



Development and Validation of an Implicit Relational Assessment Procedure (IRAP) to Measure Implicit Dysfunctional Beliefs about Caregiving in Dementia Family Caregivers

Isabel Cabrera¹  · María Márquez-González¹  · Naoko Kishita²  · Carlos Vara-García³  · Andrés Losada³ 

Accepted: 2 November 2020
© The Author(s) 2020

Abstract

Caregivers of people with dementia who endorse dysfunctional beliefs about caregiving are at high risk of experiencing higher levels of distress. These dysfunctional beliefs are presented in the form of rules, verbal statements that specify what responsibilities one should expect in order to be a “good caregiver,” and are characterized as rigid, unrealistic, or highly demanding. Previous studies relied exclusively on self-report measures when assessing such dysfunctional beliefs about caregiving. The objectives of this study were: 1) to develop and validate an Implicit Relational Assessment Procedure (IRAP) to measure implicit dysfunctional beliefs about caregiving (CARE-IRAP), and 2) considering the relatively high age of the sample, to analyze the adaptation of the IRAP for older adults, comparing the IRAP performance between older adult caregivers and middle-aged caregivers. Participants were 123 dementia family caregivers with a mean age of 62.24 ± 12.89 . Adaptations were made to the IRAP by adjusting the accuracy and response time criteria. The sample was split into middle-aged caregivers (below 60 years) and older adult caregivers (60 or older). The CARE-IRAP scores presented significant positive correlations with explicit measures of dysfunctional beliefs about caregiving and experiential avoidance in caregiving. A similar pattern of results was observed across the two age groups. The results revealed that caregivers endorse implicit dysfunctional beliefs about caregiving and offer preliminary support for the use of the IRAP as a valid measure of implicit caregiving beliefs. This exploratory study is the first to adapt the IRAP criteria to older adults, and future studies should further explore criteria suitable for this population.

Keywords Caregivers · Dementia · Dysfunctional beliefs · Implicit Relational Assessment Procedure

Caregiving of elderly persons with dementia can lead to chronic stress, resulting in negative psychological and physical consequences (e.g., Sallim, Sayampanathan, Cuttilan, & Ho, 2015). Substantial research has been conducted to identify factors leading to negative consequences of caregiving. One of the best-established models, the sociocultural stress and coping model adapted to caregiving (Knight & Sayegh,

2010), suggests that the relationship between caregivers’ stressors and negative psychological and physical consequences is often modulated by the social support available to family members, and different coping behaviors in the face of changing circumstances. The model also highlights the role of beliefs about caregiving as a critical moderator.

Beliefs about caregiving may involve verbal relations between “good caregiver” and statements that specify responsibilities one should expect of a good caregiver. These verbal relations may function as rules specifying what one should do to be a good caregiver. These rules are highly influenced by sociocultural contexts such as shared cultural values (Corcoran, 2011). One of the well-researched cultural values is familism, defined as a “strong identification and attachment of individuals to their families and strong feelings of loyalty, reciprocity, and solidarity among members of the same family” (Sabogal, Marín, Otero-Sabogal, Marín, & Perez-Stable, 1987, p. 398). Familism is considered to have a strong impact

✉ Isabel Cabrera
i.cabrera@uam.es

¹ Department of Biological and Health Psychology, Universidad Autónoma de Madrid, C /Iván Paulov, 6, 28049 Madrid, Spain

² School of Health Sciences, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK

³ Department of Psychology, Universidad Rey Juan Carlos, Madrid, Spain

on shaping caregivers' beliefs about how family members should respond to the needs of a relative with dementia (Losada et al., 2010).

Research shows that caregivers experiencing higher levels of distress are more likely to endorse dysfunctional verbal rules about caregiving, which are often characterized as rigid, unrealistic, or highly demanding (Losada et al., 2010). Following such rigid and highly demanding verbal rules may help caregivers to escape from private events (e.g., thoughts and emotions) that are experienced as aversive in the short term. In the long term, however, strict adherence to such rules may lead to an entangled life, and as a result life satisfaction may be reduced (Törneke, Luciano, & Salas, 2008). For instance, the fact that some caregivers do not ask for help may be functionally related to their belief that a good caregiver should be able to handle all caregiving demands by themselves (Mittelman, Epstein, & Pierzchala, 2003). Others may feel guilty when they express emotions such as sadness or anger because this behavior is inconsistent with the rule "a good caregiver should not complain nor express any negative feelings related to their loved ones" (Gallagher-Thompson, Solano, Coon, & Areán, 2003).

So far, the most commonly used method for assessing dysfunctional beliefs about caregiving has been self-report. The Dysfunctional Thoughts about Caregiving Questionnaire (DTCQ; Losada, Montorio, Izal, & Márquez-González, 2006) is a frequently used self-report measure that allows the assessment of different dysfunctional beliefs specific to caregiving in terms of what makes a good caregiver. Two main domains are measured by the DTCQ: 1) "perception of sole responsibility," beliefs that a good caregiver should prioritize taking care of their family member with dementia and subordinate their own well-being to this priority (e.g., "When a person takes care of a frail/sick relative, he/she should set aside his/her interests, and dedicate himself/herself completely to the care of the frail/sick relative"), and 2) "perfectionism," beliefs that a good caregiver should have high levels of emotional and behavioral self-demandingness (e.g., "A good caregiver should never get mad or lose control with the person that is being cared for").

Some studies found statistically significant correlations between self-reported dysfunctional rules about caregiving and emotional distress (Losada et al., 2006; Losada et al., 2010; Márquez-González, Losada, Izal, Pérez-Rojo, & Montorio, 2007; McNaughton, Patterson, Smith, & Grant, 1995; Stebbins & Pakenham, 2001). However, other studies do not report a significant relation between these variables (Roach, 2013; Sullivan, Beattie, Khawaja, Wilz, & Cunningham, 2016; Tandetnik, Hergueta, Negovanska, Dubois, & Bungener, 2014). A possible reason for this disparity in the results may be the exclusive reliance on self-report measures when assessing beliefs about caregiving. Self-report measures are usually obtained in conditions in which participants have

sufficient time to reflect and produce a response that coheres with culturally shaped and socially desirable rules (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). For example, some caregivers may report that they endorse a familism-driven definition because such rules are more favorably viewed in the respondents' cultural context, even when the rules are not functionally related to their behavior.

Implicit measures are not subject to such biases, because they use time pressure to capture immediate relational responses (e.g., Gawronski & De Houwer, 2014). One of the most recently developed implicit measures is the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006), a computer-based assessment tool in which participants are instructed to respond under time pressure to confirm or reject specific relations between stimuli (e.g., "good caregiver never complains—true or false?"). Response-contingent feedback is presented that is consistent with particular response biases across some blocks of trials and inconsistent with such biases across other blocks of trials. The core assumption of the IRAP is that responses would be quicker and more accurate when the relationship (true or false) between a label (e.g., good caregiver) and a target (e.g., never complain) is consistent with one's beliefs, rather than when it is inconsistent (Barnes-Holmes, Barnes-Holmes et al., 2010).

Previous research using the IRAP has demonstrated that the IRAP cannot be easily faked (McKenna, Barnes-Holmes, Barnes-Holmes, & Stewart, 2007). For example, Barnes-Holmes, Murphy, Barnes-Holmes, and Stewart (2010) explored racial bias using the IRAP and three self-report measures. In their results, the authors found a racial bias with the IRAP but not with all the explicit measures. Other studies have explored the relationship between IRAP and overt behavior. For example, Vahey, Boles, and Barnes-Holmes (2010) explored adolescents' assumptions regarding the acceptability of smoking behaviors with the IRAP, finding that adolescent smokers perceive smokers as more socially accepted than do nonsmokers. In addition, although there are currently no published empirical articles that have explored this, it is expected that the IRAP would be less vulnerable to social desirability like other implicit measures (e.g., Nosek, 2005). The IRAP has also proven to be a useful measure in clinical psychology, for example, as a predictor of depressive symptoms (Kosnes, Whelan, O'Donovan, & McHugh, 2013), or to measure emotional reactions to positive and depressing events (Hussey & Barnes-Holmes, 2012).

Therefore, the first aim of this study was to develop and test the IRAP to assess implicit dysfunctional caregiving beliefs in dementia family caregivers (CARE-IRAP). Previous IRAP studies have mainly been conducted with samples of young adults (e.g., Golijani-Moghaddam, Hart, & Dawson, 2013; Vahey, Nicholson, & Barnes-Holmes, 2015). However, caregiver samples comprise a significant percentage of older adults (mean age is around 60 years; range: 56.6–72.6;

Hopkinson, Reavell, Lane, & Mallikarjun, 2018). A significant decline processing speed is common among older adults due to age-related changes (e.g., Salthouse 1996; Salthouse & Davis, 2006), which may not have been an issue among younger samples targeted in previous IRAP studies. Therefore, some adaptations to the procedure may need to be considered. Hence, the second aim of this study was to analyze the adaptation of the IRAP for older adults, comparing the IRAP performance between older adult caregivers and middle-aged caregivers.

The IRAP typically requires participants to reach an accuracy of $\geq 80\%$ and a median response time of less than 2,000–3,000 ms across blocks (Barnes-Holmes, Barnes-Holmes et al., 2010). However, the impact of these criteria on IRAP effects has not been systematically explored (Hussey, Thompson, McEnteggart, Barnes-Holmes, & Barnes-Holmes, 2015), and there are no formal recommendations. In fact, the research team that developed the original IRAP indicates that the response latency and the accuracy criteria should be modified according to the targeted sample and the stimuli used in the task (Barnes-Holmes, Barnes-Holmes et al., 2010). Furthermore, the only published study that has used the IRAP with an older adult sample (Rezende, Bast, Huziwarra, & Bortoloti, 2020) recommended that future studies should make adjustments to use the task more successfully with this population.

Taking all this into account, our adaptation of the IRAP involved the reduction of task demands with regard to the accuracy and mean response time criteria. The simplification of the task was aimed to make the IRAP more suitable for older adult caregivers and thus avoid an excessively high attrition rate among them. To the best of our knowledge, this is the first study that has adapted the IRAP criteria to the older adult population; thus, this study must be considered exploratory.

Previous studies of dysfunctional beliefs in caregivers (e.g., Losada et al., 2006; Losada et al., 2010) and cultural norms such as familism (e.g., Sabogal et al., 1987) would lead us to expect IRAP scores indicating relations between a "good caregiver" and inflexible behaviors such as always being emotionally balanced and never complaining.

In order to analyze construct validity, a measure of experiential avoidance in the caregiving context was included in the assessment protocol (Losada, Márquez-González, Romero-Moreno, & López, 2014). Experiential avoidance (Hayes, Strosahl, & Wilson, 2011) is the tendency to deny or control aversive private experiences (e.g., emotions and thoughts), and it is considered to be associated with various psychological problems (e.g., Chawla & Ostafin, 2007). If a caregiver holds the dysfunctional verbal rule that "good caregivers should remain happy most of the time," it is likely that they display a tendency to excessively deny or control negative emotions.

Finally, the CARE-IRAP requires participants to change their answers from one block to the other and, as a result, to learn a new rule and behave in accordance with it. These switching skills are affected by the normal decline associated with aging (e.g., Allain et al., 2005). Therefore, especially for older adult caregivers, a relationship may be expected between individual differences in switching skills and the likelihood of achieving criteria in the CARE-IRAP.

Method

Participants

The researchers initially recruited 154 dementia family caregivers (63.6% women) from health and social services centers (e.g., day-care centers) for this study. The inclusion criteria were: 1) being at least 18 years old; 2) identifying as the main source of help for their relative with dementia; 3) devoting more than 1 hour daily to caregiving duties; and 4) having cared for more than 3 consecutive months. Participants were excluded if they had: 1) suspected cognitive impairment (one participant); 2) reading difficulties (four participants); 3) mobility problems in responding to the computer tasks (one participant); and 4) visual problems not corrected by eyewear (two participants). Furthermore, three participants declined to complete the computer tasks. One participant could not complete the assessment because of a fire alarm. Finally, data were not available for 19 participants due to technical issues (e.g., problems with the computers).

The final study sample thus comprised 123 caregivers (64.2% women) with a mean age of 62.24 ($SD = 12.89$, range = 33–84). Fifty-two caregivers (42.3%) were below 60 years (75% women), with a mean age of 49.9 ($SD = 7.65$, range = 33–59), and 71 (57.7%) were 60 or over (59.2% women), with a mean age of 71.28 ($SD = 7.13$, range = 60–84). The age threshold of 60 years was used because it is frequently employed in general aging studies (e.g., United Nations, 2019).

Materials

CARE-IRAP The CARE-IRAP was developed to measure implicit dysfunctional beliefs about what makes good and bad caregivers. The following adaptations were made to the original IRAP task (Barnes-Holmes et al., 2006): 1) participants were required to reach an accuracy of $\geq 70\%$ and a median response time of $\leq 5,000$ ms during the practice and test blocks; 2) the number of test blocks were reduced from six to four; 3) the number of target stimuli from each label was reduced from six to three by presenting the same target stimuli twice in each trial type in order to minimize the complexity of the task; and 4) a larger font size was used for all stimuli.

In particular, participants were asked to respond quickly and accurately in ways that may be similar or dissimilar to their own verbal rules about what makes good and bad caregivers. Table 1 presents the English translated version of rules, labels, and target stimuli used in the CARE-IRAP. The combination of two label stimuli (“good caregiver” and “bad caregiver”) and two types of target stimuli (“rigid” style and “flexible” style) generated four independent trial types (good caregiver-rigid, good caregiver-flexible, bad caregiver-rigid, bad caregiver-flexible). Figure 1 shows examples for each CARE-IRAP trial type. On each CARE-IRAP trial, one of two label stimuli (“good caregiver” or “bad caregiver”) was presented at the top of the computer screen, with one of two types of target stimuli (“rigid” style: “Never complains,” “Can do everything alone,” “Always happy,” or “flexible” style: “Can complain,” “Asks for help,” “Can be unhappy”) presented in the center. The two response options (“Press D for False” or “Press K for True”) appeared in the bottom left- and right-hand corners at the same time. In each trial, participants were required to choose between the two response options by pressing the D or K keys on the computer keyboard. If the participant gave an incorrect answer, a red X appeared in the middle of the screen until they answered correctly. After a correct response, an intertrial interval delayed the onset of the next IRAP trial by 400 ms.

The CARE-IRAP consisted of two practice blocks and four test blocks. All participants started with a consistent block (Rule A: a good caregiver is rigid and a bad caregiver is flexible) followed by an inconsistent block (Rule B: a good caregiver is flexible and a bad caregiver is rigid) for both practice and test blocks.¹ The consistent and inconsistent blocks were always presented alternately, an equal number of times each (i.e., test blocks 1 and 3 were consistent and test blocks 2 and 4 were inconsistent). After each block, participants were informed that the rule had changed in next block, as a result, the previously correct and incorrect answers would be reversed in the next block. If participants maintained beliefs pro-Rule A and anti-Rule B, a shorter response latency for the consistent blocks (Rule A) as compared with the inconsistent ones (Rule B) was expected, and this is described as the IRAP effect. In the IRAP task, it was also expected that participants would take longer to respond to the first consistent and inconsistent block, compared with the second consistent and inconsistent block, due to a learning effect. Each block comprised a random sequence of all possible pairings of label and target stimuli. Each target stimulus was presented twice for each trial type in each block, and each block therefore consisted of 24 trials. Trials were presented quasi-randomly for each

participant, so none of the four trial types were presented three times successively.

A correct response on any trial was determined by whether or not the block was consistent or inconsistent. At the beginning, participants were instructed that in this task they have to follow two rules alternatively. Rule A established that a good caregiver never complains, does everything by herself/himself, and is always happy, whereas a bad caregiver can complain, asks for help, and can be unhappy. Rule B established that a good caregiver can complain, asks for help, and can be unhappy, whereas a bad caregiver never complains, does everything herself/himself, and is always happy. In the blocks following rule A (i.e., consistent blocks), participants had to answer “true” in trials combining the label “good caregiver” and all the “rigid” targets and in those combining the label “bad caregiver” and all the “flexible” targets. On the other hand, they had to answer “false” in trials combining the label “good caregiver” and all the “flexible” targets, and in those combining the label “bad caregiver” and all the “rigid” targets. However, in the blocks with rule B (i.e., inconsistent blocks), the pattern of true and false answers was reversed. We illustrated the rules and the correct responses with printed examples of the task. Participants were asked to go as slowly as they needed in order to learn to respond according to the rules. However, they were also instructed that, once they had learned the rules, it was important to respond as quickly as possible. Participants were then given the opportunity to try a few examples, and were encouraged to ask any questions they might have before the first practice block started. Finally, we requested that participants carefully read the rules that were presented on the screen at the beginning of each block.

In the practice phase, participants were required to reach at least 70% of correct responses and a median response time $\leq 5,000$ ms. If participants did not achieve both criteria in either of the two practice blocks (i.e., first consistent and inconsistent blocks), they were given another three sets of consistent and inconsistent practice blocks. If participants failed to achieve the criteria by the end of their fourth attempt, the task automatically terminated and they did not proceed to perform the test blocks. After each block—practice and test—the program gave feedback indicating the percentage of correct responses and median response latency.

The task used the 2012 version of IRAP programming in Microsoft Visual Basic 6. All instructions and stimuli were presented in Spanish. The program recorded the reaction times, defined as the time (ms) that elapsed between the stimulus presentation and the first correct response, and accuracy for each trial.

The selection of stimuli for the CARE-IRAP was based on previous clinical studies with dementia family caregivers, in which these rigid, unrealistic, or highly demanding verbal rules were one of the main intervention targets (e.g., Gallego-Alberto, Márquez-González, Romero-Moreno, Cabrera, & Losada, 2019; Losada et al., 2015; Márquez-González et al., 2007).

¹ The order in which the IRAP blocks (consistent or inconsistent) are presented seems not to have any important significant effect (Barnes-Holmes, Barnes-Holmes et al., 2010).

Table 1 Rules and Stimuli for the CARE-IRAP

Rule A consistent	
Good caregiver is rigid and bad caregiver is flexible	
Rule B inconsistent	
Good caregiver is flexible and bad caregiver is rigid	
Label 1: “Good caregiver” (“Buen cuidador”)	Label 2: “Bad caregiver” (“Mal cuidador”)
Target 1: “Rigid”	Target 2: “Flexible”
Never complains (No se queja)	Can complain (Puede quejarse)
Can do everything alone (Puede solo)	Asks for help (Pide ayuda)
Always happy (Está siempre alegre)	Can be unhappy (Puede estar mal)
Response option 1	Response option 2
True (Verdadero)	False (Falso)

Note. Each target stimulus was presented twice in each block. The Spanish items are presented in parentheses.

Dysfunctional Thoughts about Caregiving Questionnaire (DTCQ; Losada et al., 2006) The DTCQ is a 16-item scale designed to assess beliefs about caregiving which may be obstacles to adaptive coping. The DTCQ has a two-factor structure: 1) Factor 1: Perception of sole responsibility (e.g., “It is selfish for a caregiver to dedicate time to himself/herself when a relative is frail/sick and needs care”), and 2) Factor 2: Perfectionism (e.g., “To become a good caregiver would mean not making mistakes when taking care of a frail/sick relative”). Participants are instructed to rate the degree of agreement, on a scale ranging from 0 (totally disagree) to 4 (totally agree). The reliability of this assessment instrument has been well-established (Losada et al., 2006).

Experiential Avoidance in Caregiving Questionnaire (EACQ; Losada et al., 2014) The EACQ is a 15-item scale that measures specific manifestations of experiential avoidance related to care (e.g., “I cannot bear it when I get angry with my relative”, “One should not have bad thoughts about the person you are caring for”). The scale has good psychometric properties (Losada et al., 2014).

Marlowe-Crowne Social Desirability Scale (SDS; Crowne & Marlowe, 1960) Socially desirable responding was measured with the Spanish adaptation (Ávila & Tomé, 1989) of the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960). The short form employed in this study was composed of the 10-item short version proposed by Strahan and Gerbasi (1972), which has been found to have good reliability. The answers follow a true/false format.

The Rule Shift Cards (Wilson, Alderman, Burgess, Emslie, & Evans, 1996) This task measures the ability to follow a rule in a task and to shift from one rule to another. Therefore, the task has two parts. In the first, participants are presented with red and black playing cards, one by one, and the rule is to indicate if the card is red or black. In the second part,

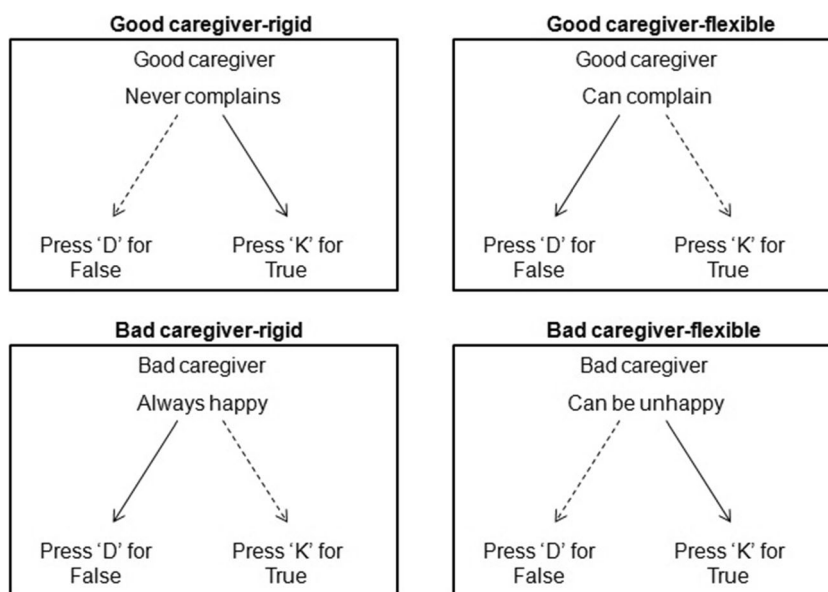
participants are presented with the same red and black playing cards, but the rule changes and participants have to indicate if the card’s color is the same or different from the previous one. Each task has 20 trials. The original task uses paper cards, but a computerized version of the task was developed for the present study using the E-Prime 2.0 software. In the first part, participants had to press the right mouse button if the card was red and the left if the card was black. In the second part, participants had to press the right mouse button if the card’s color was the same as the previous one and the left if the color was different. Rules were always presented at the top of the screen to reduce memory constraints. The card appeared in the middle of the screen, below the instructions. Following the detection of the response, the screen was cleared for 500 ms. The program recorded the accuracy data for the second part of the task, which measures the rule shift.

Different studies have found that the rule shift cards is a validated task for measuring rule learning and abilities like flexibility and inhibition (Espinosa et al., 2009; Norris & Tate, 2000).

Procedure

Potential participants from health and social services centers were contacted by phone and asked sociodemographic and other questions in order to confirm that they met the inclusion criteria. Once they agreed to participate, they were invited to an individual assessment in the collaborating centers. Informed consent was obtained from all participants at the start of the session. Participants sat at a comfortable distance from the computer screen and received instructions concerning the completion of the rule shift cards. Following completion of this task, participants then received the instructions for the CARE-IRAP. After completion of the CARE-IRAP, participants completed the DTCQ, the EACQ, and the SDS. The study was approved by the Ethics Committee of the Universidad Rey Juan Carlos (Madrid). No economic compensation was offered for participation.

Fig. 1 Example screen shots of the four CARE-IRAP trial types. A label is at the top (“good caregiver” or “bad caregiver”), a target is just under the label (e.g., “Never complains” or “Can be unhappy”), and response options (“true” and “false”) are near the bottom corners. All stimuli appear simultaneously at the onset of a trial. An incorrect selection results in a red X appearing just under the target. A correct selection clears the screen for 400 ms, followed by the appearance of a new trial (or instructions when completing a block of trials). Correct selections are indicated by the solid arrows for consistent blocks and dashed arrows for inconsistent blocks (the arrows do not appear on screen)



Data Analysis

To explore age differences in response accuracy in the rule shift cards task, an independent sample of *t*-tests was carried out. To explore the relationship between the level of achievement of the CARE-IRAP and the accuracy in the rule shift cards task, participants were split into four different groups based on their performance on the CARE-IRAP criteria: 1) participants that completed the test blocks and maintained criteria during the test blocks; 2) participants that completed the test blocks but did not maintain these criteria during the test blocks; 3) participants that did not achieve the criteria during the practice blocks and therefore did not complete the test blocks; and 4) participants that did not perform the CARE-IRAP because they could not understand the instructions. A one-way ANOVA and post-hoc tests were subsequently carried out.

The basic data from the CARE-IRAP were response latencies. These response latencies were transformed into D_{IRAP} scores using the adapted version of the D-algorithm developed by Greenwald, Nosek, and Banaji (2003; see Barnes-Holmes, Barnes-Holmes et al., 2010, for a description of this procedure). D_{IRAP} scores reduce the impact of extraneous variables such as age, motor skills, and cognitive ability (Nosek, Greenwald, & Banaji, 2007). D_{IRAP} scores from the two good caregiver trial types and the two bad caregiver trial types were collapsed to provide, respectively, a single D_{IRAP} score for the good caregiver and the bad caregiver trial types. An overall D_{IRAP} score was also calculated across all four trial types. A positive D_{IRAP} score indicates a relatively stronger relationship between good caregiver and rigid behaviors (i.e., slower responding during the inconsistent test blocks compared to the consistent test blocks), whereas a negative D_{IRAP} score indicates a relatively stronger relationship between good caregiver

and flexible behaviors (i.e., faster responding during the inconsistent test blocks compared to the consistent test blocks). Hence, positive or negative D_{IRAP} scores suggest biased beliefs, whereas a score near zero suggests the absence of bias.

To assess the internal consistency of the IRAP, split-half reliability D_{IRAP} scores for odd and even trials were calculated using Spearman-Brown corrections. This was carried out for the overall D_{IRAP} score, as well for good caregiver and bad caregiver D_{IRAP} scores. To evaluate the IRAP effect (i.e., shorter latencies for consistent relative to inconsistent test blocks) and the learning effect (i.e., shorter latencies for the second consistent and inconsistent blocks, as compared to the first consistent and inconsistent blocks), a 2 x 2 mixed repeated measures ANOVA was carried out, with the IRAP condition (consistent and inconsistent) and order (first and second) as repeated measures. To explore whether the D_{IRAP} scores differed significantly from zero, one-sample *t*-tests were employed. To assess the validity of the caregiving CARE-IRAP, correlation, analyses were performed on the D_{IRAP} scores, the DTCQ and its factors, the EACQ, and the SDS. These analyses were conducted for the overall sample, and also for the middle-aged (< 60 years) and older adult (\geq 60 years) caregivers.

Results

Attrition

Of the total of 123 caregivers, 35 (28.5%) could not meet the accuracy and response latency criteria required during the practice blocks. These participants did not complete the test blocks (8 were middle-aged and 27 older adults). Furthermore, 16 participants (13%) failed to maintain these

criteria during the test blocks and, therefore, their data were removed from the analysis (5 middle-aged and 11 older adults). Finally, 14 caregivers (11.4%, all older adults) did not perform the CARE-IRAP because they could not understand the instructions to complete the task. Therefore, our final sample was composed of 58 caregivers.

There were significant differences in the accuracy on the rule shift cards task between age groups, $t_{(119)} = 4.35, p < .001$. Older adult caregivers ($M = 13.92, SD = 3.13$) had lower accuracy levels than middle-aged caregivers ($M = 16.31, SD = 2.76$). When the relationship between the level of achievement of the CARE-IRAP and the accuracy in the rule shift cards task was explored, middle-aged caregivers demonstrated a significant difference between the performance on the CARE-IRAP criteria and the accuracy in the rule shift cards task, $F_{(2,49)} = 4.14, p = .022$. Post-hoc analyses showed that participants that completed the test blocks ($M = 16.72, SD = 2.41$) and participants that completed the test blocks but did not maintain the criteria during the test blocks ($M = 17, SD = 1.22$) had higher accuracy than those participants that did not achieve the criteria during the practice blocks ($M = 13.87, SD = 3.87$). For the older adult group, there were also significant differences between the performance on the CARE-IRAP criteria and the accuracy in the rule shift cards task, $F_{(3,65)} = 16.62, p < .001$. Post-hoc analyses showed that participants that completed the test blocks ($M = 16.58, SD = 1.83$) and participants that completed the test blocks but did not maintain the criteria during the test blocks ($M = 14.91, SD = 2.58$) had higher accuracy than participants that did not achieve the criterion during the practice blocks ($M = 13.15, SD = 2.97$) and participants that did not perform the CARE-IRAP because they could not understand the instructions ($M = 10.93, SD = 1.53$). The differences were also significant between these two last groups.

CARE-IRAP internal reliability

The split-half correlation, using the Spearman-Brown correction, for the Good caregiver D_{IRAP} score and the bad caregiver D_{IRAP} score was $r = .51$ (CI95%: .46-.56) and $r = .62$ (CI95%: .58-.65) respectively. The split-half correlation for the overall D_{IRAP} score was $r = .67$ (CI95%: .63-.70).

We replicated the split-half correlations for each age group. For the middle-aged caregiver group, the split-half correlations for the good caregiver D_{IRAP} score, the bad caregiver D_{IRAP} score, and the overall D_{IRAP} score were $r = .52$ (CI95%: .47-.56), $r = .39$ (CI95%: .32-.44), and $r = .58$ (CI95%: .54-.62), respectively. For the older adult caregiver group, the split-half correlations for the good caregiver D_{IRAP} score, the bad caregiver D_{IRAP} score, and the overall D_{IRAP} score were $r = .54$ (CI95%: .46-.56), $r = .86$ (CI95%: .85-.87), and $r = .83$ (CI95%: .81-.84), respectively.

IRAP and Learning Effects

Table 2 shows the means and standard deviations of the response latencies of the consistent and the inconsistent test blocks. The 2 x 2 ANOVA revealed a significant effect for the IRAP condition, $F_{(1,57)} = 23.73, p < .001$, and order, $F_{(1,57)} = 14.11, p < .001$, but not a significant interaction effect, $F_{(1,57)} = 0.71, p = .402$. The main effect of the IRAP condition indicated that the mean response latencies for consistent blocks were significantly shorter than the mean response latencies for inconsistent blocks. The order effect indicated that response latencies for the second consistent and inconsistent blocks were significantly shorter than those of the first consistent and inconsistent blocks, suggesting a learning effect.

We repeated the analysis for each age group. For the middle-aged caregivers, there was a significant main effect for the IRAP condition, $F_{(1,38)} = 17.75, p < .001$, and order, $F_{(1,38)} = 5.25, p = .028$, but not a significant interaction effect, $F_{(1,38)} = 1.64, p = .207$. For the older adult caregivers, there was also a significant main effect of the IRAP condition, $F_{(1,18)} = 6.05, p = .024$, and order, $F_{(1,18)} = 10.47, p = .005$, but not a significant interaction effect, $F_{(1,18)} = 0.12, p = .729$. Therefore, the same IRAP and learning effects were found for both age groups.

D_{IRAP} Scores

The mean D_{IRAP} score for good caregiver and bad caregiver trials, as well as the mean overall D_{IRAP} score for the overall sample and for both age groups, are presented in Table 2. For the overall sample, one-sample t -tests revealed that the good caregiver D_{IRAP} score, $t_{(57)} = -6.99, p < .001$, the bad caregiver D_{IRAP} score, $t_{(57)} = -4.79, p < .001$, and the overall D_{IRAP} score, $t_{(57)} = -7.22, p < .001$, were significantly different from zero. The same comparisons were carried out for each age group. For middle-aged caregivers, the good caregiver D_{IRAP} score, $t_{(38)} = -4.94, p < .001$, the bad caregiver D_{IRAP} score, $t_{(38)} = -5.18, p < .001$, and the overall D_{IRAP} score, $t_{(38)} = -6.02, p < .001$, were significantly different from zero. For older adult caregivers, only the good caregiver D_{IRAP} score, $t_{(18)} = -5.39, p < .001$, and the overall D_{IRAP} score, $t_{(18)} = -3.89, p = .001$, were significantly different from zero. The bad caregiver D_{IRAP} score, $t_{(18)} = -1.55, p = .137$, was not significantly different from zero, suggesting the absence of bias.

Correlations between Implicit and Explicit Measures

Means and standard deviations for the DTCQ, EACQ, and the SDS are presented in Table 2. As expected, there were significant positive correlations between D_{IRAP} scores and the explicit measure of dysfunctional beliefs measured with the DTCQ and the experiential avoidance in the caregiving measured with the EACQ (Table 3). In particular for the overall

Table 2 Mean and Standard Deviations for the Study Variables

	Overall sample		Middle-aged caregivers		Older adult caregivers	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
First consistent block (ms)	3341	915	3288	929	3451	901
Second consistent block (ms)	3056	837	3035	884	3097	754
First inconsistent block (ms)	3735	913	3648	857	3915	1018
Second inconsistent block (ms)	3542	810	3566	875	3494	675
Good caregiver D_{IRAP} score	0.34	0.37	0.31	0.39	0.41	0.33
Bad caregiver D_{IRAP} score	0.25	0.41	0.29	0.35	0.18	0.50
Overall D_{IRAP} score	0.30	0.31	0.30	0.31	0.29	0.33
DTCQ Total	19.10	10.03	15.94	8.59	25.57	9.84
DTCQ Factor 1	9.71	6.33	7.71	5.01	13.94	6.88
DTCQ Factor 2	9.43	4.91	8.33	4.61	11.68	4.86
EACQ	42.94	7.8	43.07	8.27	42.68	6.93
SDS	6.89	1.97	6.76	2.10	7.15	1.70

Note. EACQ = Experiential Avoidance in Caregiving Questionnaire; DTCQ = Dysfunctional Thoughts about Caregiving Questionnaire; SDS = Marlowe-Crowne Social Desirability Scale.

sample, there was a significant positive correlation between the good caregiver D_{IRAP} score and the total DTCQ, the DTCQ Factors 1 and 2, and the EACQ, as well as between the overall D_{IRAP} score and the total DTCQ, the DTCQ Factor 2, and the EACQ. For the middle-aged caregiver group, there was a significant positive correlation between the good caregiver D_{IRAP} score and the DTCQ Factor 1, and the EACQ, and between the overall D_{IRAP} score and the DTCQ Factor 1. For the older adult caregiver group, there was a significant positive correlation between the good caregiver D_{IRAP} score and the total DTCQ, the DTCQ Factor 2, and the EACQ as well as between the overall D_{IRAP} score and the DTCQ Factor 2.

Regarding the correlations between D_{IRAP} scores and social desirability measured with the SDS, none of the correlations was significant for the overall sample, neither for middle-aged nor older adult caregivers. Finally, even though we expected a significant positive correlation between the explicit measure of dysfunctional thoughts and the social desirability measures, the correlation between the SDS and the DTCQ was only significant for the older adult caregiver group, in particular with DTCQ Factor 2.

Discussion

The present study is the first to explore implicit dysfunctional beliefs about caregiving in a sample of dementia family caregivers. These beliefs about caregiving may be regarded as verbal relations, or rules that specify what responsibilities one should expect in order to be a “good caregiver” and are characterized as rigid, unrealistic, or highly demanding. The results of this study offer preliminary support for the use of the

IRAP as a valid measure of implicit caregiving beliefs in this population, as evidenced by the positive correlation between the CARE-IRAP and the explicit measures of dysfunctional beliefs and experiential avoidance, the absence of correlation between the CARE-IRAP and social desirability, and its internal reliability.

Regarding implicit dysfunctional beliefs about caregiving, the whole sample and both age groups showed shorter response latencies for consistent blocks (i.e., good caregiver should be rigid) than for inconsistent blocks (i.e., good caregiver should be flexible). This result shows that caregivers were quicker to relate the “good caregiver” label with all the target stimuli included in the “rigid” style of care (not complaining, always being happy, and not asking for help), and the “bad caregiver” label with all the target stimuli of the “flexible” style of care (complaining, expressing negative emotions, and asking for help), as compared to the opposite combination. These biases were confirmed with the D_{IRAP} scores. The mean D_{IRAP} score for “good caregiver” and “bad caregiver” trials, as well as the mean overall D_{IRAP} score, differed significantly from zero in the overall sample. Hence, the overall sample showed a bias towards the dysfunctional belief that to be a good caregiver they should do everything by themselves, never express and/or feel negative emotions nor complain about their caregiving tasks. At the same time, they also show a tendency towards the assumption that a bad caregiver is one who asks for help, expresses and/or allows himself/herself to have negative emotions and complains about his/her caregiving tasks. It is important to note that these results may reflect cultural and even religious (Catholicism-related) values and norms, such as familism or the idea of self-sacrifice, which are very present in Latin cultures such as the

Table 3 Correlations between the Variables of the Study

	Overall sample						Middle-aged caregivers						Older adult caregivers					
	DTCQ Total		DTCQ Factor 1		DTCQ Factor 2		DTCQ Total		DTCQ Factor 1		DTCQ Factor 2		DTCQ Total		DTCQ Factor 1		DTCQ Factor 2	
	SDS	EACQ	SDS	EACQ	SDS	EACQ	SDS	EACQ	SDS	EACQ	SDS	EACQ	SDS	EACQ	SDS	EACQ	SDS	EACQ
Good caregiver D_{IRAP} score	.38**	.37**	.29*	.37**	.21	.29	.33*	.18	.32*	.20	.53*	.46	.46*	.55*	.17			
Bad caregiver D_{IRAP} score	.08	-.02	.19	.16	-.09	.22	.21	.17	.18	-.07	.08	-.14	.36	.13	-.10			
Overall D_{IRAP} score	.27**	.20	.29*	.32*	.05	.30	.32*	.21	.31	.08	.33	.12	.51*	.38	.002			
DTCQ Total	-	-	-	.43**	.22	-	-	-	.53**	.10	-	-	-	.43	.44			
DTCQ Factor 1	-	-	-	.36**	.25	-	-	-	.46**	.09	-	-	-	.40	.54*			
DTCQ Factor 2	-	-	-	.40**	.13	-	-	-	.48**	.08	-	-	-	.32	.16			
EACQ	-	-	-	-	.22	-	-	-	-	.27	-	-	-	-	.10			

* $p \leq .05$. ** $p \leq .01$.

Note. EACQ = Experiential Avoidance in Caregiving Questionnaire; DTCQ = Dysfunctional Thoughts about Caregiving Questionnaire; SDS = Marlowe-Crowne Social Desirability Scale.

Spanish (e.g., Nolle, Gulbas, Kuhlberg, & Zayas, 2012). Future studies should explore cultural differences in the conceptualization of familism in care.

When D_{IRAP} scores are compared with the absence of bias in both age groups, older adult caregivers only presented biased beliefs for the overall D_{IRAP} score and the good caregiver D_{IRAP} score, whereas middle-aged caregivers also presented a bias for the bad caregiver D_{IRAP} score. Therefore, older adult caregivers seem to display a general tendency to immediately relate good caregiver to the rigid behaviors (should do everything by themselves, never express and/or feel negative emotions, nor complain about their caregiving duties), but they do not immediately relate bad caregiver to the flexible behaviors. Middle-aged caregivers, in contrast, may have biased verbal relationships both for what makes a good and a bad caregiver. Perhaps, the process of aging and the experience of caregiving increases the flexibility and acceptance of errors or absent-mindedness that can occur during the care of a sick family member. This hypothesized greater flexibility and acceptance seems to be consistent with studies finding that older people exhibit greater emotional resilience and acceptance compared to young adults when faced with emotional conflicts (e.g., Charles & Carstensen, 2010). Future studies should explore this hypothesis further.

As expected, there were significant correlations between implicit and explicit dysfunctional caregiving beliefs. These correlations were found for the overall D_{IRAP} score and the good caregiver D_{IRAP} score, but not with the bad caregiver D_{IRAP} score. It is possible that the bad caregiver D_{IRAP} score did not correlate with the explicit dysfunctional caregiving beliefs because the self-report measure (i.e., the DTCQ) mainly covers the explicit dysfunctional caregiving beliefs associated with good caregiving and not so much with bad caregiving (e.g., "Good caregivers should remain happy and in good spirits all day long to deal adequately with the daily tasks of caregiving"). The fact that the CARE-IRAP also assesses dysfunctional verbal caregiving rules about what makes a bad caregiver is another important advantage of this measure, as compared to the DTCQ.

It is interesting that caregivers with higher levels of implicit dysfunctional caregiving beliefs also presented higher levels of experiential avoidance. This was observed for both the overall sample and the two age groups, and occurred specially for the good caregiver D_{IRAP} score. These results support the hypothesis that the dysfunctional beliefs explored by the CARE-IRAP may act as rigid verbal rules that could facilitate experiential avoidance (Hayes & Gifford, 1997).

In addition, as predicted, implicit dysfunctional caregiving beliefs assessed with the CARE-IRAP did not correlate with social desirability. However, explicit caregiving beliefs correlated with social desirability, although only in the older adult caregiver group. These results point to another strength of the CARE-IRAP: its lower susceptibility to response biases compared to the DTCQ.

The internal reliability of the different D_{IRAP} scores of the CARE-IRAP were low to moderate for the whole sample and the two age groups. These internal consistencies are similar in magnitude (e.g., Barnes-Holmes, Murtagh, Barnes-Holmes, & Stewart, 2010; Drake et al., 2015) or even higher (e.g., Drake, Timko, & Luoma, 2016; Remue, De Houwer, Barnes-Holmes, Vanderhasselt, & De Raedt, 2013) than those reported in other IRAP studies.

To date, there are no formal recommendations regarding response latency and accuracy criteria for the IRAP in general, and IRAP researchers suggest changing or adapting these criteria depending on the characteristics of the study sample. However, previous IRAP studies have mainly been conducted with young participants and there is only one previous published study conducted with older adults (Rezende et al., 2020). In their study, Rezende et al. did not adapt the IRAP criteria for their sample and found a high attrition rate (46%), concluding that future studies should make adjustments to use the task more successfully in this population. In this respect, a decline in the speed of processing due to age-related changes (e.g., Salthouse, 1996; Salthouse & Davis, 2006) was expected among participants for the current study, which could affect the IRAP performance. Therefore, for the present study, we adapted response latency and accuracy criteria to the caregiver population by reducing the difficulties and the complexity of the task. That is, participants had to achieve an accuracy of $\geq 70\%$ and a median response time of $\leq 5,000$ ms during practice blocks.

However, the IRAP is a time-constraint task that prevents participants from reflecting on and manipulating their responses, as participants possibly could do in self-report measures. The adaptation made to the CARE-IRAP (i.e., an increase in the response time) may have given the participants more time to elaborate their responses in the current study. Thus, it is necessary to further explore the IRAP criteria in the older adult population in order to find suitable IRAP criteria that maintain the time constraint but also allow sufficiently accurate performance in this population. In this sense, the results of this study must be considered exploratory and future studies require to continue exploring the IRAP criteria in this population.

A result that supports the preliminary validity of the IRAP in older adults is the learning effect, also observed in other IRAP studies (e.g., Cullen, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009). This effect shows that participants achieved shorter response latencies in the second consistent and inconsistent blocks compared to the first consistent and inconsistent blocks. This learning effect was observed in the overall caregiver sample, as well as in middle-aged and older adult caregivers.

The attrition rates for the present study were high. For the overall sample, 28.5% could not meet the accuracy and response latency criteria required during the practice blocks (15.4% for the middle-aged caregiver group, and 38% for the older adult caregiver group). Moreover, 13% failed to maintain these criteria during the test blocks (9.6% for the middle-aged caregiver group

and 15.5% for the older adult caregiver group) and 11.4% did not perform the CARE-IRAP because they could not understand the instructions to complete the task (all older adults). In the published IRAP studies, there is great variability in attrition rates. Whereas in some studies, all participants reached the criteria during the practice blocks (e.g., Nicholson & Barnes-Holmes, 2012), others presented attrition rates similar to those of our study. For example, Drake et al. (2016) reported an attrition rate of 41.1%, and Chan, Barnes-Holmes, Barnes-Holmes, and Stewart (2009) reported an attrition rate of 29.1%. When we analyzed the possible causes for these high attrition rates, we found that Drake et al. (2016) used stricter criteria than other IRAP studies (i.e., reaching an accuracy $\geq 85\%$ and a median response time $\leq 2,000$), whereas Chan et al. (2009) employed a sample with a mean age of 37.6 but with an age range of 23–62 years. It may be possible that the inclusion of older adults in an IRAP study leads to a higher attrition rate. In fact, the only study (Rezende et al., 2020) that used the IRAP exclusively with older adults (mean age of 65.14 and age range of 59–79 years) required participants to reach an accuracy $\geq 80\%$ and a median response time $\leq 2,500$ and found a higher attrition rate than the present study (46%). Therefore, it seems that the use of strict criteria might increase attrition rates, especially in older adults. Despite the adjustments made to the IRAP, the attrition rate was still high in the present study; however, this rate could have been much higher if the accuracy and median response time had not been changed. Adapting the IRAP criteria thus seems to be particularly important for this population.

With regard to the attrition rates, caregivers that reached the required criteria in the practice blocks also had better performance in the rule shift cards task. Therefore, as hypothesized, there is a relationship between participants' performance in the IRAP and their switching skills (i.e., the ability to flexibly switch responses between different rules). Likewise, O'Toole and Barnes-Holmes (2009) found a significant relationship between participants' performance in the IRAP and intelligence. In the current study, switching skills were measured with a task that requires similar ability as the IRAP but uses neutral stimuli. Functional cognitive decline is normal as people age (e.g., Allain et al., 2005) and, as expected, older adult caregivers had lower levels of switching skills and higher attrition rates than middle-aged caregivers.

During the last years, there have been studies exploring the influence of different response options used in the IRAP (Maloney & Barnes-Holmes, 2016). In particular, studies have compared the use of response options of contextually related stimuli (e.g., similar/opposite, before/after, more than/less than) with relational coherence indicators (e.g., true/false, right/wrong, yes/no). The latter options cohere with patterns of relational responding in natural language within the verbal community. Maloney, Foody, and Murphy (2020) found an IRAP effect with both response options, but this effect was stronger when the response options were

contextually related. In the present study, we used “True” and “False” as the response options, and the findings demonstrated an IRAP effect similar to those found in other IRAP studies that have used relational coherence indicators (e.g., Drake et al., 2010; Hussey & Barnes-Holmes, 2012; Nicholson & Barnes-Holmes, 2012). However, future studies should explore the impact of these different response options in the IRAP.

The findings of the present study have potential clinical implications. First, the results highlight that caregivers present implicit dysfunctional beliefs about caregiving, and that these implicit beliefs are related to explicit dysfunctional beliefs and experiential avoidance. Although the present study does not analyze the dynamics of these associations, it is likely that the degree of correlation between explicit and implicit relational response varies greatly from one individual to another (e.g., Parling, Cernvall, Stewart, Barnes-Holmes, & Ghaderi, 2012). These individual differences may have important psychological implications. Furthermore, immediate relational responses (i.e., spontaneous and automatic evaluations), in contrast to carefully examined and elaborated responses, may present a different pattern of correlations with psychological and physical health indicators. The association between implicit and explicit beliefs and caregivers’ distress and health has not been explored in the present study, but this topic is a promising line of future research. Based on the results of the caregiving literature (e.g., Losada et al., 2010), it could be hypothesized that those caregivers with strong implicit dysfunctional beliefs may present high levels of emotional distress. In particular, implicit dysfunctional verbal relations may be more strongly related to psychosomatic processes and physical health indicators, as suggested by previous studies (e.g., Márquez-González, Cabrera, Losada, & Knight, 2018). In order to obtain more support for the validity of implicit dysfunctional verbal relations as measured by IRAP, studies are needed that explore the relationship between dementia caregivers’ implicit dysfunctional beliefs, as assessed with the CARE-IRAP, and mental (e.g., depression, anxiety, and burden) and physical (e.g., blood pressure) outcomes.

A second clinical implication is related to the use of the IRAP in clinical older adults. The prevalence of some mental health problems is considered to be high in older adults. For example, the estimated prevalence of anxiety disorders ranges from 3.2% to 14.2% (Wolitzky-Taylor, Castriotta, Lenze, Stanley, & Craske, 2010), and the prevalence of clinically significant depressive symptoms is 15% (Fiske, Wetherell, & Gatz, 2009) in this population. Considering the findings of the current study, future studies could explore the implicit components of psychological problems using the IRAP in this population, as has been done with young samples (e.g., Hussey & Barnes-Holmes, 2012; Kosnes et al., 2013).

However, suitable IRAP criteria for older adults should be further explored and established so that the task is more appropriate for the older population.

The main limitation of the present study was the small number of participants included in the analyses. This was due to the attrition rate, especially in the older adult caregiver group. The complexity of the IRAP and its high attrition rates may limit its application and the generalizability of the results of this study (Golijani-Moghaddam et al., 2013). Future studies should further investigate the effects of different IRAP performance criteria in order to minimize attrition rates, and continue exploring IRAP validity for older adult samples. Other implicit tasks that do not require participants to meet specific criteria, such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), may be more suitable for older people. However, the previous literature has shown that the use of strict criteria in the IRAP enables better capture of an implicit response compared to the IAT (Golijani-Moghaddam et al., 2013). The inclusion of a longer learning module in studies using IRAP may be a way to reduce the attrition rate as it may help participants to improve accuracy and to achieve faster response latencies. Along these lines, Kishita, Muto, Ohtsuki, and Barnes-Holmes (2014) and Vahey et al. (2010) included a preparation-IRAP using neutral stimuli to familiarize participants with the task before the second IRAP task, which assessed the study-specific implicit belief. Another limitation is the low to moderate internal reliability of the different D_{IRAP} scores found in this study. Although this internal reliability was similar to (e.g., Barnes-Holmes, Murtagh et al., 2010; Drake et al., 2015) or even higher (e.g., Drake et al., 2016) than previous IRAP studies, it is expected that once IRAP performance criteria for older adults are established, internal reliability would improve.

In spite of the above-mentioned limitations, this is the first study in which the IRAP has been used with dementia family caregivers, and its results offer preliminary support for the validity of the CARE-IRAP as a measure of implicit dysfunctional beliefs in dementia family caregivers. Moreover, this exploratory study is the first to adapt IRAP criteria to older adults. Even though some adjustments were made, the attrition rate in the older adult group was higher, as compared to the middle-aged group. Therefore, future studies should further explore the suitable criteria and instructional procedure for this population. The final goal is to find a balance between a low attrition rate and a time constraint that allows implicit immediate relational responses to be captured with the IRAP in older adults.

Acknowledgments We thank all the caregivers for their participation in the study and the following centers for collaborating with us in the project: AFAMSO and María Wolff.

Funding This study was funded by the Spanish Ministry of Economy and Competitiveness under Grants PSI2015-65152-C2-2-R and PSI2015-65152-C2-1-R.

Data availability The data that support the findings of this study are available from the corresponding author.

Compliance with ethical standards

Conflict of Interest The authors declare that there are no conflicts of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee of the Universidad Rey Juan Carlos (reference number: 060720166616) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

- Allain, P., Nicoleau, S., Pinon, K., Etcharry-Bouyx, F., Barré, J., Berrut, G., et al. (2005). Executive functioning in normal aging: A study of action planning using the Zoo Map Test. *Brain & Cognition*, *57*(1), 4–7. <https://doi.org/10.1016/j.bandc.2004.08.011>.
- Ávila, A., & Tomé, M. C. (1989). Evaluación de la deseabilidad social y correlatos defensivos emocionales. Adaptación castellana de la Escala de Crowne y Marlowe [Assessment of social desirability and emotional defensive correlates. Spanish adaptation of the Crowne and Marlowe Scale]. In A. Echevarría & D. Páez (Eds.), *Emociones: Perspectivas psicosociales* [Emotions: Psychosocial perspectives] (pp. 505–514). Madrid, Spain: Fundamentos.
- Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*(7), 169–177.
- Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, *60*(3), 527–542.
- Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit Relational Assessment Procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record*, *60*(1), 57–79.
- Barnes-Holmes, D., Murtagh, L., Barnes-Holmes, Y., & Stewart, I. (2010). Using the Implicit Association Test and the Implicit Relational Assessment Procedure to measure attitudes toward meat and vegetables in vegetarians and meat-eaters. *The Psychological Record*, *60*(2), 287–305.
- Chan, G., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). Implicit attitudes to work and leisure among North American and Irish individuals: A preliminary study. *International Journal of Psychology & Psychological Therapy*, *9*(3), 317–334.
- Charles, S. T., & Carstensen, L. L. (2010). Social and emotional aging. *Annual Review of Psychology*, *61*, 383–409. <https://doi.org/10.1146/annurev.psych.093008.100448>.
- Chawla, N., & Ostafin, B. (2007). Experiential avoidance as a functional dimensional approach to psychopathology: An empirical review. *Journal of Clinical Psychology*, *63*(9), 871–890. <https://doi.org/10.1002/jclp.20400>.
- Corcoran, M. A. (2011). Caregiving styles: A cognitive and behavioral typology associated with dementia family caregiving. *The Gerontologist*, *51*(4), 463–472. <https://doi.org/10.1093/geront/gnr002>.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349. <https://doi.org/10.1037/h0047358>.
- Cullen, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). The Implicit Relational Assessment Procedure (IRAP) and the malleability of ageist attitudes. *The Psychological Record*, *59*(4), 591–620.
- Drake, C. E., Kellum, K. K., Wilson, K. G., Luoma, J. B., Weinstein, J. H., & Adams, C. H. (2010). Examining the implicit relational assessment procedure: Four preliminary studies. *The Psychological Record*, *60*(1), 81–100.
- Drake, C. E., Kramer, S., Habib, R., Schuler, K., Blankenship, L., & Locke, J. (2015). Honest politics: Evaluating candidate perceptions for the 2012 US election with the Implicit Relational Assessment Procedure. *Journal of Contextual Behavioral Science*, *4*(2), 129–138. <https://doi.org/10.1016/j.jcbs.2015.04.004>.
- Drake, C. E., Timko, C. A., & Luoma, J. B. (2016). Exploring an implicit measure of acceptance and experiential avoidance of anxiety. *The Psychological Record*, *66*(3), 463–475.
- Espinosa, A. N. A., Alegret, M., Boada, M., Vinyes, G., Valero, S., Martínez-Lage, P., et al. (2009). Ecological assessment of executive functions in mild cognitive impairment and mild Alzheimer's disease. *Journal of the International Neuropsychological Society*, *15*(5), 751–757. <https://doi.org/10.1017/S15561770999035X>.
- Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in older adults. *Annual Review of Clinical Psychology*, *5*, 363–389. <https://doi.org/10.1146/annurev.clinpsy.032408.153621>.
- Gallagher-Thompson, D., Solano, N., Coon, D., & Areán, P. (2003). Recruitment and retention of Latino dementia caregivers in intervention research: Issues to face, lessons to learn. *The Gerontologist*, *43*, 45–51. <https://doi.org/10.1093/geront/43.1.45>.
- Gallego-Alberto, L., Márquez-González, M., Romero-Moreno, R., Cabrera, I., & Losada, A. (2019). Pilot study of a psychotherapeutic intervention for reducing guilt feelings in highly distressed dementia family caregivers (Innovative practice). *Dementia*. Advance online publication. <https://doi.org/10.1177/1471301219886761>.
- Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 283–310). New York, NY: Cambridge University Press.
- Golijani-Moghaddam, N., Hart, A., & Dawson, D. L. (2013). The implicit relational assessment procedure: Emerging reliability and validity data. *Journal of Contextual Behavioral Science*, *2*(3–4), 105–119. <https://doi.org/10.1016/j.jcbs.2013.05.002>.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality & Social Psychology*, *74*(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality & Social Psychology*, *85*(2), 197–216. <https://doi.org/10.1037/0022-3514.85.2.197>.

- Hayes, S. C., & Gifford, E. V. (1997). The trouble with language: Experiential avoidance, rules, and the nature of verbal events. *Psychological Science*, 8(3), 170–173. <https://doi.org/10.1111/j.1467-9280.1997.tb00405.x>.
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and commitment therapy: The process and practice of mindful change*. New York, NY: Guilford Press.
- Hopkinson, M. D., Reavell, J., Lane, D. A., & Mallikarjun, P. (2018). Cognitive behavioral therapy for depression, anxiety, and stress in caregivers of dementia patients: A systematic review and meta-analysis. *The Gerontologist*, 59(4), 343–362. <https://doi.org/10.1093/geront/gnx217>.
- Hussey, I., & Barnes-Holmes, D. (2012). The implicit relational assessment procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive & Behavioral Practice*, 19(4), 573–582. <https://doi.org/10.1016/j.cbpra.2012.03.002>.
- Hussey, I., Thompson, M., McEnteggart, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). Interpreting and inverting with less cursing: A guide to interpreting IRAP data. *Journal of Contextual Behavioral Science*, 4(3), 157–162. <https://doi.org/10.1016/j.jcbs.2015.05.001>.
- Kishita, N., Muto, T., Ohtsuki, T., & Barnes-Holmes, D. (2014). Measuring the effect of cognitive defusion using the Implicit Relational Assessment Procedure: An experimental analysis with a highly socially anxious sample. *Journal of Contextual Behavioral Science*, 3(1), 8–15. <https://doi.org/10.1016/j.jcbs.2013.12.001>.
- Knight, B. G., & Sayegh, P. (2010). Cultural values and caregiving: the updated sociocultural stress and coping model. *Journals of Gerontology. Series B, Psychological Sciences & Social Sciences*, 65B, 5–13. <https://doi.org/10.1093/geronb/gbp096>.
- Kosnes, L., Whelan, R., O'Donovan, A., & McHugh, L. A. (2013). Implicit measurement of positive and negative future thinking as a predictor of depressive symptoms and hopelessness. *Consciousness & Cognition*, 22(3), 898–912. <https://doi.org/10.1016/j.concog.2013.06.001>.
- Losada, A., Márquez-González, M., Knight, B. G., Yanguas, J., Sayegh, P., & Romero-Moreno, R. (2010). Psychosocial factors and caregivers' distress: Effects of familism and dysfunctional thoughts. *Aging & Mental Health*, 14(2), 193–202. <https://doi.org/10.1080/13607860903167838>.
- Losada, A., Márquez-González, M., Romero-Moreno, R., & López, J. (2014). Development and validation of the Experiential Avoidance in Caregiving Questionnaire (EACQ). *Aging & Mental Health*, 18(7), 897–904. <https://doi.org/10.1080/13607863.2014.896868>.
- Losada, A., Márquez-González, M., Romero-Moreno, R., Mausbach, B. T., López, J., Fernández-Fernández, V., & Nogales-González, C. (2015). Cognitive-behavioral therapy (CBT) versus acceptance and commitment therapy (ACT) for dementia family caregivers with significant depressive symptoms: Results of a randomized clinical trial. *Journal of Consulting & Clinical Psychology*, 83(4), 760. <https://doi.org/10.1037/ccp0000028>.
- Losada, A., Montorio, I., Izal, M., & Márquez-González, M. (2006). *Estudio e intervención sobre el malestar psicológico de los cuidadores de personas con demencia. El papel de los pensamientos disfuncionales* [Study and intervention on dementia caregiving distress: The role of dysfunctional thoughts]. Madrid, Spain: IMSERSO.
- Maloney, E., & Barnes-Holmes, D. (2016). Exploring the behavioral dynamics of the Implicit Relational Assessment Procedure: The role of relational contextual cues versus relational coherence indicators as response options. *The Psychological Record*, 66(3), 395–403. <https://doi.org/10.1007/s40732-016-0180-5>.
- Maloney, E., Foody, M., & Murphy, C. (2020). Do response options in the Implicit Relational Assessment Procedure (IRAP) matter? A comparison of contextual relations versus relational coherent indicators. *The Psychological Record*, 70(2), 205–214. <https://doi.org/10.1007/s40732-019-00360-7>.
- Márquez-González, M., Cabrera, I., Losada, A., & Knight, B. G. (2018). Attentional avoidant biases as mediators in the association between experiential avoidance and blood pressure in dementia family caregivers. *Aging & Mental Health*, 22(5), 669–677. <https://doi.org/10.1080/13607863.2017.1293003>.
- Márquez-González, M., Losada, A., Izal, M., Pérez-Rojo, G., & Montorio, I. (2007). Modification of dysfunctional thoughts about caregiving in dementia family caregivers: Description and outcomes of an intervention programme. *Aging & Mental Health*, 11(6), 616–625. <https://doi.org/10.1080/13607860701368455>.
- McKenna, I. M., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology & Psychological Therapy*, 7, 123–138.
- McNaughton, M. E., Patterson, T. L., Smith, T. L., & Grant, I. (1995). The relationship among stress, depression, locus of control, irrational beliefs, social support, and health in Alzheimer's disease caregivers. *Journal of Nervous & Mental Disease*, 183, 78–85. <https://doi.org/10.1097/00005053-199502000-00003>.
- Mittelman, M. S., Epstein, C., & Pierzchala, A. (2003). *Counseling the Alzheimer caregiver*. Chicago, IL: American Medical Association.
- Nicholson, E., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure (IRAP) as a measure of spider fear. *The Psychological Record*, 62(2), 263–277.
- Nolle, A. P., Gulbas, L., Kuhlberg, J. A., & Zayas, L. H. (2012). Sacrifice for the sake of the family: expressions of familism by Latina teens in the context of suicide. *American Journal of Orthopsychiatry*, 82(3), 319–327. <https://doi.org/10.1111/j.1939-0025.2012.01166.x>.
- Norris, G., & Tate, R. L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation*, 10(1), 33–45. <https://doi.org/10.1080/096020100389282>.
- Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, 134, 565–584. <https://doi.org/10.1037/0096-3445.134.4.565>.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The IAT at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York, NY: Psychology Press.
- O'Toole, C., & Barnes-Holmes, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record*, 59(1), 119–132.
- Parling, T., Cernvall, M., Stewart, I., Barnes-Holmes, D., & Ghaderi, A. (2012). Using the implicit relational assessment procedure to compare implicit pro-thin/anti-fat attitudes of patients with anorexia nervosa and non-clinical controls. *Eating Disorders*, 20(2), 127–143. <https://doi.org/10.1080/10640266.2012.654056>.
- Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M. A., & De Raedt, R. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self-versus ideal self-related cognitions in dysphoria. *Cognition & Emotion*, 27(8), 1441–1449. <https://doi.org/10.1080/02699931.2013.786681>.
- Rezende, M. C., Bast, D. F., Huziwarra, E. M., & Bortoloti, R. (2020). The Implicit Relational Assessment Procedure as a measure of hopelessness in the elderly: An exploratory study. *The Psychological Record*. Advance online publication. <https://doi.org/10.1007/s40732-020-00406-1>
- Roach, L. V. (2013). *Guilt, dysfunctional thought processes and depression in caregivers of people with dementia* [Doctoral dissertation, The University of Edinburgh]. Edinburgh Research Archive.
- Sabogal, F., Marin, G., Otero-Sabogal, R., Marin, B. V., & Perez-Stable, E. J. (1987). Hispanic familism and acculturation: What changes and

- what doesn't? *Hispanic Journal of Behavioral Sciences*, 9(4), 397–412. <https://doi.org/10.1177/07399863870094003>.
- Sallim, A. B., Sayampanathan, A. A., Cuttilan, A., & Ho, R. C. M. (2015). Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. *Journal of the American Directors Association*, 16, 1034–1041. <https://doi.org/10.1016/j.jamda.2015.09.007>.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103(3), 403.
- Salthouse, T. A., & Davis, H. P. (2006). Organization of cognitive abilities and neuropsychological variables across the lifespan. *Developmental Review*, 26(1), 31–54. <https://doi.org/10.1016/j.dr.2005.09.001>.
- Stebbins, P., & Pakenham, K. I. (2001). Irrational schematic beliefs and psychological distress in caregivers of people with traumatic brain injury. *Rehabilitation Psychology*, 46, 178–194. <https://doi.org/10.1037/0090-5550.46.2.178>.
- Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology*, 28(2), 191–193. [https://doi.org/10.1002/1097-4679\(197204\)28:2<191::AID-JCLP2270280220>3.0.CO;2-G](https://doi.org/10.1002/1097-4679(197204)28:2<191::AID-JCLP2270280220>3.0.CO;2-G).
- Sullivan, K. A., Beattie, E., Khawaja, N. G., Wilz, G., & Cunningham, L. (2016). The Thoughts Questionnaire (TQ) for family caregivers of people with dementia. *Dementia*, 15(6), 1474–1493. <https://doi.org/10.1177/1471301214553038>.
- Tandetnik, C., Hergueta, T., Negovanska, V., Dubois, B., & Bungener, C. (2014). Traduction française du «Dysfunctional Thoughts about Caregiving Questionnaire-DTCQ». Étude préliminaire auprès d'aidants de personnes atteintes de maladies neurodégénératives [French translation of the “Dysfunctional Thoughts about Caregiving Questionnaire.” A pilot study among caregivers of patients with neurodegenerative diseases]. *Journal de Thérapie Comportementale et Cognitive*, 24(1), 5–13. <https://doi.org/10.1016/j.jtcc.2013.11.001>.
- Törneke, N., Luciano, C., & Salas, S. V. (2008). Rule-governed behavior and psychological problems. *International Journal of Psychology & Psychological Therapy*, 8, 141–156.
- United Nations. (2019). *World population ageing 2019: Highlights*. New York, NY: Department of Economic and Social Affairs, Population Division, United Nations.
- Vahey, N., Boles, S., & Barnes-Holmes, D. (2010). Measuring adolescents' smoking-related social identity preferences with the Implicit Relational Assessment Procedure (IRAP) for the first time: A starting point that explains later IRAP evolutions. *International Journal of Psychology & Psychological Therapy*, 10(3), 453–474.
- Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy & Experimental Psychiatry*, 48, 59–65. <https://doi.org/10.1016/j.jbtep.2015.01.004>.
- Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *The behavioural assessment of the dysexecutive syndrome*. London, UK: Thames Valley Company.
- Wolitzky-Taylor, K. B., Castriotta, N., Lenze, E. J., Stanley, M. A., & Craske, M. G. (2010). Anxiety disorders in older adults: a comprehensive review. *Depression & Anxiety*, 27(2), 190–211. <https://doi.org/10.1002/da.20653>.

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