**Versión aceptada:**
Bravo, MJ., Galiana, L., Rodrigo, M.F., Navarro-Pérez, J.J. & Oliver, A. (2021). An adaptation of the Critical Thinking Disposition Scale in Spanish youth, *Thinking Skills & Creativity,*
https://doi.org/10.1016/j.tsc.2020.100748

# An adaptation of the Critical Thinking Disposition Scale in Spanish youth

Maria Jesús Bravo [a], Laura Galiana [b], Maria F. Rodrigo [b], José-Javier Navarro-Pérez [c], Amparo Oliver [b,*]

[a] *Department of Social Psychology, University of Valencia, Spain*
[b] *Dept. of Methodology for Behavioral Sciences, University of Valencia, Spain*
[c] *Dept. of Social Work & Social Services, Faculty of Social Sciences & IDL, Institute for Local Development, University of Valencia, Spain*

*Keywords:*
Critical thinking
Critical Thinking Disposition Scale
Psychometric validation
Multigroup confirmatory factor analysis
Thinking styles
Gender differences

A B S T R A C T

*Background:* Enhancing students' critical thinking (CT) is an essential goal of higher education. Although there are many instruments to assess its cognitive dimension, this is not so in the case of dispositional thinking. Our aim is to examine the psychometric properties of the Spanish version of the Critical Thinking Disposition Scale (CTDS; Sosu, 2013).
*Method:* Data come from 1064 students in their final years at high school and university in Spain. We report validity evidence based on its internal structure by confirmatory factor analysis (CFA) and an estimation of reliability and measurement invariance by gender. Additionally, nomological validity study with a multiple-indicator multiple-causes, MIMIC model, for the university students is provided.
*Results:* Three CFAs are reported, with evidence supporting the one-factor structure as the best representation of the data: $\chi^2(44) = 413.556$ ($p < .001$); CFI = .976; RMSEA = .089[.081,.097]; SRMR = .031; ECV = .889. With this structure, measurement invariance between boys and girls is tested and the scalar model is retained. Differences in latent means are found, favoring girls ($d =$ 0.449; $se = .065$; $p < .001$). The reliability of CTDS is adequate, with a Cronbach's alpha of .777 and a CRI value of .925. When tested in university students, the MIMIC model offers evidence for nomological validity: higher levels are found for women ($\beta = .111$, $p < .050$) and also for those with higher levels of rational thinking ($\beta = .577$, $p < .010$).
*Conclusions:* The Spanish version of the CTDS has good psychometric properties (including gender measurement invariance not yet reported in literature), and is suitable for being used to measure the disposition to critical thinking in youth. The relevance and implications of these results are further discussed.

* Corresponding author at: Av. Blasco Ibánez, 21, 46010, Valencia, Spain.
*E-mail address:* oliver@uv.es (A. Oliver).

## 1. Introduction

> 'The world needs critical thinkers now more than ever before. Not only is the sheer amount of information available to us undeniably overwhelming, but the accuracy of the information is suspect' (Butler & Halpern, 2020, p. 152)

The relevance of critical thinking has been underlined in several spheres of life and, for that reason, it is a fundamental educational objective (Davies, 2015; Dwyer et al., 2014; Ennis, 2018; Franco, 2016). To this end, developing critical thinking, as opposed to '*lazy thinking*', is key to enabling the rigorous analysis of the great amount of information we receive, allowing us to elucidate the authenticity and interpret the accuracy of the so-called 'fake news' or to appraise a situation correctly without the influence of prior opinions or knowledge (Bronstein et al., 2019; Pennycook & Rand, 2019; West et al., 2008). Several studies support the role of critical thinking as a protective mechanism as compared to what were called by Lobato et al. (2014) "epistemically unwarranted beliefs" that include pseudoscience, paranormal and conspiracy beliefs (e.g. Aarnio & Lindeman, 2005; Bensley et al., 2014, Bensley & Lilienfeld, 2017; Dyer & Hall, 2019; Wilson, 2018).

Critical thinking is a construct with multiple definitions coming from different traditions and disciplines (Moore, 2013; Taghi- nezhad et al., 2019). For instance, it has been defined as "reasonable, reflective thinking that is focused on deciding what to believe ordo" (Ennis, 1987, p. 10) or "the propensity and skills to engage in activity and 'mental activity' with reflective skepticism focused on deciding what to believe or do…" (Fasko, 2003, p.8). This second definition includes and distinguishes two dimensions within critical thinking: dispositional (the inclination or tendency to act) and cognitive (ability or skill to act) So, although the definitions of critical thinking emphasize different components of this construct, there exists a wide consensus in the need to differentiate and consider these two components of critical thinking (e.g. Davies, 2015; Ku, 2009; Toplak et al., 2014). As pointed out by Davies (2015) "skills without the disposition to use them are not much use, so critical thinking is about dispositions as well" (p. 44). There are several taxonomies of important thinking dispositions, but aspects, such as open-mindedness, intellectual curiosity, and reflective thinking, are common to all of them (see Sosu, 2013, for a review of these taxonomies).

The above definition by Ennis (1987), as pointed out by Moore (2013) in his analysis of critical thinking definitions, also outlines the rational basis of critical thinking. In this regard, critical thinking has been linked to an analytical or rational style of thinking fromthe framework of the dual process theories of reasoning (Epstein, 1998; Evans & Stanovich, 2013; Kahneman, 2011). These theories set out two opposing types of thinking: "System 1" thinking that corresponds to an automatic, fast, preconscious, holistic, non-verbal and associative style of thinking (a.k.a., intuitive or experiential thinking), compared to "System 2" corresponding to thinking that is rational, slow, analytical and logical (a.k.a., analytical or rational thinking). Various authors equate critical thinking to System 2 thinking (e.g. Bonnefon, 2016; Halpern, 2014) or conceptualize critical thinking as a subspecies of rational thinking (Stanovich, 2016). Following this framework, Stupple et al. (2017) assessed the criterion-related validity of the Critical Thinking Toolkit they proposed by using scales that measure System 2 or rational thinking, such as reflective thinking through the Cognitive Reflection Test (CRT, Frederick, 2005), but there was no evidence for the Critical Thinking Toolkit correlating with reflective thinking (CRT).

### *1.1. The Critical Thinking Disposition Scale*

As regards the measurement of both dimensions of critical thinking, whereas there are different instruments with which to measure the cognitive dimension, there are none when it comes to the dispositional. As revised by Sosu (2013), the only instrument specifically developed to measure critical thinking disposition is the California Critical Thinking Dispositional Inventory (Facione & Facione, 1992), but its reliability and fiability has been questioned (Ku, 2009) Furthermore, the measurement of critical thinking to be found in the literature is often cause for confusion; this is because, although critical thinking disposition is spoken of, there are no questionnaires specifically designed for this purpose (e.g. Noone & Hogan, 2018; Toplak et al., 2014). Other methods used to measure critical thinking disposition, though not specifically designed to this end are: the Need for Cognition scale (Cacioppo et al., 1984) and the Actively Open-minded Thinking Scale (Stanovich & West, 1997).

After reviewing several taxonomies of important thinking dispositions, Sosu (2013) developed the "Critical Thinking Disposition Scale' (CTDS) from an initial 24-item pool. After applying exploratory factor analysis (study 1), the CTDS was reduced to an 11-item instrument that measures two dispositional domains: "Critical Openness" (7 items), which "reflects the tendency to be actively open to new ideas, critical in evaluating these ideas and modifying one's thinking in light of convincing evidence", and "Reflective Skepticism" (4 items), which is "the tendency to learn from one's past experiences and be questioning of evidence" (Sosu, 2013 p. 115). Subsequently, the author obtained support from a sample of UK students for both the reliability and validity of the CTDS, as well as for the hypothesized two-factor model through Confirmatory Factor Analysis (CFA), but including the residual covariance between two items from the "Critical Openness" dimension (Study 2). The high degree of correlation observed between the two factors (.81 and .83 in both samples), leads him to hypothesize in the discussion that the two factors measure a higher-order factor, "disposition to critical thinking". The coefficient of internal consistency for the total scale was high (Cronbach's alpha .79 and .81, in Studies 1 and 2, respectively).

As far as we know, only two studies have evaluated the psychometric properties of the CTDS (Akin et al., 2015; Yockey, 2016). Akin et al. (2015) examined the validity and reliability of the Turkish version of the scale and found support for the two-factor model with correlated residuals in a sample of university students from Turkey. Similar to Sosu (2013), the correlation between the two factors was high (.87). The Cronbach's Alpha coefficients of the scale were found to be .68 for the critical openness subscale, .75 for the reflective scepticism sub-scale, and .78 for the whole scale.

In a sample of U.S. undergraduates, Yockey (2016) compared the one-factor and the two-factor models, both with and without

correlated errors between two items. Given the non-significant change in chi-square between the one and two-factor models, and the close correlations between the factors in the two-factor models (.82 and .99 in the models without and with correlated errors, respectively), he recommended the more parsimonious one-factor model with correlated errors. The alpha value for the scale was .79 equal to the one reported by Sosu (2013).

In spite of the importance of this construct in the field of education, there are no references in the literature to any studies that have evaluated the dimensionality of the CTDS by means of CFA models proving factorial invariance across gender groups. As this is a prior requisite if a comparison across gender is to be performed, the absence of this kind of previous studies calls into question the results about differences in critical thinking disposition across gender. In this scenario, the previous studies that have taken into account gender-based differences as regards critical thinking disposition in students have found no statistically significant differences (Hunter et al., 2014; Liu et al., 2019; Salahshoor & Rafiee, 2016) with the exception of Yüksel and Alci (2012) who found them to favor women. The evidence of the validity of the CTDS derived from relationships with other cognitive variables is reduced to the study carried out by Yockey (2016), in which the correlation between CTDS and the Need for Cognition Scale (Cacioppo et al. 1984) was reported ($r = .40$; $p < .001$). The Need for Cognition Scale focuses on one's general interest in engaging and enjoying effortful cognitive activity and, as pointed out by Sosu (2013), has been used to measure critical thinking dispositions although it was not developed specifically to measure this construct.

### *1.2. The present study*

The aim of this study is to establish the psychometric properties of the Spanish version of the CTDS. Thereby, the dimensionality of the questionnaire is assessed, testing several competitive CFA models (one-factor, two-factor and bifactor models). Moreover, once the best model is determined, the factorial invariance across gender groups is assessed and the corresponding estimates of the reliability of the scale are offered. Lastly, this study assesses the criterion-related validity of the scale for the university students sample. As exposedin the introduction, critical thinking has been linked to an analytical or rational style of thinking from the framework of the dual process theories of reasoning (Bonnefon, 2016; Halpern, 2014; Stanovich, 2016). So, to explore the relationships of CTD with thinking styles we included measures of these two kinds of thinking: System 1 ("Faith in intuition") and System 2 ("Need for Cognition" and "Cognitive Reflection Test"). We hypothesize that a higher disposition to thinking critically will be related with higher scores in measures of System 2, but will not be related with System 1.

## 2. Method

### *2.1. Participants*

Data came from questionnaire administration to a sample of 1064 students in the last two years at high school and university students, from public and semi-private centers of Valencia (Spain). The mean age was 18.87 years old (*SD* 4.58). 632 were female (59.4 %) and 432 were male (40.6 %). 703 (66.1 %) were high school students. Out of them, 362 were women (51.5 %) and 341 were men (48.5 %). Mean age was 16.18 years old (SD 0.85), with a minimum age of 15 and a maximum age of 25 years old. 408 (58 %) were enrolled in public schools, whereas 295 (42 %) were enrolled in private ones. 259 (36.8 %) attended high schools in urban areas, 247 (35.1 %) in metropolitan areas, and 197 (28 %) in rural areas.

Regarding the subsample of university students, it was formed by 361 (33.9 %) university students. Out of them, 270 (74.8 %) were women and 91 (25.2 %) were men. Mean age was 22.28 years old (SD 5.98), with a minimum age of 18 and a maximum age of 55 years old. 213 were at the first course (59 %), whereas 148 were in the last course of their degree (41 %). The university students were enrolled on different degrees, including: Psychology (55 %); Social Work (11 %); Teacher Training (10 %); Philosophy (10 %) and Biochemistry (6%); the rest of them were studying other degrees, such as Medicine or Engineering.

### 2.2. Procedure

The CTDS was translated into Spanish by two native speakers, working independently one from the other. In case of disagreements, a third native speaker was consulted. The Spanish adaptation of the CTDS and scoring instructions are given in Appendix A.

Data were obtained during regular lessons, in the academic years 2017–2018 and 2018–2019. All the students consented to participate in the study and received no gratification. For all the participants, informed consent was obtained before the administration of the survey. It contained a cover letter inviting the students to participate in the study and ensuring confidentiality.

University student fulfilled the surveys during regular scheduled face-to-face classes across the different degrees. As regards the high school sample, and due to the age of participants, other additional requirements were fulfilled. The participant' centers were encouraged to participate with a professional designated as an interlocutor (the gender issues coordinator). Once the parents had granted authorization, the questionnaires were administered by the coordinator in regular class time and on-site.

Data gathering was framed within a competitive research project (ref. GV/2017/208 Generalitat Valenciana (ES)) and was additionally endorsed by ethical boards and committees from the Autonomous Government's Equality and Inclusive Policies area. The requirements of Spanish Law 15/1999 of 13th December on the Protection of Personal Data were also met.

### 2.3. Measures

The Critical Thinking Disposition Scale (Sosu, 2013) measures individuals' critical thinking disposition. This scale is an 11-item, five-point Likert-type scale (ranging from 1 = *strongly disagree* to 5 = *strongly agree*). Originally, the CTDS was assumed to assesstwo dimensions: 'critical openness' (7 items) and 'reflective skepticism" (4 items). Scores from the 11 items can be added up to obtain a total score, ranging from 11 to 55.

Two additional instruments were used only in the university students' sample (the Spanish translation of this instrument, and scoring instructions are given in Appendix B and C). Thinking styles were assessed by the Rational-Experiential Inventory (REI; Pacini & Epstein, 1999)

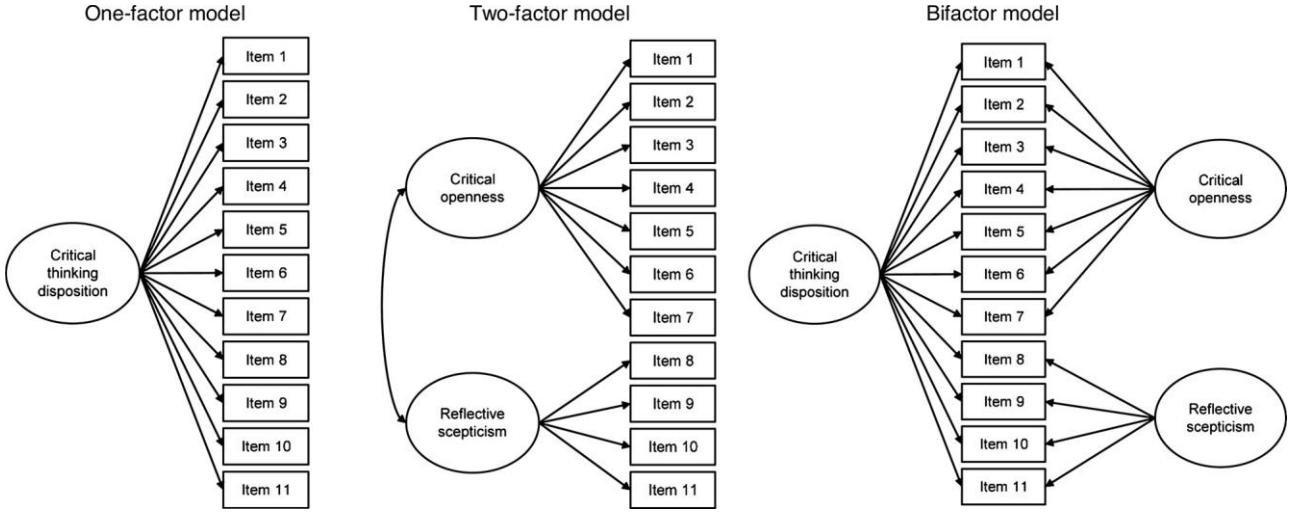


**Fig. 1.** CFA models assessed for the validation of the Critical Thinking Disposition Scale.

This inventory includes two scales that measure the preference for two types of thinking: Rational or Analytical thinking (or System 2) (a.k.a, Need for cognition) and Intuitive or Experiential thinking (System 1) (a.k.a., Faith in intuition). Both scales have 20 items, ranging in a five-point Likert-type scale (from 1= *strongly disagree*, to 5 = *strongly agree*). Examples from both scales are: '*I enjoy solving problems that require careful logical analysis*' (analytical thinking) and '*I like to rely on my intuitive impressions* (intuitive thinking). The Spanish version of this scale has showed good psychometric properties (Peñarroja et al., 2017, Sánchez et al., 2012; Witteman et al., 2009). The Cronbach's alphas in this study were .876 with a CI 95 % (.856, .894) for rational thinking, and .859 with a CI 95 % (.856, .894) for intuitive thinking in the university students' sample.

Reflective thinking was measured by the Cognitive Reflection Test (CRT; Frederick, 2005) and it refers to "the ability or disposition to resist reporting the response that first comes to mind" (Frederick, 2005, p. 35). The scale consists of three mathematical problems, in which in order to find the correct answer the subject has to inhibit an automatic, intuitive, incorrect answer. An example item is '*If it takes 5 machines* 5 min *to make 5 widgets*, *how long would it take 100 machines to make 100 widgets? ___ minutes*'. The CRT has been used extensively, and translated to different languages, including Spanish, in the heuristic and biases research in Psychology (for a review see the meta-study of Brañas-Garza et al., 2019). The CRT score was calculated as the sum of correct answers. The Cronbach's alpha was .787 with a CI 95 % (.744, .824) in the university students' sample.

### *2.4. Analysis*

First of all, we studied the factorial structure of the CTDS with competitive confirmatory factor analysis (CFA). Three CFAs were specified and tested, using Mplus 8.4 (Muthén & Muthén, 1998-2017; Muthén and Muthén, 1998), including (see Fig. 1):

- Two-factor model. This model is the original one proposed by Sosu (2013). Although this analysis included a covariance between the errors of two items in the critical openness factor, this is a depreciated practice (Bagozzi, 1983; Fornell, 1983). Therefore, no error covariance is included.
- One-factor model. This model is based on more recent evidence provided by Yockey (2016), who tested a general factor of critical thinking disposition that explained the 11 items of the CTDS and compared it with the original two-factor model (Sosu, 2013), offering evidence that supported the unidimensional structure.
- Bifactor model. This model tests the existence of three constructs: a general factor of critical thinking disposition, which explains part of the items' variance; and two specific factors, critical openness and reflective skepticism, which explain additional variance. In this bifactor model, the correlation between critical openness and reflective skepticism is due to the general factor; therefore, the correlation between them is fixed at zero. By definition, factors in a bifactor model are orthogonal.

All models were estimated with Weighted Least Squares Mean and Variance corrected (WLSMV), a recommended procedure for non-normal data of ordinal nature (Finney & DiStefano, 2006). The fit of the models to the data was assessed using several goodness-of-fit indices and literature recommendations (Kline, 2015), including: a) the chi-square statistic; b) the Comparative FitIndex (CFI); c) the Root Mean Squared Error of Approximation (RMSEA); and d) the standardized root mean squared residuals (SRMR). Following Hu and Bentler (1999), a CFI of 0.95 or higher, a RMSEA of 0.06 or lower, and a SRMR of less than 0.08 would indicate avery good fit of the model to the data.

As the competitive CFA were not nested, subjective criteria were used to compare the models. From this point of view, if a parsimonious model evidences adequate levels of practical fit, it is preferred over the more complex model. CFI differences (ΔCFI) of lower than .01 (Cheung & Rensvold, 2002) or .05 (Little, 1997) were used as cut-off criteria. Additionally, as a bifactor structure was also tested, the explained common variance (ECV) index was used (Sijtsma, 2009; Ten Berge & Socan, 2004). The ECV informs of variance specific to the general factor, by taking the ratio of variance explained by a general factor and dividing it by the variance explained by a general and specific factors, which are assumed to be uncorrelated (Reise et al., 2010; Rodriguez et al., 2016; Ten Berge & Socan, 2004). An ECV .85 suggests the instrument is sufficiently unidimensional to warrant a one-factor model (Stucky & Edelen, 2014). We also studied the performance of the items by means of IECV, in order to assess the unidimensionality at the item level. The IECV provides, in terms of Stucky et al. (2013), "the extent to which items' responses are accounted for by variation on the latent general dimension alone" (p. 51).

Reliability estimates were also offered. Cronbach's alpha and Composite Reliability Index (CRI) (McDonald, 1999) were used, both with values close to 1 for excellent reliability. Confidence intervals for Cronbach's alpha's were calculated using R 3.6.3 (R Core Team, 2020), the cocron package (v 1.0 1; Diedenhofen, 2016).

Once the best fitting model was retained (the one-factor model), the structure was tested on the groups of men and women. After gathering evidence from both groups, a measurement invariance routine was carried out, following recommendations by Van deSchoot et al. (2012). First of all, the configural or baseline model was tested. In this model, parameters are not constrained; instead,they are freely estimated for each group (men and women). Secondly, the metric invariance or weak invariance was tested. Factor loadings were constrained across groups, thus making them equal for men and women. If this model fits, the scalar invariance or strong invariance is tested, in which the intercepts of the groups are constrained. Finally, if evidence of scalar invariance is gathered, we can test latent mean differences for the groups under study. These models were compared using the CFI differences mentioned before, together with chi-square differences.

Finally, a multiple-indicators multiple-causes model (MIMIC), in which gender and age predict cognitive variables, and these cognitive variables (rational thinking, intuitive thinking, and reflective thinking), together with gender and age, predict a latent factorof critical thinking disposition, was tested on the sample of university students. The adequacy of the model, which was estimated using

WLSMV in Mplus 8.4 (Muthén & Muthén, 1998-2017; Muthén and Muthén, 1998), was assessed using the aforementioned fit indices ($\chi^2$, CFI, RMSEA, and SRMR).

## 3. Results

### 3.1. Factor structure and reliability

Three confirmatory factor analyses were specified, estimated and assessed, using the a priori structures, which are shown in Fig. 1. These models were: a one-factor model, a two-factor model, and a bifactor model. In Table 1, model fit indices for the three CFAs are shown. As shown in this table, the fit indices were excellent for the three models, except for the RMSEA of the one-factor model.

When comparing the fit of the models, negligible differences in the CFI were found between the one-factor model and the two-factor model (Δ .009). Additionally, the correlation between factors in the two-factor model was extremely high ($r$ = .933). As regards the comparison between the one-factor solution and the bifactor model, differences in CFI (Δ .019) could indicate a better fit of the bifactor model, based on the stricter cut-off criterion (Cheung & Rensvold, 2002). Evidence of unidimensionality was also studied with the ECV. The results pointed to an ECV = .889, with only items 5 and 11 with values of IECV under .85 showing relevant variance explained for the specific factors (see Table 2). However, all the items reported more than 50 % of the responses accounted for the variation in the general dimension. Taking into account this information, the one-factor model was retained as the best structure with which to represent the data.

As regards reliability, estimates were adequate. In the sample of university students, with a Cronbach's alpha of .766, CI 95 % (.728, .800) and a CRI value of .840. For high school students, the Cronbach's alpha was also .766, CI 95 % (.739, .791), and CRI was .815. Regarding women, Cronbach's alpha was .902, CI 95 % (.890, .913), and CRI was .927. Finally, for men, Cronbach's alpha was .887, CI 95 % (.870, .902), and CRI was .913.

### 3.2. Factorial invariance across gender

Once the one-factor model was retained as the best solution with which to represent the structure of the CTDS, this model was tested on the subgroups of men and women, separately. As the one-factor model adequately represented the data in both groups, the invariance routine was carried. The results indicated that the scalar model was the best representation of the data (see Table 3), as it was the most parsimonious and showed an improvement in CFI instead of a decrease. Therefore, latent mean differences between groups were tested, the results pointing to higher scores in critical thinking disposition for women ($d$ 0.449; *standard error*=.065; $p$ = < .001).

### 3.3. MIMIC model: relationships with cognitive variables

The correlations between the scores observed in CTDS and the cognitive variables were: $r$ = .13, $p$ = .008 for reflective thinking (CRT), $r$ = .50, $p$ < .001 for rational thinking and $r$ = .07, $p$ = .133 for intuitive thinking.

Finally, the *MIMIC* model in which rational thinking, intuitive thinking, and reflective thinking, together with gender and age, predicted a latent factor of critical thinking disposition, showed an adequate overall fit when tested on the sample of university students: $\chi^2(94) = 272.746$ ($p$ < .001); CFI = .906; RMSEA = .073 [.063,.083], and SRMR = .055. As shown in Fig. 2, only gender ($\beta$ = .14; $p$ < .01), age ($\beta$ = -.11; $p$ < .01) and rational thinking ($\beta$ = .59; $p$ < .01) predicted critical thinking disposition. Critical thinking disposition was higher for women, younger university students, and those with higher level of rational thinking, being rational thinking the most important predictor.

## 4. Discussion

Critical thinking is key to offset the negative effects that false beliefs or unfounded ideas have on everyday life. Accordingly, in reference to the COVID 19 pandemic, Van Bavel et al. (2020) have stated that fake news has proliferated widely on social media and that "Effort or elaboration determines the factors to which people attend when receiving messages. When effort is low, people rely on heuristics, such as source cues or argument length, whereas high effort leads them to focus on message quality" (p. 12). One group that is particularly vulnerable to these effects is the young, both male and female. This is due to their lengthy exposure to unverified sources of information as a consequence of their intensive use of new technologies. The development of a capacity to criticize during the formative stage, one which permits the individual to filter and question the quality of the information received, is a key challenge of the education system (Abrami et al., 2015; Chan, 2016; Davies, 2015; Ennis, 2018; Fong et al., 2017; Tiruneh et al., 2018). At this level,

**Table 1**
Confirmatory Factor Analyses: general fit.

| | $\chi^2$ | df | $p$ | CFI | RMSEA | IC 90% RMSEA | SRMR | ΔCFI |
|---|---|---|---|---|---|---|---|---|
| One-factor model | 413.556 | 44 | < .001 | .976 | .089 | [.081 .097] | .031 | – |
| Two-factor model | 333.525 | 43 | < .001 | .985 | .080 | [.072 .088] | .028 | .009 |
| Bifactor model | 134.977 | 33 | < .001 | .995 | .054 | [.045 .064] | .018 | .019 |

**Table 2**

Items' descriptive statistics for the Confirmatory Factor Analyses.

| | One-factor solution | Two-factor solution | | Bifactor solution | | | |
|---|---|---|---|---|---|---|---|
| | | Critical openness | Reflective skepticism | General factor | Critical openness | Reflective skepticism | IECV |
| | λ | λ | λ | λ | λ | λ | |
| Item 1 | .692 | .934 | | .937 | —.001n.s. | | 1.000 |
| Item 2 | .757 | .702 | | .712 | —.033n.s. | | .998 |
| Item 3 | .775 | .758 | | .715 | .304 | | .847 |
| Item 4 | .764 | .791 | | .766 | .211 | | .929 |
| Item 5 | .655 | .767 | | .709 | .583 | | .597 |
| Item 6 | .752 | .665 | | .653 | .116 | | .969 |
| Item 7 | .596 | .783 | | .802 | —.109 | | .982 |
| Item 8 | .798 | | .621 | .602 | | .082n.s. | .982 |
| Item 9 | .700 | | .798 | .767 | | .091 | .986 |
| Item 10 | .716 | | .729 | .689 | | .252 | .882 |
| Item 11 | .775 | | .721 | .675 | | .466 | .677 |

Notes: all factor loadings were statistically significant, except for those marked with n.s. (not statistically significant).

**Table 3**

Measurement invariance results for the one-factor model.

| | $X^2$ | df | $p$ | CFI | RMSEA | IC 90% RMSEA | SRMR | $\Delta X^2$ | Δdf | $p$ | ΔCFI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Women | 225.407 | 44 | < .001 | .986 | .081 | [.070 .091] | .032 | – | – | – | – |
| Men | 193.373 | 44 | < .001 | .979 | .089 | [.076 .102] | .036 | – | – | – | – |
| Configural model | 468.628 | 88 | < .001 | .975 | .090 | [.082 .098] | .034 | – | – | – | – |
| Metric model | 283.408 | 99 | < .001 | .988 | .059 | [.051 .067] | .041 | 26.227 | 11 | .006 | —.013 |
| Scalar model | 367.614 | 142 | < .001 | .985 | .055 | [.048 .062] | .043 | 102.680 | 43 | < .001 | —.010 |

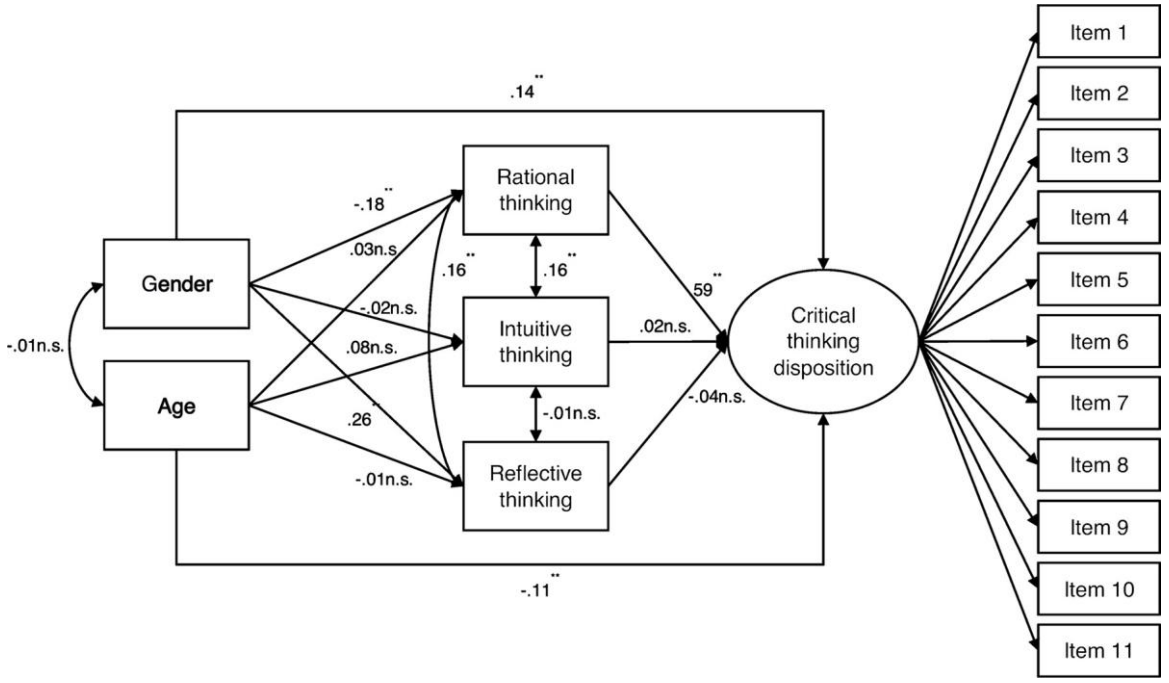


**Fig. 2.** Results of the MIMIC model.

Notes: All factor loadings were statistically significant ($p < .001$); * $p < .05$; ** $p < .01$; n.s. = not statistically significant.

any intervention has to focus on two things: on the development of skills and tools for critical thinking purposes and on the willingness to use these skills and tools because it may be through a lack of motivation or disposition that critical thinking is not used, not because the individual is not capable of it (Bensley & Lilienfeld, 2017; Ku, 2009; Ståhl and Van Prooijen, 2018; Taghinezhad et al., 2019).

Although most programs devoted to improving critical thinking outlined the necessity of promoting both components: skills and dispositions, the reality, as follows from a recent meta-analysis by Fong et al. (2017) of the literature on critical thinking and

community college student success, is that few studies measure dispositions and most of them measure only skills. The absence of suitable measures of critical thinking disposition may have contributed to this; one reason could be that, the validity of one of the most relevant scales to measure critical thinking disposition (the California Critical Thinking Dispositional Inventory) has been questionedand another may be that critical thinking disposition has been measured in several studies using scales that have not been specifically designed for that purpose. The main goal of this study has been to contribute to the development of accurate measures of this construct, in particular, through the in-depth study of the psychometric properties of the Spanish version of the CTDS, drawn up by Sosu (2013), working with a sample of young Spaniards.

As to the factorial structure of the CTDS, this study compared the two-dimension and one-dimension models proposed by Sosu (2013) and Yockey (2016), respectively, with a new bifactorial model, a comparison justified given the high degree of correlation between the two factors obtained in the previous studies. The three models fit adequately, only for the unidimensional model RMSEA performed poorly as expected in structural models with low degrees of freedom (Kenny et al., 2015). However, from our results, it canbe seen that it is the unidimensional model that provides the closest fit. It had adequate fit, it showed negligible differences in the CFI when compared to the two-factor model, and, although it showed differences in the CFI when compared to the bifactor model, evi- dence gathered with ECV pointed that the one-factor model was retained as the best structure with which to represent the data. This model implies the existence of a single general critical thinking disposition factor, capable of explaining most of the variance of every item on the scale.

These results contradict those obtained by Sosu (2013), the original author of the scale and defender of bidimensionality, although his study never tested the solution of a single factor. The results support those obtained by Yockey (2016), who found evidence in favor of the unidimensionality of the scale. When interpreting these results, the differences between the samples used should be considered,as the study by Sosu (2013) used a sample from the UK and that by Yockey (2016) worked with one from the USA, of which 54 % were Latin American in origin. As regards the estimations of reliability obtained, the values offered up by the current study are suitable and similar to those obtained in previous studies, with Cronbach's alphas ranging between .78 and .81 (Akin et al., 2015; Sosu, 2013; Yockey, 2016).

Once the best model had been determined, its measurement invariance was assessed across gender groups, which is key when the objective is to make inter-group comparisons (Kline, 2015; Millsap & Olivera-Aguilar, 2012). Our results offer evidence of said measurement invariance across gender, guaranteeing the correct use of the questionnaire as a means of assessing the differences between critical thinking disposition in men and women. The comparison between latent means in function of gender showed a higher average in the women's group than in the men's. The findings in this Spanish version support those by Yüksel and Alci (2012), who, although using the "*California Critical Thinking Dispositions Inventory*" (Facione & Facione, 1992) to measure critical thinking, found significant differences that favored women. Other previous studies, however, have found no gender-based differences (e.g. Hunter et al., 2014; Liu et al. 2019; Salahshoor & Rafiee, 2016), although this could be due to the use of other instruments to measure critical thinking or to the characteristics of the samples used. Liu et al. (2019), for example, used a questionnaire that measures intention and critical thinking skill in a sample of student nurses. Also studying a sample of student nurses, Hunter et al. (2014) makes reference to critical thinking skill and not to disposition. Salahshoor and Rafiee (2016) measure critical thinking in general, without differentiating between skill and disposition in a heterogeneous sample of language students of between 16 and 45 years old.

Finally, the relationship between the scores in critical thinking disposition and other cognitive variables was studied. The results supported the hypothesis about the relationship between the critical thinking disposition and analytical or rational thinking, as we found a correlation between CTDS and rational thinking of .50 ($p < .001$). The value of this correlation is similar to the one obtained by Yockey (2016) between CTDS and the Need for Cognition scale (Cacioppo et al., 1984) ($r$ .40; $p < .001$). It should be noted that although the rational thinking scale from the REI used in this study is not exactly the same as the Need for Cognition scale used by Yockey (2016), they are very similar, given that both are shorter versions of the original Need for Cognition scale (Cacioppo & Petty, 1982). Moreover, the relationships between CTDS and intuitive thinking was, as hypothesized, not statistically significant. To sum up, these results support the statements that, in the context of the dual process theories of thinking, equate critical thinking with system 2 rather than system 1 thinking (e.g. Bonnefon, 2016; Halpern, 2014; Stanovich, 2016).

Regarding the relationship between CTDS and reflective thinking (CRT), the correlation between the observed total scores was statistically significant but of low magnitude ($r$ .13, $p$ .008). This correlation is similar in magnitude to the ones found in Stuppleet al. (2017) between the CRT and the three factors derived from The Critical Thinking Toolkit (CriTT) they developed to measure beliefs and attitudes about critical thinking ($r$ = .05, .10 and .10). These three correlations were not statistically significant (note that the sample size in Supple et al.'s study was a great deal inferior to the one in this study). However, when including this variable as a predictor of CTD in a MIMIC model, together with gender, age, rational and intuitive thinking, the effect becomes statistically non-significant, and the biggest effect is for rational thinking ($\beta$ .59, $p < .01$). It should be noted that the CRT and the rational thinking scale share an amount of variance, given that the correlation between both measures is ∓ .25 ($p < .001$). In conclusion, although CRT has been used as a means of measuring system 2 thinking, as discussed by Stupple et al. (2017), this scale reflects an aspect of cognitive ability that is different from the dispositions, beliefs and attitudes to critical thinking measured in the CTDS or the CriTT.

This paper analyses the relationship between CTD on one hand, and System 1 and Systems 2 thinking styles on the other hand. Recently Houdé (2019) and Evans (2019) points out the relevance of a System 3. As pointed out by Houdé (2019) 'When these two systems compete (System 1 vs System 2), our brain needs a third system…for inhibiting the too fast heuristic system and activating the logical one' (p. 107). In this sense, System 3 is the inhibitory control system and metacognitive control. Future research is needed to clarified the role of CTD in the framework of the 3-System theory.

Regarding the role of demographic variables, we found the relationship between the age and the critical thinking disposition to be

low (-.11) but statistically significant. This result must be evaluated considering the low variability of the age in the sample. As far as we know, there are not previous works about the relationship between the age and the Critical thinking disposition. However, the works by Hunter et al. (2014) and Lee et al. (2020) found, in samples from students and nursing professionals, that the age was not related to the Critical thinking skills. As regards gender, and in line with the results gathered in the latent means comparisons, women showed higher levels of critical thinking when compared to men ($\beta = .14$, $p < .01$). These results, as mentioned before, agree with some of the previous literature (Yüksel & Alci, 2012).

This study has several limitations that should be considered in future research. First, this was a convenience sample including only pre-university and university students. In future studies, the CTDS should be validated using random sampling with wider age groups, such as adolescents and adults, in which to test for measurement invariance across educational levels.

## 5. Conclusion

This study enables progress to be made as regards the correct measurement of critical thinking disposition. The results lead to the conclusion that the scale developed by Sosu (2013) is unidimensional in structure, invariable across gender, and has suitable reliability estimates. In the extended analysis on nomological validity, carried out only on university students, a close relationship with the rational style of thinking is found. It has suitable psychometric properties that guarantee it may be used correctly on girls and boys, both pre-university and of university age. The CTDS can be used in future studies that assess the influence of critical thinking disposition on different aspects in both the academic and social spheres, but particularly in the field of education. The objective would be twofold: to identify students who should enhance their critical thinking disposition, and to carry out a thorough assessment of the intervention programs designed to strengthen the skills and dispositions to think critically.

## Informed consent

Informed consent was obtained from all individual participants included in the study.


## Funding

Research Project Liad@s: Prosociality and prevention of gender violence in students.
Award: GV/2017/208 with ending in 31/12/2019, Generalitat Valenciana (ES).


## CRediT authorship contribution statement

**Maria Jesús Bravo:** Conceptualization, Writing - review & editing. **Laura Galiana:** Methodology, Data curation, Formal analysis, Writing - review & editing. **Maria F. Rodrigo:** Conceptualization, Writing - review & editing. **José J. Navarro-Pérez:** Supervision, Writing - review & editing. **Amparo Oliver:** Conceptualization, Methodology, Writing - review & editing.

## Declaration of Competing Interest

The authors report no declarations of interest.

## Appendix A. Spanish translation of the CTDS

Marca con un círculo tu grado de acuerdo o desacuerdo con cada una de las siguientes afirmaciones empleando las siguientes opciones de respuesta:

| 1 Muy en desacuerdo | 2 En desacuerdo | 3 Indeciso/a | 4 De acuerdo | | | 5 Muy de acuerdo |
|---|---|---|---|---|---|---|
| 1. Generalmente trato de pensar en el panorama general durante una discusión | | | 1 | 2 | 3 | 4 5 |
| 2. A menudo utilizo nuevas ideas para modificar la forma en que hago las cosas | | | 1 | 2 | 3 | 4 5 |
| 3. Utilizo más de una fuente para encontrar la información que necesito | | | 1 | 2 | 3 | 4 5 |
| 4. Con frecuencia estoy a la búsqueda de nuevas ideas | | | 1 | 2 | 3 | 4 5 |
| 5. A veces encuentro un buen argumento que desafía algunas de mis creencias firmemente sostenidas | | | 1 | 2 | 3 | 4 5 |
| 6. Es importante entender el punto de vista de otras personas sobre un tema | | | 1 | 2 | 3 | 4 5 |
| 7. Es importante que justifique las elecciones que realizo | | | 1 | 2 | 3 | 4 5 |
| 8. A menudo reevalúo mis experiencias para poder aprender de ellas | | | 1 | 2 | 3 | 4 5 |
| 9. Por lo general, suelo comprobar la credibilidad de la fuente de información antes de hacer juicios | | | 1 | 2 | 3 | 4 5 |
| 10. Por lo general, pienso en las implicaciones más amplias de una decisión antes de llevarla a cabo | | | 1 | 2 | 3 | 4 5 |
| 11. A menudo pienso sobre mis acciones para ver si puedo mejorarlas | | | 1 | 2 | 3 | 4 5 |

## Appendix B. Spanish translation of the REI (Witteman et al, 2009). Thanks to Prof. Antonio Godoy Ávila, University of Málaga, Spain

Marca con un círculo tu grado de acuerdo o desacuerdo con cada una de las siguientes afirmaciones empleando las siguientes opciones de respuesta:

| 1 Muy en desacuerdo | 2 En desacuerdo | 3 Indeciso/a | 4 De acuerdo | 5 Muy de acuerdo |
|---|---|---|---|---|
| 1. Intento evitar las situaciones que requieren pensar mucho sobre algo | 1 | 2 | 3 | 4 | 5 |
| 2. Me gusta confiar en mis impresiones intuitivas | 1 | 2 | 3 | 4 | 5 |
| 3. Al resolver problemas en mi vida, normalmente me va bien cuando le hago caso a mis impulsos | 1 | 2 | 3 | 4 | 5 |
| 4. No soy muy bueno resolviendo problemas complicados | 1 | 2 | 3 | 4 | 5 |
| 5. Me gustan los retos intelectuales | 1 | 2 | 3 | 4 | 5 |
| 6. Cuando hay que confiar en la gente, normalmente me fío de mis impulsos | 1 | 2 | 3 | 4 | 5 |
| 7. No soy muy bueno resolviendo problemas que requieren una análisis lógico cuidadoso | 1 | 2 | 3 | 4 | 5 |
| 8. Confío en mis corazonadas | 1 | 2 | 3 | 4 | 5 |
| 9. La intuición puede ser un medio muy útil para solucionar los problemas | 1 | 2 | 3 | 4 | 5 |
| 10. Con frecuencia me dejo llevar por mi instinto al decidir un curso de acción | 1 | 2 | 3 | 4 | 5 |
| 11. No me gusta tener que reflexionar prolongadamente | 1 | 2 | 3 | 4 | 5 |
| 12. Normalmente tengo razones claras y explicables para mis decisiones | 1 | 2 | 3 | 4 | 5 |
| 13. Confío en mis primeras impresiones acerca de la gente | 1 | 2 | 3 | 4 | 5 |
| 14. No tengo una intuición muy buena | 1 | 2 | 3 | 4 | 5 |
| 15. No creo que reflexionar sea una actividad divertida | 1 | 2 | 3 | 4 | 5 |
| 16. Si confiara en mis impulsos, con frecuencia cometería errores | 1 | 2 | 3 | 4 | 5 |
| 17. Me es muy atractivo aprender nuevas formas de pensar | 1 | 2 | 3 | 4 | 5 |
| 18. Los razonamientos profundos no son uno de mis puntos fuertes | 1 | 2 | 3 | 4 | 5 |
| 19. Prefiero los problemas complejos a los simples | 1 | 2 | 3 | 4 | 5 |
| 20. No me gustan las situaciones en las que he de confiar en mi intuición | 1 | 2 | 3 | 4 | 5 |
| 21. Pensar mucho y durante mucho tiempo sobre algo me produce poca satisfacción | 1 | 2 | 3 | 4 | 5 |
| 22. Para las decisiones importantes, no creo que sea buena idea confiar en la propia intuición | 1 | 2 | 3 | 4 | 5 |
| 23. Creo que hay ocasiones en las que uno debe confiar en su propia intuición | 1 | 2 | 3 | 4 | 5 |
| 24. No puedo reflexionar bajo presión | 1 | 2 | 3 | 4 | 5 |
| 25. Resolviendo cosas lógicamente soy bastante mejor que la mayoría de la gente | 1 | 2 | 3 | 4 | 5 |
| 26. No me gustaría depender de alguien que se describe a sí mismo como intuitivo | 1 | 2 | 3 | 4 | 5 |
| 27. No tengo problemas para pensar las cosas con detenimiento | 1 | 2 | 3 | 4 | 5 |
| 28. Creo que es una locura tomar decisiones importantes basándose en impresiones | 1 | 2 | 3 | 4 | 5 |
| 29. Cuando hago juicios rápidos, probablemente no soy tan bueno como la mayoría de la gente | 1 | 2 | 3 | 4 | 5 |
| 30. Tengo una mente lógica | 1 | 2 | 3 | 4 | 5 |
| 31. Tiendo a utilizar el corazón como guía de mis acciones | 1 | 2 | 3 | 4 | 5 |
| 32. Me gusta pensar en abstracto | 1 | 2 | 3 | 4 | 5 |
| 33. Con frecuencia me doy cuenta cuando alguien acierta o se equivoca, incluso cuando no puedo explicar cómo llego a saberlo | 1 | 2 | 3 | 4 | 5 |
| 34. La utilización de la lógica es algo que me funciona al solucionar problemas de mi vida | 1 | 2 | 3 | 4 | 5 |
| 35. Normalmente no utilizo las corazonadas para ayudarme a tomar decisiones | 1 | 2 | 3 | 4 | 5 |
| 36. Cuando le hago caso a mis impulsos, pocas veces me equivoco de respuesta | 1 | 2 | 3 | 4 | 5 |
| 37. Me basta con conocer la respuesta, aunque no conozca los razonamientos en que dicha respuesta se basa | 1 | 2 | 3 | 4 | 5 |
| 38. Me gustan los problemas que requieren pensar mucho | 1 | 2 | 3 | 4 | 5 |
| 39. Supongo que mis corazonadas aciertan tanto como se equivocan | 1 | 2 | 3 | 4 | 5 |
| 40. No soy un pensador muy analítico | 1 | 2 | 3 | 4 | 5 |

## Appendix C. Spanish translation of the CRT (Brañas-Garza et al, 2019). Thanks to Prof. Pablo Brañas-Garza at Loyola Andalucia University, Córdoba, Spain.

¿Cuál es la respuesta correcta en cada una de las siguientes preguntas? Pon tu respuesta en el espacio libre

1. Un bate y una pelota cuestan 1.10€ en total. El bate cuesta 1.00€ más que la pelota. ¿Cuánto cuesta la pelota? ____ céntimos

2. Si 5 máquinas tardan 5 minutos en hacer 5 objetos, ¿cuánto tiempo tardarán 100 máquinas en hacer 100 objetos? ____ minutos

3. En un lago, hay un manto de nenúfares. Cada día, el manto duplica su tamaño. Si se tarda 48 días para que el manto cubra todo el lago, ¿cuánto tiempo se tardará en cubrir la mitad del lago? ____ días

## References


Aarnio, K., & Lindeman, M. (2005). Paranormal beliefs, education, and thinking styles. *Personality and Individual Differences*, *39*(7), 1227–1236. https://doi.org/10.1016/j.paid.2005.04.009.

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*, *85*(2), 275–314. https://doi.org/10.3102/0034654314551063.

Akin, A., Hamedoglu, M. A., Arslan, S., Akin, U., Celik, E., Kaya, C., & Arslan, N. (2015). The Adaptation and Validation of the Turkish Version of the Critical Thinking Disposition Scale (CTDS). *International Journal of Educational Researchers*, *6*(1), 31–35. https://doi.org/10.18656/jee.91401.

Bagozzi, R. P. (1983). "Issues in the application of covariance structure analysis": A further comment. *The Journal of Consumer Research*, *9*(4), 449–450. https://doi.org/10.1086/208939.

Bensley, D. A., & Lilienfeld, S. O. (2017). Psychological misconceptions: Recent scientific advances and unresolved issues. *Current Directions in Psychological Science*, *26* (4), 377–382. https://doi.org/10.1177/0963721417699026.

Bensley, D. A., Lilienfeld, S. O., & Powell, L. A. (2014). A new measure of psychological misconceptions: Relations with academic background, critical thinking, and acceptance of paranormal and pseudoscientific claims. *Learning and Individual Differences*, *36*, 9–18. https://doi.org/10.1016/j.lindif.2014.07.009.

Bonnefon, J. F. (2016). The pros and cons of identifying critical thinking with system 2 processing. *Topoi*, 1–7. https://doi.org/10.1007/s11245-016-9375-2.

Brañas-Garza, P., Kujal, P., & Lenkei, B. (2019). Cognitive reflection test: Whom, how, when. *Journal of Behavioral and Experimental Economics*, *82*, Article 101455. https://doi.org/10.1016/j.socec.2015.09.002.

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, *8*, 108–117. https://doi.org/10.1016/j.jarmac.2018.09.005.

Butler, H. A., & Halpern, D. F. (2020). Critical thinking impacts our everyday lives. In R. J. Sternberg, & D. F. Halpern (Eds.), *Critical thinking in psychology (chapter 7)* (2nded., pp. 152–172). Cambridge Univerity Press. https://doi.org/10.1017/9781108684354.008.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13.

Chan, Z. C. (2016). A systematic review on critical thinking in medical education. *International Journal of Adolescent Medicine and Health*, *30*(1). https://doi.org/10.1515/ijamh-2015-0117.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/s15328007sem0902_5.

Davies, M. (2015). A model of critical thinking in higher education. In M. Paulsen (Ed.), *Higher education: Handbook of theory and research* (vol. 30). Springer. https://doi.org/10.1007/978-3-319-12835-1_2.

Diedenhofen, B. (2016). *Package 'cocron', version 1.0-1*. Available at http://comparingcronbachalphas.org.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*, 43–52. https://doi.org/10.1016/j.tsc.2013.12.004.

Dyer, K. D., & Hall, R. E. (2019). Effect of critical thinking education on epistemically unwarranted beliefs in college students. *Research in Higher Education*, *60*(3), 293–314. https://doi.org/10.1007/s11162-018-9513-3.

Ennis, R. H. (1987). A taxonomy of critical thinking abilities and dispositions. In J. Baron, & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). W.H. Freeman.

Ennis, R. H. (2018). Critical thinking across the curriculum: A vision. *Topoi*, *37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4.

Epstein, S. (1998). Cognitive-experiential self-theory. In D. F. Barone, & M. Herser (Eds.), *Advanced personality*. Plenum Press.

Evans, J. S. B. (2019). Reflections on reflection: The nature and function of type 2 processes in dual-process theories of reasoning. *Thinking & Reasoning*, *25*(4), 383–415. https://doi.org/10.1080/13546783.2019.1623071.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685.

Facione, P. A., & Facione, N. C. (1992). *California critical thinking disposition inventory*. California: Academic Press.

Fasko, D. (2003). Critical thinking: Origins, historical development, future direction. *Critical thinking and reasoning: Current research, theory and practice* (pp. 3–20). Hampton: Press.

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course*. Greenwich, CT: Information Age.

Fong, C. J., Kim, Y., Davis, C. W., Hoang, T., & Kim, Y. W. (2017). A meta-analysis on critical thinking and community college student achievement. *Thinking Skills and Creativity*, *26*, 71–83. https://doi.org/10.1016/j.tsc.2017.06.002.

Fornell, C. (1983). Issues in the application of covariance structure analysis: A comment. *The Journal of Consumer Research*, *9*(4), 443–448.

Franco, A. (2016). What do Ode to Joy, the Nobel peace prize, umbrellas and cartoons have in common? Why critical thinking matters and how higher education moduls. *Higher Education for the Future*, *3*(1), 108–124. https://doi.org/10.1177/2347631115610231.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732.

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

Houdé, O. (2019). *3-system theory of the cognitive brain: A post-piagetian approach to cognitive development*. Routledge.

Hu, L. T., & Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Hunter, S., Pitt, V., Croce, N., & Roche, J. (2014). Critical thinking skills of undergraduate nursing students: Description and demographic predictors. *Nurse Education Today*, *34*(5), 809–814. https://doi.org/10.1016/j.nedt.2013.08.005.

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

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. NY: Guilford Press.

Ku, K. Y. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*, *4*(1), 70–76. https://doi.org/10.1016/j.tsc.2009.02.001.

Lee, D. S., Abdullah, K. L., Chinna, K., Subramanian, P., & Bachmann, R. T. (2020). Critical thinking skills of RNs: Exploring demographic determinants. *Journal of Continuing Education in Nursing*, *51*(3), 109–117.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of crosscultural data. Practical and theoretical issues. *Multivariate Behavioral Research*, *32*, 53–76. https://doi.org/10.1207/s15327906mbr3201_3.

Liu, N. Y., Hsu, W. Y., Hung, C. A., Wu, P. L., & Pai, H. C. (2019). The effect of gender role orientation on student nurses' caring behaviour and critical thinking. *International Journal of Nursing Studies*, *89*, 18–23. https://doi.org/10.1016/j.ijnurstu.2018.09.005.

Lobato, E., Mendoza, J., Sims, V., & Chin, M. (2014). Examining the relationship between conspiracy theories, paranormal beliefs, and pseudoscience acceptance among a university population. *Applied Cognitive Psychology*, *28*(5), 617–625. https://doi.org/10.1002/acp.3042.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling*. The Guilford Press.

Moore, T. J. (2013). Critical thinking: Seven definitions in search of a concept. *Studies in Higher Education*, *38*, 506–522. https://doi.org/10.1080/03075079.2011.586995.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*, 2017 (eighth edition). Los Angeles, CA: Muthén & Muthén.

Noone, C., & Hogan, M. J. (2018). Improvements in critical thinking performance following mindfulness meditation depend on thinking dispositions. *Mindfulness*, *9* (2), 461–473. https://doi.org/10.1007/s12671-017-0789-8.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972–987. https://doi.org/10.1037/0022-3514.76.6.972.

Peñarroja, V., Serrano, M. A., Gracia, E., Alacreu-Crespo, A., González, P., & Martínez-Tur, V. (2017). Rational-experiential thinking style and rational intergroup cooperation: the moderating role of intergroup conflict/Estilos de pensamiento racional-experiencial y la cooperación intergrupal racional: el rol modulador del conflicto intergrupal. *Revista de Psicologia Social*, *32*(1), 23–51. https://doi.org/10.1080/02134748.2016.1248028.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*, 544–559. https://doi.org/10.1080/00223891.2010.496477.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045.

RStudio Team. (2020). *RStudio: Integrated development for r*. Boston, MA: RStudio, PBC. http://www.rstudio.com/.

Salahshoor, N., & Rafiee, M. (2016). The relationship between critical thinking and gender: A case of Iranian EFL learners. *Journal of Applied Linguistics and Language Research*, *3*(2), 117–123.

Sánchez, E., Fernández-Berrocal, P., Alonso, D., & Tubau, E. (2012). Measuring both systems of reasoning: A study of the predictive capacity of a new version of the Rational-Experiential Inventory. *European Journal of Education and Psychology*, *5*(2), 121–132.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107–120. https://doi.org/10.1007/s11336-008-9101-0.

Sosu, E. M. (2013). The development and psychometric validation of a Critical Thinking Disposition Scale. *Thinking Skills and Creativity*, *9*, 107–119. https://doi.org/10.1016/j.tsc.2012.09.002.

Ståhl, T., & Van Prooijen, J. W. (2018). Epistemic rationality: Skepticism toward unfounded beliefs requires sufficient cognitive ability and motivation to be rational. *Personality and Individual Differences*, *122*, 155–163. https://doi.org/10.1016/j.paid.2017.10.026.

Stanovich, K. E. (2016). The Comprehensive Assessment of Rational Thinking. *Educational Psychologist*, *51*, 23–34. https://doi.org/10.1080/00461520.2015.1125787.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*(2), 342–357. https://doi.org/10.1037/0022-0663.89.2.342.

Stucky, B. D., & Edelen, M. O. (2014). Using hierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise, & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (pp. 183–206). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781315736013.

Stucky, B. D., Thissen, D., & Edelen, M. O. (2013). Using logistic approximations of marginal trace lines to develop short assessments. *Applied Psychological Measurement*, *37*, 41–57. https://doi.org/10.1177/0146621612462759.

Stupple, E. J., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y., & Aubeeluck, A. V. (2017). Development of the Critical Thinking Toolkit (CriTT): A measure of student attitudes and beliefs about critical thinking. *Thinking Skills and Creativity*, *23*, 91–100. https://doi.org/10.1016/j.tsc.2016.11.007.

Taghinezhad, A., Riasati, M. J., & Behjat, F. (2019). The effect of teaching critical thinking strategies on students' academic writing, critical thinking ability, and critical thinking dispositions. *International Journal of Foreign Language Teaching and Research*, *7*(28), 37–55. https://doi.org/10.7551/mitpress/12081.003.0005.

Ten Berge, J. M., & Socan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*, 613–625. https://doi.org/10.1007/bf02289858.

Tiruneh, D. T., De Cock, M., & Elen, J. (2018). Designing learning environments for critical thinking: Examining effective instructional approaches. *International Journal of Science and Mathematics Education*, *16*(6), 1065–1089. https://doi.org/10.1007/s10763-017-9829-z.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*(4), 1037–1048. https://doi.org/10.1037/a0034910.

Van Bavel, J., Baicker, K., Boggio, P., Capraro, V., Cichocka, A., Crockett, M., … Willer, R. (2020). *Using social and behavioural science to support COVID-19 pandemic response [Preprint]*. PsyArXiv. https://doi.org/10.31234/osf.io/y38m9.

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *The European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*(4), 930. https://doi.org/10.1037/a0012842.

Wilson, J. A. (2018). Reducing pseudoscientific and paranormal beliefs in university students through a course in science and critical thinking. *Science & Education*, *27*(1-2), 183–210. https://doi.org/10.1007/s11191-018-9956-0.

Witteman, C., van den Bercken, J., Claes, L., & Godoy, A. (2009). Assessing rational and intuitive thinking styles. *European Journal of Psychological Assessment*, *25*(1), 39–47. https://doi.org/10.1027/1015-5759.25.1.39.

Yockey, R. D. (2016). Validation study of the critical thinking dispositions scale: A brief report. *North American Journal of Psychology*, *18*(1), 101–106.

Yüksel, G., & Alcı, B. (2012). Self-efficacy and critical thinking dispositions as predictors of success in school practicum. *International Online Journal of Educational Sciences*, *4*(1).