Assessing a Derived Transformation of Functions Using the Implicit Relational Assessment Procedure Under Three Motivative Conditions

Cainã T. Gomes[1]

William F. Perez[1, 3]

João Henrique de Almeida[2]

Arthur Ribeiro[1]

Julio C. de Rose[2, 3]

Dermot Barnes-Holmes[4]

1. Paradigma – Centro de Ciências e Tecnologia do Comportamento, Brazil
2. Universidade Federal de São Carlos, Brazil
3. Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE), Brazil
4. Ghent University, Belgium

Author Note

Preparation of the manuscript was supported by the Brazilian National Research Council (CNPq, Grants 573972/2008-7,465686/2014-1) and the São Paulo Research Foundation (FAPESP, Grant 08/57705-8), both for Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE), chaired by Deisy

G. de Souza (UFSCar). JHA also was supported by a post-doc fellowship from FAPESP (2014/01874-7)

Please address correspondence to William F. Perez, Paradigma - Centro de Ciências do Comportamento, Rua Wanderley, 611, Perdizes, São Paulo – SP, CEP 05011-001, Brazil. Email: will.f.perez@gmail.com

Assessing a Derived Transformation of Functions Using the Implicit Relational Assessment Procedure Under Three Motivative Conditions

Abstract

In exploring the extent to which the Implicit Relational Assessment procedure (IRAP) may function as a measure of the derived transformation of functions, it is important to determine if it is also sensitive to particular moderating variables. This was the purpose of the current study, which focused on manipulating three different motivating conditions related to stimuli presented during the IRAP task. First, two equivalence classes were established: A1-B1-C1-D1 and A2-B2-C2-D2. These classes comprised nonsense forms (B, C and D) and two meaningful stimuli: a picture of a glass full of water (A1) and a neutral picture (A2). Derived transformation of function from the meaningful stimuli to two nonsense forms (D1 and D2) was then assessed by means of a semantic differential and an IRAP. Before assessment, participants were divided in three groups: the first had water intake; the second had pepper; the third had pepper before the semantic differential, followed by an extra dose before the IRAP testing blocks. Results suggest that the motivative conditions progressively affected both measures. Regarding the semantic differential, D1 (water) and D2 (neutral) stimuli were close to neutrality for the Water group; for the Pepper and Double Pepper groups, however, the D1 (water) stimulus had a positive valence while D2 (neutral) was neutral. In the IRAP, both the Water and Pepper group evaluated D1 as positive; nonetheless, for the Double Pepper group, IRAP scores revealed that D1 was even more positive compared to the other groups. Implications for the IRAP are discussed in terms of the DAARRE model.

Keywords: equivalence relations, transformation of function, meaningful stimuli, IRAP, semantic differential, motivation

Derived transformation of functions has been studied widely in the behavior-analytic literature for over three decades (e.g., de Rose, McIlvane, Dube, Galpin, & Stoddard, 1988; Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; Hayes, Kohlenberg, & Hayes, 1991; Perez, Fidalgo, Nico, & Kovac, 2015; Perez et al., 2017; see Dymond & Rehfeldt, 2000, for a review). The basic phenomenon involves first training participants in a series of interrelated conditional discriminations (e.g., A-B and B-C), and testing for the formation of equivalence classes (e.g., C-A) in the absence of direct reinforcement. A specific function may then be trained to one of the stimuli from the equivalence class (e.g., A=F1) and then during a test phase other members of the equivalence may also exhibit that function (e.g., C=F1) in the absence of direct training. Such derived transformation effects have been shown using a wide variety of procedures and behavioral functions (e.g., Barnes, Brown, Smeets, & Roche, 1995; de Rose et al., 1988; Bortoloti & de Rose 2009; Luciano et al., 2014; Perez, Fidalgo, et al., 2015; Perez et al., 2017). This effect has been used to help develop behavioral explanations for behaviors that appear to emerge in the absence of direct learning experiences, such as the acquisition of irrational fears and phobias (e.g., Dougher et al., 1994; Guinther & Dougher, 2015; Dymond, Bennett, Boyle, Roche, & Schlund, 2017).

Some behavior analytic researchers have argued that the derived transformation of functions provides a functional-analytic model of symbolic control. Imagine, for example, that a child learns in her classroom that “stop” “cease” and “desist” are broadly equivalent in meaning, and is then rewarded for stopping what he or she is doing when asked to “stop”. If the child also stops engaging in a particular behavior when subsequently asked to “desist” this could be an instance of a derived transformation of functions because the meaning or symbolic function of “stop” transferred via equivalence to “desist.” Of course, the foregoing remains a plausible example of how the derived transformation of functions might operate in the natural environment and thus experimental analyses are required to support the interpretation. Some of this work may be obtained from experimental studies in which the functions are established in the laboratory (e.g., Barnes et al., 1995; de Rose et al., 1988) or using stimuli that likely acquired their functions pre-experimentally (e.g., Watt, Keenan, Barnes, & Cairns, 1991; Bortoloti & de Rose, 2009; Mizael, de Almeida, Silveira, & de Rose, 2016). One issue that arises in assessing the derived transformation of functions is determining how “strong” those functions might be. An immediate difficulty in addressing this question is that the derived transformation effect must occur in the absence of differential consequences for responses that document transformation. In principle, therefore, transformation should emerge and persist in the absence of direct reinforcement. Extinction procedures have often been used to assess the strength of specific behaviors by withdrawing reinforcement for those behaviors and determining how long it takes for the behavior to reduce to some minimal level (e.g., Skinner, 1938). Using extinction procedures to assess strength of derived transformation would be difficult, however, because there would be no reinforcement to withdraw.

One alternative strategy to assess the strength of derived relational responding, including derived transformation of functions, has recently been suggested. Specifically, a procedure known as the implicit relational assessment procedure (IRAP; Barnes-Holmes, Barnes-Holmes, Power, Hayden, Milne, & Stewart, 2006; Hughes, Barnes-Holmes, & Vahey, 2012) has been used to assess the relative strength of relational responding. The procedure involves asking participants to emit a particular pattern of relational responding during some blocks of trials and then to emit the opposite pattern in other blocks of trials. The difference in the combined accuracy and response latencies across the two types of blocks is then taken as a measure of the strength of relational responding (Hussey, Barnes-Holmes, & Barnes-Holmes, 2015; see also Bortoloti & de Rose, 2012).

For illustrative purposes consider a recent study that employed the IRAP to assess derived transformation of functions from happy and fearful faces to abstract stimuli equivalent to them (Perez et al., 2019). As in previous studies of transformation of functions using emotional expressions (e.g., Bortoloti & de Rose, 2009; 2012; de Almeida & de Rose, 2015), adult participants were submitted to a delayed matching-to-sample task that aimed to established two equivalence classes between facial expressions of fear (class 1) and of happiness (class 2) and nonsense forms: A1(Fear)-B1-C1-D1; A2(Happiness)-B2-C2-D2. After relational training (AB, AC, CD) and equivalence tests (BD, DB), the participants evaluated the meaning of one nonsense stimulus from each class: D1 (equivalent to the fearful faces) and D2 (equivalent to the happy faces). The first evaluation involved a semantic differential (Osgood & Suci, 1952; Osgood, Suci, & Tannenbaum, 1957), which is an instrument comprised of multiple scales anchored by opposite adjectives (sad/happy, bad/good, ugly/beautiful etc.). A control group evaluated the facial expressions and the same nonsense stimuli without any prior relational training and testing. Results suggested that the participants who formed equivalence classes evaluated the D1 stimuli as negative and the D2 as positive. The control group evaluated the actual faces in a similar way as the experimental group evaluated the D stimuli; however, as expected, the control group failed to show any derived transformation effects to the D stimuli. The second test for transformation of functions involved an Implicit Relational Assessment Procedure (IRAP), which has been used in fear related studies (e.g., Nicholson & Barnes-Holmes, 2012; Leech, Barnes-Holmes, & Madden, 2016) and present a relatively high level of predictive validity in the clinical domain (Vahey, Nicholson, & Barnes-Holmes, 2015).

On each trial of the IRAP, one of the nonsense forms, D1 (equivalent to fear) or D2 (equivalent to happy), was simultaneously presented with negative (e.g., bad, unpleasant etc.) or positive words (e.g., good, pleasant etc.), and two relational response options, ‘True’ or ‘False’. The participants were required to respond across alternating blocks that were consistent with the equivalence training (D1-Negative/True, D1-Positive/False, D2-Negative/False, D2-Positive/True) or inconsistent (D1-Negative/False, D1-Positive/True, D2-Negative/True, D2-Positive/False). The mean response latencies on the consistent blocks were shorter compared to the inconsistent blocks. Faster responses coherent with D1(fear)-Negative and D2 (happy)-Positive indicated the transformation of emotional functions from the faces to the equivalent D stimuli, corroborating the results from the semantic differential.

In exploring the extent to which the IRAP may function as a measure of the derived transformation of functions, it seems important to determine if it is also sensitive to particular moderating variables. This was the purpose of the current study, which focused on manipulating the motivating conditions[1] for access to water by asking participants to drink water or to allow drops of liquid pepper to be dripped on the tongue of the participant. The general assumption was that the pepper would increase motivation for access to water. Thus the trial-types of the IRAP that targeted an arbitrary stimulus that was in an equivalence relation with water would produce larger positive-water bias scores for participants who were exposed to liquid pepper than those who were asked to drink water before completing the IRAP.

[1] The term “motivative conditions”, in the present study, is a general label to indicate variables affecting IRAP performances. The experimental preparations and measures used in the current study do not allow the use of other technical terms such as “motivative operations” (because there was no operant task involving water intake).

The study involved first establishing two equivalence classes: A1-B1-C1-D1 and A2-B2-C2-D2. These classes comprised nonsense forms (B, C and D) and two meaningful stimuli, one in each of the classes: a picture of a glass full of water (A1) and a picture of a wooden stool (A2). The derived transformation of functions from the meaningful stimuli to two nonsense forms (D1 and D2) was then assessed by means of a semantic differential scale (Osgood et al., 1957) and an IRAP, both using positive and negative adjectives. Before assessing derived transformation of function regarding the positive or negative valence of the nonsense stimuli in equivalence with “water” or “neutral”, the participants were divided into three groups: the first group drank some water; the second group was exposed to two drops of pepper (on the tongue); the third group were exposed to two drops of pepper before the semantic differential and two additional drops before the critical IRAP test blocks (see procedure section).

Method

Participants

Thirty-six verbally competent adults (M= 14, F= 22) ranging in age from 19 to 32 years took part in the experiment. None of them had previously participated in any research involving neither equivalence relations nor the IRAP. They were recruited by personal contacts. The participants read a term of consent (approved by the Brazilian platform for ethical committees, Plataforma Brasil, CAAE# 59350016.1.0000.5375) before the experiment began. To participate, they should report having no allergy to pepper and having tasted Tabasco pepper before without suffering any significant discomfort; volunteers that reported having gastritis or any other gastrointestinal disease, as well as piercings on their tongues were excluded from the research. The participants received no payment or compensation for their time spent in the research.

Equipment and Setting

Experimental sessions were conducted in a quiet room, free from distraction, equipped with a chair, a desk and a Hewlett Packard (model Pavilion 14V066BR) notebook. Two software presented stimuli, delivered consequences, and recorded participants' responses: the RelationalFraming 1.0 (Perez, 2014) ran the matching-to-sample (MTS) tasks during the early experimental phases; the IRAP software ran the latency-based task and calculated D_{IRAP} scores during the last experimental phase. The semantic differential was a paper-and-pencil evaluation with each stimulus and scales printed in black and white on an A4 white sheet. For group treatments, there was a 200 ml plastic cup; a bottle of *Tabasco Original* (*Capsicum frutescens L.*) delivered with a medicine dropper. The *Tabasco Original* pungency ranges from 2500-5000 Scoville heat units (Scoville, 1912). Thus, the pepper chosen for this study is mild, but with enough pungency to produce some degree of discomfort.

Procedure

The procedure comprised four different phases, as outlined in Figure 1. In Phase 1, participants performed a delayed matching-to-sample task (DMTS) to form two equivalence classes. Each class comprised a meaningful picture (a glass being filled of water for Class 1, and a stool for Class 2), and three nonsense forms. In Phase 2, participants were divided into three groups, with a different solution intake for each group: Pepper and Double Pepper groups had 0.3 ml of *Tabasco Original* and Water group had 400 ml of spring water. In Phase 3, participants rated one nonsense stimulus from each class (D1 and D2) with a semantic differential. In Phase 4 participants performed an IRAP to evaluate the same stimuli that had been rated in Phase 3. During this phase, the participants from the Double Pepper group had a second solution intake (again 0,3 ml of Tabasco Original) just before the IRAP testing blocks. The researcher remained in the room only to set up the software for data collection, to provide instructions before the beginning each phase and during solution intake, and then exited, so that the participant was alone in the room during MTS, semantic differential rating, and IRAP (except during practice blocks). The experiment lasted a single session (40 to 60 min), with breaks if necessary.

Equivalence training and testing. This phase used a DMTS task to establish two equivalence classes: A1-B1-C1-D1 and A2-B2-C2-D2. A1 was an image of a glass being filled with water, and A2 was a neutral picture obtained from the International Affective Picture System (Lang, Bradley & Cuthbert, 2008): the image of a wooden stool. The other stimuli were abstract black and white forms (randomly extracted from Dube & Hiris, 1999).

Before starting, the participants read minimal instructions on how to perform the MTS task. Each trial began with the presentation of a sample stimulus (e.g., A1 or A2) on the center of the screen. A click on the sample was followed by withdrawal of the sample and, after 1.5 s, by the onset of three comparison stimuli (e.g., B1, B2, B3) on the bottom of the screen, side-by-side. Participants selected one of the comparison stimuli by clicking with the mouse on the stimulus. A selection was followed immediately by withdrawal of all comparison stimuli. Selections of the stimulus programmed to belong to the same class as the sample (e.g., B1 with sample A1; B2 with sample A2, etc.) was considered correct and selections of any other comparison stimulus was considered incorrect. Correct selections produced the immediate presentation of the word “CORRECT” on the center of the screen for 1s, together with a sequence of ascending notes. Incorrect selections produced the immediate presentation of the word “INCORRECT” in the center of the screen for 1s and a dissonant sound. The intertrial interval (ITI) was 1s. All trials displayed three comparison stimuli, two of them from the intended classes 1 and 2, respectively, and a third (B3, C3, and D3), whose selections were never reinforced. These “null” stimuli were presented to reduce the likelihood that participants would select the correct stimulus by simply rejecting an incorrect one (sample/S- control, cf. Sidman, 1987; see also, de Rose, Hidalgo, & Vasconcelos, 2013; Perez, Tomanari, & Vaidya, 2015). The presentation of stimuli was randomized; samples could not be presented more the three trials consecutively and the same comparison stimulus could not be presented on the same location for more than three trials consecutively.

Conditional relations were trained in the following order: AB, AC and CD. Conditional relation AB was taught first, with AB trials repeating until participants met the mastery criterion of 18 consecutive correct responses. AB trial types (Sample/Comparison1Comparison2/Comparison 3, correct comparison underlined) were A1/B1B2B3 and A2/B1B2B3. AC training followed, to the same criterion; trial types were A1/C1C2C3, A2/C1C2C3. CD was training was conducted next, to the same criterion, with trial types C1/D1D2D3, C2/D1D2D3. After the conditional relations were taught separately, AB, AC and CD trials were mixed in a cumulative baseline block with all training trial types presented randomly, until participants made 36 consecutive correct responses.

Equivalence tests were conducted next. First, the participants read the following instruction: “From now on, the computer will no longer present feedback, but will keep recording your hits and errors”. There were two test blocks, each comprised of 10 trials. The first test block presented 5 trials for each of the following relations B1D1 and B2D2 (trial types were B2D2B1/D1D2D3, B2/D1D2D3); the second test block presented 5 trials to test relations D1B1 and D2B2 (trial types were D1/B1B2B3, D2/D1D2D3). There were no programmed differential consequences in the test blocks.

The criterion for equivalence class formation was 75% correct responses, i.e., 15 correct responses along the 20 trials of both blocks. If the participant did not meet criterion, the cumulative baseline block repeated, followed by retesting. Participants who did not reach criterion after retraining and retesting were thanked and debriefed, and did not advance to the following experimental phases.

Solution intake. The participants who successfully completed equivalence training and testing were divided into three groups that differed on their solution intake before rating stimuli with the semantic differential. The participants from the group Water drank two 200ml cups of water. Participants from the groups Pepper and Double Pepper received 0.3 ml medicine drops of *Tabasco Original* on their tongues, delivered by the experimenter with a medicine dropper. Participants who had had pepper intake were asked to refrain from drinking water until the end of the experiment. The experimenter told them that if they were uncomfortable with the pepper pungency, they could ask for water and/or a biscuit and quit the experiment. Before entering the room, the experimenter asked them whether they had bottles of water or other beverages in their personal belongings, and if that were the case, the experiment kept the beverages until the end of the experiment and then returned them.

Semantic differential. Participants were then instructed to rate stimuli D1 and D2 by means of a semantic differential (Osgood & Suci, 1952; Osgood et al., 1957). Each D stimulus was displayed on the top of an A4 paper sheet. Each sheet contained one of the D stimuli on the top, followed by a series of scales to the bottom. There were 13 scales comprised of seven points anchored by opposite adjectives, the Portuguese equivalents of: sad/happy, torturing/comforting, dry/wet, disgusting/delicious, beautiful/ugly, hot/cold, negative/positive, arid/refreshing, hateful/tasty, bad/good, unpleasant/pleasant, burning/freezing, distressing/enjoyable. For each scale, the participant selected a point indicating how close they judged the stimulus to be to the adjectives (e.g., good _|x|_|_|_|_|_ bad). Sheets with D1 and D2 were given in a randomized order. The position of the positive and negative adjectives was balanced across scales and the order of the polar adjectives on the sheet was randomized.

IRAP. As illustrated on Figure 2, the IRAP trials comprised the simultaneous presentation of one abstract stimulus on the top of the screen, one word on the center and two relational response options on the lower corners. The abstract stimuli were either D1 (equivalent to water) or D2 (equivalent to neutral); the words, selected from the semantic differential were adjectives, either positive (the Portuguese equivalent of positive, cold, wet, refreshing, delicious, and enjoyable) or negative (the Portuguese equivalent of negative, hot, dry, arid, disgusting, distressing). The response options were two words with fixed positions, the Portuguese equivalent of ‘True’ on the left corner and ‘False’ on the right. Participants were required to choose one of the two response options, pressing the letter ‘d’ on the keyboard for true or ‘k’ for false. Correct responses were followed by the withdrawal of all stimuli presented on the trial and a brief 400 ms ITI. Incorrect responses were followed by the presentation of a red X on the center of the screen and stimuli were not withdrawn. The trial would end and the ITI began only after the participant had emitted the correct response.

Participants were exposed to blocks of 24 trials each, in which the responses considered correct would be consistent or inconsistent with equivalence relations that were trained. During consistent blocks (see Figure 2), each of the following trial types was presented (D stimulus-Adjective/Correct response option): D1-Positive/True, D1-Negative/False, D2-Positive/False, D2-Negative/True. During inconsistent blocks, the contingencies of reinforcement for response options were reversed. Trial types were: D1-Positive/False, D1-Negative/True, D2-Positive/True, D2-Negative/False. The trial types were balanced and randomized across trials. Consistent and inconsistent blocks always alternated. Half of the participants began with a consistent block and the other half with an inconsistent block.

The IRAP comprised practice (warm up) and testing phases. The practice phase began presenting a pair of consistent/inconsistent blocks with a mastery criterion of 80% of correct responses in both blocks. After reaching criterion, the participants were exposed to another pair of consistent/inconsistent blocks to which a median latency criterion of 2000ms was added. Participants who failed to reach accuracy and latency criteria after three pairs of practice blocks were thanked, debriefed, and their data discarded. The participants who achieved accuracy and latency criteria during the practice phase went directly to the IRAP testing phase (groups Water and Pepper). Before starting the IRAP testing blocks, the participants of the Double Pepper group received another 0.3 ml of pepper directly on their tongues, exactly as they had received before the semantic differential. Testing consisted of three pairs of consistent/inconsistent blocks. There were no performance criteria for the test blocks; instead participant's data would be excluded if their accuracy fell below 75% in more than one block, or if their median latency exceeded 2000ms in any test block. At the end of the last test block, a brief message appeared ending the IRAP. Only test blocks were considered in the data analysis and to calculate the D_{IRAP} scores.

Results

A total of 36 participants initiated the experiment. Six of them did not complete all experimental phases and their data were disregarded. Four excluded participants were from the Pepper group, and did not reach the mastery criteria during IRAP practice blocks; the other two were from the Double Pepper group, and failed to maintain the accuracy and/or latency criteria during IRAP testing blocks. Each group comprised 10 participants who completed the experiment.

Table 1 presents results from equivalence training and testing. All participants acquired the conditional relations during training steps and scored relatively high during equivalence tests.

Figure 3 presents results from the semantic differential. By visual inspection, both D1 (water-equivalent) and D2 (neutral-equivalent) stimuli were close to neutrality for the Water group. For the Pepper and Double Pepper groups, however, the D1 (water-equivalent) stimulus had a positive valence whereas D2 (neutral-equivalent) was neutral. The Double Pepper group, however, showed more variability in ratings of D1 and D2, with ratings overlapping for some scales. This result was confirmed employing a Kruskall-Wallis test and a Dunn's multiple comparisons test to verify the differences among the ratings for D1 (water-equivalent) and D2 (neutral-equivalent), for the three different groups. The difference between ratings for D1 (water-equivalent) and D2 (neutral-equivalent) was statistically significant for the groups Pepper and Double Pepper (Pepper group and Double Pepper group - D1 x D2 $ps < .00001$). However, the difference for the Water group was not statistically significant (Water group – D1 x D2 $p = .0927$).

Figure 4 presents results from the IRAP. Response biases for D1 (water-equivalent) were positive for all three groups on the WATER-POSITIVE trial type (to facilitated description, D1 and D2 stimuli will be referred to as WATER and NEUTRAL on IRAP trial types). However, the D_{IRAP} score observed for the Double Pepper group was over twice the size compared to the other two groups. The Double Pepper group was also the only one that showed a relatively strong bias towards responding faster to "False" when D1 (water-equivalent) was presented with negative words. Initially, multiple t-tests were conducted to verify whether each trial type for each group was significantly different from zero. For the Double Pepper group, three of the four trial types were statistically different from zero (Trial type 1 WATER POSITIVE true, $p < .0001$; Trial type 2 WATER NEGATIVE false, $p = .0036$, Trial type 4 NEUTRAL NEGATIVE true, $p = .0489$), whereas one was not (Trial type 3 NEUTRAL POSITIVE false, $p = .2818$). A one-way repeated-measures analysis of variance (ANOVA) indicated that the differences among the four trial-types were significant for the Double Pepper group F (3, 9) = 10.705, $p < .0001$, $\eta^2_{Partial} = .47$. Uncorrected Fisher's least significant difference (LSD) tests indicated that trial types 1 (WATER POSITIVE true) and 2 (WATER NEGATIVE false) were significantly different from the remaining two trial types (all *ps* < .04), suggesting that different patterns were produced between trials in the presence of D1 compared to D2. The inferential statistics are thus broadly consistent with the descriptive analyses in supporting the conclusion that stimulus D1 (water-equivalent) had a relatively strong positive valence for the Double Pepper group.

For the Pepper group, scores for only one of the four trial types were significantly different from zero (Trial type 1, WATER POSITIVE true, $p = .0004$), whereas scores for the other trial types were not (Trial type 2 WATER NEGATIVE false, $p = .9261$, Trial Type 3 NEUTRAL POSITIVE false, $p = .064$, Trial type 4 NEUTRAL NEGATIVE true, $p = .7559$). A repeated measures ANOVA for the Pepper group proved to be significant, F (3, 36) = 6.504, $p < .001$, $\eta^2_{Partial} = .35$, and Fisher's LSD tests indicated that the trial type 1 (WATER POSITIVE true) differed significantly from the other three trial-types (all *ps* < .02).

The pattern of IRAP effects for the Water group was similar to the Pepper group, with only one of the four trial types (WATER POSITIVE true) showing scores significantly different from zero (Trial type 1, WATER POSITIVE true, p = .0007; Trial type 2 WATER NEGATIVE false, p = .1398; Trial type 3 NEUTRAL POSITIVE false, p = .3641; Trial type 4 NEUTRAL NEGATIVE true, p = .0824). A repeated-measures ANOVA was again significant, F (3, 36) = 4.525, p = .009, $\eta^2_{Partial}$ = .27, with Fisher's LSD tests indicating that trial type 3 (NEUTRAL POSITIVE false) was significantly different from trial type 1 WATER POSITIVE true $p < 0.0008$) and from trial type 4 NEUTRAL NEGATIVE true (p = 0.0463); but not from trial type 2 WATER NEGATIVE false (p > .05)

Additional analyses were conducted to make direct comparisons across the three conditions. Specifically, Four separate between-participants ANOVAs were used to compare the D_{IRAP} scores for each of the four trial types for each group (Water, Pepper, , and Double Pepper). The ANOVA for trial type 1 WATER POSITIVE true was significant F(2,27)= 7.063 p= 0.0034, η_p^2 = 0.34; the ANOVA for trial type 2 WATER NEGATIVE false was marginally significant, F (2,27) = 2.953 p= 0.06, η_p^2 = 0.17. Fisher's LSD tests for trial type 1 WATER POSITIVE true indicated that the Double Pepper group differed significantly from both Pepper (p = 0.0095) and Water (p = 0.0088) groups, but Pepper and Water did not differ from each other (p > .05). Fisher's LSD for trial-type 2 (WATER NEGATIVE false) indicated that the Double Pepper group differed significantly from the Pepper group (p = 0.00237), but did not differ from the Water group (p > 0.05). The ANOVAs for the remaining trial types 3 (NEUTRAL POSITIVE false) and 4 (NEUTRAL NEGATIVE true) were not significant (ps > .05), and thus no post-hoc tests were conducted. Overall, the conclusions arising from the inferential statistics were broadly consistent with the descriptive statistics presented in Figure 4, with the Double Pepper group yielding the strongest response biases for the two water trial-types (WATER POSITIVE true and WATER NEGATIVE false).

Discussion

The present study aimed to explore the extent to which the IRAP, as a measure of derived transformation of functions, is sensitive to motivating variables. As motivating conditions, participants were divided into three groups that differed regarding water or pepper intake, before they evaluated a nonsense symbol equivalent to water, by means of a semantic differential and an IRAP task with positive and negative words. Results indicate that the motivating conditions progressively affected the measures obtained in the semantic differential and in the IRAP. Semantic differential ratings for stimuli D1 (equivalent to water) and D2 (equivalent to neutral) were close to neutrality for the Water group. For the Pepper and Double Pepper groups, however, the stimulus equivalent to water (D1) showed a positive valence, whereas the valence of stimulus D2 (equivalent to neutral) was neutral. The IRAP scores revealed positive biases for both the Water and Pepper groups for D1, with the Double Pepper group producing the largest positive bias scores compared to the other groups. The findings of the current study are generally consistent with previous research showing that the IRAP may be used as a measure of derived transformation of function (Perez et al., 2019). The findings of the current study are unique, however, in showing that the relative sizes of the IRAP effects, produced via derived transformation, appeared to be sensitive to the impact of a motivating variable that was employed to increase the positivity of a stimulus that was in an equivalence class with a picture of water.

The current findings are also consistent with previous research showing that multiple variables may affect the transformation of stimulus functions. Among these variables, for instance, are the types of relational training (MTS vs. DMTS, Bortoloti &

de Rose, 2009), the number of stimuli in each class (Bortoloti & de Rose, 2009), overtraining of baseline relations (Bortoloti et al., 2013), and the training structure (Bortoloti & de Rose, 2011). The present findings point to the potentially important role played by the motivating context in which tests for derived transformation are conducted.

The current findings appear to be relevant to a recent conceptual analysis of the IRAP in terms of the differential arbitrarily applicable relational responding effects (DAARRE) model (e.g., Finn, Barnes-Holmes, & McEnteggart, 2018). The basic argument behind the model is that IRAP effects may arise from the interaction among various properties of the IRAP task. For example, the relational coherence between label and target stimuli (the C_{rel} properties of the stimuli) may be partly responsible for an IRAP effect; thus, in the present study, participants may have found it easier to respond quickly when a water-related stimulus was coordinated with positive words. In addition, coherence among the individual functions of the label, target, and response option stimuli (the C_{func} properties) might also contribute an IRAP effect. In the current case, the participants may have found it easier to respond quickly when the water-related stimulus, the positive words, and the response option “True” all possessed appetitive and/or orienting functions. The current findings do not allow us to determine whether one or some combination of these variables was responsible for the scores. Nevertheless, the fact that the IRAP effect for the water-positive trial-type was consistently larger than the effect for the neutral-negative trial-type (across each of the three conditions) suggests that the second source of control played some role in each case. Perhaps a replication of the current study could employ response options with Cfunc properties lower in positive versus negative valence (e.g., “similar” and “different” rather than true and false; see Maloney and Barnes-Holmes, 2016).

In conducting future research that aims to extend the current findings it may be interesting to add some form of water-consumption behavioral approach task, to determine whether explicit and implicit responses to the derived water stimuli are correlated to actual water consumption. For example, participants could be provided with access to actual water immediately after the IRAP to determine how much liquid they consumed. Alternatively, or in addition, perhaps a measure of relative "thirst" could be included in the task by means of a Likert scale or a thirst related questionnaire. An additional issue to consider is that there was no baseline measure of the sensitivity to the pepper, but it was possible to observe its effect in the stimuli functions. On balance the participant were assigned completely randomly. Nonetheless it was not formally checked, and future studies might include a baseline check for pepper sensitivity.

In closing, it seems important to emphasize that the current study is only one of a small number that has reported that derived transformation effects may be revealed using the IRAP as a measure. As noted by Perez et al. (2019), the use of the IRAP in this context may control for experimenter demand effects in derived transformation research in a way that other measures do not. The fact that the current study also showed that the relative size of the IRAP effects was sensitive to a motivational variable serves to highlight the potential benefits of using the IRAP in future research to explore the impact of motivational variables on derived transformation of function effects more generally.

Compliance with Ethical Standards

Conflict of interest. The authors declare they have no conflict of interest.

Funding. The present research was supported by the São Paulo Research Foundation (FAPESP 2016/05935-6, 2017/10037-0, 2008/57705-8) and by the Brazilian National Research Council (CNPq, Grants 573972/2008-7 and 465686/2014-1); the present study is also part of the research program of the Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE).

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 and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The research is approved by the Brazilian platform for ethical committees (Plataforma Brasil, CAAE# 59350016.1.0000.5375).

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

Availability of Data and Materials. Data and materials can be provided via e-mail to the first author or via ResearchGate (https://www.researchgate.net)

References

Barnes, D., Brown, M., Smeets, P. M., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three to six-year-old children. *The Psychological Record*, *45*, 617–638. https://doi.org/10.1007/BF03395151

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*, 169–177.

Bortoloti, R., & de Rose, J. C. (2009). Assessment of the relatedness of equivalent stimuli through a semantic differential. The Psychological Record, 59, 563–590. https://doi.org/10.1007/BF03395682

Bortoloti, R., & Coelho de Rose, J. C. (2011). Avaliação do efeito de dica semântica e da indução de significado entre estímulos abstratos equivalentes. *Psicologia: Reflexão e Crítica*, *24*, 381-393. http://dx.doi.org/10.1590/S0102-79722011000200020

Bortoloti, R., & de Rose, J. C. (2012). Equivalent stimuli are more strongly related after training with delayed matching than after simultaneous matching: A study using the Implicit Relational Assessment Procedure (IRAP). *The Psychological Record*, *62*, 41-54. doi: 10.1007/BF03395785

Bortoloti, R., Rodrigues, N. C., Cortez, M. D., Pimentel, N., & de Rose, J. C. (2013). Overtraining increases the strength of equivalence relations. *Psychology & Neuroscience*, *6*, 357. http://dx.doi.org/10.3922/j.psns.2013.3.13

de Almeida, J. H., & de Rose, J. C. (2015). Changing the meaningfulness of abstract stimuli by the reorganization of equivalence classes: Effects of delayed matching. *The Psychological Record*, *65*, 451-461. doi: 10.1007/s40732-015-0120-9

de Rose, J. C., Hidalgo, M., & Vasconcellos, M. (2013). Controlling relations in baseline conditional discriminations as determinants of stimulus equivalence. *The Psychological Record, 63*, 85-98. https://doi.org/10.11133/j.tpr.2013.63.1.007

de Rose, J. C., McIlvane, W. J., Dube, W. V., Galpin, V. C., & Stoddard, L. T. (1988). Emergent simple discriminations established by indirect relations to differential consequences. *Journal of the Experimental Analysis of Behavior*, *50*, 1-20. doi: 10.1901/jeab.1988.50-1

Dougher, M., Augustson, E., Markham, M., Greenway, D., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 62*, 331-351. doi:10.1901/jeab.1994.62-331

Dymond, S., Bennett, M., Boyle, S., Roche, B., & Schlund, M. (2017). Related to anxiety: arbitrarily applicable relational responding and experimental psychopathology research on fear and avoidance. *The Behavior Analyst*, *41*, 189-213. doi: 10.1007/s40614-017-0133-6

Dube, W. V., & Hiris, E. J. (1999). MTS software documentation. *Waltham, MA: EK Shriver Center*.

Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23*, 239-254. https://doi.org/10.1007/BF03392013

Finn, M., Barnes-Holmes, D., & McEnteggart, C. (2018). Exploring the single-trial-type-dominance-effect on the IRAP: Developing a differential arbitrarily applicable relational responding effects (DAARRE) model. *The Psychological Record*, *68*, 11–25. https://doi.org/10.1007/s40732-017-0262-z.

Guinther, P. M., & Dougher, M. J. (2015). The clinical relevance of stimulus equivalence and Relational Frame Theory in influencing the behavior of verbally competent adults. *Current Opinion in Psychology, 2*, 21-25. https://doi.org/10.1016/j.copsyc.2015.01.015

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2008). *International affective picture system (IAPS): Technical manual and affective ratings.* NIMH Center for the Study of Emotion and Attention, 1*, 39-58.*

Luciano, C., Valdivia-Salas, S., Ruiz, F. J., Rodrıguez-Valverde, M., Barnes-Holmes, D., Dougher, M. J., & Gutierrez-Martinez, O. (2014). Effects of an acceptance/defusion intervention on experimentally induced generalized avoidance: A laboratory demonstration. *Journal of the Experimental Analysis of Behavior, 101*, 94–111. doi: 10.1002/jeab.68

Hayes, S. C., Kohlenberg, B. K., & Hayes, L. J. (1991). The transfer of specific and general consequential functions through simple and conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 56*, 119-137. doi: 10.1901/jeab.1991.56-119

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science, 1*, 17-38. https://doi.org/10.1016/j.jcbs.2012.09.003

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

Leech, A., Barnes-Holmes, D., & Madden, L. (2016). The implicit relational assessment procedure (IRAP) as a measure of spider fear, avoidance, and approach. *The Psychological Record, 66*, 337–349. https://doi.org/10.1007/s40732-016-0176-1.

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*, 395-403. doi: 10.1007/s40732-016-0180-5

Mizael, T. M., de Almeida, J. H., Silveira, C. C., & de Rose, J. C. (2016). Changing racial bias by transfer of functions in equivalence classes. *The Psychological Record, 66*, 451-462. doi: 10.1007/ s40732-016-0185-0

Nicholson, E., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure (IRAP) as a measure of spider fear. *The Psychological Record*, *62*, 263-277. https://doi.org/10.1007/BF03395801

Osgood, C. E., & Suci, G. J. (1952). A measure of relation determined by both mean difference and profile information. *Psychological Bulletin*, *49,* 251-262. http://dx.doi.org/10.1037/h0062981

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana, IL: University of Illinois Press.

Perez, W. F. (2014). *RelationalFraming* (1.0) [Computer software]. São Paulo, SP: Paradigma–Centro de Ciências do Comportamento

Perez, W. F., de Almeida, J. H., & de Rose, J. C. (2015). Transformation of meaning through relations of sameness and opposition. *The Psychological Record*, *65*, 679-689. https://doi.org/10.1007/s40732-015-0138-z

Perez, W. F., de Almeida, J. H., de Rose, J. C., Dorigon, A. H., de Vasconcellos, E. L., da Silva, M. A., ... & Barnes-Holmes, D. (2019). Implicit and explicit measures of transformation of function from facial expressions of fear and of happiness via equivalence relations. *The Psychological Record*, *69*(1), 13-24. doi.org/10.1007/s40732-018-0304-1

Perez, W. F., Fidalgo, A. P., Kovac, R., & Nico, Y. C. (2015). The transfer of Cfunc contextual control through equivalence relations. *Journal of the Experimental Analysis of Behavior, 103,* 511-523. https://doi.org/10.1002/jeab.150

Perez, W. F., Tomanari, G. Y., & Vaidya, M. (2015). Effects of select and reject control on equivalence class formation and transfer of function. *Journal of the Experimental Analysis of Behavior*, *104*, 146-166. https://doi.org/10.1002/jeab.164

Perez, W. F., Kovac, R., Linares, I. M., Barbosa, S. F., Gomes, C. T., dos Santos, G. A., & Almeida, J. H. D. (2017). Utilizando o AAQ-II e o IRAP para prever a ocorrência de esquiva experiencial. *Revista Brasileira de Terapia Comportamental e Cognitiva*, *19*, 34-48. https://doi.org/10.31505/rbtcc.v19i3.1052

Scoville, W. L. (1912). Note on capsicums. Journal of Pharmaceutical Sciences, 1, 453-454. Sidman, M. (1987). Two choices are not enough. *Behavior Analysis*, *22*, 11-18.

Skinner, B. F. (1938). *The Behavior of Organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

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. doi.org/10.1016/j.jbtep.2015.01.004

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*, 33-50. https://doi.org/10.1007/BF03395092

Table 1.

Results on Equivalence Training and Testing. Number of trials to met mastery criteria on training (AB, AC and CD) and number of correct responses/total test trials during equivalence tests (BD and DB).

Group	Participant	Training				Testing	
		AB	AC	CD	Mixed	BD	DB
Water	P1	48	24	39	36	10/10	10/10
	P2	19	37	25	36	5/10	10/10
	P3	18	21	20	36	10/10	10/10
	P4	27	25	38	36	10/10	10/10
	P5	43	18	22	36	10/10	10/10
	P6	29	19	30	38	5/10	10/10
	P7	22	18	22	36	10/10	10/10
	P8	20	18	21	36	8/10	10/10
	P9	24	21	19	36	10/10	10/10
	P10	31	27	25	36	10/10	10/10
Pepper	P1	22	21	18	54	10/10	10/10
	P2	28	34	36	54	10/10	10/10
	P3	22	20	22	54	10/10	10/10
	P4	22	47	20	54	10/10	10/10
	P5	37	20	27	54	9/10	10/10
	P6	21	19	24	54	9/10	10/10
	P7	24	24	21	54	10/10	10/10
	P8	26	19	21	54	10/10	10/10
	P9	28	19	19	84	10/10	10/10
	P10	39	19	25	54	9/10	10/10
Double Pepper	P1	27	24	27	36	10/10	10/10
	P2	42	29	26	42	10/10	9/10
	P3	22	25	27	36	10/10	10/10
	P4	29	20	19	49	9/10	10/10
	P5	23	21	25	36	10/10	9/10
	P6	22	20	18	36	10/10	10/10
	P7	19	18	18	36	10/10	10/10
	P8	18	21	18	36	10/10	10/10
	P9	28	24	21	47	10/10	9/10
	P10	25	26	23	38	10/10	10/10

Figure 1. Outline of the procedure.

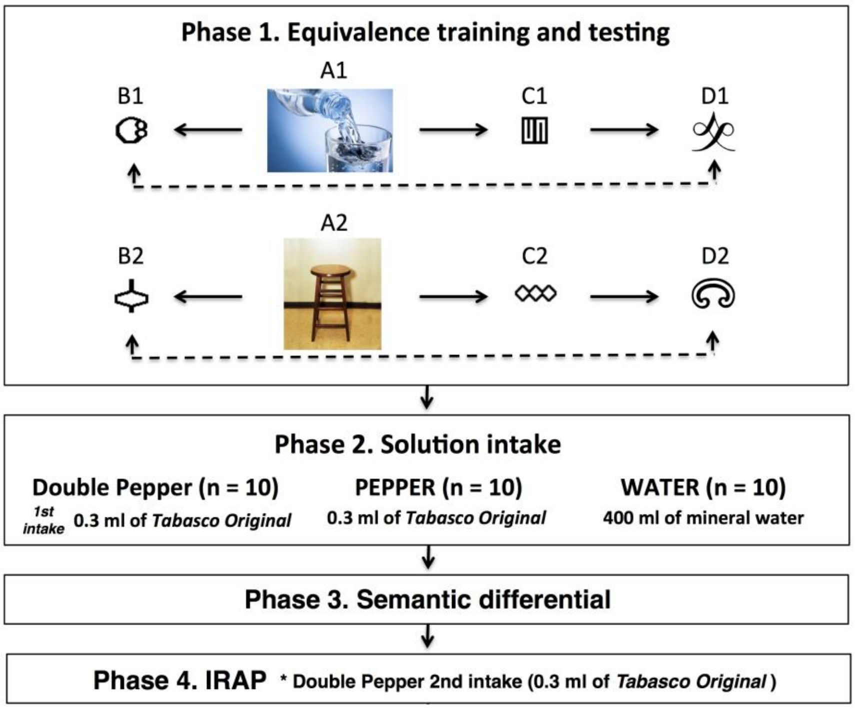

Figure 2. Schematic representation of the four IRAP trial types. D stimuli (D1 or D2) are presented on the top of the screen, the positive or negative words on the center and the response options on the lower corners. The correct response option in each trial type is alternated across consistent and inconsistent blocks.

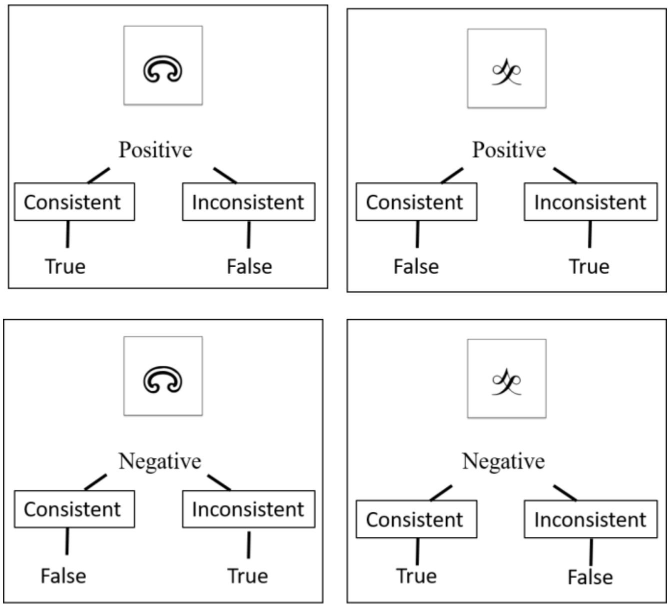

Figure 3. Semantic differential profiles for each group (Water, Pepper and Double Pepper). Median values for D1 (black dashed line) and D2 (gray solid line) for each of the 13 bipolar scales.

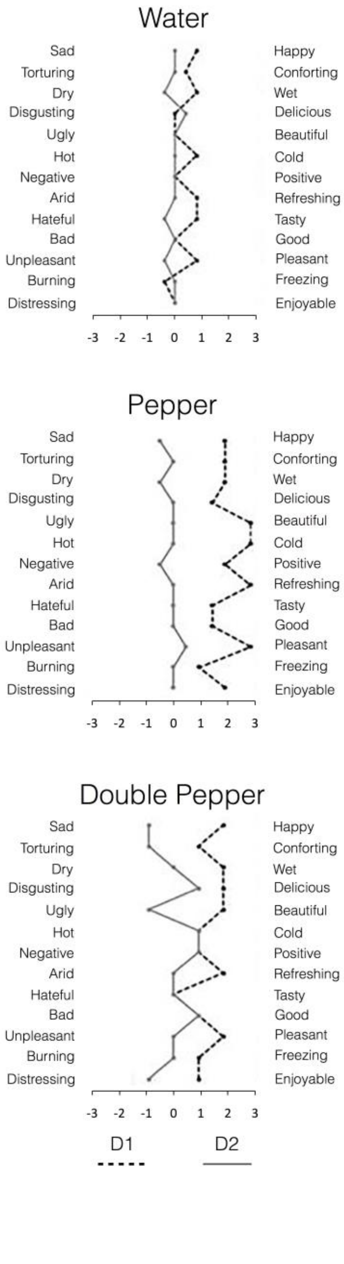

Figure 4. Results from the IRAP. D_{IRAP} scores for each IRAP trial type. To facilitate description, trial types involving D1 (water-equivalent) stimulus were labels as Water – Positive and Water-Negative; trial types involving D2 (neutral-equivalent) stimulus were labeled as Neutral-Positive and Neutral-Negative; "T" and "F" stands for the response options "True" and "False".