master disciplinary academic language" could be reasons for overseas students to plagiarise (ibid.). Chou [Chou 2010] also observed that plagiarism by students coming from a

Chinese background may not be intentional mainly due to the fact that, throughout their education, students are required to present, cite, and memorise articles, text-book passages, proverbs, and other text and are rewarded for their memorising skills. Chou argues that originality is viewed differently by Chinese and Western writers — Chinese writers believe that when writing about the work of others they must present the content exactly as it was written by the original author. Such practices are not acceptable in UK educational institutions. Importantly, because students memorise large amounts of text, they may not remember the source and may not provide citations when they write about a topic. Chou further argues that cultural values impact on the concept of plagiarism and it may be “unreasonable” and “unfair” to view it as intentional by English Learning Language Students (*ibid.*). This view contrasts with the findings of Barrett and Malcolm [Barrett and Malcolm 2006] mentioned above.

Although cultural influences have been cited by many authors, such as Sowden [Sowden 2005], Leask [Leask 2006] and Adiningrum and Kutieleh [Kutieleh and Adiningrum 2011] as being contributory factors for student plagiarism, this is not yet clearly established. Liu [Liu 2005], in a critique of the positions taken by authors such as Sowden, notes that in the Chinese tradition, although there is a greater preponderance of memorising and quoting, attribution is still required and plagiarism is not regarded as acceptable. Liu notes that the purpose of memorisation is not to copy, it is to help the learner remember good writing style. Stone [Stone 2008] further identifies the relatively late development of intellectual property rights in China as another factor linked to plagiarism there.

A particular problem when considering plagiarism of source-code is that the distinction between acceptable code reuse and unacceptable plagiarism is often indistinct [Simon and Sheard 2015]. Due to the object-oriented nature of many programming languages, reuse of code may be encouraged, and within a commercial context such reuse would often be considered good practice. Furthermore, IDEs typically provide templates for programmers to use, and educators may give students partially complete programs for students to “fill in the gaps”. These aspects of the software development process undoubtedly complicate the plagiarism agenda, but the view of an academic institution is likely to be very clear - the student must graduate with an understanding of plagiarism which is consistent with academic norms and basic (copyright) legislation. If, at a later time, the student legitimately shares code as part of a team development process, then it becomes the responsibility of the employer to train the student as to acceptable source-code re-use consistent with the software development process being undertaken.

1.3. Plagiarism in Greece and South Cyprus

At the moment there has been very little research published on plagiarism in Cypriot and Greek HEIs. Kokkinaki et al. [Kokkinaki et al. 2015] conducted a survey to determine the awareness and perceptions of university students studying in South Cyprus. Their findings revealed that there is lack of clear understanding of plagiarism attributed to poor education on the topic of plagiarism, deficiencies in university plagiarism policies, procedures, and penalties. They emphasise that there is a need for appropriate mechanisms for informing faculty and students about these definitions in order to ensure that the concept of plagiarism is understood. Foltýnek and Čech [Foltýnek and Čech 2012] observed that anti-plagiarism policies are (partly) ignored by Cypriot educators, and that most Cypriot students are unaware of penalties for plagiarism. Sidera-Sideri [Σιδερά-Σιδέρη 2003] gathered the perceptions of Greek students studying in a UK university on what constitutes plagiarism, and found that although these students have a good understanding of plagiarism concepts, they ex-

perience great difficulty avoiding it due to lack of good writing skills. Within Greece itself, Timm [Timm et al. 2008] notes that cheating in exams has been claimed by academic staff as “both endemic and casual”, and remarks that there is an almost complete absence of policies on plagiarism, and a lack of interest on the part of academic staff — “we found no policies or penalties in place in any of the institutions we visited”. Without doubt a plagiarism policy may confuse or even seem unfair to Asian, Chinese and Greek students, or to those students who come from a similar educational background where memorisation is enforced, and understanding should be shown to these students. A particular concern raised by Kambouri is fact that Greece and South Cyprus have many similarities in their educational systems, curricula, assessment, and policies [Kambouri 2012].

The literature survey revealed that limited work exists on plagiarism in Cyprus HEIs. This paper extends our previous studies which discuss the perceptions of home students studying in UK [Joy et al. 2011], [Joy et al. 2013] and China [Zhang et al. 2013] by including results gathered from Cyprus home students. In addition, this paper builds upon the existing literature on perceptions of Cypriot students on plagiarism, in comparison to students from other backgrounds, as discussed in Section 2.

2. CONTRIBUTION AND SIGNIFICANCE

Our study is the first study that gathers the perceptions of Southern Cypriot students on source-code plagiarism, and also the first study to compare perceptions of Southern Cypriot students to students of different cultural backgrounds on the topic of source-code plagiarism. Furthermore, our research contributes to the very limited literature that currently exists on the issue of plagiarism in Cyprus. The aim of the study is to identify those topics which are well understood, and those topics which are not properly understood across the different groups of students, providing useful information to educators from UK, China, and South Cyprus who wish to better inform their students on the issues of plagiarism and source-code plagiarism. In addition, such information can be useful to all educators teaching students of various educational backgrounds, since many of the topics were commonly misunderstood regardless of educational background. By applying the same instrument that has already been used to explore student understanding of source-code plagiarism in the UK and in China, equivalent data captured in South Cyprus will help address this issue [Joy et al. 2011], [Zhang et al. 2013], [Joy et al. 2013]. In conclusion, this paper combines the results of a survey undertaken in three different countries to investigate the differences in understanding of issues of plagiarism based on educational background.

3. METHODOLOGY

Previous studies by the authors have identified the perceptions of students studying in the UK [Joy et al. 2011], [Joy et al. 2013] and China [Zhang et al. 2013]. These two studies used the same questionnaire, written in different languages, and published as an on-line survey. The questions were informed by an extensive literature review, as discussed by Joy and colleagues [Cosma and Joy 2008], [Joy et al. 2009], [Joy et al. 2011], and are grouped into the following five categories.

- (1) Self-plagiarism and source-code reuse;
- (2) Copying text from books and online sources;
- (3) Stealing or paying other people to produce work; falsification as opposed to plagiarism;
- (4) Collusion due to inappropriate collaboration; and
- (5) Converting code to another programming language.

The aim of this study was to determine whether educational background influences students' understanding and approach to plagiarism. Therefore, data were collected from students studying in their home country. To ensure that all responses were from home students one of the demographic questions asked students to select whether they were home or overseas student from a list of options. Responses were filtered such that only UK home-student responses, Chinese home-student, and Cypriot home-student responses were considered for the statistical analysis. The survey consisted of six demographic questions and 15 scenarios each addressing one of the five source-code plagiarism topics mentioned above. The overall aim of the research was to compare the responses of UK, China, and Cypriot students in order to determine whether there exist any significant differences on their perceptions on the five topics.

A total of 615 responses were gathered from UK home students at 18 universities throughout the UK, and a total of 159 responses were gathered from students at 31 universities throughout China. The same survey was translated to the Greek language by native Greek language speakers, and completed by 210 students studying at 6 different HEIs in Cyprus, offering university level Computing degrees. The original (UK) survey was conducted in May/June 2008, the China survey in December 2012, and the Cyprus Survey was conducted from November 2013 until February 2014.

The UK survey was distributed to all computing departments in all universities, through a mailing list of contacts supplied by the Higher Education Academy. The China survey was sent to known contacts at a representative selection of 30 universities (it was expected that requests sent to universities where there was no existing relationship would be ignored). The Cyprus survey was distributed to academics teaching computing modules at all Cyprus public and private Universities, and was also distributed to academics in two other well-established HEIs. Academics and university administration contacts were emailed and encouraged to forward a 'prepared email' containing instructions and the link to the on-line survey to all their undergraduate computing students.

4. ANALYSIS OF RESPONSES

A statistical analysis was performed with SPSS, using a range of tests including frequency statistics, cross-tabulation tests and t-tests for comparing the mean scores, in order to appropriately determine statistical significance of the results. The alpha value of $\alpha = 0.01$ was used and results with statistical significance $p > 0.01$ were rejected. Of the UK respondents, 22.1% did not state their institution. Of the remaining 77.9%, the majority (60.8%) came from established ("Pre 1992") universities, 15% from the Open University (the UK's main distance learning institution), 12.3% from Oxbridge colleges, and 11.9% from the new ("Post 1992") universities. English universities accounted for the majority of respondents (58.2% of those who stated their institution), with Scottish and Northern Irish universities accounting for the remainder (20.5% and 21.3% respectively).

The Chinese respondents came from 30 institutions throughout the country in 14 provinces, including regional teaching universities and nationally renowned researched institutions. The provinces most highly represented were Henan (54.1%), Shanghai (12.6%) and Hebei (11.3%). 13 institutions are included in the "Project 211" list of 112 nationally important key universities, 6 in the "Project 985" list of 39 top-ranking universities and 3 in the recent "Project 2011" (Higher Education Innovative Capacity Improvement Scheme) list.

South Cyprus has a total of two public universities, one based in Nicosia and one in Limassol; and four private universities three of which are based in Nicosia (with campuses in other cities) and one in the city of Paphos. The Cyprus respondents who participated in the survey came from four universities (one public and three private

universities) and two well-established private colleges offering undergraduate university degrees. All participating HEIs are “Post 1992” and offer accredited undergraduate and postgraduate programmes. The degrees offered by the private HEIs are accredited by the Cyprus Evaluation Committee of Private Universities. The majority of the respondents (72.6%) studied at institutions located in Nicosia, whereas 11.4% of students studied in Larnaca, 2.7% in Limassol, followed by those students (8.7%) who studied in a university with campuses in Limassol and Nicosia, and finally 4.6% of students did not specify their institution.

4.1. Overall scores and differences of overall scores based on educational background

For each of the 15 scenarios, the students were asked to select one of the possible responses:

- a. Yes, definitely;
- b. I think it is, but I am not completely sure;
- c. I don’t know;
- d. I think it isn’t, but I am not completely sure;
- e. No, definitely not.

For each scenario, there was one correct response, which was agreed by at least four academics who were experienced in plagiarism detection and were familiar with university disciplinary processes. This ensured that each scenario was either very clearly plagiarism or was clearly not plagiarism, and that there were no ambiguous scenarios.

During analysis each student’s response was allocated a mark in the range $\{1, 0.5, 0, -0.5, -1\}$, where a 1.0 mark was given to a correct response, responses b and d were given a 0.5 or -0.5 mark based on the correct answer, and a -1.0 mark was given to an incorrect response. Response c was worth 0 in all cases. The coding allowed the calculation of an “average mark” for each respondent on a scale from -1.0 to +1.0, where a score of +1.0 indicates that they answered each question correctly. The frequency distributions of the average score variable is shown in Fig. 1.

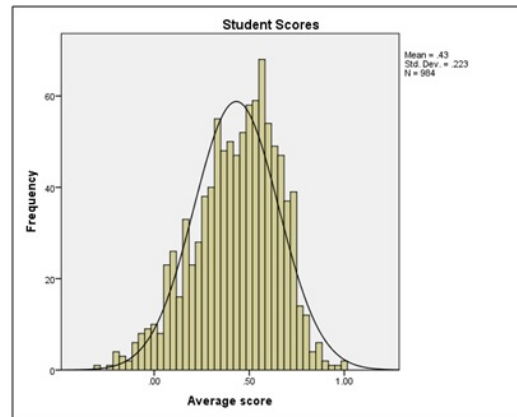


Fig. 1. Frequency distributions of average scores variables

Table I shows some descriptive statistics about the survey scores obtained by the different groups. Statistical t-tests revealed significant differences between the mean

scores of the UK and China groups ($t=19.669$, $df=230.355$, $p=0.00$), China and Cyprus ($t=-5.647$, $df=358.631$, $p=0.00$), and UK and Cyprus ($t=12.624$, $df=306.004$, $p=0.00$).

Table I. Descriptive statistics of scores based on group

	N	Mean	SD	Min	Max
UK Students	615	0.5274	0.16979	-0.17	0.97
Chinese Students	159	0.2065	0.18637	-0.20	0.73
Cypriot Students	210	0.3238	0.21156	-0.30	1.00
Total	984	0.4321	0.22253	-0.30	1.00

4.2. The impact of informing students about plagiarism on their understanding of plagiarism

In order to remind students about the meaning of “plagiarism”, a definition of the term (taken from Wikipedia) was included in the survey. A total of 80.2% of students were informed about plagiarism during their current course (Table II), and 92.3% claim that they understand what plagiarism is (Table III). The results of a cross tabulation test (Table IV), revealed that, of the students who reported they were not informed about plagiarism, 21.5% do not understand what plagiarism is, and 78.5% claim to understand what plagiarism is. Of those students who were informed about plagiarism, 4.3% do not understand it and 95.7% understand it. A t-test determined there were no statistically significant differences in the mean scores of the two groups – informed and uninformed.

Table II. Informed vs. uninformed: Have you, during your current degree course, been informed about what plagiarism is?

	No	Yes	Total
UK Students	22 (3.6%)	593 (96.4%)	615
Chinese Students	89 (56.0%)	70 (44.0%)	159
Cypriot Students	84 (40.0%)	126 (60.0%)	210
Total	195 (19.8%)	789 (80.2%)	984

Table III. Understanding of plagiarism: Do you think you understand what plagiarism is?

	No	Yes	Total
UK Students	12 (2.0%)	603 (98.0%)	615
Chinese Students	45 (28.3%)	114 (71.7%)	159
Cypriot Students	19 (9.0%)	191 (91.0%)	210
Total	76 (7.7%)	908 (92.3%)	984

Interestingly, when looking at each group separately, a large number of China (56.0%) and Cyprus (40.0%) students claimed not to have been informed about plagiarism. The mean score of the group which were informed about plagiarism was $\bar{x} = 0.47$ ($N=789$, $std=0.206$), and the mean score of those students who reported not to have been informed was much lower $\bar{x} = 0.27$ ($N=195$, $std=.209$). The difference between the scores of the two groups is significant ($t=12.549$, $df=294.440$, $p=0.00$). Of those students who have been informed about plagiarism, 98.3% of UK, 77.1% of Chinese students, and 93.7% of Cyprus students claim to understand plagiarism. Of those students who said they had not been informed 90.9% of UK, 67.4% of China, and 86.9% of Cyprus students claim to understand plagiarism. Only a total of 1.7% of UK, 22.9% of China, 6.3% of Cypriot students who have been informed about plagiarism claim not to understand it. A very high proportion of students claim to understand plagiarism, regardless of whether they have been informed. However, the mean scores of those who reported

Table IV. Informed and understand

			Understand	
			No	Yes
UK Students	Informed (%)	No	9.1% (N=2)	90.9% (N=20)
		Yes	1.7% (N=10)	98.3% (N=583)
Chinese Students	Informed (%)	No	32.6% (N=29)	67.4% (N=60)
		Yes	22.9% (N=16)	77.1% (N=54)
Cypriot Students	Informed (%)	No	13.1% (N=11)	86.9% (N=73)
		Yes	6.3% (N=8)	93.7% (N=118)
Total	Informed (%)	No	21.5% (N=42)	78.5% (N=153)
		Yes	4.3% (N=34)	95.7% (N=755)

not to have been informed ($\bar{x} = 0.27$) and of those who have been informed ($\bar{x} = 0.47$), show a general lack of understanding of the issues of plagiarism which does not match students' perception of their understanding.

These results emphasize the importance of educating students about plagiarism and that this significantly impacts of their understanding on plagiarism, regardless of cultural/educational background. The following section discusses findings related to the different topics of plagiarism as stated in Section 3.

4.3. The issue of self-plagiarism

Self-plagiarism occurs when a student reuses parts of an assignment previously submitted for academic credit into a different graded assignment without providing adequate acknowledgement. In programming assignments self-plagiarism is not such a straight forward issue due to the Object-Oriented nature of many programming languages [Cosma and Joy 2008]. A study by Halupa and Bolliger [Halupa and Bolliger 2013] revealed that many institutions did not have policies on self-plagiarism in place, and that academics did not understand what constitutes self-plagiarism. Moreover, it was found that academics believed that their students did not understand self-plagiarism and that students should be educated on the concept. The questionnaire consisted of two scenarios on source-code self-plagiarism (see Table V) describing cases where a student borrowed source-code, authored as part of a previous programming assignment, and incorporated that code into a new assignment which was submitted for grading. The assumption is that students do not perceive copying their own work (which they submitted as part of an assignment and for which they received academic credit) as plagiarism. A summary of the responses is shown in Table VI.

Table V. Topic 1 scenarios

No.	Scenario summary	Plagiarism?
1a	Student copied source-code from his own previous work and has acknowledged the fact.	No
1b	Student copied source-code from his own previous work and has not acknowledged the fact.	Yes

Table VI. Topic 1 responses (C=% correct, I=%incorrect, U= %uncertain)

	UK			China			Cyprus		
	%C	%I	%U	%C	%I	%U	%C	%I	%U
1a	89.1	1.0	9.9	66.7	2.5	30.8	73.3	2.9	23.8
1b	3.4	81.6	15.0	3.8	30.8	60.4	3.8	60.0	36.2

The mean scores for scenario 1a were (UK $\bar{x} = 0.915$, China $\bar{x} = 0.701$, Cyprus $\bar{x} = 0.740$), and scenario 1b (UK $\bar{x} = -0.817$, China $\bar{x} = -0.642$, Cyprus $\bar{x} = -0.652$).

Overall, students received a higher score for scenario 1a ($\bar{x} = 0.843$) than for scenario 1b ($\bar{x} = 0.5068$).

The difference between the means of UK and Chinese students was statistically significant for both scenarios 1a ($t=4.977$, $df=184.155$, $p=0.00$) and 1b ($t=-3.624$, $df=217.350$, $p=0.00$) and the mean difference of the UK and Chinese student scores for scenario 1a was $d=0.2142$ and for 1b was $d=-0.1756$. Similarly, statistically significant differences were determined between the scores of UK and Cypriot students for scenarios 1a ($t=4.786$, $df=258.723$, $p=0.00$) and 1b ($t=-3.915$, $df=319.802$, $p=0.00$). The difference of the UK and Cyprus in mean score for scenario 1a was $d=0.1750$ and for 1b was $d=-0.1647$. No statistically significant differences were determined between the scores of China and Cyprus students for scenario 1a ($t=-0.727$, $df=333.908$, $p=0.468$), and 1b ($t=0.186$, $df=334.555$, $p=0.852$). The difference of the China and Cyprus mean score for scenario 1a was $d=-0.0392$ and for 1b was $d=0.0109$. Although most students from all groups answered scenario 1a correctly, and understand that copying your own work and appropriately acknowledging it, does not constitute plagiarism, when the scenario was changed to copying your own work without providing acknowledgement (1b), the majority of students from all groups answered incorrectly and many were uncertain. These figures are shown in Table VI, and strongly suggest that most students are confused about the concept of self-plagiarism regardless of ethnic or educational background.

This is an interesting result, since it suggests that the concept of self-plagiarism is unfamiliar to students, and that there is a perception that code written by a student is theirs to use as they please. On the one hand, the copyright on a student's code usually belongs to the author, and this is not disputed. On the other hand, an institution will not allow a student to "double count" their work and have the same code marked twice in different contexts (although such an institution might set similar assignments which could cause inadvertent self-plagiarism to occur). We hypothesise that guidance given by institutions on plagiarism tends not to be clear on this specific issue.

4.4. Source-code reuse (copying, using, gaining inspiration) and acknowledgement

Table VII shows a summary of the scenarios, and Table VIII shows the results of each group for each scenario.

Table VII. Topic 2 scenarios

No.	Scenario summary	Plagiarism?
2a	Copied code and did not provide acknowledgement (references)	Yes
2b	Used code and provided incorrect references.	Yes
2c	Used code and provided acknowledgement.	No
2d	Gained inspiration from books and did not provide acknowledgement	No
2e	Used code from books and provided acknowledgement.	No

Table VIII. Topic 2 responses (C=% correct, I=%incorrect, U= %uncertain)

	UK			China			Cyprus		
	%C	%I	%U	%C	%I	%U	%C	%I	%U
2a	82.1	2.4	15.4	25.8	17.6	56.6	61.9	7.6	30.5
2b	11.2	36.9	51.9	10.7	15.1	74.2	29.5	10.0	60.5
2c	76.3	5.2	18.5	57.9	0.6	41.5	51.9	5.2	42.9
2d	47.8	14.3	37.9	34.6	10.7	54.7	56.7	6.7	36.7
2e	93.8	0.8	5.4	64.2	1.3	34.6	79.5	1.0	19.5

82.1% of UK students answered scenario 2a correctly, compared to only 61.9% of Cyprus and 25.8% of Chinese students, clearly demonstrating a misconception about plagiarism by the scores by the Cypriot and Chinese students.

Scenario 2b described a case where a student used code and provided incorrect references. This is plagiarism, because the students copied the code from somewhere without acknowledging the source, and it is also falsification, because the student was not honest about the sources of information. The largest proportion of students across all three groups appeared confused about whether this scenario described plagiarism, and surprisingly only 29.52% of Cyprus, 10.7% of China, and 11.2% of UK responses were correct. These results show the confusion amongst students about providing correct and accurate acknowledgements when using the work of others.

Scenario 2c describes a case where a student used code from a source and provided appropriate acknowledgement. Clearly, this case does not describe plagiarism because the student used and acknowledged the source. The responses from Cypriot and Chinese student were rather low - only 51.9% of Cyprus, 57.9% of China responses were correct. 76.3% of UK responses were correct.

Scenario 2d describes a case where a student read several books for gaining inspiration on how to produce his programming solution, and did not provide any acknowledgements. Since the student did not copy, or use, someone else's programming solution then this case does not describe plagiarism. Only 56.67% Cyprus, 34.67% China, and 47.8% UK answered correctly. Clearly, there is a lot of misunderstanding about when it is appropriate to cite.

Scenario 2e describes a case where a student has used code from a book and provided acknowledgement of the source. This scenario does not describe plagiarism, and as expected most of the students from all groups answered correctly. However, the number of Cypriot students and Chinese students were relatively low in comparison to those of UK. A total of 79.52% of Cypriot students, 64.2% of Chinese students, and 93.8% of UK students answered correctly. Interestingly, nearly all Cypriot and Chinese students who did not answer correctly responded "uncertain" (34.6% and 19.52% respectively). Furthermore, each group appeared to be more confident that acknowledged code from a book was acceptable compared to acknowledged code from a less specific source.

Comparing the mean scores of UK and Chinese students, the results reveal significant differences for scenarios 2a ($t=-2.595$, $df=178.687$, $p=0.10$), 2c ($t=-4.464$, $df=220.738$, $p=0.00$) and 2e ($t=-4.814$, $df=174.943$, $p=0.00$), but none for scenarios 2b ($t=-1.408$, $df=250.323$, $p=0.160$) and 2d ($t=-1.815$, $df=234.552$, $p=0.71$). Comparing the mean scores of UK and Cypriot students, the results do not reveal significant differences for any of the scenarios. Comparing the mean scores of Chinese and Cypriot students, the results reveal significant differences for scenarios 2c ($t=4.675$, $df=366.144$, $p=0.00$) and 2e ($t=4.226$, $df=304.534$, $p=0.00$) only.

Of these five questions, where significant differences exist between two specific groups, the reason is not that one group responded incorrectly, rather that a substantial number of students in one of the groups were uncertain. This suggests that the general issue of source-code re-use is not adequately addressed in the guidance – formal or otherwise – given to those students.

4.5. Stealing and copying code from another student without consent

Paying someone to produce code, or obtaining code illegitimately (e.g. stealing it from a fellow student; taking it without consent from another student) and submitting it for gaining academic credit, without acknowledging the original author is a twofold academic offence — cheating and plagiarism [Cosma and Joy 2008]. Paying someone else to do the work or stealing another's solution is cheating, and submitting that work for academic credit without acknowledging the original author constitutes plagiarism.

Providing fake references (fabrication), providing false references which exist but do not match the source-code that was copied (falsification), and modifying the program output in any way (for example to make it seem as if it works, or as if it works in a particular way) are all academic offences which may co-occur with plagiarism [Cosma and Joy 2008].

Table IX. Topic 3 scenarios

No.	Scenario summary	Plagiarism?
3a	Copying code from another student (without consent) and modifying it by making minor adaptations before submitting it	Yes
3b	A student pays another student to produce a graded assignment	Yes
3c	Student finds another students uncollected printout which contains the solution to the assignment he is working on, and copies this solution and submits it.	Yes
3d	Student modifies own programs output to make it look as if the program works as required by the assignment specification	No

Table X. Topic 3 responses (C=% correct, I=%incorrect, U= %uncertain)

	UK			China			Cyprus		
	%C	%I	%U	%C	%I	%U	%C	%I	%U
3a	70.6	4.4	25.0	57.2	3.1	39.6	77.1	1.9	21.0
3b	80.8	7.8	11.4	16.4	28.9	57.4	16.2	40.5	43.3
3c	80.8	8.0	11.2	43.4	4.4	52.2	50.0	6.2	43.8
3d	80.8	16.8	2.4	45.9	5.7	48.4	51.9	7.5	40.6

Table IX shows the four scenarios which were devised to determine the students' understanding on issues related to plagiarism when it co-occurs with cheating. Table X shows the results of each group for each scenario. Scenarios 3a, 3b, and 3c all describe acts of plagiarism, and 3b also describes a case of cheating, since the student paid someone to produce the code. Both scenarios 3a and 3c involve cases where a student takes another students work without the other students consent — hence these scenarios exclude inappropriate collaboration and only constitute plagiarism. Scenario 3b is not inappropriate collaboration because the student did not collaborate with someone else to produce the solution.

A total of 77.1% Cyprus, 70.6% UK, and only 57.2% Chinese students answered scenario 3a correctly. Comparing the responses of students it appears that more Chinese and Cypriot students perceive copying from another student as plagiarism, but not when copying from books. For the China group, 60.4% answered correctly, 3.1% answered incorrectly, and the proportion of uncertain students was high (36.5%).

Clearly, Cypriot and Chinese students are uncertain about what actions constitute plagiarism, and that the form of cheating described in scenario 3b is also plagiarism because it is essentially submitting someone else's work. Such form of plagiarism was correctly recognized by most UK students (80.8%). Scenario 3c describes a case where a student finds another student's uncollected printout which contains the solution to the assignment he is working on, and copies this solution and submits it. Similarly to the responses gathered for scenario 3b, a large proportion of Chinese and Cypriot students were uncertain whether the copying of a discarded printout (scenario 3c) constitutes plagiarism, whereas most UK students answered correctly.

The falsification described in 3d was correctly identified as not being plagiarism by 80.8% of UK students, whereas only 51.9% of Cyprus and 45.9% of Chinese students were certain. Falsification is a form of cheating, and it may be that the non-UK students were unsure whether they had understood the exact definition of "plagiarism" correctly.

4.6. Collusion due to inappropriate collaboration

Collaborating with other students and submitting similar solutions when the assignment requires students to work on their own constitutes plagiarism (or collusion) [Cosma and Joy 2008]. Topic 4 contained three scenarios describing cases of inappropriate collaboration which resulted in plagiarism. These scenarios are shown in Table XI. Table XII shows the results of each group for each scenario.

Table XI. Topic 4 scenarios

No.	Scenario summary	Plagiarism?
4a	Students collaborate on an assignment which requires that they work alone, and submit similar solutions.	Yes
4b	Students from different groups collaborate on an assignment which requires that they work in their own group, exchange parts of their code and submit their solutions.	Yes

Table XII. Topic 4 responses (C=% correct, I=%incorrect, U= %uncertain)

	UK			China			Cyprus		
	%C	%I	%U	%C	%I	%U	%C	%I	%U
4a	34.1	21.1	44.7	8.2	32.7	59.1	13.8	32.9	53.3
4b	32.4	20.3	47.3	13.8	17.0	69.2	12.4	23.8	63.8

The results show a common failure to understand that inappropriate collaboration results in plagiarism, and only 34.1% of UK, 13.8% of Cypriot, and 8.2% of Chinese students answered correctly. Scenarios 4a and 4b are similar, students were required to work alone and collaborated with other students to produce a solution. The results revealed similar findings to topic 2, students, regardless of background, are uncertain and confused about actions which constitute plagiarism. It is pedagogically valuable for students to share ideas while discussing assignments, and most academics encourage students to do so as long as they do not share their work and do not copy each others solutions [Cosma and Joy 2008]. Student collaboration which results in them submitting similar solutions, on assignments which require them to work on their own, is considered as plagiarism.

4.7. Cross-language source-code plagiarism

Cross-language source-code plagiarism occurs when students take code from one programming language and convert it to a similar programming language [Flores et al. 2011]. Whether translation of source-code across different languages constitutes plagiarism depends on the similarity between the languages and the effort required by the student to perform the conversion. Cross-language source-code plagiarism should not be confused with taking ideas and inspiration from programs, which does not constitute plagiarism [Cosma and Joy 2008]. Nor should implementing an algorithm using multiple languages in a team setting be misunderstood as plagiarism (this may happen as part of a collaborative learning process). Two scenarios (Table XIII) were devised to determine the students' understanding of the issue of conversion of code between languages.

Table XIII. Topic 5 scenarios

No.	Scenario summary	Plagiarism?
5a	Converting code to another programming language without providing acknowledgement.	Yes
5b	Converting code to another programming language and providing acknowledgement.	No

Table XIV. Topic 5 responses (C=% correct, I=%incorrect, U= %uncertain)

	UK			China			Cyprus		
	%C	%I	%U	%C	%I	%U	%C	%I	%U
5a	24.2	25.5	50.2	15.1	22.0	62.9	35.2	16.7	48.1
5b	89.4	2.8	7.8	60.4	3.1	36.5	66.2	3.8	30.0

Scenario 5a describes a plagiarism case where a student takes Java source-code from a book, converts it to C++, incorporates it into their assignment without referencing the origin of the Java code. The responses to the scenario demonstrate that students are confused about actions which constitute plagiarism, including converting code to another language without making appropriate references of this fact. As with natural-language text, converting text from one language to another constitutes plagiarism, and this also applies to source-code (including source-code comments, and other documentation). Scenario 5b describes a scenario where a student takes Visual Basic source-code, converts it into Java and incorporates it into their assignment but provides a note of this fact. Since acknowledgement was provided then this scenario does not constitute plagiarism. The majority of students from all groups answered correctly, although a large number of students from China (60.4%) and Cyprus (66.2%) were uncertain. As apparent in other scenarios, students from China and Cyprus show more confusion as to actions that constitute plagiarism than UK students.

5. SURVEY LIMITATIONS

Any study on plagiarism is limited by the honesty of the respondents, and by the willingness of institutions to engage with such research. The numbers of respondents providing data for these surveys has by necessity been restricted and the respondents self-selected, and for other practical reasons the three surveys were spaced across six years rather than concurrent. Notwithstanding these limitations, the results are sufficient to highlight similarities and differences of student perceptions in the three countries. In addition to providing useful information to academics engaging with students from these different places, the results will form a useful starting point for future studies, especially with the growing understanding by institutions globally of the necessity to combat student plagiarism.

6. DISCUSSION

Problems in understanding plagiarism can be experienced by any student regardless of their cultural/educational background. This paper presents evidence that Cypriot and Chinese students may struggle more than UK students to understand plagiarism, and identifies factors which contribute to source-code plagiarism which are common to all three groups.

The results reveal that students who were properly informed better understood what actions constitute plagiarism, although most students claimed to understand plagiarism well. Some topics, however, were still unclear among students regardless of their educational background and whether they were informed about plagiarism.

Self-plagiarism is a topic which is widely misunderstood, and results show that a very low percentage of students across all three countries actually understood the concept of self-plagiarism. Similarly, students from all three countries show a lack of understanding of collusion, although UK students understand this issue better than the others (the study does not provide the reason for this, but we surmise that a combination of cultural factors referred to in the Introduction, together with a robust embedding of group work in UK curricula, may offer an explanation).

Other common areas of misunderstanding among all three groups of students are: using and appropriately referencing code; gaining inspiration from books and provid-

ing acknowledgement; appropriate and inappropriate collaboration; and converting code to another programming language and providing acknowledgement.

Overall, the statistical analysis revealed that UK students had a better knowledge of source-code plagiarism than the other two groups, with significant differences between the mean survey scores of each group. The survey results further revealed that a lower proportion of Chinese and Cypriot students (44% and 60% respectively) are informed about plagiarism by their institutions in their current degree studies, compared to the 96.4% of UK students.

The results presented in this paper provide evidence that there is a problem with understanding source-code plagiarism in Cypriot institutions, and analysis of the data collected from South Cyprus suggests how the results may be explained by the Cypriot educational context. This is the first study which gathers Cypriot student perceptions on source-code plagiarism, and contributes to the small volume of literature that currently exists on the issue of plagiarism in South Cyprus.

The literature, as well as our findings, provides evidence that students from China experience difficulties in understanding and even accepting that plagiarism is a problem — our results show that many students who were informed about plagiarism did not understand it.

Areas of misunderstanding only among Chinese home students are copying code and providing acknowledgement (references), and copying code from another student (without consent) and modifying it by making minor adaptations before submitting it.

A specific area of misunderstanding shared by both Chinese and Cypriot home students is reuse of code and providing acknowledgement. Due to the nature of programming, any reused code (including code produced as part of another assignment, and open-source code) should be appropriately acknowledged within the source-code documentation (and clearly stated above the code in the form of a comment), and this is a clear message that has not been fully understood.

The distinction between plagiarism and other forms of cheating is sometimes subtle, and the failure of the Chinese students (and, to a lesser extent, the Cypriot ones) to regard “paying another student to produce a graded assignment” as plagiarism suggests that a future study may seek to explore students’ perceptions of the different forms of cheating.

Further issues common to both the Chinese and Cypriot groups include taking another student’s work without consent and submitting it as if it is their own work, and modifying a program’s output to make it look as if the program works as required by the assignment specification.

7. CONCLUSION

Our findings contribute to the growing body of literature that demonstrates plagiarism to be a serious on-going area of misunderstanding in Higher Education internationally. Identifying differences in understanding between different groups can support targeted interventions which may help address specific areas of confusion for particular groups. It is important for educators to inform students early and with sufficient clarity and examples about plagiarism, embedding and reinforcing plagiarism education consistently throughout the curriculum. The research reported here suggests that international students may need extra support to introduce and explain the policy that applies in the current institution of study.

Equally important is the finding that certain aspects of plagiarism cause confusion across international boundaries, despite the fact that countries such as the UK place considerable emphasis on informing students about plagiarism and in detecting and punishing infringements. Our research has focused on source-code plagiarism, revealing that aspects such as appropriate attribution of code and the boundary between

permissible discussion and collusion remain unclear for many students. If genuine misunderstanding remains even where there is a strong emphasis on plagiarism education, it is necessary to ask whether these issues are really as clear-cut as institutions might like to believe or whether an underlying lack of consensus and clarity is undermining efforts to reduce source-code plagiarism.

8. FUNDING

This work was supported by the Higher Education Reform Research Project of Henan Agricultural University (Xiao Zheng Jiao [2012] No. 3).

REFERENCES

- Jan Bamford and Katerina Sergiou. 2005. International students and plagiarism: an analysis of the reasons for plagiarism among international foundation students. *Investigations in university teaching and learning* 2, 2 (2005), 17–22.
- Ruth Barrett and James Malcolm. 2006. Embedding plagiarism education in the assessment process. *International Journal for Educational Integrity* 2, 1 (2006).
- Krishna Bista. 2011. Academic dishonesty among international students in higher education. *To improve the academy: Resources for faculty, instructional, and organizational development* 30 (2011), 159–172.
- Cara Bradley. 2011. *Plagiarism Education and Prevention: A subject-driven case-based approach*. Chandos, Oxford, UK.
- I-Chia Chou. 2010. Is plagiarism a culture product: The voice of a Chinese-speaking ELL student. *The International Journal–Language Society and Culture* 31 (2010), 37–41.
- Daniela Chuda, Pavol Navrat, Bianka Kovacova, and Pavel Humay. 2012. The issue of (software) plagiarism: A student view. *Education, IEEE Transactions on* 55, 1 (2012), 22–28.
- Georgina Cosma and Mike Joy. 2008. Towards a definition of source-code plagiarism. *Education, IEEE Transactions on* 51, 2 (2008), 195–200.
- Enrique Flores, Alberto Barrón-Cedeño, Paolo Rosso, and Lidia Moreno. 2011. *Natural Language Processing and Information Systems: 16th International Conference on Applications of Natural Language to Information Systems, NLDB 2011, Alicante, Spain, June 28-30, 2011. Proceedings*. Springer, Berlin, Chapter Towards the Detection of Cross-Language Source Code Reuse, 250–253. http://dx.doi.org/10.1007/978-3-642-22327-3_31
- Tomáš Foltýnek and František Čech. 2012. Attitude to plagiarism in different European countries. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* 60, 7 (2012), 71–80.
- Colleen Halupa and Doris U Bolliger. 2013. Faculty Perceptions of Student Self Plagiarism: An Exploratory Multi-university Study. *Journal of Academic Ethics* 11, 4 (2013), 297–310.
- Niall Hayes and Lucas D Introna. 2005. Cultural values, plagiarism, and fairness: When plagiarism gets in the way of learning. *Ethics & Behavior* 15, 3 (2005), 213–231.
- Chris Ireland and John English. 2011. Let them plagiarise: Developing academic writing in a safe environment. *Journal of Academic Writing* 1, 1 (2011), 165–172.
- Mike Joy, Georgina Cosma, Jane Sinclair, and Jane Yin-Kim Yau. 2009. A taxonomy of plagiarism in computer science. *EDULEARN09 Proceedings* (2009), 3372–3379.
- Mike Joy, Georgina Cosma, Jane Yin-Kim Yau, and Jane Sinclair. 2011. Source Code Plagiarism—A Student Perspective. *Education, IEEE Transactions on* 54, 1 (2011), 125–132.
- Mike Joy, Jane Sinclair, Russell Boyatt, Jane Yin-Kim Yau, and Georgina Cosma. 2013. Student perspectives on source-code plagiarism. *International Journal for Educational Integrity* 9, 1 (2013).
- Maria Kambouri. 2012. The Educational System in Cyprus. In *Proceedings of WIEGA*, Vol. 5. 57–67.
- Angelika Kokkinaki, Catherine Demoliou, and Melpo Iacovidou. 2015. Students’ perceptions on plagiarism and relevant policies in Cyprus. *International Journal for Educational Integrity* 11, 1 (2015), 1–11.
- Salah Kutieleh and Tatum Syarifah Adiningrum. 2011. How different are we? Understanding and managing plagiarism between East and West. *Journal of Academic Language and Learning* 5, 2 (2011), A88–A98.
- Betty Leask. 2006. Plagiarism, cultural diversity and metaphor—implications for academic staff development. *Assessment & Evaluation in Higher Education* 31, 2 (2006), 183–199.
- Dilin Liu. 2005. Plagiarism in ESOL students: Is cultural conditioning truly the major culprit? *ELT journal* 59, 3 (2005), 234–241.

- Stephen Marshall and Maryanne Garry. 2005. How well do students really understand plagiarism?. In *Proceedings of the 22nd annual conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE)*. 457–467.
- Jane Mattisson. 2010. Teaching international students in English. A matter of Culture. *Porta Linguarum: revista internacional de didáctica de las lenguas extranjeras* 13 (2010), 165–178.
- Catherine Montgomery. 2010. *Understanding the international student experience*. Palgrave Macmillan, London.
- Sandra Nadelson. 2007. Academic misconduct by university students: Faculty perceptions and responses. *Plagiarism: Cross Disciplinary Studies in Plagiarism, Fabrication, and Falsification* 2, 2 (2007).
- Chris Park. 2003. In other (people's) words: Plagiarism by university students—literature and lessons. *Assessment & Evaluation in Higher Education* 28, 5 (2003), 471–488.
- Alastair Pennycook. 1996. Borrowing others' words: Text, ownership, memory, and plagiarism. *TESOL quarterly* 30, 2 (1996), 201–230.
- Siwei Qi. 2008. *A Descriptive Study of Plagiarism in the Writing Assignments of First-year Chinese Students in the American English Language Institute at Stephen F. Austin State University*. ProQuest, Ann Arbor, MI.
- Ιουλία Σιδερά-Σιδέρη. 2003. Έλληνες φοιτητές, λογοκλοπή και συμπαιγνία= Greek students perception of plagiarism and collusion. (2003).
- Simon and Judy Sheard. 2015. Academic Integrity and Professional Integrity in Computing Education. In *ITiCSE 15: Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*. 237–241.
- Colin Sowden. 2005. Plagiarism and the culture of multilingual students in higher education abroad. *ELT Journal* 59, 3 (2005), 226–233.
- Charles R Stone. 2008. What plagiarism was not: Some preliminary observations on classical Chinese attitudes toward what the West calls intellectual property. *Marquette Law Review* 92 (2008), 199.
- Peter Tennant, Gill Rowell, and Fiona Duggan. 2007. Academic misconduct benchmarking research project: part I The range and spread of penalties available for student plagiarism among UK higher education institutions. *Higher Education Academy, York*, [online] Available at: <http://www.jiscpas.ac.uk/AMBeR/index> (2007).
- Aanja Timm, Niall Hayes, Lucas Introna, and Edgar Whitley. 2008. Report on educational practices in China. Student Diversity and Academic Writing (SDAW) Project. (2008).
- Tish Eshelle Twomey, Holly White, and Ken Sagendorf. 2009. *Pedagogy, not policing: Positive approaches to academic integrity at the university*. Graduate School Press, Syracuse University, Syracuse, NY.
- Nadya Yakovchuk. 2008. Towards understanding the root causes of plagiarism among non-native speaker students. (2008).
- Dongyang Zhang, Mike Joy, Georgina Cosma, Russell Boyatt, Jane Sinclair, and Jane Yau. 2013. Source-code plagiarism in universities: a comparative study of student perspectives in China and the UK. *Assessment & Evaluation in Higher Education* ahead-of-print (2013), 1–16.