# Life is good, but death ain't bad either: Counter-intuitive implicit biases to death in a normative population

Ian Hussey[1], Tarah Daly[1], & Dermot Barnes-Holmes[2]

[1]National University of Ireland Maynooth, Maynooth; [2]Gent University, Belgium

*Accepted manuscript – The Psychological Record*

## Author Note

IH was assisted by a Government of Ireland scholarship from the Irish Research Council.

Correspondence should be addressed to Ian Hussey, Department of Psychology, John Hume building, National University of Ireland Maynooth, Kildare, Republic of Ireland.

ian.hussey@nuim.ie

Abstract

The current study explored implicit attitudes to life and death in a student population using both the Implicit Association Test (IAT) and the Implicit Relational Assessment Procedure (IRAP). The IAT was similar to one used in previously published researched in the context of the prospective prediction of suicide and self-harm. Two IRAPs were employed, one that assessed relational responses specific to death and life with respect to self, and a second that assessed relational responses specific to evaluations of death and life. The IAT replicated previous results found in normative populations. The IRAPs indicated "prolife" biases, as expected. However, they also failed to demonstrate the presence of strong "antideath" biases, and in one case a specific "death–positive" bias was found. The results observed on the explicit measures did not readily explain the absent or "prodeath" effects observed on the IRAPs. Indeed, participants reported a normative level of anxiety and fear of death. Implications for the study of implicit attitudes to death using the IRAP are considered.

*Keywords*: Implicit Relational Assessment Procedure, Relational Frame Theory, death

# Life is good, but death ain't bad either: Counter-intuitive implicit biases to death in a normative population

The concept of death plays a central role in much of human culture, including religion and the search for meaning in life (see Kastenbaum, 2000; Neimeyer, Wittkowski, & Moser, 2004). Indeed, a body of research demonstrates that reminders of our own mortality impact our behavior in important ways, such as on our conceptualization of self and personal values (see Burke, Martens, & Faucher, 2010 for meta analysis). From a behavior-analytic perspective, the conceptualization of death is inherently interesting, because death itself cannot be experienced or consequated and is instead constructed through the metaphor of sleep. It is therefore somewhat difficult to account for our conceptualization of death in terms of direct contingencies (Hayes, 1992). Relational Frame Theory (RFT: Hayes, Barnes-Holmes, & Roche, 2001) represents an attempt to account for such "emergent" or "derived" responding that occur in the absence of a direct history of reinforcement (see Hughes & Barnes-Holmes, in press for overview).

RFT's core concept of "arbitrarily applicable relational responding" has demonstrated much utility in *conceptualizing* and *modeling* complex verbal behaviors, such as death and suicide (Hayes, 1992), self (Foody, Barnes-Holmes, & Barnes-Holmes, 2012; McHugh, Barnes-Holmes, & Barnes-Holmes, 2004) and metaphor (Foody et al., 2014). However, more recently, RFT researchers have also become increasingly interested in *assessing* such relational responses "in flight", as they are emitted (see Hussey, Barnes-Holmes, & Barnes-Holmes, 2015). Biases in the strength of relational responding are more frequently referred to as "implicit attitudes" (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009), and are assessed using such procedures as the Implicit Association Test (IAT: Greenwald, McGhee, & Schwartz, 1998) and the Implicit

Relational Assessment Procedure (IRAP: Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). The IRAP, in particular, was created with the explicit intention of assessing the strength or persistence of relational responding.

A recent meta-analysis demonstrated that the IRAP has been used to explore a variety of clinically relevant domains, including self-esteem, depression, OCD, and substance abuse (Vahey, Nicholson, & Barnes-Holmes, 2015). This study represents the first study to explore the utility of the IRAP in exploring attitudes to death. While we initially planned to group together a series of studies on this topic for publication, our initial findings produced a counter-intuitive effect that seemed worth of dissemination at this point. Additionally, as will be elaborated upon in the discussion, our results may have implications for a more systematic analysis of the IRAP as a measure of strength of relational responding itself.

Previous research has shown that IATs that target implicit attitudes to self and death (i.e., using the categories self, others, death, & life) are prospectively predictive of self-harm and suicide attempts over and above established risk factors. These include clinical judgment, self-report measures of impulsivity and hopelessness, a variety of routine risk assessment tools such as the Manchester Self-Harm Rule (Cooper et al., 2006) and the SAD PERSONS assessment (Patterson, Dohn, Bird, & Patterson, 1983), and the individual's own self forecast (see Nock et al., 2010). However, it is important to note that the IAT does not permit the identification of individual, independent response biases. Indeed, the creators of the IAT have been assiduous in specifying that it is a measure of the relative (rather than absolute) strength of associations between categories. This has been argued for both conceptually (e.g., Greenwald et al., 1998; Nosek, Greenwald, & Banaji, 2007) and empirically (e.g., Nosek, Greenwald, & Banaji, 2005; Pinter & Greenwald, 2005) on numerous occasions. For example, these authors point out that an

IAT that includes the stimuli "self", "others", "life", and "death" (e.g., Nock et al., 2010) must be interpreted as a measure of the relative strength of associations between all four categories. For example, while somewhat verbose, these biases should be interpreted as either towards "self and life relative to others and death" or "self and death relative to others and life". Specifically, whereas the IAT presents participants with all four categories on each trial (e.g., self, other, life, and death) and examines the relative ease of categorization (e.g., self–life and others–death vs. self–death and others–life), the IRAP only ever presents exemplars from one category in a pair (e.g., either self or others and life or death) on each trial, and requires participants to respond in opposing directions across blocks (e.g., "similar" vs. "different"). As such, four separate bias scores are produced, one for each "trial-type" (e.g., self–life, self–death, others–life, and others–death). This ability to separate out the strength of individual category pairings may serve to uncover subtle effects that might be obscured within the IAT's overall bias score (Barnes-Holmes, Barnes-Holmes, et al., 2010).

This study therefore tests the assumption made with previous research, whether explicitly or implicitly, that normative participants do indeed demonstrate specific "self–life" and/or "self–*not*-death" biases (i.e., Dickstein et al., 2015; Harrison, Stritzke, Fay, Ellison, & Hudaib, 2014; Nock et al., 2010; Price, Nock, Charney, & Mathew, 2009; Price et al., 2014; Randall et al., 2013; Tang, Wu, & Miao, 2013; Violanti, Mnatsakanova, & Andrew, 2013). Participants completed both a death–identity IAT and a death–identity IRAP that was created from the same stimuli. It was expected that a sample of normative participants would demonstrate overall "self–life/others–death" effects on the IAT, but that effects on the IRAP would load onto a two specific trial-types: an assertion of "self–life" and a rejection of "self–death".

A second IRAP was also included to explore implicit evaluations of death. While previous research on implicit attitudes to death has explored the associations between the concepts of (a) self and evaluation (i.e., "self-esteem": Creemers, Scholte, Engels, Prinstein, & Wiers, 2013; Dickstein et al., 2015; Franck, De Raedt, Dereu, & Van den Abbeele, 2007; Glashouwer et al., 2010; Price et al., 2009; Price et al., 2014); and (b) self and death (i.e., "death–identity": Dickstein et al., 2015; Harrison et al., 2014; Nock et al., 2010; Price et al., 2009; Price et al., 2014; Randall et al., 2013; Tang et al., 2013; Violanti et al., 2013); no work has explored the association between death and evaluation. This is somewhat surprising, given the centrality of evaluations within many psychological theories that focus on death. For example, Terror Management Theory (Greenberg, Pyszczynski, & Solomon, 1986) is predicated on the assumption that awareness of ones own mortality is highly aversive, and argues that this serves as the motivator for humans' ubiquitous need for meaning and self-esteem. Comparably, many theories of non-normative attitudes to death (e.g., suicidality) attempt to account for how life and/or self acquires aversive properties in the face of unbearable psychological suffering, often with explicit reference to an acquired loss of fear of death (e.g., Interpersonal Theory of Suicide: Joiner, 2005; Integrated Motivational-Volitional Model of Suicidal Behavior: O'Connor, 2011). This idea that death is aversive or evaluated negatively within 'normative' individuals, and that decreased fear of death is associated with suicidality, has been supported by a relatively large number of studies using self-report questionnaires (e.g., Ribeiro et al., 2014). In summary, given the centrality of evaluations of death to many theories, and the recent emphasis on the relative utility of implicit measures to traditional self-report methods (see Randall, Colman, & Rowe, 2011), it is therefore somewhat surprising that no research to date has examined evaluations of

death using implicit measures. Consequently, a second IRAP was included, that assessed evaluations of death (i.e., life, death, positive, & negative).

Finally, in addition to the death–identity IAT, death–identity IRAP and death–evaluation IRAP, a number of self-report measures were included in order to explore the self-report correlates of such implicit attitudes. Self-report measures of depression and hopelessness were included, due to their known association with non-normative attitudes to death (i.e., suicidality: see Brown, Beck, Steer, & Grisham, 2000). A number of additional exploratory self-report measures were included to assess beliefs in the afterlife, fear of death, and the relationship between an individual and their aversive mental content.

## Method

### Participants

Forty-two undergraduate students (25 female, 17 male) aged between 18 and 51 years old ($M$ = 27.8, $SD$ = 9.6) were recruited from the student population at the National University of Ireland Maynooth. Participants completed the study individually in an experimental cubicle in the Department of Psychology. Inclusion criteria were self-reported fluent English, normal or corrected to normal vision, age 18–65, and full use of both hands. Participants reported having completed between 0 and 9 IRAPs previous to the current study ($M$ = 2.60, $SD$ = 3.08). It should be noted that, due to ethical approval constraints, self-reports of history of suicidal behavior were not collected. As such, the current study was conducted within a normative sample with an uncertain history of suicidal behaviors, rather than a strictly “non-suicidal” sample. A range of self-report measures was employed to establish the normativity of the sample (see below).

**Measures**

Self-report measures included an assessment of individuals' attitudes to death and dying using the Death Anxiety Scale (Templer, 1970), and belief in what comes after death using the Belief in the Afterlife Scale (Osarchuk & Tatz, 1973). Both depressive symptoms and hopelessness were assessed, based on their known relationship with suicidal behaviors, using the Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995) and Beck Hopelessness Scale (Beck, Weissman, Lester, & Trexler, 1974), respectively. Finally, we included the Acceptance and Action Questionnaire II, a measure of psychological flexibility, as an exploratory measure (Bond et al., 2011).

**Acceptance and Action Questionnaire II.** The AAQ-II is a 7-item measure of psychological flexibility (Bond et al., 2011). That is, the ability to change or persist in behavior in the service of valued ends (e.g., "Worries get in the way of my success"). Psychological flexibility is a key process within Acceptance and Commitment Therapy (Hayes, Strosahl, & Wilson, 1999). Participants respond on a 7-point Likert scale from 1 (never true) to 7 (always true), with possible scores ranging from 7 to 49. Internal consistency was excellent in the current sample (Cronbach's $\alpha$ = .90).

**Beck Hopelessness Scale.** The Beck Hopelessness Scale is a 20-item self-report measure of an individual's hopelessness over the past week (e.g., "My future seems dark to me"), and has been shown to be a significant predictor of death by suicide in longitudinal studies (Beck, Steer, Kovacs, & Garrison, 1985; Beck et al., 1974). Each item has a binary (True/False) response format and is scored from 0–20, where higher scores represent greater levels of hopelessness. Internal consistency was good ($\alpha$ = .86).

**Belief in the Afterlife Scale.** This 7-item self-report questionnaire assesses individuals' beliefs in the afterlife (e.g., "There must be an afterlife of some sort": Osarchuk & Tatz, 1973). Participants respond using a 1 (disagree) to 10 (agree) scale and is scored from 0 to 70, where higher scores represent greater belief in the afterlife. It should be noted that specific beliefs about the afterlife (e.g., whether it will involve punishment or reward) are not examined, only whether participants believe death to be extinction of self. Internal consistency was good in the current sample ($\alpha$ = .89).

**Death Anxiety Scale.** This 17-item self-report questionnaire assesses fear and anxiety around the act of dying and the finality of death using items such as "I am very much afraid to die" and "I often think about how short life really is" (Templer, 1970). It employs a true/false answer format. This scale was employed over other measures of death anxiety on the basis that it is one of the most frequently employed in previous research (see Iverach, Menzies, & Menzies, 2014 for review). Internal consistency was found to be good in the current sample ($\alpha$ = .70).

**Depression Anxiety and Stress Scale-21.** This 7-item depression subscale asks participants about depressive symptoms in the past week (e.g., "I felt that life was meaningless"), and uses a 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time) response format. This subscale has been shown to correlate highly with other well-established depression scales such as the Beck Depression Inventory II ($r$ = .74: Lovibond & Lovibond, 1995). Internal consistency was excellent in the current sample ($\alpha$ = .92).

**Death Identity IAT.** The Implicit Association Test (Greenwald et al., 1998) is a computer-based test that assesses reaction time biases. The current study employed the death–identity stimulus set developed by Nock and colleagues (2010) that has been employed in multiple previous studies (e.g., Dickstein et al., 2015; Harrison et al., 2014; Price et al., 2014,

2009; Randall et al., 2013; Tang et al., 2013; Violanti et al., 2013). One small modification was made to the stimulus set: due to ethical constraints on asking university students about suicide, and in order to assess attitudes to death generally rather than suicide specifically, the stimulus "suicide" was substituted for the word "death". The stimuli employed in the four categories (i.e., self, others, life, and death) are presented in Table 1. All procedural details of the IAT were standard (see Nosek et al., 2007 for methodological review).

**IRAPs.** The structure of the IRAP employed in the current study was as described by Barnes-Holmes, Barnes-Holmes and Stewart (Barnes-Holmes, Barnes-Holmes, et al., 2010); see below). The task was programmed in Visual Basic 6, and the "2012" version of the program was used (Barnes-Holmes & Hussey, 2012). Several parameters were identical across both IRAPs, and thus will be summarized here. Participants were provided with up to 4 pairs of practice blocks. If participant met both the mastery criteria on both blocks in a pair of practice blocks (i.e., accuracy ≥ 80% and median latency ≤ 2000ms) they were immediately moved to the test blocks. If they failed to meet the mastery criteria after four practice block pairs, the task ended. In accordance with the majority of previous research, three pairs of test block pairs were completed (e.g., Hussey & Barnes-Holmes, 2012; Rönspies et al., 2015). The latency feedback message was set to "!". The location of the response options remained static throughout the task.

***Death–Identity IRAP.*** The stimuli employed within the IRAP were drawn from those used within the IAT in order to make the two measures maximally comparable. The IRAP program required an even number of stimuli per category. As such, one additional stimulus was added to each category (i.e., "me", "others", dead" and "living": see Table 1). The four trial-types were therefore "self–death", "self–life", "others–death", and "others–life". The response options were set to "Similar" and "Different". Specific responding rules were presented on

screen before each block. Responding rule A was "Please answer as if you associate yourself with death and others with life," and responding rule B was "Please answer as if you associate yourself with life and others with death". The order in which individuals were exposed to these blocks (i.e., rule A first vs. rule B first) was counterbalanced between participants in both IRAPs.

Table 1. *Stimuli used in the Death–Identity IAT and IRAP*

| Label 1: self | Label 2: others | Target 1: death | Target 2: life |
|---|---|---|---|
| Myself | Them | Death[1] | Alive |
| My | They | Die | Live |
| Mine | Theirs | Funeral | Thrive |
| I | Their | Lifeless | Survive |
| Self | Other | Deceased | Breathing |
| Me[2] | Others[2] | Dead[2] | Living[2] |

*Note.* [1]Substituted the word "Suicide" as used by Nock et al. (2010); [2]Used in IRAP only.

***Death–Evaluation IRAP.*** This IRAP attempted to target relational responding around the valence of life and death. Labels 1 and 2 were set to "living" and "dying", respectively, target 1 stimuli were set to positive-valence-high-arousal words (i.e., enjoyable, exciting, lovely, great, pleasant, and satisfying), and target 2 stimuli were set to negative-valence-high-arousal words (i.e., awful, distressing, hurtful, horrible, painful: see Table 2). The four trial-types were therefore "life–positive", "life–negative", "death–positive", and "death–negative". The response options were set to "True" and "False". These words were selected following consultation of the Affective Norms for English Words battery (ANEW: Bradley & Lang, 1999) and the agreement of two researchers acquainted with the research area. Responding rule A was "Please answer as if life is pleasant and death is painful" and responding rule B was "Please answer as if life is painful and death is pleasant". Due to their greater complexity, the internal consistencies of the two IRAPs are reported in the results section.

Table 2. *Stimuli used in the Death–Evaluation IRAP*

| Label 1: life | Label 2: death | Target 1: positive | Target 2: negative |
|---|---|---|---|
| Living | Dying | Enjoyable | Awful |
| | | Exciting | Distressing |
| | | Great[1] | Hurtful |
| | | Lovely | Horrible |
| | | Pleasant | Painful |
| | | Satisfying | Upsetting |

*Note:* Label stimuli appear at the top of the screen and target stimuli in the middle of the screen.
[1] Vernacular usage of the word "great" in Ireland is most frequently as a synonym of positively valenced words such as "excellent", rather than denoting quantity or rank (cf. "substantial" and "prominent").

**Procedure**

All experimental sessions were conducted in individual experimental cubicles. This was done in a one-to-one setting with a trained researcher. Written informed consent was obtained from the participant prior to participation, followed by a verbal assessment by the researcher of all inclusion criteria. Participants completed the self-report measures first, followed by the death–identity IAT, death–identity IRAP, and death–evaluation IRAP. It should be noted that previous research suggests that the order of completion of self-report and implicit measures has minimal impact on the results of either (Nosek et al., 2005), therefore we elected to follow the common convention of presenting the self-report measures first. The order of the three implicit measures was fully counterbalanced between participants. Upon completion of all tasks, participants were fully debriefed and thanked for their time. No remuneration was offered.

**Death–Identity IAT**

The IAT consisted of two instruction screens followed by seven blocks, each of which consisted of a number of trials. The pre-block instruction screens contained the following written instructions:

"For this portion of the study, words will appear one at a time in the middle of the screen. Classify those words into groups which will be designated with labels appearing on the top half of the screen. All words belonging to the groups on the left will be classified with the "e" key. All words belonging to the groups on the right will be classified with the "i" key. Classify the words as quickly as possible while making as few mistakes as possible. Accuracy and speed are both important. Pay close attention to the group labels, they will change from block to block. Direct any questions to the experimenter."

"For the next portion of this study, you will be asked to classify words into the categories of DEATH and LIFE, as well as words related to ME and NOT ME. The words related to each of the categories are shown below. Remember, when the word in the center corresponds to the category on the left, you will use the "e" key, and when the word in the center corresponds to the category on the right, you will use the "i" key. Classify the words as quickly as possible while making as few mistakes as possible."

Table 3. *The length and content of the IAT blocks (adapted from Nosek et al., 2007, p.268).*

| Block | Number of trials | Categories assigned to the "e" key | Categories assigned to the "i" key |
|---|---|---|---|
| 1 | 20 | Self | Other |
| 2 | 20 | Death | Life |
| 3 | 20 | Self + death | Other + life |
| 4 | 40 | Self + death | Other + life |
| 5 | 40 | Other | Self |
| 6 | 20 | Other + death | Self + life |
| 7 | 40 | Other + death | Self + life |

*Note:* The order of presentation of the category pairings in blocks 3, 4, 6 and 7 (i.e., self–death/others–life vs. self–life/others–death) was counterbalanced between participants.

Before each block, the message "Check categories - Press space bar when ready" appeared at the bottom of the screen. A trial was defined as the time in milliseconds from the onset of a stimulus to the emission of a correct response. The numbers of trials in each block and the classes of stimuli presented in each are presented in Table 3. The stimulus categories (self, other, death, life) remained on the top left and top right of the screen throughout each block. Self and other related words were presented in white, whereas death and life related words were presented in green. Each trial presented the to be categorized stimulus in the middle of the screen. Participants responded using the "e" and "i" keys. If an incorrect response was emitted, a red "X" was displayed below the stimulus, and a correct response was required before proceeding to the next trial. After each trial the stimulus in the middle of the screen was cleared for an inter-trial interval of 250ms.

**IRAPs**

Participants were verbally instructed in how to complete the IRAP in several stages using a prewritten script (Hussey, 2015), which was developed as part of an ongoing effort to minimize attrition rates from the task. Therefore, in contrast to much previously published research (e.g., Nicholson & Barnes-Holmes, 2012), no additional written or onscreen instructions were provided. The experimenter's verbal instructions for the death–identity IRAP contained the following key points, which were delivered before the participant completed the first practice block. If a participant indicated a lack of clarity around any point, as the researcher worked through the script, that point was reiterated and clarified to the participant's satisfaction. The instructions for the death–evaluation IRAP were identical other than the specific stimuli that were referred to.

1. Participants were instructed that they would be presented with pairs of words related to “self”, “others”, “death” and “life”, and would be asked to respond to those pairs as being “Similar” or “Different”.
2. They were informed that, unlike a questionnaire that asked for their subjective opinion, this behavioral task simply required that they follow a rule, and this rule would be provided on screen.
3. Next, they were instructed that the rule would swap after each block, that there were only two rules, and that they would be reminded of the rule for the following block on screen.
4. It was emphasized that they were to initially go as *slowly* as they needed to get as many trials as possible ‘right’ according to the rule, and that they would naturally become faster with practice. Furthermore, it was emphasized to each participant that they must learn how to be accurate before they could learn to go both quickly and fluently. Once they had learned to be accurate they should then naturally learn to speed up.
5. Finally, they were then informed that they would complete pairs of practice blocks until they learned to meet accuracy and speed criteria that would be presented at the end of the block. Once these were met on both blocks within a pair, they would then complete three pairs of test blocks.

The IRAP task consisted of up to three pairs of practice blocks and exactly three pairs of test blocks. Each block included (a) a preblock rule screen; (b) 24 trials; and (c) a postblock feedback screen. The preblock rule screen contained the responding rule for the forthcoming block (rule A or rule B, see above), as well as the instructions “Try to get as many as possible

'right' according to the rule. If you go over time on any trial "!" will appear. If you get one wrong an "X" will appear – press the correct response to continue." Each trial presented participants with one label stimulus at the top of the screen, one target stimulus in the middle of the screen, and the response options on the bottom left and bottom right of the screen. The correct response option differed depending on the trial-type and alternated between the blocks, in line with the rules presented before that block. Participants responded using the "d" and "k" keys. If an incorrect response was emitted, a red "X" was displayed below the stimulus, and a correct response was required before proceeding to the next trial. After each trial the stimulus in the middle of the screen was cleared for an intertrial interval of 400 ms. The postblock feedback screen displayed both the participant's percentage accuracy and median latency performance on the previous block and the mastery criteria (i.e., accuracy ≥ 80% and median latency ≤ 2000 ms).

**Data processing**

The primary datum produced by both the IAT and IRAP is reaction time in millisecond from the onset of a trial to the first correct response. Effects on both the IAT and IRAP are defined as the latency difference between the two blocks in a block pair (e.g., rule A block vs. rule B block). The IAT provides one overall bias score (e.g., "life–self/death–others"), whereas the IRAP provides four, one for each of the trial-types (e.g., "life–self", "life–others", "death–self", and "death–others"). The effects on both measures were quantified using the *D* score, which has been found to limit the impact of extraneous variables such as responding speed and age (Greenwald, Nosek, & Banaji, 2003). Briefly, *D* scores were calculated as follows. First, latencies above 10,000 ms were removed. Then, for each pair of blocks, *D* is equal to the difference between mean block A reaction times and mean block B reaction times, divided by the standard deviation of the reaction times from both rule A and rule B blocks. In the current article,

reaction times from rule B blocks were always subtracted from rule A blocks. There are some differences in the calculation of *D* between the IAT and IRAP, which will now be discussed.

**IAT data processing.** *D* scores were calculated separately for the practice block pair (blocks 3 & 4) and test block pair (blocks 6 & 7) before being averaged to create a final *D* score. Previous research has demonstrated that excluding participants' IAT data based on accuracy or latency (e.g., accuracy ≥ 80%, median latency ≤ 2000ms) criteria does not significantly increase the reliability or validity of the measure (Nosek et al., 2007). As such, no such data exclusion criteria were applied here. Positive *D* scores represented quicker responding on rule A blocks (i.e., where self was coordinated with death and others was coordinated with life) relative to rule B blocks (i.e., where self was coordinated with life and others was coordinated with death). This can be interpreted as a "self–death/others–life" effect. In contrast, negative *D* scores represented the opposite pattern of responding, where responding on rule B blocks was quicker than on rule A blocks. This can be interpreted as a "self–life/others–death" effect.

**IRAP data processing.** As with the majority of previous research using the IRAP, one *D* score for each of the IRAP's four individual trial-types (Barnes-Holmes, Barnes-Holmes, et al., 2010). Practice block data were not included in the analysis. Accuracies on the IRAP are typically lower than the IAT. As such, in order to ensure that IRAP effects were derived from performances that involved the targeted patterns of stimulus control, *D* scores were excluded if a participant failed to maintain the mastery criteria in the test blocks. Although a number of specific exclusion strategies could in principle be applied, the current study adopted the method employed by Nicholson and Barnes-Holmes (2012) because it equally balances the two goals of removing unwanted performances and minimizing attrition. Following these authors' approach,

*D* scores from IRAP test blocks that failed to meet criteria were excluded from the analysis in the following manner:

1. If accuracy on one or both test blocks within a pair was < 78% and/or median latency was > 2000ms, then the four *D* scores from that test block pair were excluded.
2. If only one of a participant's three test block pairs were excluded in this manner, the final *D* scores were calculated by averaging the *D* scores across the two remaining test block pairs. *D* scores for three participants were calculated on this basis.
3. If more than one of a participant's three test block pairs were excluded in this manner, all of the data from that particular IRAP was excluded from the analysis. IRAP data for three participants were removed on this basis.

Following convention, *D* scores for two of the four trial-types were inverted. In this case, trial-types 3 and 4 on both the death–identity and death–evaluation IRAPs were inverted so as to create a common axis within each IRAP. Specifically, for the death–identity IRAP, the others–life and others–death trial–types were multiplied by -1. Positive *D* scores on this IRAP therefore represented "death" or "*not*-life" effects, whereas negative *D* scores represented "life" or "*not*-death" effects, depending on whether the stimuli within that trial-type referred to life or death, respectively. For example, if a positive *D* score was found on trial-type 1 this would be interpreted as a "self–death" effect (i.e., participants responded to self and death with "similar" more rapidly than with "different"); in contrast, if a negative *D* score was found it would be interpreted as a "self–*not*-death" effect (i.e., participants responded to self and death with "different" more rapidly than with "similar"). Axes were therefore arranged in a similar manner to the IAT in order to allow for comparison. That is, on both measures, positive scores represented prodeath effects and negative scores represented prolife effects.

A comparable set of inversions was applied to the death–evaluation IRAP: *D* scores for the "death–positive" and "death–negative" trial-types were multiplied by -1. This IRAP employed "True" and "False" as response options and, as such, positive *D* scores represented "positive" or "*not*-negative" effects, whereas negative *D* scores represented "negative" or "*not*-positive" effects (see Hussey, Thompson, McEnteggart, Barnes-Holmes, & Barnes-Holmes, 2015 for an article length discussion of the inversion and interpretation of IRAP trial-type effects).

## Results

### Self-report measures

The sample reported normative levels of depression ($M = 4.6$, $SD = 4.8$), hopelessness ($M = 4.7$, $SD = 4.0$), psychological flexibility ($M = 19.9$, $SD = 8.4$), and death anxiety ($M = 8.1$, $SD = 2.0$). The sample was roughly equally divided by their belief in the afterlife ($N = 12$ low belief, $N = 12$ ambivalent belief, $N = 20$ high belief; $M = 35.5$, $SD = 17.3$). Results from the self-report measures therefore indicated that the sample could be considered to represent normative levels of psychopathology and attitudes to death.

### Death identity IAT

The sample produced an overall "self–life/others–death" effect on the IAT, as hypothesized ($M = -0.34$, $SD = 0.43$: see Figure 1). That is, participants were faster to pair self with life and others with death than self with death and others with life. A one sample *t*-test showed that this effect was significantly different from zero, $t(41) = -5.2$, $p < 0.001$. Participants were therefore faster to categorize stimuli when self was coordinated with life and others was coordinated with death relative to when self was coordinated with death and others was coordinated with life. However, as discussed previously, the nature of the IAT precludes the

ability to determine which specific relational responses drove this effect (i.e., self–life, self–death, others–life, others–death). IAT *D* scores were not correlated with age, depression, hopelessness, psychological flexibility, belief in the afterlife, or death anxiety (all *p*s > .05, see Table 4).

When IAT *D* scores were binarized using zero as a cutoff score, as employed by previous research (Nock et al., 2010; Randall et al., 2013), 32 participants were shown to have overall "self–life/others–death" effects (i.e., *D* scores ≤ 0) and 7 participants were shown to have overall "self–death/others–life" effects (i.e., *D* scores > 0). The majority of the sample (82%) therefore demonstrated effects on the IAT that were consistent with what previous research has classified as low risk of attempted suicide or self-harm, as might be expected within a normative sample.

Table 4. *Correlations between the implicit and self-report measures.*

| | DI-IAT | DASS-D | BHS | AAQ-II | BIA | TDA | Age |
|---|---|---|---|---|---|---|---|
| Death–identity IAT | 1.0 | .19 | .23 | .27 | -.09 | .18 | .03 |
| *Death–identity IRAP* | | | | | | | |
| Self–death | .04 | -.09 | .00 | .05 | .14 | -.05 | -.12 |
| Self–life | -.16 | -.07 | .03 | .15 | -.10 | .07 | -.30 |
| Others–death | .14 | -.22 | -.14 | -.31 | .01 | -.15 | .15 |
| Others–life | .12 | -.15 | -.06 | -.26 | .14 | -.33* | .17 |
| *Death–evaluation IRAP* | | | | | | | |
| Living–positive | -.27 | .05 | .00 | .21 | .17 | .04 | .02 |
| Living–negative | -.38* | -.14 | -.15 | -.15 | -.21 | .03 | .08 |
| Dying–positive | .01 | .05 | .03 | .06 | .12 | .09 | .08 |
| Dying–negative | -.07 | .07 | .26 | .05 | .01 | -.01 | .06 |

*Note:* DI-IAT = death–identity IAT, DASS-D = Depression Anxiety Stress Scales' depression subscale, BHS = Beck Hopelessness Scale, AAQ-II = Acceptance and Action Questionnaire II, TDA = Templer Death Anxiety scale. For ease of interpretation, correlations among the IRAP trial-types are not included here.
* $p < .05$

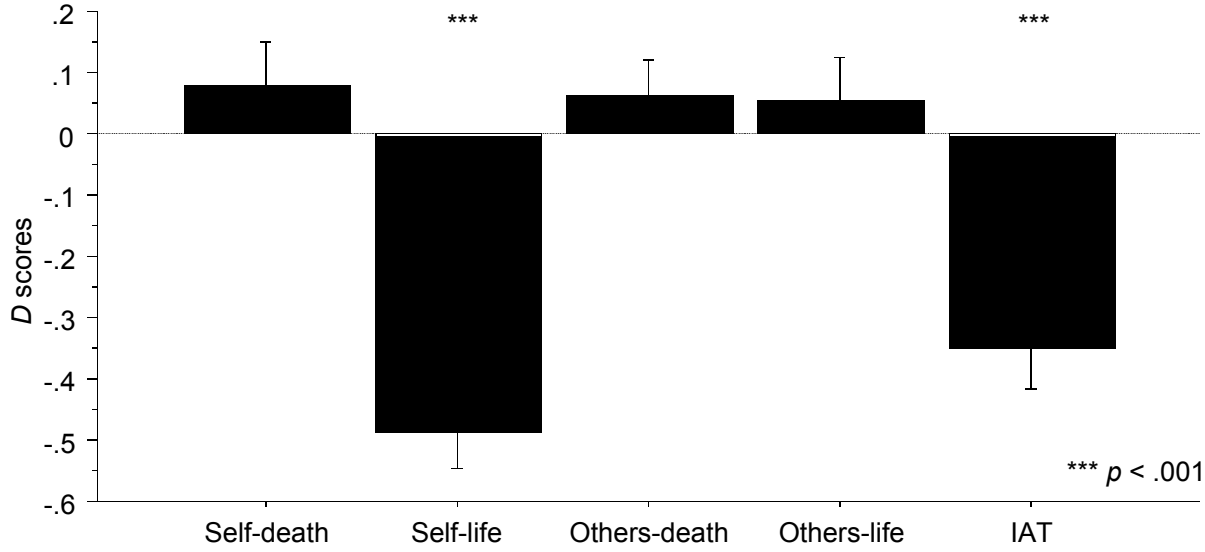


*Figure 1.* Performance on the death identity IRAP and IAT. Positive scores represent "death" or "*not*-life" biases, whereas negative scores represent "life" or "*not*-death" biases. Bars 1–4 represent mean *D* scores on the IRAP trial-types. Bar 5 represents mean IAT *D* scores. Error bars represent standard errors.

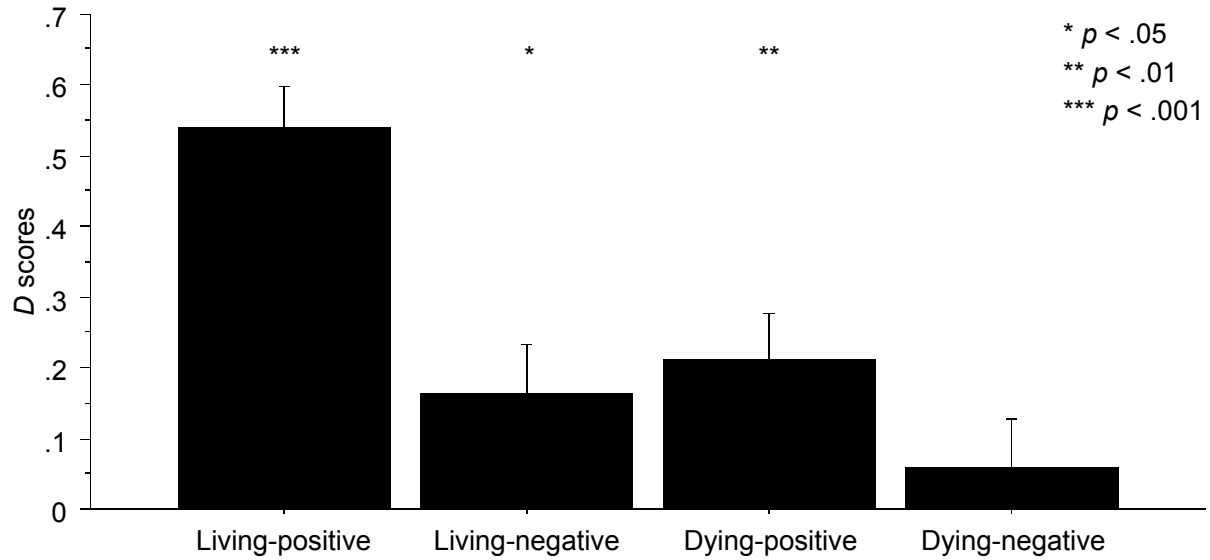


*Figure 2.* Performance on the death evaluation IRAP. Positive scores represent "positive" or "*not*-negative" evaluative biases, whereas negative scores represent "negative" or "*not*-positive" evaluative biases. Bars represent mean *D* scores on each IRAP trial-type. Error bars represent standard errors.

**Death identity IRAP**

Mean scores for the four death–identity IRAP trial-types and the death–identity IAT are presented in Figure 1, which demonstrates participants' strong "self–life" biases (i.e., participants were faster to respond to "self" and "life" with "similar" relative to "different"). No strong effects were found on the other three trial-types (i.e., self–death, others–life & others–death). It should be noted that the IAT is included in this figure only for the purposes of visual comparison, and was not included in the following ANOVA. A within subjects ANOVA confirmed significant differences between the IRAP trial-types $F(3, 40) = 15.9$, $p < .0001$. Follow up one sample *t*-tests demonstrated a significant "self–life" effect ($p < 0.0001$), but no other effects (all *p*s > .27: see Figure 1). This result therefore provides evidence for the assumptions made with previous research, whether explicitly or implicitly (e.g., Nock et al., 2010; Dickstein et al., 2015) that normative participants do indeed demonstrate "self–life" effects *specifically*, as opposed to "self–*not*-death" biases.

IRAP *D* scores were not correlated with age or self-reported depression, hopelessness, psychological flexibility, belief in the afterlife, or death anxiety (see Table 4), with one exception: death anxiety and the "others–life" trial-type ($p < .05$). Given the large number of correlations performed and the lack of a systematic pattern, however, this effect should be interpreted with extreme caution.

**Relationship between the death–identity IAT and IRAP**

Responses on the death–identity IAT demonstrated a "self–life/others–death" effect. In contrast, the results of the death–identity IRAP suggest that the IAT effect may have been driven by a "self–life" bias specifically, given that the biases on the other three trial-types were all relatively weak. A dependent *t*-test showed that the magnitude of the IRAP's "self–life" trial-

type did not differ significantly from the IAT *D* score, $t(38) = 1.4$, $p = .16$. On balance, however, a series of correlations failed to find any correlations between the IAT and any of the IRAP trial-types, including the "self–life" trial-type (*r*s = -.16 to .14).

**Death evaluation IRAP**

Mean scores for the four death–identity IRAP trial-types and the death–identity IAT are presented in Figure 2. Participants' demonstrated strong "life–positive" biases (i.e., they were faster to respond to "life" and "positive" with "true" relative to "false"). Smaller "life–*not*-negative" and "death–positive" biases were also found. No biases were found on the "death–negative" trial-type. A within subjects ANOVA demonstrated significant differences between the IRAP trial-types, $F(3, 40) = 11.4$, $p < .0001$. Follow up one sample *t*-tests demonstrated significant "life–positive", "life–*not*-negative" and "death–positive" effects (all *p*s < .03: see Figure 2). Effects on the "death–negative" trial-type were not significantly different from zero ($p = .38$). As such, effects on the life related trial-types are relatively intuitive for a normative sample: participants confirm that life is positive and reject that it is negative. However, the "death is positive" effect is both unexpected and counterintuitive. That is, the sample responded "True" more quickly than "False" on the "death–positive" trial-type, despite reporting normative levels of depression, hopelessness, and death anxiety on the self-report measures. The absence of a "death–negative" effect is equally surprising. Specifically, a correlation matrix indicated that none of the IRAP trial-types were significantly correlated with age, depression, hopelessness, psychological flexibility, belief in the afterlife, or death anxiety (*r*s = -.21 to .26: see Table 4). Finally, it should be noted that a significant correlation was found between the "life–negative" trial-type and the death–identity IAT ($r = -.38$, $p = .02$). This indicates that a bias towards

rejecting that life is negative on the IRAP was associated with a "self–life/others–death" bias on the IAT.

## Discussion

In the current study, participants were exposed to the death–identity IAT, which has previously been used to study attitudes to death and their relation to suicidality (e.g., Nock et al., 2010; Dickstein et al., 2015). In addition, participants were asked to complete two versions of the IRAP, one that targeted death–identity and one that targeted death–evaluations. The IRAPs were designed to provide information on the strength of specific relational responses rather than overall biases, as is the case with the IAT. The key findings in this largely exploratory study were as follows. The "self–life/others–death" bias found in previous studies using the IAT with normative samples was replicated in the current results (i.e., Dickstein et al., 2015; Harrison et al., 2014; Tang et al., 2013; Violanti et al., 2013). Intuitively, one might expect that such biases reflect a general positivity to and preference for life over death. Indeed, such attitudes are reflected in many psychological theories (e.g., need for meaning in life, Greenberg et al., 1986; suicidality, Joiner, 2005). Somewhat unexpectedly, therefore, although both IRAPs employed in the current research produced biases that reflected strong "pro" life responses, they also produced either weak "anti" or "pro" death responses. Indeed, one trial-type ("death–positive" on the death–evaluation IRAP) was in a significant "pro" death direction. How might we explain this curious implicit bias that reflects the lack of an aversive response to death, or even more bizarrely, a bias towards it?

At this stage, it is worth emphasizing that this was a normative sample of participants, and that the IRAP and IAT effects, in general, failed to correlate with, or even produce effects that were consistent with any of the self-report measures (i.e., death anxiety, belief in the

afterlife). As such, it would be difficult to explain these implicit biases as reflecting the participants' self-reported beliefs and attitudes towards death or dying or as indicators of non-normative attitudes to death, such as suicidality (i.e., hopelessness, depression). Indeed, as discussed in the introduction, positive evaluations of death in normative participants run contrary to a wide range of psychological theories of death and suicidality, which assume that death is aversive and/or negatively evaluated (e.g., Joiner, 2005; O'Connor, 2011). It should also be noted that such findings are not readily explained via cultural differences, for example, given that Irish samples have been shown to have similar attitudes to death as participants from other cultural backgrounds (Weafer, 2014).

Given that we cannot appeal readily to something specific or unusual about the sample of participants, perhaps we should consider the impact of the stimuli employed within the IRAPs. Specifically, it is possible that the death related words employed were not strongly valenced, particularly for a sample of young college students. In other words, death and mortality may be so far removed from the immediate psychological awareness of a young person that these words lack the salience to produce relatively strong IRAP effects (i.e., on the death–identity IRAP). Thus, it might be useful to consider the possible effects of using death related words that are less abstract to a young person, in the sense that they refer to more proximal adverse events. For example, the word "cancer" would likely have been more evocative of specific functions (e.g., pain, frailty, fear: see Barnes-Holmes, Keane, Barnes-Holmes, & Smeets, 2000) because these may have been observed directly in a family member who contracted the disease. Of course, while this explanation may account for the absence of death biases on one IRAP (i.e., death–identity), it is harder to explain the presence of a "death–positive" effect on the other IRAP (i.e., death–evaluation).

One alternative explanation for these patterns of bias might be found in the stimuli that were included in the IRAPs. First, it should be noted that the death related stimuli differed between the death–identity and death–evaluation IRAPs, both in the number of exemplars and the specific stimuli employed (see Tables 1 & 2). It is possible that this difference may contribute to any differences between them. Possibly more importantly, death can have a variety of valence functions depending on the context. For example, death could involve a lonely, painful decline, or it could involve a romantic (e.g., Romeo and Juliet) or heroic (e.g., Martin Luther King) demise. Furthermore, in this era of computer gaming, death may be associated with excitement and reversible at the press of a button. In retrospect, it is therefore difficult to know exactly what functions of death the stimuli evoked within the IRAPs. Perhaps future studies might seek to narrow the range of likely functions that their stimulus sets target. For example, future work might at minimum specify a relevant deictic relation. That is, *whose* death is being referred to (personal death, death of a specified other, death of an unspecified other, etc.)?

The absence of correlations between the death–identity IAT and IRAP is also worth commenting upon. This result was somewhat surprising, given that the two measures employed highly similar stimuli. On balance, previous research has reported mixed results in this regard, with some studies reporting correlations between implicit measures and others not (Bosson, Swann, & Pennebaker, 2000; see also Golijani-Moghaddam, Hart, & Dawson, 2013 for review). Of course, the two tasks do bear important methodological differences. For example, as a relative measure, the IAT presents all four categories on each trial, whereas the IRAP presents only one of the four possible pairings of these categories per trial. Thus, the lack of a strong and consistent relationship between the measures, which is hardly unique in the literature, should not cause excessive concern for the current study. It should also be noted that the lack of correlation

between the IAT and IRAP should not necessarily be interpreted as problematic for either measure, given that meta analyses have shown both to be have relatively high levels of predictive validity in clinically relevant domains (see Fazio & Olson, 2003; Vahey et al., 2015 for reviews).

We would recommend that the unexpected effects for the two death trial-types found in the current study (i.e., presence of a "death-positive" bias and absence of a "death-negative" bias) should be further explored in future research. There are at least two ways in which this might be pursued. First, future work might pose questions about the relevant effects in terms of understanding relational response biases in the context of suicidal behaviors. For example, one might compare evaluations of death on the IRAP between normative individuals and those with a history of suicidal behavior (e.g., ideation and/or attempts). Second, future research might seek to better understand the behavioral processes involved in IRAP performances themselves (e.g., Hussey, Ní Mhaoileoin, et al., 2015). In so doing, we would be in a better position to explain how unexpected or counterintuitive patterns of bias emerge. This latter strategy would require us to move beyond the notion of the IRAP as a measure of "implicit attitudes" and towards understanding it as a measure of natural verbal relations (see Barnes-Holmes, Barnes-Holmes, & Hussey, in press; Hussey, Barnes-Holmes, et al., 2015).

For illustrative purposes, consider that the largest bias on the death–evaluation IRAP was produced on the "life–positive" trial-type, whereas the "death–negative" trial-type was relatively weak (and non-significant) by comparison. One possible explanation is that the valence of the stimuli presented within the "life–positive" trial-type was more easily associated with the "True" response option, whereas the valence of the stimuli presented for the "death–negative" trial-type was more easily associated with the "False" response option (assuming that "True" is more positively valenced than "False"). If this was the case, then any response bias towards "True",

when confirming that death is negative, may have been reduced somewhat by a competing bias to associate negatively valenced stimuli with the negatively valenced response option ("False").

A related explanation might appeal to a general positivity bias to which the IRAP may be sensitive (see Barnes-Holmes, Murphy, Barnes-Holmes, & Stewart, 2010, pp.75-76). For example, all things being equal, in natural language interactions speakers tend to emphasize the positive over the negative, reporting for instance that a glass is half full rather than half empty (see Dodds et al., 2015, for evidence that this effect is observed across numerous languages). Given that the IRAP was specifically designed to capture differential probabilities (or biases) in patterns of verbal or relational responding that are found in natural language (Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2008) it seems reasonable to assume that such biases may also be reflected in IRAP performances. Indeed, one would hope so if the IRAP is to be considered a measure of the response patterns found in natural language (see O'Shea, Watson, & Brown, 2015 for empirical evidence to support this claim). In fact, it might even be important to capture such positivity biases if the variance they create helps to increase the prediction of criterion variables by the IRAP (see Vahey et al., 2015). In this context, it is worth noting that unpublished research from our group using a Life-Death IRAP similar to those employed in the current study successfully predicted self-harm and suicide ideation using a known-groups design (Hussey & Barnes-Holmes, 2013). Perhaps the interaction, or response competition, between general positivity biases and specific response biases towards the positive and negative aspects of life and death were jointly responsible for producing the observed level of predictive validity. In any case, future research into this intriguing possibility certainly seems warranted.

**Conflict of Interest:** The authors declare that they have no conflict of interest.

References

Barnes-Holmes, D., Barnes-Holmes, Y., & Hussey, I. (in press). Relational Frame Theory: Finding its historical and intellectual roots and reflecting upon its future development. In S. C. Hayes, D. Barnes-Holmes, R. D. Zettle, & A. Biglan (Eds.), *Handbook of contextual behavioral science*. New York, NY: Wiley-Blackwell.

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*, 527–542.

Barnes-Holmes, D., Hayden, E., Barnes-Holmes, Y., & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study. *The Psychological Record*, *58*(4), 497–516.

Barnes-Holmes, D., & Hussey, I. (2012). Implicit Relational Assessment Procedure (Version 2012). Maynooth, Ireland. Retrieved from IRAPresearch.org

Barnes-Holmes, D., Keane, J., Barnes-Holmes, Y., & Smeets, P. (2000). A derived transfer of emotive functions as a means of establishing differential preferences for soft drinks. *The Psychological Record*, *50*(3), 493–511.

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*, 57–66.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide: a 10-year prospective study of patients hospitalized with suicidal ideation. *American Journal of Psychiatry*, *142*(5), 559–563. http://doi.org/10.1176/ajp.142.5.559

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: the hopelessness scale. *Journal of Consulting and Clinical Psychology*, *42*(6), 861–865.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., … Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy*, *42*(4), 676–688. http://doi.org/10.1016/j.beth.2011.03.007

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: the blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*(4), 631–643.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. Technical report C-1*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Brown, G. K., Beck, A. T., Steer, R. A., & Grisham, J. R. (2000). Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. *Journal of Consulting and Clinical Psychology*, *68*(3), 371.

Burke, B. L., Martens, A., & Faucher, E. H. (2010). Two decades of Terror Management Theory: A meta-analysis of mortality salience research. *Personality and Social Psychology Review*, *14*(2), 155–195. http://doi.org/10.1177/1088868309352321

Cooper, J., Kapur, N., Dunning, J., Guthrie, E., Appleby, L., & Mackway-Jones, K. (2006). A clinical tool for assessing risk after self-harm. *Annals of Emergency Medicine*, *48*(4), 459–466. http://doi.org/10.1016/j.annemergmed.2006.07.944

Creemers, D. H. M., Scholte, R. H. J., Engels, R. C. M. E., Prinstein, M. J., & Wiers, R. W. (2013). Damaged self-esteem is associated with internalizing problems. *Frontiers in Psychology*, *4*. http://doi.org/10.3389/fpsyg.2013.00152

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*(3), 347–368. http://doi.org/10.1037/a0014211

Dickstein, D. P., Puzia, M. E., Cushman, G. K., Weissman, A. B., Wegbreit, E., Kim, K. L., … Spirito, A. (2015). Self-injurious implicit attitudes among adolescent suicide attempters versus those engaged in nonsuicidal self-injury. *Journal of Child Psychology and Psychiatry*, Advance online publication. http://doi.org/10.1111/jcpp.12385

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., … Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, *112*(8), 2389–2394. http://doi.org/10.1073/pnas.1411678112

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*(1), 297–327. http://doi.org/10.1146/annurev.psych.54.101601.145225

Foody, M., Barnes-Holmes, Y., & Barnes-Holmes, D. (2012). The role of self in Acceptance and Commitment Therapy. In L. McHugh & I. Stewart (Eds.), *The self and perspective taking: research and applications* (pp. 125–142). Oakland, CA: New Harbinger.

Foody, M., Barnes-Holmes, Y., Barnes-Holmes, D., Törneke, N., Luciano, C., Stewart, I., & McEnteggart, C. (2014). RFT for clinical use: The example of metaphor. *Journal of Contextual Behavioral Science*, *3*(4), 305–313. http://doi.org/10.1016/j.jcbs.2014.08.001

Franck, E., De Raedt, R., Dereu, M., & Van den Abbeele, D. (2007). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(1), 75–85. http://doi.org/10.1016/j.jbtep.2006.05.003

Glashouwer, K. A., de Jong, P. J., Penninx, B. W. J. H., Kerkhof, A. J. F. M., van Dyck, R., & Ormel, J. (2010). Do automatic self-associations relate to suicidal ideation? *Journal of Psychopathology and Behavioral Assessment*, *32*(3), 428–437. http://doi.org/10.1007/s10862-009-9156-y

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. http://doi.org/10.1016/j.jcbs.2013.05.002

Greenberg, J., Pyszczynski, T., & Solomon, S. (1986). The causes and consequences of a need for self-esteem: A terror management theory. In R. F. Baumeister (Ed.), *Public self and private self* (pp. 189–212). Springer New York.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

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 and Social Psychology*, *85*(2), 197–216. http://doi.org/10.1037/0022-3514.85.2.197

Harrison, D. P., Stritzke, W. G. K., Fay, N., Ellison, T. M., & Hudaib, A.-R. (2014). Probing the implicit suicidal mind: Does the death/suicide Implicit Association Test reveal a desire to die, or a diminished desire to live? *Psychological Assessment*, *26*(3), 831–840. http://doi.org/10.1037/pas0000001

Hayes, S. C. (1992). Verbal relations, time, and suicide. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 109–118). NV: Context Press.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Press.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An experiential approach to behavior change*. Guilford Press.

Hughes, S., & Barnes-Holmes, D. (in press). Relational Frame Theory: The basic account. In S. C. Hayes, D. Barnes-Holmes, R. D. Zettle, & A. Biglan (Eds.), *Handbook of contextual behavioral science*. New York, NY: Wiley-Blackwell.

Hussey, I. (2015, January 9). IRAP Experimenter Script. Retrieved from http://irapresearch.org/wp/downloads-and-training/

Hussey, I., & Barnes-Holmes, D. (2012). The implicit relational assessment procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive and Behavioral Practice*, *19*(4), 573–582. http://doi.org/10.1016/j.cbpra.2012.03.002

Hussey, I., & Barnes-Holmes, D. (2013). *Developing new tools to predict and understand suicide*. Presented at the Irish Higher Education Authority “Making an Impact” awards, Dublin, Ireland. Retrieved from http://www.hea.ie/ga/node/577

Hussey, I., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). From Relational Frame Theory to implicit attitudes and back again: clarifying the link between RFT and IRAP research. *Current Opinion in Psychology*, *2*, 11–15. http://doi.org/10.1016/j.copsyc.2014.12.009

Hussey, I., Ní Mhaoileoin, D., Barnes-Holmes, D., Ohtsuki, T., Kishita, N., Hughes, S., & Murphy, C. (2015). *The IRAP is non-relative but not a-contextual: Changes to the contrast category influence men's de-humanization of women*. Manuscript under review.

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*.

Iverach, L., Menzies, R. G., & Menzies, R. E. (2014). Death anxiety and its role in psychopathology: Reviewing the status of a transdiagnostic construct. *Clinical Psychology Review*, *34*(7), 580–593. http://doi.org/10.1016/j.cpr.2014.09.002

Joiner, T. E. J. (2005). *Why people die by suicide*. Harvard University Press.

Kastenbaum, R. (2000). *The psychology of death*. Springer Publishing Company.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. http://doi.org/10.1016/0005-7967(94)00075-U

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *Psychological Record*, *54*(1), 115–144.

Neimeyer, R. A., Wittkowski, J., & Moser, R. P. (2004). Psychological research on death attitudes: an overview and evaluation. *Death Studies*, *28*(4), 309–340. http://doi.org/10.1080/07481180490432324

Nicholson, E., & Barnes-Holmes, D. (2012). Developing an implicit measure of disgust propensity and disgust sensitivity: Examining the role of implicit disgust propensity and sensitivity in obsessive-compulsive tendencies. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(3), 922–930. http://doi.org/10.1016/j.jbtep.2012.02.001

Nock, M. K., Park, J. M., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science*, *21*(4), 511–517. http://doi.org/10.1177/0956797610364762

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality & Social Psychology Bulletin*, *31*(2), 166–180. http://doi.org/10.1177/0146167204271418

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York, NY: Psychology Press.

O'Connor, R. C. (2011). The integrated motivational-volitional model of suicidal behavior. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, *32*(6), 295–298. http://doi.org/10.1027/0227-5910/a000120

Osarchuk, M., & Tatz, S. J. (1973). Effect of induced fear of death on belief in afterlife. *Journal of Personality and Social Psychology*, *27*(2), 256–260. http://doi.org/10.1037/h0034769

O'Shea, B., Watson, D. G., & Brown, G. D. (2015). Measuring implicit attitudes: a positive framing bias flaw in the Implicit Relational Assessment Procedure (IRAP). *Psychological Assessment*, Advance online publication.

Patterson, W. M., Dohn, H. H., Bird, J., & Patterson, G. A. (1983). Evaluation of suicidal patients: the SAD PERSONS scale. *Psychosomatics*, *24*(4), 343–345, 348–349. http://doi.org/10.1016/S0033-3182(83)73213-5

Pinter, B., & Greenwald, A. G. (2005). Clarifying the role of the "other" category in the self-esteem IAT. *Experimental Psychology*, *52*(1), 74–79. http://doi.org/10.1027/1618-3169.52.1.74

Price, R. B., Iosifescu, D. V., Murrough, J. W., Chang, L. C., Al Jurdi, R. K., Iqbal, S. Z., … Mathew, S. J. (2014). Effects of ketamine on explicit and implicit suicidal cognition: a randomized controlled trial in treatment-resistant depression. *Depression and Anxiety*, *31*(4), 335–343. http://doi.org/10.1002/da.22253

Price, R. B., Nock, M. K., Charney, D. S., & Mathew, S. J. (2009). Effects of intravenous ketamine on explicit and implicit measures of suicidality in treatment-resistant depression. *Biological Psychiatry*, *66*(5), 522–526. http://doi.org/10.1016/j.biopsych.2009.04.029

Randall, J. R., Colman, I., & Rowe, B. H. (2011). A systematic review of psychometric assessment of self-harm risk in the emergency department. *Journal of Affective Disorders*, *134*(1–3), 348–355. http://doi.org/10.1016/j.jad.2011.05.032

Randall, J. R., Rowe, B. H., Dong, K. A., Nock, M. K., & Colman, I. (2013). Assessment of self-harm risk using implicit thoughts. *Psychological Assessment*, *25*(3), 714–721. http://doi.org/10.1037/a0032391

Ribeiro, J. D., Witte, T. K., Van Orden, K. A., Selby, E. A., Gordon, K. H., Bender, T. W., & Joiner, T. E. (2014). Fearlessness about death: The psychometric properties and construct validity of the revision to the Acquired Capability for Suicide Scale. *Psychological Assessment*, *26*(1), 115–126. http://doi.org/10.1037/a0034858

Rönspies, J., Schmidt, A. F., Melnikova, A., Krumova, R., Zolfagari, A., & Banse, R. (2015). Indirect measurement of sexual orientation: Comparison of the Implicit Relational Assessment Procedure, viewing time, and choice reaction time tasks. *Archives of Sexual Behavior*, Advance online publication. http://doi.org/10.1007/s10508-014-0473-1

Tang, J., Wu, S., & Miao, D. (2013). Experimental test of escape theory: Accessibility to implicit suicidal mind. *Suicide and Life-Threatening Behavior*, *43*(4), 347–355. http://doi.org/10.1111/sltb.12021

Templer, D. I. (1970). The construction and validation of a death anxiety scale. *The Journal of General Psychology*, *82*(2), 165–177. http://doi.org/10.1080/00221309.1970.9920634

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 and Experimental Psychiatry*, *48*, 59–65. http://doi.org/10.1016/j.jbtep.2015.01.004

Violanti, J. M., Mnatsakanova, A., & Andrew, M. E. (2013). Suicidal ideation in police officers: Exploring an additional measure. *Suicidology Online*, *4*, 33–41.

Weafer, J. A. (2014). *Irish attitudes to death, dying and bereavement*. Irish Hospice Foundation. Retrieved from http://hospicefoundation.ie/publications/