Assessment of Undergraduates' Real-World Outcomes of Critical Thinking in Everyday Situations

Amanda R. Franco
Institute of Education, University of Minho, Braga, Portugal

Patrício S. Costa
School of Medicine, University of Minho, Braga, Portugal

Heather A. Butler
California State University Dominguez Hills, Carson, CA, USA

Leandro S. Almeida
Institute of Education, University of Minho, Braga, Portugal

Abstract

Critical thinking is a kind of "good" thinking that integrates a set of cognitive skills and dispositions to use those skills with knowledge to increase the chances of success in academic settings, job market, and daily life. The impact of critical thinking on life events, in face of everyday decisions and challenges, is still unclear, and further research is needed. In this exploratory study, a sample of 230 first-year students of a Bachelor's Degree or a Master's Degree in Portugal completed an experimental Portuguese version of the *Real-World Outcomes*, a self-report inventory measuring everyday negative life events that are mediated by a lack of critical thinking. Based on exploratory factor analysis results and theoretical premises, changes were made to the Portuguese version of the inventory that was administered, and items were aggregated into six dimensions, creating a new version that is more familiar to Portuguese young adults in college. This original proposal of the inventory presents six types of negative life events resulting from a lack of critical thinking: health neglect, mismanagement, slackness, poor impulse control, academic negligence, and rashness. Both limitations and future potentialities of this version are presented.

Corresponding Author:
Amanda R. Franco, Instituto de Educação, Universidade do Minho, Campus de Gualtar, Braga 4704-553, Portugal.
Email: amanda.hr.franco@gmail

Introduction

There is a multiplicity of definitions for critical thinking (CT). Yet, CT can be defined simply as having a set of cognitive skills and strategies, and the disposition to use those skills and knowledge to increase the chances of success in one's favor (Halpern, 2014). According to the literature, someone capable of this kind of ''good'' thinking (Franco, Butler, & Halpern, 2015) has a higher possibility of success in academic settings, in the job market, and in daily life (Butler, 2012; Dwyer, Hogan, & Stewart, 2012). A small number of studies have explored the relationship between CT and behavior in real-life settings and situations, but more empirical data are needed to understand more precisely the impact that being a critical thinker has on everyday decisions and challenges. Furthermore, if CT does impact the outcomes of everyday decisions and actions, the impact may vary according to individual characteristics, such as age or education. According to Howenstein, Bilodeau, Brogna, and Good (1996), the number of years of education is a significant factor to explain CT. Butler (2012) presents data reinforcing this possibility; in her study, it was years of education, not age, that predicted CT. Nevertheless, these findings may reflect individual differences in CT, not life events or outcomes, so they must be considered with caution. That said, age may impact the quantity and type of life experiences an individual is exposed to and may contribute significantly to our ability to predict life outcomes.

In light of this gap in the literature, there is a line of research focused on CT and life outcomes using the *Real-World Outcomes* (RWO; Butler, 2012), a self-report inventory designed to measure everyday decision-making, more specifically, the inventory measures the proportion of negative life events experienced by the respondent, which is assumed to be mediated by a lack of CT given the negative impact on the individual. The RWO inventory assesses a broad range of everyday behaviors concerning diverse areas of life, such as academic, interpersonal, health, political, legal, or financial. The negative outcomes have different levels of severity and result (presumably) from poor reasoning at that time.

The present study aimed to explore the psychometric qualities of the RWO inventory. For this purpose, the *Theory of Planned Behaviour* may aid our understanding of students' performance and better contextualize the (negative) outcomes experienced by each individual. This explanatory model of social behavior is focused on how information is processed and decisions are made, namely, for the prediction of behavioral intentions, given the correlation between the intention to perform a given behavior and the effective performance

of that specific behavior (Ajzen, 2011). The *Theory of Planned Behaviour* hypothesizes that the intention to perform derives from attitudes, subjective norms, and perceptions of behavioral control. Attitudes refer to the positive or negative assessment made about a given behavior. Subjective norms refer to the social pressure towards performing or preventing a given behavior. Perceptions of behavioral control refer to how easy or difficult the subject perceives the performance of a given behavior (Ajzen, 1991). The context has an impact on the attitudes, subjective norms, and perceptions of control that shall be activated, and also, on the prominence they will gain. Thus, cognitive processing varies according to circumstances and context (Ajzen, 2011), and behavior is not always the result of reasoned action. Individuals do not, in face of each decision, ''carefully and systematically review all available information before they form an intention to engage in a behaviour'' (Ajzen, 2011, p. 1121). Behavior may be the product of conscious processes, but it may also be the product of more automatic processes where awareness is lacking (Wood, Quinn, & Kashy, 2002). Indeed, Kahneman (2011) claims that there are two modes of thinking: one that is intuitive, immediate, and fast; one that is logical, deliberate, and time-consuming. Only decisions that are considered important and actions that occur in face of novel situations require the kind of deep thinking associated with the second mode; everyday decisions and routine behaviors do not (Ajzen, 2011). For the present study, such considerations about the fallibility of thinking could be helpful to explain how undergraduates make decisions and behave in ways that quite often precipitate negative outcomes for themselves. Overall, the RWO inventory makes it possible to detect if a person experienced a given situation and, more importantly, if in face of that situation that person made a ''good'' decision, thus preventing the occurrence of a negative outcome.

RWO of CT: Cross-cultural studies using the RWO

According to Butler (2012), a higher quality of CT should be associated with a lower frequency of negative outcomes resulting from everyday life decisions, on the grounds that CT is ''good'' thinking used deliberately across situations and problems to tackle challenges and accomplish goals, to prevent negative outcomes (Halpern, 2014). In Butler's (2012) original study, community college students ($n = 35$) and state university students ($n = 46$) from different courses, and community adults ($n = 50$) were assessed. Participants' mean age was 27.2 years ($SD = 13.16$), and a majority was female (66.4%). According to her findings, scores on the RWO inventory were predicted by scores on the *Halpern Critical Thinking Assessment* (HCTA; Halpern, 2012), with participants who scored higher on the CT test reporting a lower proportion of negative RWO, $r(131) = -.38$, $p < .001$, with 14% of explained variability. This study was replicated in Ireland by Dwyer et al. (2012). A sample of 74 undergraduate

Psychology majors were considered, ranging in age from 18 to 25 years, and a majority was female (64.9%). Participants were assigned to either an experimental ($n = 43$) or a control ($n = 31$) group; students in the experimental group participated in a free six-week CT course online. Both groups responded to the RWO inventory and the HCTA test at pre-testing, and to the HCTA test once again after the CT course. Participants' performance on the CT test predicted scores on the RWO inventory, $r(70) = -.28$, $p = .019$, hence replicating the findings from the original study (Butler et al., 2012), even though accounting for only 7.8% of explained variability. As for the study conducted in the Netherlands, by de Bie and Wilhelm (2014), a sample of 240 Dutch freshmen and sophomore students majoring in Communication or Psychology was assessed. The majority of the participants were female (80%) and ranged in age from 18 to 32 years ($M = 20.5$, $SD = 2.07$). In contrast with findings from the USA and Ireland, the relationship between the RWO inventory and the HCTA test scores was not significant, which could be due to social desirability, cultural differences, or a more limited range of life experiences experienced by college students because of their age.

In keeping with this line of research, we have been conducting research aimed at understanding the impact that CT has on the everyday decision-making outcomes of college students in Portugal. The present paper explores the psychometric properties of a Portuguese version of the RWO inventory and builds on a previous study (Franco & Almeida, 2015). In future research, our goal is to explore whether scores on the RWO inventory can be predicted by scores on the HCTA test, similarly to the studies conducted in the USA, Ireland, and the Netherlands, with the goal of considering interesting cross-cultural implications. For the time being, we present a reformulation of the Portuguese preliminary version of the RWO inventory, with items that are more familiar to Portuguese college students, aggregated according to life areas impacted by the negative outcomes on the grounds of statistical analyses. Even though the RWO inventory comprises a diversity of distinct negative outcomes that are assumed to result from a lack of CT, such an effort to find a dimensionality in this inventory was made because CT has a contextual nature. It is "applied" thinking, very dependent on personal skills and dispositions, but also, the context itself, to which are summoned particularities of knowledge, situation, and/or circumstances. For this reason, it is relevant to analyze if the behaviors resulting in negative outcomes for the performer that are assessed by the RWO inventory are more common in (one) specific area(s), especially since we refer to a very homogeneous population, with very particular characteristics, experiences, and routines, such as young adults pursing a college education at a public university. By grouping items according to context and analyzing the structure of this inventory, it becomes possible to identify particular domains in which CT may be applied or may lack given the particularities of a specific situation. Finally, both limitations and potentialities of the RWO inventory are identified and reflected

upon, with the goal of providing helpful insights for further research using this inventory, namely in the field of CT. Overall, this study can contribute to research concerning CT and its impact on the RWO in everyday life not only in Portugal but also in a cross-cultural panorama. Next, we present the Portuguese study in further detail.

Method

Participants

A total of 239 students attending a public university located in the north of Portugal completed the RWO inventory. Data from nine participants were disregarded: three participants showed a high number of missing answers (nine or four missing answers each), two participants failed to respond to item 17 (concerning the occurrence of sexual intercourse) completely, and four participants failed to respond to item 36 (referring to having voted in the most recent elections). The remaining 230 participants ranged in age from 17 to 48 years ($M = 21.8$, $SD = 5.56$), and a majority of the sample was female ($n = 188$, 81.7%). This convenience sample was composed of students in their freshmen year of a Bachelor's Degree ($n = 128$, 55.7%) or a Master's Degree ($n = 102$, 44.3%) in a variety of majors (e.g., Biological Sciences, Biomedical Sciences, Communication, Computer Science, Economics, Education, Engineering, Foreign Languages, Literatures and Linguistics, Management, Medicine, Physics, Psychology, etc.) in the scientific field of Social Sciences and Humanities ($n = 116$, 50.4%), or Science and Technology ($n = 114$, 49.6%).

Instrument

The RWO inventory is an adaptation of the *Decision Outcomes Inventory* (de Bruin, Parker, & Fischoff, 2007), an instrument designed to assess decision-making competence from adults' everyday decisions and behaviors. An experimental Portuguese version of the RWO inventory was used after undergoing a process of linguistic and cultural translation/adaptation to make it familiar, relevant and culturally adapted to Portuguese college students (see Franco & Almeida, 2015). A few items were adapted, others were created, and others eliminated (in this latter case, since they showed to be infrequent for a young age cohort). Such decisions were chiefly made on the grounds of a study concerning the preliminary Portuguese version (see Franco & Almeida, 2015). This Portuguese experimental version of the RWO inventory was composed of 33 items sets plus 9 individual items; overall, there are 42 dichotomous "Yes or No" neutral statements that describe daily life events (e.g., *Gone shopping for food or groceries*); then, referring to each neutral statement, sub-items that describe negative outcomes that derive from that situation in particular

(e.g., *Threw out food or groceries you had bought because they went bad*) are presented. The respondent indicates whether each daily life event, and each negative outcome, has (in which case "Yes" is selected) or has not (in which case "No" is selected) been experienced in the past year. Similarly to the original English version, scores on each item ranged from 0 (the respondent did not experience that daily life event, or negative outcome) to 1 (the respondent did experience that daily life event, or negative outcome), and a higher RWO total score indicates a greater proportion of negative outcomes.

Procedures

The administration of the RWO inventory took approximately 15 minutes per participant. Participants were approached through their teachers, who kindly spared a few minutes of class to present the study goals, ask for students' voluntary participation, and guarantee the principles of informed consent and confidentiality.

Data analysis followed the following steps: (i) analysis of the frequency distribution of items; (ii) analysis of the tetrachoric correlation matrix for items; (iii) performance of an exploratory factor analysis (EFA) using the Principal Component Analysis extraction method, with an Oblimin rotation; and (iv) analysis of internal consistency, using the Kuder-Richardson Formula 20 (KR-20), given the dichotomous scale of the items. The significance cutoff point considered for statistical analyses was $p < .05$. All statistical analyses were conducted using the statistical software IBM SPSS for Windows (version 22.0).

Results

First, we analyzed the frequency distribution of items. For each item, respondents indicated whether each daily life event, and each negative outcome, had (in which case "Yes" is selected) or not (in which case "No" is selected) been experienced in the past year. From the analysis of the items' response frequency, a set of items does not contribute to variability, seeing that a high percentage of participants did not experience that situation and/or negative outcome in particular.

Following, we conducted an analysis of the tetrachoric correlation matrix for items. In this analysis, we did not consider the neutral statements that describe daily life events (e.g., 3a. *Gone shopping for food or groceries*), but only the sub-items, i.e., the items that describe actual negative outcomes that derive from that situation in particular (e.g., 3b. *Threw out food or groceries you had bought because they went bad*). Aside from determining whether there was an inherent dimensionality to the RWO inventory by analyzing which items were correlated, this analysis was also conducted to identify which items would be relevant to maintain in the Portuguese version of the RWO inventory, as well as which

items needed to be reformulated and/or eliminated or even merged (item 42b and 42c, since they were highly correlated: $r^{tet} = .88$), in order to create a more structured instrument with items that are both pertinent and common to the average college student. Only the items that showed at least 10% of explained variance (with correlations $\geq .30$) were retained, since they meet the minimal level of practical significance (Hair, Anderson, Tatham, & Black, 1998). Nevertheless, not all items that assured this criteria were kept, since they (i) correlated to just a few other items, (ii) did not correlate to other items at all, or (iii) were considered only fairly relevant to maintain, seeing that they did not concern very common outcomes which could provide information about college students to a large extent.

Based on the tetrachoric correlation matrix, and following the previous elimination and merging of items, we conducted an EFA (extraction method: Principal Component Analysis, using the Oblimin rotation method) to examine how the items we decided to maintain aggregated amongst them. This would enable us to test our hypothesis regarding the dimensionality of the RWO inventory; therefore, the existence of particular life domains in which impulsivity or routine would take over consciousness and good judgment, i.e., CT. The Kaiser criteria was used for factor extraction; only factors with eigenvalues greater than one were retained for interpretation. Results from the EFA showed a five-dimension solution with satisfactory internal consistency: KR-20 index ranging between .518 and .694, accounting for 51.3% of explained variance (cf. Table 1).

Based on the results from the EFA, we opted to remove items 8c and 10c, since they did not load on any of the five dimensions that emerged and their removal increased the Cronbach's alpha if deleted (.588 and .705, respectively). Also, items 26 and 40b were eliminated, since their deletion increased the internal consistency of the dimension they belonged to (.523 and .757, respectively). Lastly, a few items were removed because they did not group in a dimension that made theoretical sense, nor did they seem to add relevant information for a Portuguese college population (4m, 21, 31b, 32b, 36c, 36d, and 42b + 42c).

Given their content and correlations, we aggregated items in six dimensions, each created by a set of items that had grouped in the same factor in the EFA and that had a common nature. Each set seems to represent a type of negative outcome: (i) *health neglect*—neglect concerning one's health, or lack of knowledge or information when making decisions concerning health; (ii) *mismanagement*—improvidence and poor management of time and everyday chores; (iii) *slackness*—carelessness and inattention concerning goods and finances; (iv) *poor impulse control*—behavior that is harmful to oneself and/or other people, such as substance abuse, reckless driving, or aggressiveness; (v) *academic negligence*—negligence or carelessness that affect different aspects of academic life, such as classes, performance, and/or grades; and finally (vi) *rashness*—imprudent decision-making, resulting from unawareness or from being misinformed.

Table 1. Factors resulting from the EFA (Principal Component Analysis, Oblimin rotation).

Items	F1	F2	F3	F4	F5
26	.734	−.107	.153	−.225	−.165
42bc	.588	.201	.081	.197	−.309
4m	.547	.125	−.346	.107	.195
32b	.518	.093	−.292	.247	.169
31b	.425	.404	−.181	.096	−.040
21	.290	.248	−.159	−.026	.004
30b	−.023	.798	.053	−.137	−.025
30c	−.011	.706	.016	.070	.000
19b	.158	.576	.031	−.046	.218
29b	.094	.534	−.130	−.019	.049
4n	−.200	.474	−.308	−.060	−.375
2b	.107	.466	.218	.245	−.229
3b	−.127	.449	−.003	.216	−.247
1b	.211	.420	.030	.056	−.251
4d	.006	.359	−.208	.315	.165
18b	.137	−.241	−.847	.050	−.197
4j	.049	.099	−.737	.093	.117
18c	−.127	.056	−.698	.233	.004
40d	.316	−.093	−.621	.239	.116
34	−.184	−.138	−.606	.058	−.285
4i	−.008	.259	−.567	−.233	.082
6f	.163	−.015	−.546	−.296	−.322
4l	−.072	.219	−.458	.264	−.041
10c	.095	.150	−.278	.098	.180
36c	−.120	−.058	−.081	.889	−.033
36d	.042	−.148	−.016	.828	−.043
40b	.139	.223	−.061	.424	−.006
6h	.022	.045	−.168	−.088	−.831
6c	.094	−.041	.013	.236	−.689
6g	.050	.375	−.179	−.125	−.562
17c	−.115	.275	−.327	−.015	−.495
8c	−.288	.345	−.207	−.022	.489
Kuder-Richardson 20	.518	.616	.694	.574	.536
Eigenvalues	7.58	2.71	2.33	2.05	1.74
% Cumulative explained variance	23.7	32.2	39.4	45.8	51.3

Hence, we created the final version of the Portuguese RWO inventory (cf. Table 2), with fewer and more precise items, which seem more familiar to Portuguese young adults pursing a college degree. In this final version, all items were reformulated to the first person, so participants could more easily identify with the inventory when thinking about the situation and responding. Also, we decided not to include the neutral items in this final version of the RWO inventory, since we considered their inclusion would contribute for time consumption and, very likely, fatigue when responding to the inventory.

Table 2. Final version of the Portuguese RWO, by dimension.

Dimension	Item	
Health neglect	1	I ate too much food too often.
	2	I ate unhealthy food too often.
	3	I been out in the sun and decided not to wear sunscreen.
Mismanagement	4	I spent so much time watching television it affected college negatively.
	5	I repeatedly arrived late to class.
	6	I forgot to do a class assignment.
Slackness	7	I bought new clothes or shoes I never wore.
	8	I threw out food or groceries I had bought because they went bad.
	9	I returned a book I borrowed from the library without having read it at all.
Poor impulse control	10	I drank so much alcohol I vomited.
	11	I drank so much alcohol I could not remember parts of the night (I "blacked out").
	12	I smoked cigarettes.
	13	I hit something with my car.
Academic negligence	14	I skipped an important class to do something fun (e.g., go shopping).
	15	I posted something to a social networking website during class.
	16	I cheated on an exam.
	17	I went out with friends instead of studying for a test/an exam.
Rashness	18	I texted while driving.
	19	I got a parking ticket.
	20	I ran a stop sign or traffic light.
	21	I had unprotected sex.

Overall, the final Portuguese version has a total of 21 Yes-No items aggregated in six dimensions representing a type of negative outcome.

Discussion

This study must be considered in a longitudinal time frame. Our present goal of translating, adapting, and validating a Portuguese version of the RWO inventory anticipates future research using both the RWO inventory and the HCTA test to examine the relationship between CT and everyday life decisions and outcomes of college students. The RWO inventory was chosen given its potential to measure students' everyday behaviors and outcomes in a multiplicity of situations in their real lives. The HCTA test is an instrument that uses both elaboration and multiple-choice items to measure CT in everyday scenarios. Hence, it allows for a more comprehensive way to measure CT, with scenarios that are familiar to students because they encounter them in their everyday lives (Franco, Almeida, & Saiz, 2014).

From the analysis of the correlations between sub-items (i.e., items that describe negative outcomes), those that are relevant to maintain in the Portuguese version of the RWO inventory, as well as items that needed to be changed or eliminated, were identified. The goal was to obtain a structured inventory with items that are truly relevant and familiar to Portuguese undergraduates. This goal is especially relevant because CT is a cognitive construct that is very dependent on personal skills and dispositions, but it is also reliant on the context itself. By grouping items and analyzing the dimensionality of the RWO inventory, particular domains in which CT is used or neglected emerged. Six types of items emerged from our data, which we designated as: *health neglect*, *mismanagement*, *slackness*, *poor impulse control*, *academic negligence*, and *rashness*. These types of outcomes that emerged shed some light on the daily worries and everyday experiences of young adults pursuing a college degree. There is the academic life sphere, which assumes a huge part of the life of an individual who is attending university, perhaps as a freshman who has just recently moved away from her/his parents' house and who is now starting to experience independent decision-making and assuming more adult responsibilities, and whose main occupation is to study. Then, the social dimension of going to college becomes evident, with all its challenges about spending time with other young adults and being under their influence, and having to manage self, time, tasks, goods, or personal finances, and having to make tough decisions about what to believe in, and how to behave. From the analysis of the EFA and according to a qualitative analysis, items were eliminated, creating a final version of the RWO inventory with sets of three to four items divided according to the type of negative outcome.

Such decisions made by college students, as well as the skills and will to think about them, can be better understood in light of the *Reflective Judgment Model*

by King and Kitchener (2004), which is similar to CT models. According to this model, cognitive processes undergo a process of development, which is why reasoning becomes more and more complex by late adolescence and adulthood. An individual's assumptions about the nature of her/his beliefs and knowledge change, making her/him gradually more "qualified" to think about ill-structured problems, blurry issues, and tricky questions, progressing from a (more or less) dualist (i.e., "*black or white* thinking") to a (more or less) relativist (i.e., "*shades of grey* thinking") approach, and yet, being able to opt for one or the other according to the matter in hand. In fact, there are no right or wrong answers; there are answers that are better or worse—according to the circumstances and situation. Concerning the participants in this study, as is the case with the majority of university students, they are at stage four of these reasoning stages, "quasi-reflective reasoning," where the nature and justification of knowledge is still very personal and situational (King & Kitchener, 2004). In other words, personal beliefs, significant others, and contextual cues have an impact on reasoning and decision-making. This may help to understand how these students make decisions, why they behave like they do, and which variables produce their everyday life outcomes. Both context and circumstances do seem to have an impact on cognitive processing and may explain why individual action is not always (or even often) directed by a careful and rational process of decision-making. On the contrary, everyday decisions and daily actions may be very much dependent of a multiplicity of variables, such as intuitive thinking or emotions (Kahneman, 2011), or even past behavior (Ajzen, 1991, 2011). Rather than deliberate intentions, it may be habit (which is guided by automatic processing modes that occur astray from awareness and consciousness) that directs action. And while action guided by habit might demand less cognitive effort and arouse minimal levels of anxiety, habitual behavior should be performed purposefully, in order to assure success, emotional commitment (Wood et al., 2002), or even fewer negative outcomes. In the frame of the present study, for the kind of real-life events and experiences evaluated by the RWO inventory, it may be that convenience or expediency is running the show when it comes to perform routine decisions and actions, rather than rationality. Overall, these findings give strength to one of the core characteristics of CT: it is applied thinking. It happens in a real-life context, according to a given set of real-life circumstances, in the face of what is known and can be done in that moment in time by that person in particular (Franco et al., 2015; Halpern, 2014).

Final considerations

The present study has limitations, the first one concerning the lack of strong data supporting the validity of the RWO inventory. The answer to this question would be that we believe that this inventory can become a relevant instrument to grasp everyday outcomes of students. Moreover, the RWO inventory has

been used in especially pertinent research about CT and the assessment of CT using the HCTA test (Butler, 2012; Butler et al., 2012; de Bie & Wilhelm, 2014; Dwyer et al., 2012) to analyze CT in students' everyday lives. Bearing in mind the comprehensive format and relevance of the HCTA test, as well as the possibilities that arise from linking the RWO inventory and the HCTA test, and also, our validation study of the HCTA in Portugal (Franco, Costa, & Almeida, in press), we consider that this inventory is useful after undergoing necessary adjustments in further research. Indeed, this is an exploratory study, and future studies are needed, namely to analyze the psychometric characteristics of the revised version of the RWO inventory we proposed here on the grounds of statistical analyses and theoretical assumptions. Moreover, in future studies, we intend to correlate this revised version of the RWO inventory with the Portuguese version of the HCTA test (Franco, Costa, & Almeida, under review), seeing that such a contribution would add to both instruments' validation. It is of utmost importance that both instruments are validated in Portugal, seeing that this is a glaring gap for researchers and practitioners who (wish to) study and work in the field of CT and lack instruments to measure this construct. Also, perhaps a reformulation of the RWO's response range would be of interest to better understand how common these (negative) outcomes really are. For instance, instead of a "yes or no" dichotomous grading system, perhaps a wider response range would allow participants to report the frequency with which they experience certain negative life outcomes. Nevertheless, some of the items describe situations that do not happen on a daily basis (e.g., *I purchased herbal remedies to enhance my thinking or memory*), while others can happen on a more frequent basis (e.g., *I ate too much food or/and unhealthy food too often*). Hence, the reasoning behind why we decided to keep a dichotomous range of response.

Declaration of Conflicting Interests

Funding

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology and Health*, *26*(9), 1113–1127.

Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, *26*(5), 721–729.

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., & Almeida, L. S. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills & Creativity*, *7*(2), 112–121.

De Bie, H., & Wilhelm, P. (2014). *The Halpern critical thinking assessment: Towards a Dutch appraisal of critical thinking*. Retrieved from http://essay.utwente.nl/64857/

de Bruin, W. B., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*(5), 938–956.

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.

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.

Franco, A. H. R., Butler, H. A., & Halpern, D. F. (2015). Teaching critical thinking to promote learning. In D. Dunn (Ed.), *The Oxford handbook of undergraduate psychology education* (pp. 65–74). New York, NY: Oxford University Press.

Franco, A. R., Almeida, L. S., & Saiz, C. (2014). Pensamiento crítico: Reflexión sobre su lugar en la enseñanza superior. [Critical thinking: Reflection about its place in higher education.]. *Educatio Siglo XXI*, *32*(2), 81–96.

Franco, A. R., Costa, P. S., & Almeida, L. S. (in press). Translation, adaptation, and validation of the Halpern Critical Thinking Assessment to Portugal: Effect of disciplinary area and academic level on critical thinking. *Annals of Psychology*.

Franco, A. R., Costa, P. S., & Almeida, L. S. (under review). Do critical thinkers drink too much alcohol, forget to do class assignments, or cheat on exams? Using a critical thinking measure to predict college students' real-world outcomes. *Psychological Studies*.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). New Jersey, NJ: Prentice-Hall.

Halpern, D. F. (2012). *Halpern Critical Thinking Assessment [Measurement instrument]*. Mödling, Austria: Schuhfried.

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). New York, NY: Psychology Press.

Howenstein, M. A., Bilodeau, K., Brogna, M. J., & Good, G. (1996). Factors associated with critical thinking among nurses. *Journal of Continuing Education in Nursing*, *27*(3), 100–103.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

King, P. M., & Kitchener, K. S. (2004). Reflective judgment: Theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, *39*(1), 5–18.

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday Life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, *83*(6), 1281–1297.

Author Biographies

Amanda R. Franco, holds a Master's degree in Educational Psychology, and a PhD in Science of Education, both from University of Minho, Braga, Portugal. Her PhD thesis concerns the assessment of college students' critical thinking. Her main publications are in the field of critical thinking, cognitive assessment, and higher education, yet her research interests also involve transversal skills, lifelong learning, and citizenship. She has recently been awarded a Post-doctorate grant to lead a research project at University of Aveiro, Aveiro, Portugal.

Patrício S. Costa, PhD, is Assistant Professor of Biostatistics at School of Medicine, University of Minho, Braga, Portugal, and of Research Methodologies and Statistics at Faculty of Psychology and Education Sciences, University of Porto, Oporto, Portugal. He is also coordinator of *DOMP, SA the Statistics Department*, an opinion and market research company. He is reviewer and has published in several international journals of education, marketing, political science, and neurosciences.

Heather A. Butler received her PhD in Psychology from Claremont Graduate University, California, USA. She is an Assistant Professor in the Psychology department at California State University Dominguez Hills. She has a number of research interests that are grounded in human cognition (such as critical thinking, advanced learning technologies, and cognitive bias in the legal system).

Leandro S. Almeida holds a PhD in Psychology from University of Porto, Oporto, Portugal, and he is a Full Professor and researcher at the Education Research Center (CIEd) at University of Minho, Braga, Portugal. He is the leading researcher of many studies in the field of cognition and learning, and the author/co-author of psychological assessment tests, as well as of cognitive training or learning promotion programs. He coordinates the Observatory of Students' Academic Paths at University of Minho.