



# Revising the core knowledge confusions scale: a measure of logical error associated with cognitive and personality traits

Brenton M. Williams<sup>1</sup> · Matthew Browne<sup>1</sup> · Matthew Rockloff<sup>1</sup> · George Stuart<sup>2</sup>

Accepted: 2 January 2024 / Published online: 16 January 2024  
© The Author(s) 2024

## Abstract

The Core Knowledge Confusions scale (CKC) was designed to predict paranormal and spiritual beliefs, alternative health beliefs and the degree to which people assign meaning to events. It measures the likelihood of individuals' tendency to accept ontologically confused content as literally true and has been used to investigate beliefs such as the paranormal. However, the psychometric properties of the CKC have not been tested, limiting the scale's practical utility. The CKC was revised and tested, internal consistency was checked, and associations to paranormal belief were assessed using a panel survey of 1010 Australian residents ( $M_{\text{age}} = 52$ ,  $SD = 17.69$ , Female = 56%). Using structural equation modelling and regression, the revised scale (CKC-R) content deviated from previous findings. However, it demonstrated acceptable internal, construct, and divergent validity. External correlates of the CKC-R were aligned with expectations: associating with high verbal knowledge, an intuitive cognitive style, and the absorption personality trait (i.e., a tendency to experience altered states of consciousness). The CKC-R provides researchers with a validated measure predicting paranormal belief that is associated with both cognitive and personality-based traits. Interpretation of the CKC-R as a measure of ontological error is less clear and requires further investigation.

**Keywords** Ontological confusion · Logical error · Absorption · Verbal knowledge · Beliefs

## Introduction

The Core Knowledge Confusions scale (CKC; Lindeman & Aarnio, 2007) was designed to test the assertion that paranormal beliefs could be defined by ontological mistakes. It was assumed that people understand the world in sets of hierarchical ontologies: trees are living inanimates with a certain set of properties, and these differ from houses which are inanimate objects. From this, it is assumed that our ontological understanding results in beliefs about what is and is not possible and leads people to be more or less likely to believe in things like the paranormal (e.g. ghosts). The CKC claimed to measure ontological errors using metaphor classification tasks (Lindeman & Aarnio, 2007). People are asked to classify items such as *a house knows its history* as either

better understood as literal or metaphoric. As houses cannot know, classifying the item as literal is interpreted as an ontological error. The CKC was able to significantly explain variation in paranormal beliefs (Lindeman & Aarnio, 2007) and has since been used to explain others such as religiosity (Lindeman et al., 2015), and supernatural attributions (Barber, 2014; Svedholm et al., 2010). However, the underlying assumptions of the CKC are not necessarily sound and further investigation of the scale and its properties would provide more evidence about its relationship to ontologies (Lindeman & Aarnio, 2007), or as more recently claimed cognitive bias (Lindeman et al., 2022). In the present study, we aim to revise and test the psychometric properties of the CKC so that further work can be done to evaluate the scale and its usefulness in predicting beliefs.

The CKC targets a very specific set of ontological errors, those that are core knowledge-related. Core knowledge is gained through development alone, intuitively, without the need for formal education and is defined by its commonality across cultures (children learn the same things across different countries) and across time (people exhibit the same knowledge during the 1800's and 1900's for example)

✉ Brenton M. Williams  
b.williams@cqu.edu.au

<sup>1</sup> Health & Applied Sciences, CQUniversity School of Medical, University Drive, Branyan, QLD 4670, Australia

<sup>2</sup> School of Psychology, Charles Sturt University, Panorama Avenue, Bathurst, NSW 2795, Australia

(Carey, 2000; Carey & Spelke, 1996). Knowing that a toy bear does not feel thirsty is something children will learn without instruction, so an understanding that inanimate objects do not have the properties of animates can be considered core knowledge. However, understanding the scientific properties of light requires formal education and cannot be considered core knowledge. The CKC identifies six core knowledge domains (such as mental states having physical properties) and uses them as subscales of the general core knowledge confusion trait (Lindeman & Aarnio, 2007). While these have been defined, it is unknown if these subscales have any practical implications for the CKC and associated beliefs.

Measurement of core knowledge (or ontological) confusions is achieved by the CKC using predicate metaphors to obscure the targeted error (Lindeman & Aarnio, 2007). Participants are not asked to categorise *rocks' live* as literal or metaphorical, but *rocks live a long time*. The target measure, the association between *living* and a *rock* is obscured through the inclusion of time which introduces the possibility that *live* might be understood not as living, but rather to imply a long existence. To be clear, the CKC doesn't claim to measure explicit ontological beliefs: when asked explicitly participants do not believe in the ontological mistakes they have endorsed: they do not believe that houses can know (Lindeman et al., 2008, 2011). However, this raises the question of what aspect of individual characteristics the scale is measuring through its classification tasks.

The use of metaphors in communication is not uncommon, they are almost unavoidable (Geary, 2011), yet assumptions made about how they are understood and the language implications are not without debate. The CKC is based on the assumption that we store and understand information in ontologies, and a literal/metaphorical distinction is representative of those ontological categorisations (Lindeman & Aarnio, 2007). In this sense, the CKC adheres to a view taken by cognitive metaphor theory (Lakoff & Johnson, 1980) that metaphors are grounded in primary experiences. That is, we understand metaphors because they use experiential language, grounded in a base-level perception such as movement or sight. *Love burns like the sun* can be understood as we physically understand what burning is or might be like. However, multiple criticisms have been raised about these views including whether or not meaning is transferred via grounding in a primary experience (for a review see: Madsen, 2016). Sound and colour metaphors can easily be found to reference the others' domain which creates issues for determining primacy (Williams, 1976). For example colour (or sight) is sound: *it was a loud shirt*; sound is colour: *it was a bright sound*. Understanding one is not easily explainable by its use of knowledge from the other domain as primary and assumptions of cognitive metaphor theory become less stringently applicable: if we don't need

a primary domain to understand metaphor it implies that the metaphor / literal distinction may not reflect how people understand language (Madsen, 2016). While the CKC has been acknowledged not to measure ontologies (Lindeman et al., 2008, 2011), it was designed assuming their use and necessity for making a literal/metaphoric categorisation (Lindeman & Aarnio, 2007) leaving it open to criticisms about these assumptions.

Regardless of theoretical criticisms, allowing for the definition of core knowledge confusions used by the CKC as externally derived (and not necessarily representative of cognitive organisation), classifying metaphors containing core knowledge confusions as literal has been shown to predict multiple beliefs. It is not the aim of the present study to evaluate and critique these assumptions (for a broad critique see: Madsen, 2016) but to generate a valid replication of the CKC in English that can be used to explore its properties further. Knowing more about the properties of the CKC is of interest in understanding what individual differences might lead some to believe in logically questionable or faith-based beliefs and not others. The current barrier remains an appropriate verification of the CKC in Australia. Alterations to the original Finnish CKC were made in a North American study (Barber, 2014) where the wording of some items was changed to make better sense in English. Unfortunately, this implementation did not fully replicate the statistical structure found in the original studies. For example, several items clustered with non-related subscales and the structural equation model produced only a marginal fit (Barber, 2014). Nevertheless, the altered CKC used by Barber predicted significant variation in events having purpose and paranormal beliefs in both men (78%) and women (57%; Barber, 2014). These results are consistent with the original Finnish work, Svedholm et al. (2010) for example, found 49% of variation in paranormal belief was predicted by the CKC. Barber's results support cross-cultural use of the CKC outside of Finnish studies, and the scale more generally appears to have promise for assessing an underlying cognitive trait with explanatory power for a range of real-world folk beliefs. In the present study, we seek to check the psychometric properties of a revised English-language CKC scale and confirm associations with relevant external measures in a cross-sectional survey. In doing so, we also hope to gain better insight into what the CKC is measuring, and whether the construct merits further study.

To inform expectations about the desirable properties of a revised scale, we turn to previous studies. Verbal intelligence has been associated with CKC outcomes in research on susceptibility to 'pseudo-profound bullshit'; i.e., phrases that have the structure of insightful commentary, but no meaning. Pennycook et al. (2015) included an English translation of the CKC as well as measures of intelligence. Significant correlations were reported for both verbal intelligence and the

CKC with paranormal beliefs (verbal intelligence:  $r = -0.26$ ,  $p < 0.01$ ; CKC:  $r = 0.38$ ,  $p < 0.01$ ;  $N = 187$ ) and between verbal intelligence and the CKC ( $r = -0.33$ ,  $p < 0.001$ ). However, since multivariate analyses were not conducted it is unclear whether the relationship to paranormal beliefs holds after controlling for verbal intelligence. As verbal knowledge influences metaphor understanding (Carriedo et al., 2016; Prat & Just, 2011; Stamenković et al., 2019, 2020), CKC outcomes may reflect differences in verbal intelligence rather than a separate trait related to ontological reasoning. That is, errors on the CKC may reflect a poor understanding of written statements. If the CKC scale offers little or no additional predictive value when predicting paranormal beliefs over and above verbal intelligence measures, this would tend to discredit the utility of CKC as a meaningful independent trait.

Personality traits provide another source of influence on cognitive biases (Ahmad, 2020; Kumar et al., 2021) potentially leading to ontological confusion. One relevant trait is absorption which is a measure of an individual's imagination, affective states, and object relations (Tellegen, 1981). It was initially designed as a measure predicting an individual's level of hypnotisability and was interpreted as "a capacity for absorbed and self-altering attention" (Tellegen & Atkinson, 1974, p.276). Being high in absorption may also result in individuals attributing personal characteristics to the object of attention (Tellegen & Atkinson, 1974) and lends itself to interpreting the world in "unconventional or idiosyncratic" ways (Tellegen & Atkinson, 1974, p.275). Accordingly, when assessing statements such as *a house knows its history*, individuals high in absorption may attribute their personal characteristic (knowing) to the object (a house), resulting in a literal interpretation of the metaphoric statement.

Absorption has been associated with a wide range of effects ranging from religious belief (Luhmann et al., 2010), supernatural experiences (Lifshitz et al., 2019; Luhmann et al., 2021; Maji & van Elk, 2018; Wilson & Barber, 1982), decreases in agency and increases in automaticity (Bregman-Hai et al., 2020), disruptions in aspects of control (Barrett & Keil, 1996), a focus on negative emotions in music (Hall et al., 2016), synaesthesia (Chun & Hupé, 2016), and memory errors (Nichols & Loftus, 2019). Several outcomes involve ontologies and directly relate to the supernatural. For example, absorption correlates with experiencing God as a person ( $r = 0.66$ ; Luhmann et al., 2010), daily spiritual experiences ( $r = 0.34$ ) and hearing voices when alone ( $r = 0.43$ ,  $p < 0.001$ ; Luhmann et al., 2013). These correlations are similar in magnitude to those reported for the CKC on related measures: supernatural purpose ( $r = 0.42$ ,  $p < 0.05$ ; Lindeman et al., 2015) and religion ( $r = 0.35$ ,  $p < 0.05$ ; Lindeman et al., 2015). The common link to religious and spiritual beliefs further supports the position that absorption and ontological confusions are associated with

one another and might possibly indicate some degree of construct overlap.

Although reliability and validity tests of the CKC have been generally positive, it has not been established that the CKC provides additional value when evaluated in conjunction with alternative measures of related constructs. Anthropomorphism for example incorrectly assigns attributes to targets which is very similar to the measures used in the CKC. The CKC is built around the attribution of either life, psychological states, or movement to categories of inanimate objects (e.g. the sky hears the thunder), mental states (e.g. grief moves in the stomach), and 'force' which can be interpreted as something which, when applied can change the motion of an object (e.g. force aims to influence; Lindeman & Aarnio, 2007). These can largely be classed as anthropomorphic: *Trees aim to move upwards*, or *the earth wants water* for example. These ontological confusions are similar to questions from general anthropomorphism scales such as the Individual Differences in Anthropomorphism Questionnaire (IDAQ; Waytz et al., 2010). This asks for the various degrees to which different objects and animals hold attributes such as free will. Thus, at face value, at least some of the content of CKC overlaps with the construct of anthropomorphism, the degree of overlap requires further investigation.

Anthropomorphism was omitted in initial CKC research and subsequent work has not checked the potential for construct overlap although arguments have been made about their differences (see Lindeman et al., 2015). Anthropomorphism measures (such as the IDAQ) do not generally cover mental states or physical forces and their association with characteristics such as agency, whereas the CKC does (Lindeman et al., 2015). This difference is reflected in research findings: anthropomorphism does not specifically predict a belief in God but does predict more general paranormal beliefs (Willard & Norenzayan, 2013). Hence, Lindeman et al. (2015) claim that anthropomorphism is not necessarily a good predictor of *all* spiritual and paranormal beliefs, appears to be justified. Nevertheless, the CKC and anthropomorphism both measure a range of ontological errors and anthropomorphism may represent a specific subset of ontological confusion. Thus, it is worthwhile to test the benefit of the CKC whilst controlling for anthropomorphism, to establish its usefulness as a broader measure of faulty ontologies.

Lastly, the CKC has been associated with intuitive thinking (Lindeman & Aarnio, 2007) and negatively with analytical thinking (Pennycook et al., 2015). When people think analytically about their answers, it appears to decrease the likelihood of making an ontological error and categorising items as literal. This assertion is consistent with participants' explicit denial of belief in the ontological mistakes made when asked directly (Lindeman et al., 2008); people do not believe the mistakes that they make, but when they

are not thinking critically people who believe in paranormal beliefs are more likely to make ontological errors. Hence, a revision of the CKC should evaluate the relationship between the revised scale and analytical, or intuitive thinking preferences.

To reiterate some of the points above, the present study aimed to test and validate an English revision of the CKC (the CKC-R) in the context of existing related measures. Candidate CKC items were tested with the following hypotheses: CKC-R factor structure would adhere to the original hierarchical model reported in previous CKC research (see Barber, 2014; Lindeman & Aarnio, 2007; Svedholm et al., 2010). Criterion validity was tested by comparing the variance of paranormal belief explained by both the CKC-R and the CKC short form. Given the CKC-R is longer than the CKC short form, we evaluated whether this yields greater measurement precision, which was hypothesised to be reflected in better predictive ability in paranormal belief than the original CKC short form. Testing the divergent validity was performed by evaluating relationships to anthropomorphism, analytical thinking, verbal knowledge, absorption, and bullshit susceptibility. Analytical thinking and high verbal fluency were hypothesised to correlate negatively with the CKC-R. In contrast, absorption, bullshit receptivity, and anthropomorphism were hypothesised to correlate positively with the CKC-R. To evaluate the usefulness of the CKC-R, as an independent cognitive trait, it should explain unique variations in outcomes in multivariate regression predicting paranormal belief including measures of anthropomorphism, cognitive style, absorption, and verbal knowledge. While anthropomorphism is hypothesised to overlap with the CKC-R, the other traits are predicted to be influential in response to CKC-R items and provide insight into ontological mistakes.

## Methodology

### Participants

Participants were sourced by a commercial panel provider (Qualtrics) and received payment from the provider in the form of redeemable vouchers. Participation was restricted to anyone over 18 years old, residing in Australia. A total of 1618 participants completed the survey. Screening and quality checking for issues such as speeded responding and partial responses removed 608 participants. Due to the survey length and the necessity for imputing many responses, partial responses were not retained for analysis. All survey items were required to answer, hence there was no random missing data, rather abandoned sessions. While these respondents may constitute characteristically different respondents than those who completed the survey, it was

deemed unnecessary and impractical to impute responses for inclusion. The final sample of 1010 participants retained in the analysis consisted of 56% female, and 42% male, with a mean age of 52 ( $SD = 17.69$ ). Ethics approval was obtained from the Human Research Ethics committee of CQUniversity (0000022768).

### Materials

The CKC-R included 47 statements sourced from Barber (2014), supplemented with newly generated candidate items adhering to the format of the original CKC. Candidate items were generated using commonly referenced subjects (such as *sleep*). These were paired with appropriate target items (in this example *the weary*) and a believable, yet metaphorical bridge inserted to form a predicate metaphor that, if interpreted literally would qualify as a core knowledge confusion, e.g. *sleep welcomes the weary*. Eight potential items were created and are shown in the results section (Table 2). Candidate items replaced poor contributors to the statistical model reported by Barber (2014). Scale items are grouped to cover the domains originally identified as core knowledge-related confusions: artificial entities have attributes of animates, force has attributes of animates, inanimate organisms have attributes of animates, lifeless entities have attributes of animates, and mental states have physical attributes (Lindeman & Aarnio, 2007). Items classed as core-knowledge confusions ( $n = 34$ ) consisted of predicate metaphors such as *a house knows its history*. Participants were asked to rate on a Likert-type scale to what degree items could be interpreted literally (0 = metaphorical, 5 = literal), conforming to prior CKC use (Barber, 2014; Lindeman et al., 2015; Lindeman & Aarnio, 2007; Svedholm et al., 2010). Distractor items consisted of both literal statements (e.g. “a drawing in pencil can be erased”,  $n = 9$ ) and metaphors not classed by Lindeman as core knowledge confusions ( $n = 4$ ) such as “A good memory is a mine”.

The CKC short form was used to compare the performance of the best-performing predictors of paranormal belief from the original CKC (Lindeman et al., 2015) to the CKC-R. CKC scores were created by summing item scores with the results ranging from 0 – 14. Higher scores indicate a greater tendency to interpret metaphoric statements containing ontological errors as literal. Cronbach's alpha for the CKC short form has been reported as 0.85 (Lindeman et al., 2015), in this study it was 0.83.

The Paranormal and Supernatural Beliefs Scale (Dean et al., 2021) was used to measure paranormal beliefs. It was selected as the most appropriate measure to use as it represents a modern, up-to-date, and specific measure of paranormal and supernatural beliefs when compared to earlier scales such as the paranormal beliefs scale (Tobacyk, 2004; for full reporting see Dean et al., 2021). The scale



presents participants with 13 items, such as “it is possible to be reincarnated”. Participants indicate the extent to which they agree with the statements on a Likert-type scale ranging from 1 (strongly disagree) to 4 (strongly agree). Two items are reverse scored, and the total scale score is created by summing the items, with scores ranging from 13 to 52. Higher scores indicate greater belief in the paranormal. The original scale reported Cronbach’s alpha of 0.91 (Dean et al., 2021), in this study it was 0.87.

Verbal knowledge was measured using the ten-item version of the Wordsum Plus test (Cor et al., 2012). This scale has previously been used in similar research on the detection of pseudo-profound bullshit which also included the use of the CKC: both were shown to correlate (Pennycook et al., 2015). Participants are presented with a word in capitals and asked to select an alternative closest in meaning. For example, APPARITION was presented with “ghost, insurrection, apparent, farce, apparel, or don’t know”. Correct answers were allocated a score of one and items were summed to create the overall verbal knowledge score. Scores range between zero and 10, where higher scores indicate greater verbal knowledge. Cronbach’s alpha was previously reported to be 0.787 (Cor et al., 2012). Using our dataset, Cronbach’s alpha (0.80) indicated reasonable internal consistency.

Absorption was measured using the Modified Tellegen Absorption Scale (MODTAS; Jamieson, 2005) and has been shown to associate with similar religious and paranormal type beliefs as the CKC (see introduction for expanded discussion). The MODTAS consists of five factors (synaesthesia, aesthetic involvement, imaginative involvement, extrasensory perception [ESP], and altered states of consciousness). A total of 34 statements such as “I feel as if my mind could envelop the whole world” are evaluated. Respondents are asked to rate how often they experienced the presented feelings on a Likert-type scale ranging from 0 (never) to 5 (very often). Items are summed to create the overall score, resulting in a maximum score of 170. Higher scores indicate greater levels of absorption. Cronbach’s alpha was reported as 0.94 (Jamieson & Loi, 2014) and consistent (also 0.94) in this study.

Anthropomorphism was measured using the Individual Differences in Anthropomorphism Questionnaire (IDAQ; Waytz et al., 2010), a commonly used, relatively recent measure of anthropomorphism. The IDAQ uses fifteen items to measure generalised anthropomorphism directed towards artificial objects, animals, and natural formations such as mountains. Participants were asked to rate each item, such as “To what extent does a fish have free will?” on a Likert-type scale between 0 (not at all) and 10 (very much). Scale items were summed, resulting in a total possible score of 150, where higher scores indicate a greater tendency to anthropomorphise. Cronbach’s alpha for the IDAQ was reported as 0.82 (Waytz et al., 2010) and calculated in this study as 0.90.

Cognitive style was measured using the Cognitive Reflection Test 2 (CRT-2; Thomson & Oppenheimer, 2016). The CRT-2 was validated as an alternative test to the original Cognitive Reflection Test (Frederick, 2005) and measures participants’ preference for analytical, or intuitive-based thinking. The alternate test was used as an attempt to leverage the validated avoidance of mathematical knowledge associated with the original scale (Thomson & Oppenheimer, 2016). Five questions are presented that have an intuitive, yet incorrect response, such as: “A farmer had 15 sheep and all but 8 died. How many are left?”. Each incorrect answer attracts a score of one and items are summed, resulting in a maximum score of five. Higher scores indicate a greater preference for intuitive thinking. Cronbach’s alpha has been reported as low for both the original (0.62) and CRT-2 versions (0.51) (Thomson & Oppenheimer, 2016). The consistency of the CRT-2 in this dataset was similarly low (0.56).

Susceptibility to bullshit was measured using the Bullshit and Profoundness Susceptibility scale (Pennycook et al., 2015). The bullshit susceptibility scale presents individuals with seven meaningless sentences such as “the hidden meaning transforms the abstract beauty”. The profoundness scale measures responses to seven items that have accepted meaning to them, such as “Imagined pain does not hurt less because it is imagined”. For this study, only the susceptibility to bullshit scale was used as a measure of an individual’s tendency to find meaning where there is none. Responses on the susceptibility to bullshit scale are measured using a Likert-type scale ranging from 1 (not at all meaningful) to 6 (very meaningful), with a scale range of 7–42. The scale was found to have good internal consistency previously (0.82; Pennycook et al., 2015), and this is found to be similar in this dataset with a Cronbach’s alpha of 0.88.

## Results

Analysis was conducted using R (version 4.1.1). Participant demographics are presented in Table 1. Characteristics including education levels, income and country of birth broadly represent the national norms (Australian Bureau of Statistics, 2021). Candidate CKC-R item characteristics are presented in Table 2. All items were approximately normally distributed, and skew and kurtosis statistics were moderate. Correlations (supplied in supplementary materials 1) between items indicate most CKC-R items are significantly related to others, one indication the data is related and suitable for SEM analysis (Rosseel, 2012). Sampling adequacy, an indicator of shared common variance between items, was checked via Kaiser–Meyer–Olkin (KMO) and was found to be acceptable (0.96), and Bartlett’s test of sphericity was significant ( $\chi^2(528) = 12715.63, p < 0.001$ ). Thus, there was

**Table 1** Participant demographics

| Gender association                        | Count | %      |
|---|-------|--------|
| Male                                      | 434   | 42.97% |
| Female                                    | 568   | 56.24% |
| Non-Binary / Third gender                 | 3     | 0.30%  |
| Prefer not to say                         | 4     | 0.40%  |
| Self-described                            | 1     | 0.10%  |
| Highest Education Level Obtained          |       |        |
| Year 11 or below                          | 199   | 19.70% |
| Year 12                                   | 157   | 15.54% |
| Certificate III/IV                        | 169   | 16.73% |
| Diploma or Advanced Diploma               | 105   | 10.40% |
| Bachelor degree                           | 209   | 20.69% |
| Graduate Diploma or Graduate Certificate  | 56    | 5.54%  |
| Post Graduate                             | 103   | 10.20% |
| Prefer not to say                         | 12    | 1.19%  |
| Is Religion important in your daily life? |       |        |
| Yes                                       | 350   | 34.65% |
| No  | 660   | 65.35% |
| Income (Before tax, annual)               |       |        |
| \$0—\$19,999                              | 113   | 11.19% |
| \$20,000—\$39,000                         | 277   | 27.43% |
| \$40,000—\$59,999                         | 167   | 16.53% |
| \$60,000—\$79,999                         | 134   | 13.27% |
| \$80,000—\$99,999                         | 87    | 8.61%  |
| \$100,000—\$119,999                       | 59    | 5.84%  |
| \$120,000—\$139,999                       | 37    | 3.66%  |
| \$140,000—\$159,999                       | 42    | 4.16%  |
| \$160,000 or more                         | 43    | 4.26%  |
| Prefer not to say                         | 51    | 5.05%  |
| Born in Australia?                        |       |        |
| No  | 319   | 31.58% |
| Yes                                       | 691   | 68.42% |

sufficient relationship between items to assess their association with a latent variable. All variables were scaled with a mean of zero, and a standard deviation of one, enabling direct comparisons for the remainder of the analyses.

We performed a confirmatory factor analysis using structural equation modelling (SEM) consistent with previous explorations of the CKC (Barber, 2014; Svedholm et al., 2010). Structural equation modelling (SEM) using ordinary least squares was performed to test the predicted model of relationships between items, the global latent trait, and paranormal belief. The hierarchical latent trait model was specified for the CKC and its nominal domains, as shown in Fig. 1. All subscales were defined as uncorrelated components of the global trait (i.e., nil residual covariation), reflecting the originally reported scale structure (Lindeman & Aarnio, 2007; Svedholm et al., 2010). Weak items were evaluated and removed from the model, applying previously

proposed methodologies and rationale (Browne et al., 2018). This approach minimises error terms between items and both the global latent trait, and the outcome variable (paranormal beliefs), and maximises item contribution to the overall variance explained. Error terms were based on the  $r^2$  and item residuals, calculated as  $(1-r^2)^2 + (1-(\text{sum of squared residuals}))^2$ . These values were calculated for each of the items at each iteration. The item with the largest error term was removed in each iteration, and overall model goodness of fit indicators were checked to indicate when iterations should cease. Goodness of fit measures were acceptable after eight iterations. The procedure for this process can be obtained via Williams (2022).

The first iteration of SEM output found the subscale *living inanimates have properties of animates* to have no significant relationship ( $\beta = 0.07$ ,  $p = 0.07$ ) with the common latent trait. The subscale, including all constituent items, was dropped from further analysis. After further item removals, the final model reported acceptable model fit statistics: goodness of fit = 0.95, RMSEA = 0.04 (95% CI: 0.03—0.04,  $p = 1$ ), standardised root mean square residual = 0.03. Omega was obtained using the R package semTools and Cronbach's alpha using package Cronbach. Internal consistency was acceptable for both methods of evaluation (Omega = 0.85, Cronbach's alpha = 0.95). Final CKC-R items along with their standardised factor loadings are presented in bold text, in Table 2 with CKC-R and subscale correlations presented in Table 3. Model paths are shown in Fig. 1. The CKC-R model structure therefore partially met expectations since the subscale *living inanimates have properties of animates* and did not significantly relate to the global latent trait of core knowledge confusions. Strong correlations found between the remaining CKC-R subscales indicate they measure the same global latent trait and overall, the CKC-R exhibits both acceptable model fit and consistency.

Criterion validity was tested by comparing the CKC-R to the original CKC short form in both scale consistency and prediction of paranormal belief. The SEM model for the CKC short form did not meet goodness of fit guidelines (GFI = 0.90, RMSEA = 0.08 (95% CI: 0.08—0.09,  $p = < 0.001$ ), standardised root mean square residual = 0.07). However, scale consistency was acceptable for both scales (CKC short form: Omega = 0.87, Cronbach's alpha = 0.83; CKC-R Omega = 0.85, Cronbach's alpha = 0.95). Correlations between both versions of the CKC and other measures are presented in Table 4. As predicted, both versions of the CKC positively correlate with paranormal belief. The hypothesis that the additional length of the CKC-R would result in a more precise measurement of paranormal beliefs was tested using a two-tailed dependent groups t-test; the correlation between the CKC-R and paranormal beliefs was expected to be significantly higher than that of the CKC. The 95% confidence

**Table 2** CKC-R item characteristics

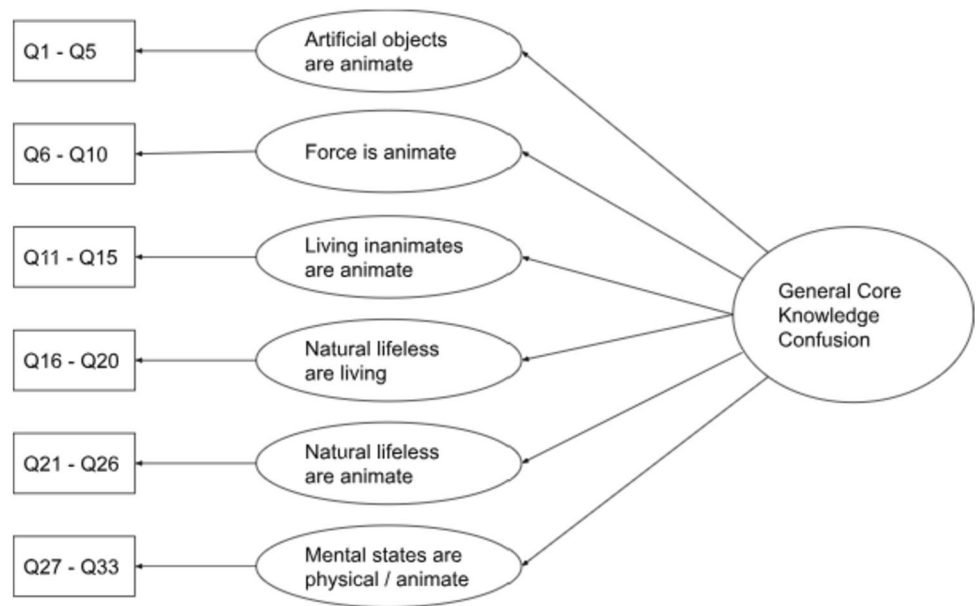
| Subscale and Item                                     | Mean | SD   | Skew  | Kurtosis | SE  | $\beta^{****}$ |
|---|------|------|-------|----------|-----|----------------|
| Artificial Animate                                    |      |      |       |          |     | .93            |
| 1 Furniture wants a home**                            | 2.32 | 1.40 | .63   | -.92     | .04 | .61            |
| 2 A house knows the people who live in it**           | 1.95 | 1.24 | 1.03  | -.13     | .04 | .81            |
| 3 A house misses its inhabitants when they are gone** | 1.91 | 1.24 | 1.12  | .03      | .04 | .79            |
| 4 A house senses its environment**                    | 2.05 | 1.27 | .91   | -.36     | .04 | .82            |
| 5 A house knows its history**                         | 2.04 | 1.30 | .96   | -.34     | .04 | .79            |
| Force Animate   |      |      |       |          |     | .91            |
| 6 Force wants to act**                                | 2.52 | 1.34 | .35   | -1.01    | .04 | .63            |
| 7 Force tries to intervene**                          | 2.48 | 1.30 | .33   | -1.02    | .04 | .66            |
| 8 Force lives in the universe**                       | 2.72 | 1.39 | .17   | -1.18    | .04 | .50            |
| 9 Force knows the direction it is going in**          | 2.60 | 1.40 | .28   | -1.21    | .04 | .58            |
| 10 Force knows there is a human present**             | 2.20 | 1.32 | .68   | -.76     | .04 | .76            |
| Inanimate Animate                                     |      |      |       |          |     |                |
| 11 Plants know the difference of the seasons**        | 3.59 | 1.33 | -.59  | -.77     | .04 |                |
| 12 Flowers want light***                              | 4.01 | 1.19 | -1.13 | .40      | .04 |                |
| 13 Plants want to face the sun*                       | 3.82 | 1.25 | -.85  | -.25     | .04 |                |
| 14 Mushrooms value darkness*                          | 3.46 | 1.40 | -.50  | -.98     | .04 |                |
| 15 Flowers like bees*                                 | 3.46 | 1.43 | -.49  | -1.04    | .04 |                |
| Natural Living  |      |      |       |          |     | .96            |
| 16 A rock lives a long life**                         | 2.53 | 1.52 | .41   | -1.32    | .05 | .50            |
| 17 The moon is alive at night**                       | 2.44 | 1.48 | .49   | -1.21    | .05 | .59            |
| 18 Stones live in the forest**                        | 2.57 | 1.44 | .34   | -1.21    | .05 | .47            |
| 19 Stars live in the sky**                            | 3.42 | 1.52 | -.48  | -1.24    | .05 |                |
| 20 Water lives in rivers and lakes**                  | 3.49 | 1.45 | -.53  | -1.06    | .05 |                |
| Lifeless Animate                                      |      |      |       |          |     | .99            |
| 21 Earth wants water***                               | 3.35 | 1.43 | -.39  | -1.11    | .04 |                |
| 22 The planets know things***                         | 2.21 | 1.32 | .68   | -.77     | .04 | .72            |
| 23 The sky hears the thunder***                       | 2.13 | 1.36 | .83   | -.67     | .04 | .72            |
| 24 The sea embraces swimmers*                         | 2.47 | 1.38 | .41   | -1.10    | .04 | .53            |
| 25 The sky darkens in anger*                          | 1.98 | 1.28 | 1.02  | -.23     | .04 | .70            |
| 26 The sun touches the morning*                       | 2.39 | 1.41 | .54   | -1.05    | .04 |                |
| Mental Material                                       |      |      |       |          |     | .90            |
| 27 A plan lives in nature***                          | 2.50 | 1.36 | .35   | -1.10    | .04 | .58            |
| 28 A mind touches the other***                        | 2.17 | 1.31 | .74   | -.67     | .04 | .71            |
| 29 Fear poisons man***                                | 2.55 | 1.41 | .37   | -1.15    | .04 | .54            |
| 30 Misery seeks company*                              | 2.32 | 1.37 | .61   | -.88     | .04 | .65            |
| 31 Sleep welcomes the weary*                          | 2.82 | 1.50 | .11   | -1.40    | .05 |                |
| 32 Knowledge is in our hands*                         | 2.88 | 1.52 | .07   | -1.44    | .05 |                |
| 33 Knowledge lives within people*                     | 3.29 | 1.40 | -.35  | -1.10    | .04 |                |

\* Newly generated items; \*\* Items from Barber (2014); \*\*\* Items from CKC (Lindeman & Aarnio, 2007). \*\*\*\*  $\beta$  entry for subscale relates to estimate relating to global ontological confusion trait,  $\beta$  for items shows item standard estimate to subscale's-R estimate to Paranormal Beliefs = .44. All retained items  $p < .001$

interval for  $r.CKC-R - r.CKCshort = 0.03 - 0.09$  (Diedenhofen & Musch, 2015; Zou, 2007) supports the prediction; the CKC-R presents a model with appropriate fit statistics and is a stronger predictor of paranormal beliefs when compared to the original CKC short form.

Divergent validity was assessed through the evaluation of CKC-R relationships to anthropomorphism, analytical thinking, verbal knowledge, absorption, and bullshit receptivity. Scale characteristics and correlations are presented in Table 4 above and are supportive of the

**Fig. 1** Hierarchical SEM model used to test CKC-R items



**Table 3** Correlations between CKC-R and CKC-R subscales

|                  | Artificial Animate | Force Animate | Natural Living | Lifeless Animate | Mental Material |
|------------------|--------------------|---------------|----------------|------------------|-----------------|
| Force Animate    | .85                |               |                |                  |                 |
| Natural Living   | .89                | .88           |                |                  |                 |
| Lifeless Animate | .92                | .90           | .95            |                  |                 |
| Mental Material  | .84                | .83           | .87            | .89              |                 |
| CKC-R            | .93                | .91           | .96            | .99              | .90             |

N= 1010; All  $p < .001$

**Table 4** Scale characteristics and correlations

|            | Mean  | SD    | Skew | Kurtosis | Scale Correlations |            |        |         |         |            |         |
|------------|-------|-------|------|----------|--------------------|------------|--------|---------|---------|------------|---------|
|            |       |       |      |          | BS                 | Absorption | IDAQ   | WordSum | CRT-2   | Paranormal | CKC-SF  |
| CKC-R      | -.15  | .76   | .55  | -.53     | .49***             | .40***     | .62*** | -.47*** | -.25*** | .42***     | .91***  |
| BS         | 21.32 | 7.93  | .23  | -.49     |                    | .53***     | .52*** | -.32*** | -.21*** | .45***     | .46***  |
| Absorption | 39.57 | 20.38 | .37  | -.31     |                    |            | .49*** | -.17*** | -.08    | .50***     | .39***  |
| IDAQ       | 56.28 | 21.58 | .41  | -.35     |                    |            |        | -.39*** | -.14*** | .45***     | .56***  |
| WordSum    | 6.17  | 2.61  | -.37 | -.67     |                    |            |        |         | .32***  | -.27***    | -.40*** |
| CRT-2      | 1.44  | 1.19  | .37  | -.91     |                    |            |        |         |         | -.14***    | -.24*** |
| Paranormal | 30.22 | 7.18  | -.24 | -.23     |                    |            |        |         |         |            | .41***  |
| CKC-SF     | 36.56 | 10.58 | .26  | -.25     |                    |            |        |         |         |            |         |

N= 1010; CKC-SF=CKC short form; \*\*\*  $p < .001$

hypothesised relationships. To further test relationships, a conjoint factor analysis (ConFA) using an oblimin rotation (allowing correlations between factors) was run. The model exhibited a reasonable model fit (RMSEA = 0.04,  $p = 1$ , SRMR = 0.06). Factor loadings are presented in Table 5. Table 4 shows both the anthropomorphism and

susceptibility to bullshit scales were moderately, positively correlated with the CKC-R. The ConFA output suggests the IDAQ shows a strong loading on all the CKC-R subscales. Similarly, verbal knowledge and absorption were both correlated to the CKC-R in the expected directions, but the ConFA identifies they have weak loadings on



**Table 5** Factor Loadings from SEM Analysis of CKC-R subscales and comparison scales

|           | IDAQ | Absorption | WordSum | CRT  | Bullshit |
|-----------|------|------------|---------|------|----------|
| ArtAnim   | .96  | .30        | -.13    | -.06 | .43      |
| ForceAnim | .90  | .33        | -.11    | -.06 | .44      |
| NatLiving | .91  | .29        | -.11    | -.05 | .42      |
| LifelAnim | 1.09 | .32        | -.15    | -.07 | .48      |
| MentalMat | .85  | .30        | -.11    | -.05 | .43      |

*ArtAnim* Artificial are Animate, *ForceAnim* Force has properties of animates, *NatLiving* Natural lifeless are living, *LifelAnim* Lifeless objects are Animate, *MentalMat* Mental states have material properties

CKC-R subscales. Overall, the CKC-R appears to be differentiated from verbal knowledge, absorption, susceptibility to pseudo-profound bullshit, and analytical thinking measures. However, the CKC-R appears significantly similar to anthropomorphism (IDAQ).

Linear regression was used to evaluate the usefulness of the CKC-R when predicting paranormal beliefs in a model including anthropomorphism, verbal knowledge, absorption, and analytical thinking style. Both a standard linear regression and a robust linear regression were compared: results did not materially differ, and the output of the standard linear regression is summarised in Table 6. Using `calc.relimp` from R library `relaimpo`, the CKC-R was the third most important predictor ( $r^2 = 0.07$ ,  $p < 0.001$ ). The most important was absorption ( $r^2 = 0.15$ ,  $p < 0.001$ ), and anthropomorphism was the second highest contributor to the model ( $r^2 = 0.08$ ,  $p < 0.001$ ). Verbal knowledge was a small, yet significant contributor to the model whereas analytical thinking did not meet statistical significance. To measure the individual significance of variables, partial eta-squared (calculated using R package effect size) and the change in model  $R^2$  when individual variables were removed are presented in Table 5. These indicate that absorption has the most unique contribution to the model. Other variables have significant overlap and little unique contributions when explaining variation in paranormal beliefs. While the effect of the CKC-R on paranormal beliefs is statistically significant, most of the

relationship between the CKC-R and paranormal beliefs is not unique, sharing common variance with anthropomorphism, absorption, and verbal knowledge.

## Discussion

This research tested the psychometric properties of a revised Core Knowledge Confusions Scale: the CKC-R. Candidate replacements for items previously found to be suboptimal (see Barber, 2014) were generated according to the specifications of the original CKC; items were created using predicate metaphors including core knowledge-related ontological errors. Although the final CKC-R structure deviates from the original CKC, broader psychometric properties met most expectations. The SEM model fit, defining the CKC-R was acceptable. Evaluations of alpha and omega for both the CKC-R and CKC short form indicate they are similarly consistent and, as predicted, the CKC-R is the stronger of the two when predicting paranormal belief. The CKC-R was also correlated to the related traits of absorption, verbal knowledge, and cognitive style in the expected directions. These relationships were moderate and interpreted as support for CKC-R divergent validity when combined with conjoint factor analysis (CFA) although the CFA also highlighted significant overlaps between the CKC-R and anthropomorphism. The CKC-R was able to explain unique variations in paranormal belief when controlling for other measures, although the effect size was small.

The structural equation model used to evaluate the CKC-R model met goodness of fit guidelines, resulting in a scale with acceptable consistency as evaluated using both Cronbach's alpha and Omega. However, the final model structure only partially met expectations. The subscale '*living inanimates have properties of animates*' (e.g. plants want to face the sun) was dropped from the model based on a lack of association with the principle trait; a deviation from previous findings (see Barber, 2014; Lindeman & Aarnio, 2007; Svedholm et al., 2010). This deviation suggests a potential for cultural differences when evaluating core knowledge. The original CKC (Lindeman & Aarnio, 2007) purported to test *core* knowledge, in essence, knowledge that one gains

**Table 6** Regression output predicting paranormal and supernatural beliefs using standardised variables

|            | $\beta$ | 95% CI |      | $t$   | $p$    | $\eta^2$ | DeltaR <sup>2</sup> |
|------------|---------|--------|------|-------|--------|----------|---------------------|
| CKC-R      | .14     | .07    | .21  | 4.06  | < .001 | .02      | .01                 |
| WordSum    | -.08    | -.14   | -.02 | -2.48 | .013   | .01      | .00                 |
| IDAQ       | .15     | .09    | .22  | 4.38  | < .001 | .02      | .01                 |
| Absorption | .35     | .29    | .41  | 11.58 | < .001 | .12      | .09                 |
| CRT        | -.03    | -.08   | .02  | -1.20 | .232   | .00      | .00                 |

$F(5,1004) = 97.82$ ,  $p < .001$ ,  $R^2 = .33$ ,  $N = 1010$ ,  $\eta^2 =$  eta-squared calculated using simultaneous entry (anova type 2); DeltaR<sup>2</sup> = change in model R<sup>2</sup> when the variable is removed

without formal instruction (Carey & Spelke, 1996; Spelke, 2017; Spelke & Kinzler, 2007). However, both biology students and professors classify living inanimates as animates when quick, intuitive responses are obtained (Goldberg & Thompson-Schill, 2009): a core knowledge violation according to the CKC. Given *core* ontological knowledge is not dependent on domain expertise both biology students and professors should not make errors when classifying living inanimates such as trees. However, it takes years of experience before professors perform substantially better than students at the categorisation task and expert knowledge is required for the categorical error to be resolved (Goldberg & Thompson-Schill, 2009). Thus, biological categorisations as measured by the inanimate animate subscale appear to reflect learned knowledge rather than core knowledge, which would explain the poor psychometric properties of the living inanimate subscale observed here.

Further evaluation of the CKC-R structure suggests that the sub-domains are not statistically distinct, relative to the general trait of making categorical errors. Correlations between all subscales were strong and positive (all  $r > 0.83$ ,  $p < 0.001$ ). Additionally, there was an extremely strong ( $r = 0.99$ ) correlation between the lifeless animate subscale and the global latent trait. Dropping all other subscales would only slightly decrease the performance of the CKC-R, and the subscale itself should be considered as a possible short version of the CKC-R. When making core knowledge-related ontological errors, domains appear to be approached similarly: participants who made one kind of ontological error are very likely to make other kinds of ontological errors irrespective of the referenced ontology.

Hypotheses around criterion validity were partially supported. Scale consistency was found to be acceptable for both the CKC-R and CKC short form, assessed using omega and alpha. In this dataset, the revised CKC-R proved to be a better fit for the hierarchical model tested as the CKC short-form SEM model did not meet goodness of fit guidelines. As hypothesised, the additional length of the CKC-R (when compared to the short form) appears to yield a stronger predictive relationship with paranormal beliefs possibly due to the additional coverage and item quality. Despite the change in structure, the CKC-R retained appropriate characteristics, partially meeting expectations and demonstrating criterion validity.

Results supported the view that the CKC-R is different to scales that have been associated with the CKC in prior research. Verbal knowledge, absorption, and cognitive style were all significantly correlated with ontological confusion in the expected directions. However, CFA analysis highlighted only small loadings of these scales on CKC-R subscales and indicated divergence between scales. As predicted, higher levels of absorption, a greater focus and a tendency to become engrossed in the subject of attention

was also associated with making more ontological errors. This effect may be interpreted using embodied semantics. Theories of embodied semantics suggest a role for the sensorimotor system in understanding abstract concepts (i.e. metaphors) such as those used by the CKC-R (e.g. Gallese & Lakoff, 2005; Meteyard et al., 2012); and moderate forms of the hypothesis are considered supportable (Meteyard et al., 2012). In this view, both external cues and sensory-motor systems influence individuals' generation of understanding. For example, when we engage in an interaction such as buying a drink, there are cues, such as numerals indicating a price, available to assist with understanding; these are not present when we *think* about buying a drink. Understanding is reached by re-creating or simulating experience (Meteyard et al., 2012). Given that the creation of categorical understanding is to some degree dependent on experience, higher levels of absorption may reflect an increased emphasis on sensory-motor roles (embodiment) when understanding abstract concepts. Neurological responses related to semantic embodiment may differentiate CKC-R results in much the same way as verbal intelligence is related to responses in N400 activity.

The moderate negative correlation between verbal knowledge and the CKC-R is also consistent with previous findings (see Pennycook et al., 2015). Greater verbal knowledge is associated with a better understanding of metaphors (Carriedo et al., 2016), and according to our results, it also influences performance on the CKC-R. However, this association raises the possibility of measurement error in the attempt to measure core knowledge-related ontological errors. Verbal skills were significant predictors of both the CKC-R and paranormal beliefs in this study, suggesting verbal understanding is important in both categorical understandings and paranormal beliefs. However, verbal skills are influenced by education (Beck & McKeown, 2007), and education should not influence the *core* knowledge-related categorical understanding targeted by the CKC-R. Any potential change in item wording is complicated by the explicit directive that core knowledge-related metaphors are necessarily abstract (Lindeman & Aarnio, 2007). Given the IDAQ is similarly related to both paranormal beliefs and the CKC-R, and IDAQ items are less abstract, the requirement for abstractness appears questionable. Whether the results indicate verbal skills are related to actual ontological beliefs (the belief that *a knife can live*), or to interpretations of language that the CKC-R is predicated on (is “a knife *lives* in the cupboard” literal or metaphorical?) is unclear. The role for verbal knowledge in providing evidence for these interpretations may be clarified by further research using more direct items such as ‘to what degree does a house know its history?’.

Scales argued to overlap with the CKC-R (anthropomorphism and bullshit susceptibility) were both positively

correlated with the CKC-R as hypothesised. Consistent with previous findings the CKC-R and bullshit susceptibility are positively correlated (see Pennycook et al., 2015). CFA output also supports the suggestion that the bullshit susceptibility scale is associated with the CKC-R through moderate loadings on all CKC-R subscales. Seeing meaning in pseudo-profound statements tends to be associated with more ontological errors. However, the relationship is only moderate suggesting ontological confusion and finding meaning in pseudo-profound statements are different, yet overlapping constructs. Anthropomorphism (the IDAQ) was also related to the CKC-R as predicted. Although the extent of this overlap had not previously been tested, results (both correlational and CFA output) indicate a significant overlap between the two scales. Factor loading output from the CFA highlights that while all subscales of the CKC-R are strongly associated to the IDAQ, the *lifeless are animate* subscale has a loading of greater than one and suggests the constructs are very similar. Anthropomorphism is not always associated with beliefs such as religion (Lindeman et al., 2015) and why this is the case deserves more attention through detailed comparisons between the two scales. It appears there is significant overlap between the two and arguments about the utility of either scale over the other remain questionable.

Arguments about the difference between the CKC-R and anthropomorphism have previously focused on religion and may be explained either through evaluating item wording, or the different domains of ontologies covered in each scale. Whereas judgments about metaphors are used by the CKC-R (To what degree is *a house knows its history* a literal statement?), direct statements are used by the IDAQ (*to what degree does a mountain think?*). As verbal knowledge was significantly correlated with both the IDAQ and CKC-R, verbal knowledge may influence the interpretation of scale items and create measurement errors between the two scales. This may explain the theoretical divergence between CKC and IDAQ associations to religion. Alternatively, the minimal divergence may reflect differences in ontological coverage. The CKC-R covers disembodied categories such as mental states and the IDAQ does not. This appears to be reflected in CFA output where the IDAQ had the lowest (although still high) loading on the mental states' subscale of the CKC-R. Subsequently, the CKC-R may be interpreted as a more complete representation of general core knowledge related ontological confusion; an argument previously inferred but not tested (Lindeman et al., 2015). This difference in coverage of ontological categories may explain the divergence between the IDAQ and CKC-R seen more generally, however, it is considered less likely given our prior suggestion that ontological domains are treated similarly and the high loading of the IDAQ in our CFA output. Scale coverage differences may be explored by adding missing ontological categories such as mental states to the IDAQ.

Any such studies are likely to benefit from the inclusion of a broad measure of religiosity covering both religious belief and behaviour-related content that was not available in this study. Such an inclusion may be able to find insight into why despite their similarity, anthropomorphism is less consistently associated with religious beliefs whereas the CKC is.

As expected, the CKC-R was a significant contributor to the multivariate regression predicting paranormal beliefs but the effect is small. Individual correlations of the IDAQ and the CKC-R to paranormal belief were very similar (IDAQ:  $r=0.45$  and CKC-R  $r=0.42$ ) and the CKC-R proved to be a similar but less significant predictor of paranormal beliefs than the IDAQ when both were included in multiple regression. Overall, absorption was the most important predictor of paranormal belief, explaining 15% of the variation and CKC-R explained 7% of the variation in paranormal beliefs. Further analysis indicated most of this relationship was not unique to the CKC-R. Removing the CKC-R from the linear regression indicated only 2% of the variance explained in the initial model was unique to the CKC-R. Hence, this research provides some evidence for the utility of the CKC-R and the complexity of underlying relationships that go some way to explaining why it is that interpreting statements as literal or metaphorical predicts beliefs such as paranormal belief.

## Limitations

Low internal consistency for the CRT-2 suggests the potential for problems such as inattention in data collection. Since other scale consistency measures were acceptable, this is not considered likely, nor problematic for our overall conclusions: the issue of consistency appears restricted to the CRT-2 and is consistent with reports from the CRT-2 development (Thomson & Oppenheimer, 2016). Using a combined seven-item evaluation of analytical thinking (see Pennycook et al., 2015) would have been a more appropriate choice for measuring analytical thinking. Because of this issue, our results are on the lower border of expectations when comparing previous findings between analytical thinking and the CKC. Our findings, along with previous studies report the correlation of analytical thinking to the CKC to be significant, negative and either small (Lindeman & Aarnio, 2007) or moderate (Pennycook et al., 2015). This limited relationship between analytical thinking and ontological confusion may be explained by examining the theoretical relationship between cognitive style and core knowledge. It is plausible that analytical thinking is leveraged for knowledge or decisions that are difficult, requiring reference to educationally obtained information. Conversely, intuitively gained knowledge may be less likely to trigger conflict detection processes associated with analytical thinking (De Neys & Glumicic, 2008; Neys, 2014). Any increased relationship found between the CKC and analytical thinking

(e.g., Barber, 2014) may be the result of CKC inclusions of non-core knowledge-related errors that may be more suited to analytical processing. This study did not directly test this suggestion and further analysis of the dropped subscale and its differences from retained measures may provide some insight. Regardless, our results are aligned with previous findings suggesting either the CKC-R is consistent with Lindeman and Aarnio (2007) and has a weak relationship to analytical thinking, or that the measure of analytical thinking we used was flawed, or likely both. Future research may include alternative cognitive style measures and include tests of non-core knowledge-related errors to test the relationships.

The CKC-R faces other limitations. Firstly, the CKC-R suffers from a lack of clarity in construct definition. As touched on in the introduction, the measurement of ontologies may not reflect cognitive organisation. Additionally, the CKC-R does not offer a clear explanation for the cognitive errors it measures. Initial interpretations of the CKC suggest that the scale was a simple reflection of ontological error that was associated with magical thinking (Lindeman & Aarnio, 2007). However, it has since been interpreted as a cognitive bias (Lindeman et al., 2022) and results here suggest an association with both cognitive performance and personality traits. Discussions on learning and conceptual change (in understanding) give credence to interpreting the CKC-R as a reflection of bias towards intuitive models of understanding. One view of conceptual change allows for both the intuitive and analytical accounts to co-exist and may have a positive influence on such aspects as speed when making decisions (Shtulman & Lombrozo, 2016). For example, intuitive (yet incorrect) claims about why giraffes have long necks may be wrong but ultimately lead to quicker decisions about giraffes' development, with little negative outcomes (for an expanded discussion see Shtulman & Lombrozo, 2016). Given reasonable evidence that conceptual change requires at some point overlapping and inconsistent information and that this may be useful in heuristically guided decisions, the CKC-R may measure the extent to which individuals utilise heuristics in decision making. Future studies may broaden our understanding by testing the association of the CKC-R to existing measures of cognitive bias or test it against beliefs not generally associated with magical thinking.

Although no causal inferences can be made from this study, results expand the understanding of personal influences resulting in core knowledge-related ontological mistakes and support CKC-R use. Intuitively (not learned) ontological understandings such as those tested by the CKC-R are significantly influenced by personality and cognitive performance, with little influence from cognitive style. Whether these core knowledge-related errors result in general logical errors (i.e., those not associated with core knowledge) is not clear. However, given the lack of association to the

biological inanimate scale, argued here to be unrelated to core knowledge, the suggestion is the CKC-R appears specifically useful for predicting beliefs with magical thinking-type associations. This view is consistent with its historical implementation. Vaccination avoidance, for example, is associated with magical health beliefs (Bryden et al., 2018) and is predicted by a composite measure of ontological confusion (Lindeman et al., 2022). However, such use of composite measures required to cover core knowledge theorised subscales appears unnecessary. Making errors in inanimate related metaphors is also associated with responses in metaphors using mental states as the subject; adding additional questions should have little practical effect on the measure of core knowledge confusion.

## Summary

This research sought to revise and test the properties of the core knowledge confusion scale. The result is a scale with similar properties, one significant structural deviation, and several theoretical implications. The deviation, dropping the '*inanimates have the characteristics of animate*' items, suggests that measuring core knowledge is complex and influenced by the degree to which common, metaphoric language is thought of as literal. Further to this, the use of subscales appears to provide little additional benefit when measuring the global latent trait of ontological confusion. Ontological mistakes do not appear to be significantly influenced by their domain (mental states have physical properties vs artificial entities have properties of living animates for example). Hence, when it comes to measuring the core trait of ontological confusion, the use of composite scales to ensure all core domains are covered (see Lindeman et al., 2022) appears unnecessary. Results also suggest that ontological confusions are avoided by those who display high verbal skills and low trait absorption. The relationship with verbal knowledge could imply a level of measurement error in the CKC-R. Verbal skills are increased with education, which is not supposed to be associated with core knowledge. Hence, there is some question as to the level of measurement error in the 'core knowledge' construct. Furthermore, the moderately strong positive correlation and CFA loadings between anthropomorphism and the CKC-R suggest a significant overlap between the two scales. The relationship to an outcome variable, paranormal beliefs, is very similar for both scales. We are left to suggest that claiming to measure ontological errors is a complicated and questionable interpretation of the CKC-R. Perhaps item wording is more important than content differences between the CKC-R and IDAQ scales and this warrants further investigation. Nevertheless, the expected relationships were exhibited when



testing the psychometric properties and supporting the use of the CKC-R.

Our study finds that both personality and cognitive performance measures are significantly associated with literal interpretations of metaphoric statements. It appears to be useful when predicting beliefs containing logical flaws based on core knowledge-related ontological confusions. Whether these core knowledge-related errors generalise to broader categorical mistakes is yet to be tested. However, while the association of the CKC-R to ontologies themselves is questionable, the CKC-R predicts paranormal beliefs as expected, and may provide further insight into other beliefs affecting personal and social well-being.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12144-024-05611-7>.

**Authors' contributions** All authors contributed to the study conception and design. Material preparation and analysis were performed by Brenton Williams. The first manuscript draft was written by Brenton Williams and all authors contributed to the submitted version. All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions This RHD candidature was supported under the Commonwealth Government's Research Training Program/Research Training Scheme. I gratefully acknowledge the financial support provided by the Australian Government. Additional funds were provided by internal Institution funding, I gratefully acknowledge their support.

**Data availability** Data for this study is available via repository: Williams (2022), alternatively the corresponding author can be contacted for access (b.williams@cqu.edu.au).

**Code availability** Code for reproducing data analysis can be obtained via <https://doi.org/10.5281/zenodo.7229134> or by contacting the corresponding author (b.williams@cqu.edu.au).

## Declarations

**Ethics approval** The methodology for this study was approved by the Human Research Ethics committee of CQUniversity (0000022768).

**Consent to participate** Informed consent was obtained from all individual participants included in the study.

**Consent to publish** All participants consented to publication of anonymous, summarised results.

**Conflict of interest** Brenton Williams declares that he has no conflicts of interest in relation to this manuscript.

Matthew Browne has received grants from the Victorian Responsible Gambling Foundation, the New Zealand Ministry of Health the NSW Dept of Industry and Trade, the Department of Social Services, the Alberta Gambling Research Institute and Gambling Research Australia. He declares that he has no conflicts of interest in relation to this manuscript.

Matthew Rockloff has received research grants from the Queensland Treasury, the Victorian Treasury, the Victorian Responsible Gambling Foundation, the New Zealand Ministry of Health, the NSW Dept of Industry and Trade, the Department of Social Services, the Alberta

Gambling Research Institute and Gambling Research Australia. He declares that he has no conflicts of interest in relation to this manuscript. George Stuart declares that he has no conflicts of interest in relation to this manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmad, F. (2020). Personality traits as predictor of cognitive biases: Moderating role of risk-attitude. *Qualitative Research in Financial Markets*, 12(4), 465–484. <https://doi.org/10.1108/QRFM-10-2019-0123>
- Australian Bureau of Statistics. (2021). *2021 Australia, Census All persons QuickStats* | Australian Bureau of Statistics. <https://www.abs.gov.au/census/find-census-data/quickstats/2021/AUS>
- Barber, J. (2014). Believing in a purpose of events: cross-cultural evidence of confusions in core knowledge. *Applied Cognitive Psychology*, 28(3), 432–437. <https://doi.org/10.1002/acp.3003>
- Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a nonnatural entity: anthropomorphism in God concepts. *Cognitive Psychology*, 31(3), 219–247. <https://doi.org/10.1006/cogp.1996.0017>
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, 107(3), 251–271. <https://doi.org/10.1086/511706>
- Bregman-Hai, N., Kessler, Y., & Soffer-Dudek, N. (2020). Who wrote that? Automaticity and reduced sense of agency in individuals prone to dissociative absorption. *Consciousness and Cognition*, 78, 102861. <https://doi.org/10.1016/j.concog.2019.102861>
- Browne, M., Rockloff, M., & Rawat, V. (2018). An SEM algorithm for scale reduction incorporating evaluation of multiple psychometric criteria. *Sociological Methods & Research*, 47(4), 812–836. <https://doi.org/10.1177/0049124116661580>
- Bryden, G. M., Browne, M., Rockloff, M., & Unsworth, C. (2018). Anti-vaccination and pro-CAM attitudes both reflect magical beliefs about health. *Vaccine*, 36(9), 1227–1234. <https://doi.org/10.1016/j.vaccine.2017.12.068>
- Carey, S. (2000). The origin of concepts. *Journal of Cognition and Development*, 1(1), 37–41. [https://doi.org/10.1207/S15327647JCD0101N\\_3](https://doi.org/10.1207/S15327647JCD0101N_3)
- Carey, S., & Spelke, E. (1996). Science and core knowledge. *Philosophy of Science*, 63(4), 515–533. <https://doi.org/10.1086/289971>
- Carriedo, N., Corral, A., Montoro, P. R., Herrero, L., Ballestrino, P., & Sebastián, I. (2016). The Development of metaphor comprehension and its relationship with relational verbal reasoning and executive function. *PLoS One*, 11(3), e0150289. <https://doi.org/10.1371/journal.pone.0150289>
- Chun, C. A., & Hupé, J.-M. (2016). Are synesthetes exceptional beyond their synesthetic associations? A systematic comparison of creativity, personality, cognition, and mental imagery in synesthetes and controls. *British Journal of Psychology*, 107(3), 397–418. <https://doi.org/10.1111/bjop.12146>



- Cor, M. K., Haertel, E., Krosnick, J. A., & Malhotra, N. (2012). Improving ability measurement in surveys by following the principles of IRT: The Wordsum vocabulary test in the general social survey. *Social Science Research, 41*(5), 1003–1016. <https://doi.org/10.1016/j.ssresearch.2012.05.007>
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*(3), 1248–1299. <https://doi.org/10.1016/j.cognition.2007.06.002>
- Dean, C. E., Akhtar, S., Gale, T. M., Irvine, K., Wiseman, R., & Laws, K. R. (2021). Development of the paranormal and supernatural beliefs scale using classical and modern test theory. *BMC Psychology, 9*(1), 98. <https://doi.org/10.1186/s40359-021-00600-y>
- Diedenhofen, B., & Musch, J. (2015). Cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS One, 10*(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*(4), 25–42. <https://doi.org/10.1257/089533005775196732>
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*, 455–479.
- Geary, J. (2011). *I is an other: The secret life of metaphor and how it shapes the way we see the world*. Harper Collins.
- Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental “Roots” in mature biological knowledge. *Psychological Science, 20*(4), 480–487. <https://doi.org/10.1111/j.1467-9280.2009.02320.x>
- Hall, S. E., Schubert, E., & Wilson, S. J. (2016). The role of trait and state absorption in the enjoyment of music. *PLoS One, 11*(11), e0164029. <https://doi.org/10.1371/journal.pone.0164029>
- Jamieson, G. (2005). *The modified tellegen absorption scale: A clearer window on the structure and meaning of absorption*. Retrieved May, 8 2019, from: <https://rune.une.edu.au/web/handle/1959.11/3807>
- Jamieson, G., & Loi, N. (2014). *An empirical test of tellegen's model of absorption: instrumental and experiential sets and the phenomenology of trance induction*. Retrieved from: [https://www.researchgate.net/publication/262451945\\_An\\_Empirical\\_Test\\_of\\_Tellegen%27s\\_Model\\_of\\_Absorption\\_Instrumental\\_and\\_Experiential\\_Sets\\_and\\_the\\_Phénoménology\\_of\\_Trance\\_Induction](https://www.researchgate.net/publication/262451945_An_Empirical_Test_of_Tellegen%27s_Model_of_Absorption_Instrumental_and_Experiential_Sets_and_the_Phénoménology_of_Trance_Induction)
- Kumar, V., & DudaniLatha, R. K. (2021). The big five personality traits and psychological biases: An exploratory study. *Current Psychology, 30*(10), 1007–1014. <https://doi.org/10.1007/s12144-021-01999-8>
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. The University of Chicago Press.
- Lifshitz, M., van Elk, M., & Luhrmann, T. M. (2019). Absorption and spiritual experience: A review of evidence and potential mechanisms. *Consciousness and Cognition, 73*, 102760. <https://doi.org/10.1016/j.concog.2019.05.008>
- Lindeman, M., & Aarnio, K. (2007). Superstitious magical and paranormal beliefs: An integrative model. *Journal of Research in Personality, 41*(4), 731–744. <https://doi.org/10.1016/j.jrp.2006.06.009>
- Lindeman, M., Cederström, S., Simola, P., Simola, A., Ollikainen, S., & Riekk, T. (2008). Sentences with core knowledge violations increase the size of N400 among paranormal believers. *Cortex, 44*(10), 1307–1315. <https://doi.org/10.1016/j.cortex.2007.07.010>
- Lindeman, M., Svedholm, A. M., Takada, M., Lönnqvist, J.-E., & Verkasalo, M. (2011). Core knowledge confusions among university students. *Science & Education, 20*(5–6), 439–451. <https://doi.org/10.1007/s11191-009-9210-x>
- Lindeman, M., Svedholm-Häkkinen, A. M., & Lipsanen, J. (2015). Ontological confusions but not mentalizing abilities predict religious belief, paranormal belief, and belief in supernatural purpose. *Cognition, 134*, 63–76. <https://doi.org/10.1016/j.cognition.2014.09.008>
- Lindeman, M., Svedholm-Häkkinen, A. M., & Riekk, T. J. J. (2022). Searching for the cognitive basis of anti-vaccination attitudes. *Thinking & Reasoning, 0*(0), 1–26. <https://doi.org/10.1080/13546783.2022.2046158>
- Luhrmann, T. M., Nusbaum, H., & Thisted, R. (2010). The absorption hypothesis: learning to hear god in evangelical Christianity. *American Anthropologist, 112*(1), 66–78. <https://doi.org/10.1111/j.1548-1433.2009.01197.x>
- Luhrmann, T. M., Nusbaum, H., & Thiste, R. (2013). “Lord, teach us to pray”: prayer practice affects cognitive processing. *Journal of Cognition & Culture, 13*(1/2), 159–177. <https://doi.org/10.1163/15685373-12342090>
- Luhrmann, T. M., Weisman, K., Aulino, F., Brahinsky, J. D., Dulin, J. C., Dzikoto, V. A., Legare, C. H., Lifshitz, M., Ng, E., Ross-Zehnder, N., & Smith, R. E. (2021). Sensing the presence of gods and spirits across cultures and faiths. *Proceedings of the National Academy of Sciences, 118*(5), e2016649118. <https://doi.org/10.1073/pnas.2016649118>
- Madsen, M. W. (2016). Cognitive metaphor theory and the metaphysics of immediacy. *Cognitive Science, 40*(4), 881–908. <https://doi.org/10.1111/cogs.12320>
- Maij, D. L. R., & van Elk, M. (2018). Getting absorbed in experimentally induced extraordinary experiences: Effects of placebo brain stimulation on agency detection. *Consciousness and Cognition, 66*, 1–16. <https://doi.org/10.1016/j.concog.2018.09.010>
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48*(7), 788–804. <https://doi.org/10.1016/j.cortex.2010.11.002>
- Neys, W. D. (2014). Heuristics, biases and the development of conflict detection during reasoning. In H. Markovits (Ed.), *The Developmental Psychology of Reasoning and Decision-Making* (pp. 99–111).
- Nichols, R. M., & Loftus, E. F. (2019). Who is susceptible in three false memory tasks? *Memory, 27*(7), 962–984. <https://doi.org/10.1080/09658211.2019.1611862>
- Pennycook, G., Cheyne, J. A., Barr, N., Fugelsang, J. A., & Koehler, D. J. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making, 10*(6), 15.
- Prat, C. S., & Just, M. A. (2011). Exploring the neural dynamics underpinning individual differences in sentence comprehension. *Cerebral Cortex, 21*(8), 1747–1760. <https://doi.org/10.1093/cercor/bhq241>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Shtulman, A., & Lombrozo, T. (2016). Bundles of contradiction: A coexistence view of conceptual change. In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 49–67). Oxford University Press. Retrieved May, 8 2019, from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.713.776&rep=rep1&type=pdf>
- Spelke, E. S. (2017). Core knowledge, language, and number. *Language Learning and Development, 13*(2), 147–170. <https://doi.org/10.1080/15475441.2016.1263572>
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96. <https://doi.org/10.1111/j.1467-7687.2007.00569.x>
- Stamenković, D., Ichien, N., & Holyoak, K. J. (2019). Metaphor comprehension: An individual-differences approach. *Journal of Memory and Language, 105*, 108–118. <https://doi.org/10.1016/j.jml.2018.12.003>
- Stamenković, D., Ichien, N., & Holyoak, K. J. (2020). Individual differences in comprehension of contextualized metaphors. *Metaphor and Symbol, 35*(4), 285–301. <https://doi.org/10.1080/10926488.2020.1821203>

- Svedholm, A. M., Lindeman, M., & Lipsanen, J. (2010). Believing in the purpose of events—Why does it occur, and is it supernatural? *Applied Cognitive Psychology*, 24(2), 252–265. <https://doi.org/10.1002/acp.1560>
- Tellegen, A. (1981). Practicing the two disciplines for relaxation and enlightenment: Comment on “Role of the feedback signal in electromyograph biofeedback: The relevance of attention” by Qualls and Sheehan. *Journal of Experimental Psychology: General*, 110(2), 217–226. <https://doi.org/10.1037/0096-3445.110.2.217>
- Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ('absorption'), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology*, 83(3), 268–277. <https://doi.org/10.1037/h0036681>
- Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1), 15.
- Tobacyk, J. J. (2004). A revised paranormal belief scale. *International Journal of Transpersonal Studies*, 23(1), 94–98. <https://doi.org/10.24972/ijts.2004.23.1.94>
- Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3), 219–232. <https://doi.org/10.1177/1745691610369336>
- Willard, A. K., & Norenzayan, A. (2013). Cognitive biases explain religious belief, paranormal belief, and belief in life's purpose. *Cognition*, 129(2), 379–391. <https://doi.org/10.1016/j.cognition.2013.07.016>
- Williams, B. (2022). *CKC R Data Analysis*. GitHub. <https://github.com/brenwil/ckc-r-data-analysis>
- Williams, J. M. (1976). Synaesthetic adjectives: a possible law of semantic change. *Language*, 52(2), 461–478. <https://doi.org/10.2307/412571>
- Wilson, S. C., & Barber, T. X. (1982). The fantasy-prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena. *PSI Research*, 1(3), 94–116.
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, 12(4), 399–413. <https://doi.org/10.1037/1082-989X.12.4.399>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Preregistration: this research was not preregistered with any independent, or institutional registry.