

Real-World Outcomes and Critical Thinking: Differential Analysis by Academic Major and Gender¹

Amanda Franco²
Universidade do Minho, Braga,
Portugal

Leandro da Silva Almeida
Universidade do Minho, Braga,
Portugal

Abstract: The Real-World Outcomes is an inventory that measures everyday problematic behaviors that represent decisions where critical thinking is presumably absent; assessing the negative outcomes of poor daily decisions helps to infer the degree of critical thinking that mediates everyday reasoning. In the present paper, we describe the process of translation and cultural adaptation of this inventory to Portuguese. We present evidence of its administration to 259 college students concerning reliability, and differences based on academic major and gender. No statistically significant differences were found, either due to academic major or gender. Results suggest the value of this instrument to assessing daily decision making and life outcomes, and also, to estimate the quality of critical thinking in everyday life.

Keywords: cognitive ability, competence, evaluation, critical thinking

Real-World Outcomes e Pensamento Crítico: Análise Diferencial por Área e Género

Resumo: O Real-World Outcomes é um inventário que avalia comportamentos problemáticos do dia-a-dia que representam decisões onde se presume que o pensamento crítico está ausente; ao avaliar os resultados negativos de decisões quotidianas pobres, permite inferir o grau de pensamento crítico que medeia o raciocínio no quotidiano. Neste artigo, descreve-se o processo de tradução e adaptação cultural deste inventário para a língua Portuguesa. Apresentam-se dados da sua administração a 259 estudantes universitários relativos à consistência interna, e as diferenças de acordo com a área académica e género. Não foram encontradas diferenças estatisticamente significativas, nem decorrentes da área académica, nem do género. Os resultados sugerem a relevância deste instrumento para avaliar a tomada de decisão no dia-a-dia e seu impacto na vida dos sujeitos, e também, para estimar a qualidade do pensamento crítico usado na vida quotidiana.

Palavras-chave: aptidão cognitiva, competência, avaliação, pensamento crítico

Real-World Outcomes y Pensamiento Crítico: Análisis Diferencial por Área y Género

Resumen: El Real-World Outcomes es un inventario que evalúa comportamientos problemáticos del cotidiano que representan decisiones en que se presume la ausencia de pensamiento crítico; al evaluar los resultados negativos de malas decisiones cotidianas, permite inferir el grado de pensamiento crítico que media el razonamiento en el cotidiano. En el presente artículo, describimos el proceso de traducción y adaptación cultural del inventario para la lengua Portuguesa. Se presentan datos de su administración a 259 estudiantes universitarios relativos a consistencia interna, y a diferencias en función de la área científica de estudios y del género. No fueron encontradas diferencias estadísticamente significativas entre los participantes, ni decorrentes de la área, ni del género. Los resultados sugieren la relevancia de este instrumento para evaluar la toma de decisión en el cotidiano y su impacto en la vida de los sujetos, y además, para estimar la cualidad del pensamiento crítico utilizado en la vida diaria.

Palabras clave: aptitud cognitiva, competencia, evaluación, pensamiento crítico

Introduction to Critical Thinking

We live in an era in which making reasonably accurate predictions about the future seems to be particularly hard, whether it is about where one will live, or what line of work one will pursue. Yet, people are expected to make complex and successful decisions while taking into consideration the immensity of available information. People are expected to prepare for the unknown. And

this entails “a new pedagogical challenge if not urgency in the contemporary age” (Barnett, 2012, p. 65). The responsibility of Education – of higher education institutions, in particular – seems to be to equip students with the tools they need to face such uncertainty. To do so, higher education must offer students the possibility of a curriculum that not only is aware of present and future complex decisions and unpredictable challenges, but more importantly, a curriculum that clearly links personal, academic, and real life (Franco, Butler, & Halpern, in press). Even so, higher education institutions may be failing in a core mission: to prepare students to be lifelong learners (Kember, Leung, & Ma, 2007). Here, we enter the realm of *critical thinking* (CT).

¹ Support: This work was supported by Fundação para a Ciência e a Tecnologia under Grant SFRH/BD/76372/2011 (QREN - POPH Programme - Typology 4.1 - Advanced Training).

² Correspondence address:

Amanda Franco. Instituto de Educação - Universidade do Minho. Campus de Gualtar. 4710-057. Braga, Portugal. E-mail: amanda.hr.franco@gmail.com

There is an array of possibilities to describe it (Sternberg, Roediger, & Halpern, 2007), and yet, CT may be defined plainly as having a set of cognitive skills, and being in the disposition to use those skills to tilt the odds (i.e., chances for success) in one's favor (Halpern, 2014). Both CT skills and the disposition to actually develop and use those skills are crucial, which is the reason why lacking the former (one can have the skills and not feel inclined to use them) or the latter (one may feel the disposition to be a critical thinker, yet lack the skills) may be problematic (Stupnisky, Renaud, Daniels, Haynes, & Perry, 2008). From the literature review (Almeida & Franco, 2011; Ennis, 1994; Facione, 2011; Halpern, 2014; Paul, 2005), a critical thinker's thinking is by its nature curious enough to be tempted by questions; logical enough to analyze the validity of the reasons and conclusions of an argument; scientific enough to develop and test hypotheses, and let the derived conclusions guide behavior; organized enough to delineate a course of action, and sufficiently determined to commit to those guidelines; open enough to participate, to debate, and to compromise; empirical enough to search for valid and reliable evidence to support claims; flexible enough to consider (divergent) ideas, perspectives, or alternatives; epistemically modest enough to be proven wrong, to recycle knowledge, to reformulate beliefs (that are very often disguised as facts), and to readjust strategies; deliberate enough to transfer knowledge and skills between topics or situations.

CT benefits academic performance, helping students to organize, monitor, and evaluate their performance (Paul, 2005; Phan, 2010). According to Stupnisky et al. (2008), college freshmen who are more inclined to think critically display stronger perceptions of academic control, with a positive impact on their academic experience. If CT is presented as being important throughout schooling, it seems to play a key-role in higher education, in that college students are expected to develop a deeper approach to learning. This helps them to gain meaningful knowledge and develop useful tools to start a professional career in their near future (Barnes, 2005), and to deal with a reality that is defined by ambiguity and complex, ill-defined problems (Kek & Huijser, 2011).

Moreover, this combination of skills and dispositions creates a rational and reasonable way of thinking that is helpful in life's everyday subjects and settings (Halpern, 1999). CT entails the consideration of diverse topics, the search for divergent sources of information, and the anticipation of possible different results. On the grounds of this information, it is possible to make independent decisions that are well-founded and conscientious, and to act in conformity with the alternative that is most likely to be advantageous (Ku, 2009). These CT skills are an everyday need, valuable to making *good* decisions in many domains of life (Butler, 2012), and to avoiding being at the mercy of others, whether it is media's amalgamation of facts and fiction; politics' half truths; (so called) expert's hazardous

health advice; or researcher's (allegedly) scientific data on the paranormal (Halpern, 2014).

It is such a connection that brings together CT and real life, that adds value to the study of this topic, and that stresses its magnitude and broad applicability. In that it is applied thinking, CT must be studied in real-life situations, in people's real life.

Critical Thinking and Real-World Outcomes

To analyze the association between CT and real life, more specifically, the impact of CT on real life, the Real-World Outcomes (RWO) (Butler & Halpern, 2011) has been used as external criteria in the Halpern Critical Thinking Assessment (HCTA) (Halpern, 2012a) validation studies. Through theoretical integration and assessment of CT and real world, this line of research opens promising possibilities for the comprehension of the real relevance of CT.

The RWO is a self-report inventory that measures a wide range of everyday problematic behaviors that represent daily decisions where CT seems to be lacking. By assessing the negative outcomes of poor daily decisions, the RWO can help to infer the degree of CT that mediates everyday reasoning. This instrument is an adaptation of the Decision Outcomes Inventory (DOI) (Bruine de Bruin, Parker, & Fischhoff, 2007). The aim of its authors was to analyze the links between decision making competence and the decision outcomes people experience in the real world. In light of their findings, higher decision making competence is related to better decision making outcomes. According to them, decision making skills have been researched in isolation, preventing a better understanding of the relationship between the different skills, the association between decision making competence and other cognitive abilities, and more important, between decision making competence and real-world outcomes.

In turn, the HCTA represents a turning point in the scenario of evaluations. The HCTA presents 25 daily life scenarios (concerning health, social welfare, publicity, education, or politics) to assess five dimensions of CT: verbal reasoning; argument analysis; hypothesis testing; likelihood and uncertainty; decision making and problem solving. The HCTA is innovative because it combines open-answer items (assessing free recall) and forced choice items (assessing recognition). It is also innovative, because it captures the two essential components of CT: CT *skills* are displayed when taking the test, which reveals the *disposition* to commit to the test and use CT skills to answer each item. Finally, because the HCTA provides a situational assessment, as it assesses not only respondents' CT, but also the actual use that respondents make of their CT to overcome daily challenges. The HCTA makes the assessment of CT comprehensive, and links cognition to real-life situations and settings.

Research articulating CT and real-life outcomes (the HCTA and the RWO) is both current and relevant. In a study by Butler (2012), the HCTA and the RWO were administered to a

sample of community adults and students in community college and university. Participants with a higher HCTA score have fewer negative outcomes in their daily life, when compared to participants with a lower HCTA score. Moreover, in a study by Dwyer, Hogan, and Stewart (2012), graduate students' performances on the HCTA predicted the RWO results.

These findings highlight the links between CT and real-world outcomes, suggesting that critical thinkers will more easily avoid negative outcomes in daily life decisions. CT skills transfer to other sorts of content and situations outside the content/situation where they were originally developed (Butler, 2012). As a result, we infer the crucial role that Education has in stimulating students' CT skills and dispositions, hence, in improving the quality of decisions in their lives (Saiz & Rivas, 2010). The role of higher education, in particular, is to equip students with the (higher order) thinking skills that are needed for them to learn to understand, and thus, thrive throughout life (Halpern, 2014). Here, CT is crucial (Kember et al., 2007). In a study by Dwyer et al. (2012), students who participated in a course to develop CT had higher HCTA scores when compared to students in the control group. In other studies using the HCTA (Butler, 2012; Butler et al., 2012), the respondents who scored higher on this CT test had more years of education. Moreover, in a study by Cosgrove (2011), it is suggested that students are more inclined to develop and use the CT skills and dispositions if these are taught in a deliberate and explicit way by instructors. According to Pascarella et al. (2014), college students who are exposed to "diversity experiences" – i.e., experiences where new and unfamiliar situations/challenges are presented, which require unique approaches and different ways of thinking – show higher gains in their CT. Such empirical evidence stresses the importance of CT, and the need to develop it. This is done by formally addressing CT in the curriculum, throughout the years of schooling, and across countries. CT does not flourish spontaneously; it requires deliberate and explicit teaching, continued practice, consistent feedback.

Considering the broad relevance of CT, in addition to its clear impact on daily life, hence, the importance of stimulating college students' skills and dispositions to think critically and soundly in everyday (personal, academic, etc.) situations, we are conducting research seeking to translate the HCTA into Portuguese, and then adapt and validate it for Portuguese. However, this is a more sophisticated endeavor to be completed and presented in the future. Yet, it is within this line of research that the present study, concerning the translation and adaptation of the RWO, and the assessment of undergraduates, can be located. A better understanding of the links connecting college students' CT to real life is needed, and this is an area in which little research has been undertaken. Research shows that CT has a positive impact on academic and work settings, but its impact on other domains of everyday life is still lacking (Butler, 2012). In fact, although the focus of the present study is the RWO,

this inventory is a smaller component in a broader line of research. Nonetheless, work involving the RWO is extremely relevant, since CT would lose its meaning if removed from real life, and real-life settings and situations. CT is applied thinking; CT is situational thinking; CT is real-life thinking.

Given the huge differences between European countries and the USA, we first sought to obtain a successful "cultural translation" of this inventory, in order to better understand college students' (disruptive) experiences and (negative) outcomes in daily life. Thus, this can also help to develop considerations regarding the quality of CT which presumably influences their daily decisions and behavior. Evidence concerning the reliability of the Portuguese version of the RWO is presented, in addition to differences in the total RWO score according to students' academic major and gender. The literature referring to the impact of academic major on the development of CT seems to lack consensus (Li, Long, & Simpson, 1999). Some authors report that a college education alone seems to have a positive impact on CT (Pascarella, Bohr, Nora, & Terenzini, 1996), but it also potentiates a variety of many other positive effects on students (Pascarella & Terenzini, 2005). This is the reason it is essential to conduct grounded, comprehensive, systematic, and valid research to better understand "the net or unique impact of the postsecondary experience on students" (Pascarella, 2006, p. 508).

Other authors claim that cognitive skills develop differently according to academic major (Kim & Sax, 2011); others state that what is determinant for the development of CT is the teacher's approach (Brint, Cantwell, & Saxena, 2012), or well-structured classroom instruction that fosters a more active, participative, and reflexive approach to learning from students (Pascarella, Wang, Trolian, & Blaich, 2013). As for the association between CT and gender, CT may not be directly dependent on gender (Butler et al., 2012), perhaps due to its strong situational nature, making it influenced by, and adapted to, contextual settings.

Besides presenting the translation and adaptation process of the RWO, its consistency, and a differential analysis according to academic major and gender, considerations are made regarding the value of this instrument for assessing life outcomes and the quality of decision making, and more important, the degree of CT that seems to mediate undergraduates' daily decisions and everyday life outcomes. In the long run, this study is a stepping-stone to assuring cultural validity of the RWO in Portugal, being also a contribution for cross-cultural research concerning CT using the HCTA and the RWO. Also, we anticipate future research with the HCTA and the RWO to compile more knowledge concerning CT and its impact, namely, in higher education.

Method

Participants

The RWO was administered to 259 undergraduate students, 166 (64.1%) female, ranging in age from 18 to

50 years ($M = 21.54$, $SD = 6.25$). This convenience sample was composed of freshmen ($n = 181$, 70.2%), sophomores ($n = 51$, 19.8%), and juniors ($n = 26$, 10.1%), taking an academic major in the humanities ($n = 137$, 52.9%) or engineering ($n = 122$, 47.1%). Students were attending one of two public universities: one in the north ($n = 208$, 80.3%) and the other in the center ($n = 51$, 19.7%) of Portugal.

There are differences in the distribution of participants according to academic major: women prevail in the humanities (87.6%), a tendency in Portuguese universities (Table 1).

Instrument

A Portuguese version of the *Real-World Outcomes - RWO* was used in this study, composed of 40 dichotomous statements that describe daily life events (e.g., “Gone shopping for food”); one or more items that describe negative outcomes from that situation in particular are provided (e.g., “Threw out food I had bought because I let the expiration date go by”). Respondents indicate if, in the past year, they have/have not experienced each daily life event, and if they have/have not experienced each of its possible negative outcomes. Scores on each item range from 0 to 1, whether the respondent did not or did, respectively, experience that daily life event/negative outcome. A higher RWO total score equals a higher number of negative outcomes. For all items, the complete RWO is published elsewhere (Butler et al., 2012).

Procedure

Data collection. Permission was obtained to use the RWO in our study, followed by its translation to Portuguese by two university teachers and a professional translator, and then by its back-translation into English. The Portuguese version of the RWO was compared to its original version in English, in order to identify and correct possible inconsistencies. This phase required a closer collaboration with the RWO’s authors, given the need to clarify the specificity of some items, to find a more suitable equivalent in Portuguese. The translation and/or adaptation of some sets of items was more intricate to accomplish for one of several different reasons: (a) the specificity of a few words or expressions required Portuguese equivalents that would be understandable enough, to fully capture the real meaning of an item (e.g., “Pulled an ‘all-nighter’”); (b) linguistic and/or cultural Portuguese specificities required that some of the

items suffer minor adaptations, in order to convey authenticity to the situations described and, thus, help students to identify with them (e.g., “Wasn’t able to get inside my home because I locked my keys inside”); (c) cultural differences between Portugal and the USA called for some adaptations to make the items accurate (e.g., “Quit a job without giving at least 2 weeks’ notice” – labor laws differ from country to country), better adapted to the cultural reality of this country (e.g., “I believe in some conspiracy theories” – the term *conspiracy theory* is rather unusual in Portugal), or appropriate (e.g., “Had more than \$5,000 in credit card debt” – this sum of money was considered excessive, in light of the exchange rate, and also, economic and financial dissimilarities between this Iberian country and the USA). Following its cultural translation, both the Portuguese and original versions of the RWO were compared and standardized; the goal was to secure linguistic equivalence, and at the same time, respect linguistic singularities and cultural features. A qualitative analysis of the RWO was undertaken resorting to the *think aloud* method (Ericsson & Simon, 1993), with a group of 14 undergraduate students enrolled in the third year of a psychology major, to ensure that the items were clear, comprehensible, and relevant. The Portuguese experimental version of the RWO was administered in the course of regular classes. We presented our study goals, asked for students’ voluntary participation, and guaranteed their anonymity. Instructions were read aloud, and participants spent about 20 minutes to complete the RWO.

Data analysis. Quantitative analyses were undertaken using the statistical software IBM SPSS for Windows (version 22.0).

Ethical Considerations

This project follows the ethical requirements of the EFPA, European Federation of Psychologists’ Associations, and of the Portuguese Psychologists’ Association. All ethical principles were respected, guaranteeing to the participants the principles of informed consent, voluntary participation in the present study, and confidentiality of their responses.

Results

To examine the internal consistency of the RWO, we first analyzed variation in item difficulty, estimated by the frequency of participants who said they had (*Yes*) experienced each negative outcome described in the RWO in the previous year.

The most frequent negative outcomes experienced by participants in our study (with a frequency of around 50% or higher), were: *Forgotten a birthday of someone close to me and only remembered it the next day or later* (60.3% of undergraduates); *Threw out food I had bought because I let the expiration date go by* (50.4% of undergraduates); *Missed a class because I fell asleep and didn’t hear my alarm or because I forgot to set it* (49.2% of undergraduates); and *Forgotten to do a class assignment* (45.9% of

Table 1
Characterization of the Participants by Academic Major, Gender, and Age

Academic major	Gender		Age	
	Female	Male	$M (SD)$	Min - Max
Humanities	$n = 120$ (87.6%)	$n = 17$ (12.4%)	22.19 (6.68)	18-50
Engineering	$n = 46$ (37.7%)	$n = 76$ (62.3%)	20.82 (5.68)	18-49

undergraduates). On the other hand, (a few of) the most infrequent negative outcomes (with a frequency below 2%) were: *Been kicked out of a rented apartment/house/single room before the lease ran out* (0.4% of undergraduates); *Been detained in jail overnight for any reason* (0.4% of undergraduates); *Had more than 1,000 Euros in credit card debt* (1.2% of undergraduates); and *Received a DUI* (abbreviation for Driving Under the Influence of alcohol and/or drugs) *for drunk driving* (1.2% of undergraduates). Such results must be interpreted according to our participants' ages. The more common negative outcomes are familiar to the average student – such as missing a class, or having to throw out food because it went bad – and may be common situations in the life of young adults who have (just recently, perhaps,) left their parents' house and moved away to attend university, students who go live on their own or with friends their age, and who now begin to take on more adult responsibilities.

As for the less frequent negative outcomes, they concern more severe situations related to problems with the law, substance abuse, or violation of a rental contract. Given their young age, students may not have (yet) had the opportunity to experience such severe negative events. As for the low frequency of the item *Had more than 1,000 Euros in credit card debt* (experienced by 1.2% of our sample), this may be explained by cultural differences; unlike what may happen in the USA, in Portugal, it is common to use money or a debit card, not a credit card, to pay for daily expenses.

When conducting reliability analyses, one set of items did not meet the .20 cutoff, referred to by some authors as the minimum criteria for corrected item-total correlation, to determine the item's internal validity (Almeida & Freire, 2010), which may be used as a rule of thumb in an initial phase of developing psychological assessment instruments for new constructs (Nunnally & Bernstein, 1994). Before raising the possibility of the reduced reliability of the Portuguese version of the RWO, the nature of these items and the age group of our sample should be considered. Up to 50% of the participants are 18 or 19 years old, and 90% of our sample is aged between 18 and 28 years. Items with very low internal consistency involve extremely severe situations (e.g., “Had my driver's license confiscated by the police”), or to situations that most young people have not yet experienced because there is a higher likelihood that they will happen to an adult who has lived longer and has a wider range of life experiences (e.g., “Been kicked out of a rented apartment, house or single room before the lease ran out”). There is a possibility of data being somehow influenced by our sample's age characteristics.

If considered independently, none of these “problematic” items seemed to greatly improve the Cronbach's alpha coefficient of the inventory if deleted. Nevertheless, given their low reliability, but also because they describe severe situations/outcomes that do not appear to be common in the life of younger individuals (e.g., 7b - “Had my driver's

license confiscated by the police”), but especially since the RWO is to be administered to undergraduates and we intended to retain items that are familiar to young students in university, we opted to remove the items that did not meet a .15 criteria, which is not so restraining, from subsequent analyses (given the initial phase of exploratory analysis using the RWO): 7; 8b; 13; 20; 29; 30; 31; 33; 34; 37. Item 5e was retained, due to its academic nature, although it did not meet that criterion. All items related to academic situations and outcomes were maintained, since they are more familiar to college students. After deleting the 10 low-reliability items, we obtained a 31-statement RWO (with 43 items describing negative outcomes), used in further analyses.

To estimate the reliability of this 43-item inventory, we divided the complete instrument into two parts: the first half was composed of the 22 even numbered items, and the second half was formed by the remaining 21 odd numbered items. We divided each part according to even and odd numbered items so that one half was an alternate of the other half (split-half method). The Cronbach's alpha coefficients for each part of the RWO were appropriate ($\alpha = .64$), and there was also correlation between the two halves ($\alpha = .70$). According to Nunnally and Bernstein (1994), in initial stages of the development of new tests and constructs, values ranging between .60 and .70 are acceptable. The corrected Spearman-Brown correlation coefficient was .83, which is considered fit.

We conducted a *t* test to explore differences in the RWO total score according to academic major and gender. No statistically significant differences were found across academic majors, $t(215) = -1.02, p = .31, d = .14$, nor according to undergraduates' gender, $t(215) = 1.14, p = .26, d = .16$. Our findings suggest that such variables may not be a differentiating factor of undergraduate students when it comes to daily negative outcomes, as assessed by the RWO. Yet, though we found no statistical significance for academic major and gender differences ($p = n.s.$), the direction of the differences between participants in each group is noteworthy.

When comparing undergraduates according to academic major, engineering students scored higher on the RWO ($M = 6.32, SD = 3.59$) than undergraduates from the humanities ($M = 5.82, SD = 3.61$). As for differences related to gender, male undergraduates scored higher ($M = 6.42, SD = 3.67$) when compared to their female peers (Table 2). Given the inverse direction of the RWO scoring (higher scores reveal a higher frequency of negative outcomes), our data indicates that male-engineering undergraduates experienced daily negative outcomes more frequently than female-humanities undergraduates. Although lacking statistical significance, these findings may be interesting if we bear in mind that, in Portuguese universities, engineering is mostly studied by male students, while women typically occupy the degree courses integrated in the humanities. Hence, gender could have an impact on the academic major differences concerning real-world outcomes.

Table 2
RWO Total Score According to Academic Major and Gender

	<i>n</i>	Min - Max	<i>M</i> (<i>SD</i>)	Skewness	Kurtosis	α
Humanities	119	0-18	5.82 (3.61)	.72	.48	.81
Engineering	98	0-17	6.32 (3.59)	.75	.29	.76
Female	140	0-18	5.84 (3.57)	.67	.31	.80
Male	77	0-17	6.42 (3.67)	.82	.46	.78
Total	217	0-18	6.04 (3.60)	.72	.36	.79

Discussion

The three main goals of our study using the RWO, a measure of everyday problematic behaviors/negative outcomes that derive from decisions where CT seems to have been absent, were to: (a) translate and “culturally adapt” the RWO to the Portuguese language; (b) administer this experimental version to college students and estimate its reliability; (c) collect empirical data to offer some considerations concerning undergraduates’ everyday negative outcomes – and, presumably, CT mediating their daily decisions – according to academic major and gender. The results are expected to contribute to developing knowledge in the field of CT, given the existing line of studies using measures of CT and of real-world outcomes, and seeing that research involving the RWO is truly relevant in the broader field of CT, which would lose its significance if detached from real-life settings and situations. Also, by assessing and analyzing the real outcomes of CT, it is possible to gain a better understanding of college students’ personal, academic, and professional experiences. From such information, it is also possible to develop more efficient intervention strategies.

Concerning the translation of the RWO from English to Portuguese, our main purpose was to assure a “cultural conversion” of the items, making them properly clear and comprehensible, but mostly, making it possible for Portuguese college students to effortlessly relate to them (a). From our data, and according to the perceptions of a sample of students interviewed for this purpose, a Portuguese version of the RWO that is faithful to the original version, and “familiar” to Portuguese young adults pursuing a college degree, was accomplished.

Using this experimental Portuguese version of the RWO, a sample of undergraduates was assessed, first to analyze the reliability of this instrument (b). After eliminating a set of 10 problematic items, we obtained an inventory (with a total of 43 items) with appropriate internal validity (Cronbach’s $\alpha = .79$). We assume that the deleted items showed low reliability given the nature of these items and the age group of our sample. Similar to what happened in a study conducted by Butler (2012), the HCTA and RWO results were positively correlated, although this relation may not always be significant with samples of students in universities. The author questions if this derives from

the content of some items, such as divorce, detention in jail, check-bouncing, or credit card debt, which are more probable to have happened to adults.

As for the items deleted in the present study, given the current nature of this inventory (a check-list of outcomes that does not yet have a definite structure or a clear dimensionality), and since these items do not describe situations/outcomes that seem to be central in the lives of traditional undergraduates (e.g., plane trips, taking the wrong bus or train, using a credit card, etc.), this option does not anticipate difficulties. In fact, de Bie and Wilhelm (2014) undertook a similar study aiming the translation, adaptation, and validation of the RWO for The Netherlands, and in order to obtain a properly fit instrument, in terms of linguistic and cultural concerns, items were changed/removed, to achieve a precise cultural adaptation.

Finally, we examined if there were differences in undergraduates’ real-world outcomes related to their academic major and gender (c). Our data indicate that there are no statistically significant differences between participants, either due to their academic major or their gender.

First, academic major does not seem to be a variable that differentiates students’ everyday negative outcomes. If we refer exactly to CT, the literature seems to lack consensus concerning the impact of academic major on the development of CT (Li et al., 1999). Some authors report that it is not a factor *per se* when analyzing differences in students’ CT skills (Pike & Killian, 2011). For example, in a study by Pascarella et al. (1996), higher education was found to have a positive and increasing (i.e., throughout an academic year) impact on the quality of CT of students pursuing a two or four-year undergraduate degree, regardless of academic major. Then, there is research showing that cognitive skills do indeed develop differently according to academic major; for example, Kim and Sax (2011) indicate that the development of cognitive skills is less evident in the fields of the social sciences/humanities.

However, this association was mediated by a positive impact of student-teacher interaction: students in academic majors characterized by positive communication and supportive interaction with teachers, and a well-structured program that highlights skills related to CT, were the ones who showed higher gains in the development of their thinking skills. These results are analogous to data from research linking the development of CT to well-structured

classroom instruction that stimulates a deep approach to learning (Pascarella et al., 2013). In fact, a debate focused on which academic major is known for having *the best* critical thinkers may be a one-dimensional debate. According to Brint et al. (2012), what may originate, or exacerbate, the hypothesized gap between students from different academic majors concerning the quality of their CT skills is, not so much the academic field they are enrolled in, but mostly the teaching approaches to which they are exposed in that academic major.

These authors suggest that there are a few variables that make a difference. On the one hand, students in the social sciences/humanities tend to have contact with pedagogical techniques that call for more interaction with teachers and classmates, as well as for a higher level of participation in class. These situations are more favorable to the development of CT skills. On the other hand, science/engineering majors seem to be more challenging and have higher expectations, a reason why students are more committed to their academic learning, which is favorable to a better academic performance and better CT skills. Overall, perhaps it is a combination of both internal and external factors that encourage students' use of CT skills and the disposition to be a critical thinker, and it is difficult to determine what triggers the cycle: if it is an internal drive or external forces. Nevertheless, given the apparent lack of consensus concerning the association between academic major and CT, differential analyses using real-world outcomes of CT will, consequently, require further research.

Second, no gender differences were found in our study. Empirically driven data from the general literature show a set of differences between women and men in regard to cognitive ability, resulting from both biological and social variables (Halpern, 2012b). However, CT may not be one of those cognitive abilities that is highly susceptible to sex-based differences (Butler et al., 2012), perhaps because it is strongly influenced by, and adapted to, the contextual setting that is faced at that moment in time. In light of our findings, and considering the links between *good* thinking and life outcomes, in regard to differences in CT skills and everyday outcomes according to college students' gender, we shall avoid the dualistic debate focused on "who has *better* CT, men or women?"

There are a few limitations to be overcome in future research. It would be relevant to assess a sample where women and men are equally represented. Assessing this sample with both the HCTA and the RWO would also be advisable, in order to offer more thorough considerations regarding the quality of undergraduates' CT and their life outcomes as a result.

Despite its limitations, this study is a preliminary step in the study of the reliability of the RWO in Portugal, calling for further research in the future. It makes a contribution to the assessment field, also to CT assessment in higher education, and to cross-cultural research. The RWO could

be a relevant instrument to better understand the quality of daily life decisions, in this case of college students, and to develop improved intervention efforts. More importantly, the RWO can be used in cross-national research to assess the predictive ability of the HCTA, bringing European countries and the USA closer in this research field.

Despite (or perhaps because of) the immensity of information available, the variety of sources of information, and the diversity of opportunities made available that democratize knowledge and encourage further education, Socrates spoke truly when he stated "All I know is that I know nothing", more than 2,000 years ago. His quote is still accurate, given the daunting load of information that exists today, which cannot all possibly be transformed into knowledge by each one of us. We must be able to make the most of what we know and of what we manage to learn, and to optimize such data when dealing with real-life situations and making daily life decisions. Here, Education must prepare students for the real world. There are countless links connecting the quality of CT to outcomes in daily real-life situations, problems, or tasks. Hence, it is key, in today's world, to give CT the prominence it needs in the curriculum, making the changes that are needed to improve college students' thinking and performance. This should be done not simply within university walls, but especially outside those walls, amid real-life settings, in the face of real-world challenges.

References

- Almeida, L. S., & Franco, A. H. R. (2011). Critical thinking: Its relevance for education in a shifting society. *Revista de Psicologia (Lima)*, 29(1), 175-195.
- Almeida, L. S., & Freire, T. (2010). *Metodologia de investigação em psicologia e educação* (5th ed.). Braga, Portugal: Psiquilíbrios.
- Barnes, C. A. (2005). Critical thinking revisited: Its past, present, and future. *New Directions for Community Colleges*, 2005(130), 5-13. doi:10.1002/cc.191
- Barnett, R. (2012). Learning for an unknown future. *Higher Education Research & Development*, 31(1), 65-77. doi:10.1080/07294360.2012.642841
- Brint, S., Cantwell, A. M., & Saxena, P. (2012). Disciplinary categories, majors, and undergraduate academic experiences: Rethinking Bok's "underachieving colleges" thesis. *Research in Higher Education*, 53(1), 1-25. doi:10.1007/s11162-011-9227-2
- Butler, H., & Halpern, D. F. (2011). *Real-World Outcomes*. Unpublished instrument.
- Butler, H. A. (2012). Halpern critical thinking assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, 26(5), 721-729. doi:10.1002/acp.2851
- Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. S. (2012). Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, 7(2), 112-121. doi:10.1016/j.tsc.2012.04.001

- Cosgrove, R. (2011). Critical thinking in the Oxford tutorial: A call for an explicit and systematic approach. *Higher Education Research & Development*, 30(3), 343-356. doi: 10.1080/07294360.2010.487259
- de Bie, H., & Wilhelm, P. (July, 2014). *The Halpern Critical Thinking Assessment: Towards a Dutch appraisal of critical thinking*. Paper presented at the Ninth Conference of the International Test Commission, San Sebastián, Spain.
- Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, 92(5), 938-956. doi:10.1037/0022-3514.92.5.938
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning*, 7(3), 219-244. doi:10.1007/s11409-012-9092-1
- Ennis, R. H. (July, 1994). *The nature of critical thinking: An outline of critical thinking dispositions and abilities*. Paper presented at the Sixth International Conference on Thinking, Cambridge, MA.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Facione, P. A. (2011). *Critical thinking: What it is and why it counts*. Millbrae, CA: Measured Reasons & The California Academic Press.
- Franco, A. H. R., Butler, H. A., & Halpern, D. F. (in press). Teaching critical thinking to promote learning. In D. Dunn (Ed.), *The Oxford handbook of undergraduate psychology education*. New York, NY: Oxford University Press.
- Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New Directions for Teaching & Learning*, 1999(80), 69-74. doi:10.1002/tl.8005
- Halpern, D. F. (2012a). *Halpern Critical Thinking Assessment*. Moedling, Austria: Schuhfried (Vienna Test System).
- Halpern, D. F. (2012b). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.
- Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). New York, NY: Psychology Press.
- Kek, M. Y. C. A., & Huijser, H. (2011). The power of problem-based learning in developing critical thinking skills: Preparing students for tomorrow's digital futures in today's classrooms. *Higher Education Research & Development*, 30(3), 329-341. doi:10.1080/07294360.2010.501074
- Kember, D., Leung, D. Y. P., & Ma, R. S. F. (2007). Characterizing learning environments capable of nurturing generic capabilities in higher education. *Research in Higher Education*, 48(5), 609-632. doi:10.1007/s11162-006-9037-0
- Kim, Y. K., & Sax, L. J. (2011). Are the effects of student-faculty interaction dependent on academic major? An examination using multilevel modeling. *Research in Higher Education*, 52(6), 589-615. doi:10.1007/s11162-010-9209-9
- Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4(1), 70-76. doi:10.1016/j.tsc.2009.02.001
- Li, G., Long, S., & Simpson, M. E. (1999). Self-perceived gains in critical thinking and communication skills: Are there disciplinary differences? *Research in Higher Education*, 40(1), 43-60. doi:10.1023/A:1018722327398
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Pascarella, E. T. (2006). How college affects students: Ten directions for future research. *Journal of College Student Development*, 47(5), 508-520. doi:10.1353/csd.2006.0060
- Pascarella, E. T., Bohr, L., Nora, A., & Terenzini, P. T. (1996). Is differential exposure to college linked to the development of critical thinking? *Research in Higher Education*, 37(2), 159-174. doi:10.1007/BF01730114
- Pascarella, E. T., Martin, G. L., Hanson, J. M., Trolan, T. L., Gillig, B., & Blaich, C. (2014). Effects of diversity experiences on critical thinking skills over 4 years of college. *Journal of College Student Development*, 55(1), 86-92. doi:10.1353/csd.2014.0009
- Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco, CA: Jossey-Bass.
- Pascarella, E. T., Wang, J. S., Trolan, T. L., & Blaich, C. (2013). How the instructional and learning environments of liberal arts colleges enhance cognitive development. *Higher Education*, 66(5), 569-583. doi:10.1007/s10734-013-9622-z
- Paul, R. (2005). The state of critical thinking today. *New Directions for Community Colleges*, 2005(130), 27-38. doi:10.1002/cc.193
- Phan, H. P. (2010). Critical thinking as a self-regulatory process component in teaching and learning. *Psicothema*, 22(2), 284-292.
- Pike, G. R., & Killian, T. S. (2011). Reported gains in student learning: Do academic disciplines make a difference? *Research in Higher Education*, 42(4), 429-454. doi:10.1023/A:1011054825704
- Saiz, C., & Rivas, S. F. (2010). Melhorar el pensamiento crítico contribuye al desarrollo personal de los jóvenes? In H. J. Ribeiro & J. N. Vicente (Eds.), *O lugar da lógica e da argumentação no ensino da Filosofia* (pp. 39-52). Coimbra, Portugal: Universidade de Coimbra.
- Sternberg, R. J., Roediger, H. L., III, & Halpern, D. F. (2007). *Critical thinking in psychology*. New York, NY: Cambridge University Press.
- Stupnisky, R. H., Renaud, R. D., Daniels, L. M., Haynes, T. L., & Perry, R. P. (2008). The interrelation of first-year college students' critical thinking disposition, perceived academic control, and academic achievement. *Research in Higher Education*, 49(6), 513-530. doi:10.1007/s11162-008-9093-8

Amanda Franco is a Ph.D. student at the Instituto de Educação, Universidade do Minho.

Leandro da Silva Almeida is a Professor at the Instituto de Educação, Universidade do Minho.

Received: Jan. 28, 2015

1st Revision: Mar. 12, 2015

Approved: Mar. 25, 2015

How to cite this article:

Franco, A., & Almeida, L. S. (2015). Real-World Outcomes and critical thinking: Differential analysis by academic major and gender. *Paidéia (Ribeirão Preto)*, 25(61), 173-181. doi:10.1590/1982-43272561201505