



Deictic Framing Performance Fails to Generalize to Other Perspective-Taking Tasks

Tokiko Taylor¹  · Rebecca J. Sargisson¹  · Timothy L. Edwards¹ 

Accepted: 22 May 2023
© The Author(s) 2023

Abstract

The behavioral processes underlying perspective taking have not been studied extensively. One approach to understanding and enhancing perspective taking, deictic framing, has been proposed. Proponents of this approach have suggested that deictic framing is a core property of perspective taking. A training protocol based on deictic framing has been developed and tested, but researchers generally evaluate the effectiveness of the protocol using tasks that have a similar format to the training protocol. Little research has examined the protocol's effectiveness for improving performance in different perspective-taking tasks. We investigated generalization of the performance of three groups of university students trained with a deictic-framing protocol (or not) and tested with two other perspective-taking tasks: a visuospatial perspective taking using a cupboard containing a range of objects and a version of the implicit relational assessment procedure specifically designed to measure perspective taking. The first group was trained with the original verbal protocol with deictic expressions; the second group with the same protocol involving nondeictic words; and the third group was merely exposed to deictic expressions as a control condition. The results suggested that deictic framing is not fundamental to perspective taking, as the performance of the two experimental groups was not significantly different from the control group's performance. Identification of specific stimulus functions involved in successful perspective taking and how those functions can be established should be addressed in future research.

Keywords Theory of mind · Perspective taking · Deictic framing · Relational frame theory

Perspective taking is an ability of critical importance that helps people establish and maintain relationships, negotiate deals, predict the actions of others, and achieve a wide variety of other valuable outcomes. Perspective taking has been defined as an ability to perceive another person's situation by "spontaneously adopt[ing] the psychological point of view of others" (Davis, 1983, p. 114). Premack and Woodruff (1978) coined the term "theory of mind" (ToM), theorizing that people rely on mental attribution to predict what other people may think, believe, feel, or otherwise experience.

Under the rubric of relational frame theory (RFT; Hayes et al., 2001), some researchers have proposed "deictic framing" as a core component of ToM (Barnes-Holmes et al., 2004). RFT was developed as an extension of Skinner's (1957)

original functional account of verbal behavior to include "the action of framing events relationally" (Hayes et al., 2001, p. 43). The theoretical framework is based on "a generalized pattern of arbitrarily applicable relational responding (AARR); that is, relational responses that are not based solely on the formal properties of the stimulus relations" (Barnes-Holmes & Harte, 2022, p. 3). According to RFT, AARR is characterized by three defining relational properties: mutual entailment (e.g., if $A > B$, then $B < A$), combinatorial entailment (e.g., if $A > B$ and $B > C$, then $C < A$), and the transformation of those stimulus functions (e.g., if C elicits a response, A will elicit a response of greater magnitude as a result of its relationship with C). These properties are direct outcomes of contingencies of reinforcement involving specific relational stimuli as antecedents across multiple exemplars (i.e., the relational responses are generalized operants; see Hayes & Barnes-Holmes, 2004; Healy et al., 2000; Palmer, 2004a, b).

There are multiple types of AARR (Hayes et al., 2001). Relational responding in coordination, for example, entails a quality of sameness or stimulus equivalence (e.g., A "is"

✉ Tokiko Taylor
tokitaylor@gmail.com

¹ School of Psychology, University of Waikato,
Hamilton 3216, New Zealand

B, or A “is similar to” B)¹. Another example of relational responding is a frame of distinction (e.g., A “is different from” B). Hayes et al. proposed that children learn such frames through questions presented by a verbal community, sometimes explicitly through educational programs (for instance, “one of these things is not like the others; one of these things does not belong; can you guess which thing is not like the others?”). Responding to such exercises in everyday life is considered to be a prerequisite for children eventually developing derived relational responding of similarity or distinction among arbitrary stimuli (i.e., regardless of differences in their physical features) in the presence of specific contextual stimuli, which indicate when a type of relational response is likely to be effective. The application of the correct relational response under the relevant contextual control is reinforced by the verbal community, or by some other outcome of the relational response (such as the ability to solve a problem). The specific contextual stimuli, which are sometimes explicit (e.g., “is similar to” or “is different from”) and sometimes subtle (e.g., contexts of causal relational responding such as “X causes Y”), evoke AARR, which can then be applied to novel stimuli in different contexts. One proposed type of relational responding, deictic relational framing, specifies a relation from the perspective of a speaker, and is controlled by deictic cues such as I versus You, Here versus There, and Now versus Then. Deictic relational framing is thought to generate a “constant division” (Hayes et al., 2001, p. 124) between the speaker (e.g., from one’s stance of being always Here and Now) and the person or object being spoken about (e.g., being always There and Then); “I am here now, but you were here then” or “You and I are both here now, but I was here then” (p. 124). Some researchers have suggested that deictic framing is fundamental to perspective taking and ToM, and that understanding this type of relational responding will advance our understanding of perspective taking (Dymond & Barnes, 1997; Hayes, 1984; McHugh et al., 2004a).

Barnes-Holmes Protocol

RFT researchers have developed a perspective-taking protocol (sometimes referred to as the Barnes-Holmes protocol [BH protocol], named after the creator of the list) based on the concept of deictic framing (see Hayes et al., 2001; McHugh et al., 2004a, 2004b). The protocol has 62 verbal tasks that include deictic components (e.g., I, you, here, there, now, then). The tasks increase in three levels

of complexity from simple to reversed to double-reversed (McHugh et al., 2004a, 2004b). In a simple task, the participant only needs to identify contextual information provided in a sentence. For example, they are told, “I have a red brick, you have a green brick” and are then asked, “What do I have? What do you have?” In a reversed task, a conditional statement such as “If I were you and you were me” is provided in addition to the simple task, so the correct answer would be “I have a green brick” and “you have a red brick.” Lastly, a double-reversed task might be: “I am sitting here on a black chair and you are sitting there on a blue chair, if I were you and you were me, and here were there and there were here, where would you be sitting?” The correct answer is “here on a blue chair.”

McHugh et al. (2004a) conducted one of the first studies testing the protocol with people from different age groups (five groups in total: 18–30 years [adulthood]; 12–14 years [adolescence]; 9–11 years [late childhood]; 6–8 years [middle childhood]; and 3–5 years [early childhood]) to investigate whether the performance data were consistent with the developmental profile commonly found in ToM studies (i.e., the rate of error responses decreasing as a function of age). In the simple trials, all participants performed with high accuracy, except for those in the early childhood group, who performed with low accuracy. Other childhood groups (6–8 and 9–11 years) produced more errors (50%–70%) in the reversed and double-reversed trials compared to the adult group (30%–40%). The early childhood group did not perform well on complex trials, with a mean error rate of 80%–90%. In addition, of the three deictic components, participants’ performance on Now–Then was lower than on Here–There and I–You trials. The authors pointed out that their findings correspond with the progression of development of ToM.

RFT Deictic Framing Training Studies

The BH protocol has been used in various empirical investigations for assessing and training both typically developed children aged 4–7 years (Davlin et al., 2011; Heagle & Rehfeldt, 2006; Montoya-Rodríguez & Cobos, 2016; Weil et al., 2011) and atypically developed children and adults, including those considered to have deficiencies in perspective-taking ability. This includes individuals with autism aged 6–18 (e.g., Barron et al., 2018; Belisle et al., 2016; Gilroy et al., 2015; Jackson et al., 2014), high-functioning autism aged 6–18 (e.g., Lovett & Rehfeldt, 2014; Rehfeldt et al., 2007; Tibbetts & Rehfeldt, 2005), schizophrenia (O’Neill & Weil, 2014; Villatte et al., 2010), and social anxiety disorder (Janssen et al., 2014).

Some researchers have investigated whether learning the BH protocol leads to higher performance in other ToM tasks, such as Howlin et al.’s (1999) Level 3

¹ According to RFT, Sidman’s stimulus equivalence (Sidman & Tailby, 1982) is an example of relational frames, such as the “frames of coordination.”

seeing-leads-to-knowing task (sensory attending or seeing is understood as knowing), the Level 4 task for an ability to predict actions based on true belief, or the Level 5 task to teach an ability to predict whether people can act on the basis of a false belief. Other ToM tasks have also been used, including the Unexpected Contents task (or sometimes called the “Smarties” or “M&M” task),² Hinting task,³ and Theory of Mind Inventory (ToMI).⁴ However, training with the BH protocol does not always influence other ToM measures. To date, few studies have reported generalization of the skills acquired from training with the BH protocol over to different ToM tasks. For example, following training with the BH-protocol, O’Neill and Weil (2014) demonstrated an improvement in performance by all three participants with schizophrenia in both the Unexpected Contents Task and the Hinting Task (although one of the three participants was unable to complete the final session of the Hinting Task and the other two participants scored less than 80%). In contrast, improvements were not consistently observed in Howlin et al.’s (1999) Level 3 to 5 ToM tasks following training with the BH protocol (Jackson et al., 2014; Lovett & Rehfeldt, 2014; Montoya-Rodríguez & Cobos, 2016; O’Neill & Weil, 2014; Weil et al., 2011). None of Jackson et al.’s (2014) participants (three children with autism) improved their scores on the ToM tasks after the BH protocol training. In Weil et al.’s (2011) study, only one of three normally developing children scored 100% correct in the ToM tasks after deictic training; the remaining two children showed some improvements but their performance remained at chance levels. Lovett and Rehfeldt (2014) used the Social Language Development Test-Adolescent (SLDT-A) and ToMI to examine the effect of mastering the BH protocol among three young adults with Asperger syndrome. The SLDT-A scores increased for two participants between the pre- and posttraining probe trials; however, the scores for ToMI did not change for any of the three participants.

² This task is designed to test the participant’s ability to identify a false belief. For example, a box of crayons is shown to participants who are asked what they think is inside. Then, the experimenter reads the following statement: “If I put the pencils in a chalk box and you are not here, you would think the chalk box contains chalk.” The participant is then prompted to answer true or false.

³ The ToM Hinting task has 10 short scenarios describing interactions between two characters. For example, “Stephanie says to her friend Nicole: ‘I can’t afford the repairs on my car. Could you lend me some money?’ Nicole answers: ‘I have to have my car repaired too.’” Then, the participant is asked “What does Nicole really mean when she says this?” Then a second hint is provided; “Nicole then says, ‘The repairs to my car are going to be very expensive.’” The participant is again asked, “What does Nicole really mean when she says this?”

⁴ ToMI includes 48 items assessing inferences about emotion and belief based on caregiver reports that reflect the caregiver’s evaluation of the perspective-taking ability of adolescents with autism.

Critical Analysis of the BH Protocol Deictic Frame

The term deixis comes from the Greek word meaning pointing out or drawing attention to (Harman, 1990). According to the Oxford English Dictionary, the word deictic is defined as an “expression whose meaning is dependent on the context in which it is used such as here, you, me, that one there, or next Tuesday.” If a deictic word such as there is presented without additional contextual information, it will not make sense to a listener (i.e., they will not be able to respond effectively). Deictic statements must be accompanied by additional contextual cues, such as the presence of an item of discussion when the word there is spoken. These cues may also be provided directly by the speaker (e.g., a pointing gesture to a broken window when saying that in a crowded workshop). Without these other cues, the listener would have difficulty responding appropriately to the verbal statement.

Deictic cues were described as a critical element of perspective-taking ability in the original RFT conceptual work explaining perspective-taking ability (Hayes et al., 2001). In addition, some RFT researchers have specified that deictic cues function as “deep grammar,” influencing all other framing to produce what we know as perspective taking (Hayes et al., 2001, p. 124). Hayes et al. (2001) explained “deictic” relational framing as follows:

The frames of I and You, Here and There, and Now and Then . . . are unlike most of the other relational frames in that they do not appear to have formal or non-arbitrary counterparts. . . . Frames of perspective have no simple nonverbal counterpart, and must be taught through demonstration and multiple exemplars without any use of formal properties. For that reason, they are sometimes called “deictic” relations—literally, demonstrative relations that must be “shown directly”—but these relations are anything but direct. (p. 122)

This description suggests that RFT researchers use “deictic” somewhat differently to how linguists use the term. The conceptual framework is also somewhat unclear. Hayes et al. (2001) emphasized the strictly arbitrary nature of this particular relational frame but, in the same work, stated that we can replace the deictic cues with nondeictic nouns (i.e., it is fine to use “Emily” instead of “you” or “Burger King” in place of “there”) in the deictic training protocol. This appears to contradict the conceptual view of proponents of deictic framing and linguists with respect to the notion of “deictic expressions,” whose meaning changes depending on the context in which they are used (e.g., “there” will change in meaning depending on where the speaker is pointing or looking but “Hamilton, New Zealand” will not change).

In some of the studies where BH protocols were used as an intervention to improve participants’ perspective-taking ability, researchers modified the text of the protocols, replacing some of the deictic expressions with familiar character names

and their actions from stories or cartoons familiar to participants, such as Cinderella and SpongeBob (e.g., “You are waiting for recess and Cinderella is dancing at the ball, what are you doing? What is Cinderella doing?”; Davlin et al., 2011; Gilroy et al., 2015; Heagle & Rehfeldt, 2006; Montoya-Rodríguez et al., 2017). Participants in these studies were mainly examined for the demonstration of stimulus generalization by increasing their response accuracy in the slightly modified BH protocol tasks, which were provided using stories different to the one used in the training session. Some questions arise over the validity and utility of the concept of deictic framing and the BH protocol. What is controlling perspective taking if the deictic words, which are defined as the main contextual cues, can be swapped with other nondeictic words? If the BH protocol contributes to the acquisition of “deep grammar,” why do many studies fail to show response generalization to different ToM tasks? Moreover, if we cannot rely upon other ToM tasks to provide evidence of a central role of deictic framing in perspective taking, how can we evaluate this hypothesis?

Guinther (2017) pointed out what appears to be a theoretical flaw in the concept of deictic framing: deictic framing does not entail any relational component that constitutes relational responding. Guinther noted that, “in contrast to the relatively straightforward entitlement and transformation markers of other relational framings, it is at present unclear how I–you, here–there, and now–then deictic relational framings are to be functionally identified under tightly controlled conditions” (p. 449). The three types of entailments explained earlier (mutual entailment, combinatorial entailment, and transformation of stimulus functions) do not seem to be applicable to deictic framing. For example, there appears to be no mutual entailment between “I” and “You” unless in the context of some additional framing, such as coordination framing (e.g., once we learn the unusual relation “I” am “You,” then the relation “You” are “Me” may be derived). The functionality associated with the deictic expressions could be defined by other relational framings, but nothing would be entailed by the deictic expressions themselves. For example, “I am to the left of you and you are to the right of me,” indicates the spatial cues (i.e., “left of” and “right of”) that are required to derive other relations (the relative location) between the stimuli, and “I” and “You” just happen to be the stimuli participating in the frame. In addition, without being specified by “if . . . then” conditional framing, we would not be able to do perspective taking as it is defined in many tasks, including the BH protocol, to predict or intuit the knowledge or behavior of others (e.g., If you were me and I were you, what would you be doing and what would I be doing?). It appears that the work the participant is doing in such a protocol is solving the if–then puzzle, which can be done by attending to the reversal cues. Substituting nonsense stimuli for “I,” “You,” “Here,” or “There,” does not seem to fundamentally change the puzzle. However, an empirical analysis is required to test this hypothesis. Clarification is required to find out whether

deictic framing has been conceptualized coherently within RFT, and to discover the role of deictic expression in perspective taking from a functional language perspective.

The current study may help to clarify some of the questions regarding the theoretical assumptions of the concept of deictic framing, and shed light on the features of the protocol that are responsible for its effectiveness (or lack thereof) in improving perspective taking. To investigate the effectiveness of the deictic components of the protocol, we compared the effects of training in three different groups:

- 1) BH+ Group, a group of university students who were trained using the existing BH protocol (see Appendix A);
- 2) BH– Group, a group of students who were trained using a modified version of the BH protocol with specific nouns instead of deictic expressions (see Appendix B); and
- 3) A control group of students, who were exposed to parts of the original BH protocol with deictic expressions, but without the problem-solving part of the protocol. The members of the control group were simply asked to select a deictic word used in a given sentence (see Appendix C).

The dependent variables were a tally of participants’ correct responses on an activity indicating their visuospatial perspective-taking, a “cupboard task” that involved moving items in a cupboard in response to instructions from someone who could only see some items in the cupboard, and a version of the implicit relational assessment procedure (IRAP) specifically designed to measure perspective taking (self vs. others). For the latter task, the performance of each group (i.e., number of correct responses and latency) was recorded and analyzed to determine if the training produced any differences between the groups. Raven’s Progressive Matrices (RPM) were also administered before introducing participants to the BH protocol to evaluate the degree to which the results of this test were predictive of performance in the other measures, including performance with the BH protocols. Given that, conceptually, deictic expressions do not appear to play a critical role in the BH protocol, if the BH protocol is effective in producing general perspective-taking improvements, we predicted that the experimental groups (the BH+ and BH– groups) would perform better than the control group on both of the cupboard task and the IRAP task but that the performance of the experimental groups would not differ significantly from each other.

Method

Power Analysis

A priori power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007) to estimate the minimum sample size required to test the study hypothesis. Results

indicated the required sample size to achieve 80% power for detecting a medium effect (Cohen's $f^2 \geq .15$; Cohen, 1988), at a significance criterion of $\alpha = .05$, was $N = 45$ for a mixed ANOVA on the analysis of the first data set using a cupboard, and $N = 30$ for the second data set analysis of the result measured by IRAP.

Participants

A total of 98 university students was recruited. The participants were randomly assigned to each group: (1) 33 participants to a deictic (original Barnes-Holmes protocol) group (BH+); (2) 33 participants to a nondeictic Barnes-Holmes protocol group (BH-); and (3) 32 participants to a control group. Students were recruited through advertisements on bulletin boards around the University of Waikato. Participants chose either course credit (1% for each hour of participation) or entry in a draw to win one of five \$50 vouchers that could be spent at a local store upon completion of the experiment.

Experimental Tasks

Raven's Progressive Matrices

Raven's Progressive Matrices (RPM) are 60 visual-puzzle tasks measuring analogical reasoning and problem solving, and are generally used for assessing "intelligence" (Raven, 1984). A participant is presented with geometrical figures, which each have one part missing, and then selects an option from six alternatives. We provided participants with a booklet of problem items and a piece of paper where they filled in their answers to the multiple-choice questions. Validity studies indicate that correlations with the RPM and several subtests of the Wechsler Adult Intelligence Scale–Third Edition range between 0.75 and 0.88 (Lezak et al., 2004). Internal consistency was found to be 0.89 and split-half reliability was 0.91 (Cotton et al., 2005). In another study done in Kuwait, the test–retest reliability ($n = 969$ aged from 8 to 15 years) was between 0.88 and 0.93 (Abdel-Khalek, 2005). RPM is commonly used as a measure to explore the effects of relational training on cognitive abilities in RFT research (see Janssen et al., 2014; Thirus et al., 2016; Villatte et al., 2010); we administered RPM to evaluate the intellectual ability of participants in each group so that we could ensure that the groups were matched for their intellectual ability.

Visuospatial "Cupboard" Perspective-Taking Task

The visuospatial perspective-taking cupboard task was implemented in the manner described by Keysar et al. (2003) and Ferguson and Cane (2017) to assess adults' perspective-taking abilities. We designed our task to investigate the extent to which participants behaved in accordance with information

that the person giving instructions had access to, which is demonstrative of one form of perspective taking. The task involved a cupboard, a piece of furniture containing 16 cells (in a 4-x-4 layout), with the view of some of the cells occluded from one side of the cupboard but not the other (See Figure 1).

From one side—the open side—a participant could see all the objects placed in the cells, but from the other side, only a subset of the cells could be seen. Five cells were occluded from the experimenter's (i.e., director's) view, and the other 11 cells were mutually perceptible from both the participant's and the director's view. Each cell was 16 cm high, 18 cm wide, and 15 cm deep. The participant sat on the open side of the cupboard and followed a set of instructions given by the director, which instructed them to pick a target object and move it to another cell in the cupboard. The participants could view all the objects placed in the cupboard, but some cells were occluded from the director's view. Thus, the accuracy of the participant's performance in perspective taking was determined by their ability to follow instructions correctly according to the information to which the director has access (i.e., they had to take into account which cells were visible to both them and the director, and which cells only they could see). To perform well on this task, participants had to correctly respond to the trials with critical paired items that had ambiguous names (e.g., an instruction contained a word such as "candle" but there were two candle-like objects in the array: a glass jar candle in an open cell, and a pillar candle in an occluded cell). For comparison, we included trials with baseline objects (i.e., a range of objects unrelated to each other) to show participants' performance under normal conditions without any ambiguous pairings (e.g., given a glass jar candle and a yoyo, in a baseline trial the participant would respond correctly by reaching for the candle when asked to move the object called "candle"). Some objects were displayed in open cells, and some were in occluded cells. In the first session, participants started with four critical trials of ambiguous objects where eight objects with ambiguous names were placed in the cupboard. In Table 1, the list of the trial orders from 1 to 4 in the left column indicates which pairs of objects



Fig. 1 Shelf Used to Display Items in the Visuo-Spatial "Cupboard" Task

were used in each trial with the ambiguous objects. The trials with unambiguous objects followed the first critical session, in which eight unrelated objects were randomly placed in the cupboard (see Table 1 in the right column, Trial order 5 to 8). The remaining two sessions were provided in the order described above, with Trials 9 to 12 using ambiguous objects, and Trials 13 to 16 using unambiguous objects. Two separate video cameras (one positioned at the right side of the participant, and the other one behind them) captured the movement of the participant's arms (e.g., movements like reaching and moving an object to the left, right, above, or below its current position among the 4-x-4 squares of the cupboard).

Implicit Relational Assessment Procedure

The IRAP was developed under the rubric of RFT (Barnes-Holmes et al., 2006) to detect relational framing based on a latency response measure. It is a computer-based task where participants respond to a series of paired stimuli (often words and images) by selecting true or false in response to whether a pair of stimuli are consistent or inconsistent with the participant's historically coherent relational network (i.e., those relational responses that they are likely to have learned to a high degree of fluency). The participants are encouraged to respond both quickly and accurately within specific time limits to meet an accuracy threshold. The IRAP consists of practice blocks (usually three paired blocks of consistent and inconsistent trials, each block containing about 24 trials) and test blocks (six paired blocks of the consistent and inconsistent trials, containing 24 trials each). All response latencies are recorded in the IRAP and are used to calculate the D-IRAP score. The score is a normalized index of raw IRAP response latency, which is similar to Cohen's effect size (Cohen's *D*). The outcome data are used to evaluate whether the observed differences in average response latency between consistent and inconsistent blocks are large enough to conclude that there is a difference between the two. The larger the *D* score value, the bigger the effect size (e.g., 0 is no effect, and above 0.8 indicates a large effect size).

The IRAP has been used to investigate perspective-taking abilities, specifically for comparing responding to "self" versus responding to "other" stimuli (Barbero-Rubio et al., 2016; Kavanagh et al., 2018). In the perspective-taking IRAP test, participants are exposed to a sample stimulus that is either "I," or an experimenter's name, such as "Mary," which represents the perspective of "other." Another stimulus is presented at the bottom of the screen, showing a description of an action such as "standing near the desk," "sitting on the chair," or "staring at the computer monitor." The participants answer "Yes" or "No" in accordance with arranged rule-following conditions that are either consistent (pro-self-perspective) or inconsistent (pro-other-perspective). Participants in these experiments were slower to take the perspective of others, supporting the generally held notion that people are

faster at responding from a pro-self perspective, and slower when responding from the perspective of others.

We used Open Source IRAP software (<https://doi.org/10.17605/OSF.IO/KG2Q8>). During each trial, the software displayed the text "I" or "Tokiko" (the experimenter's name) at the top of the screen. Below the sample words was 1 of 12 words and phrases describing the activities of either "I" (i.e., the participant) or "Tokiko," and these sample and action words appeared simultaneously. The six action words belonging to "I" were "seated," "participant," "with keyboard," "looking at screen," "here," and "blue Post-it." The six words describing "Tokiko" were "standing up," "experimenter," "holding a pen," "holding a notebook," "there," and "pink Post-it." These were the same words used by Barbero-Rubio et al. (2016). In the lower left and right corners of the screen, the software displayed "PRESS 'd' FOR [Yes/No]" and "PRESS 'k' for [Yes/No]" with the position of the words "Yes" and "No" randomly assigned to the left or right in each trial.

Prior to each block of 24 trials, the program displayed the message "answer as if you were you and Tokiko were Tokiko" (consistent block) or "answer as if you were Tokiko and Tokiko were you" (inconsistent block) in the center of the screen and, below the message, "press the spacebar to proceed." The 24 trials contained four trial types: I / I's actions ("I" sample stimulus and words descriptive of the participant), I / Tokiko's actions, Tokiko / Tokiko's actions, and Tokiko / I's actions. The specific stimuli corresponding with each of the four trial types were randomly selected for each trial. In each trial, the participants were presented with the trial stimuli and, if a response did not occur within 2000ms, feedback in the form a red exclamation mark ("!") appeared in the center of the screen. The participant could still respond at any time after the "!" had appeared. The presentation of latency feedback was programmed to start from the second pair of blocks in the practice phase. In all trials, if an incorrect response (i.e., a response that did not correspond with the rule for that block) occurred, a red "X" appeared in the center of the screen, and the trial stimuli remained in place until the correct response occurred. The inter-trial interval was 400ms.

Interobserver Agreement

A second observer watched video footage of each participant's performance on the cupboard task. The video footage was selected by the original rater, who randomly selected seven participants from each of the three groups. These random selections made up 21% of the total participants (21 of 98). A block of probe trials completed by each participant had a total of 16 probes, and the responses were recorded as either correct (1) or incorrect (0). Using the data from the second observer, inter-observer agreement was calculated on a trial-by-trial basis, by dividing the total number of agreements by the total number of test trials offered, then multiplying by 100.

Agreement for the correct or incorrect responses made by each participant ranged from 94% to 100%. For the final analysis, we used only the scores that were agreed upon by both raters.

Procedure

The experimental sessions were conducted in a room in the presence of an experimenter. We asked participants to attend two separate sessions. In the first session (Day 1), RPM was administered. In the second session (Day 2), participants completed the visuospatial “cupboard” perspective-taking task and the IRAP. Between the two sessions, all participants completed a series of online training tasks that were available through the University’s online learning management system (Moodle). We instructed participants to complete the training tasks any time, within 3 days after the first session. The duration of each task was approximately 15 min or less, with each containing 15 questions. Participants completed four of these tasks in total. After completion, the correct answers were shown to the participant.

On Day 1, all three groups followed the same procedures. Upon a participant’s arrival at the experimental room, they took a seat, received a briefing of the experimental information by the experimenter, and signed a consent form. The participant took the RPM test and was handed a sheet to fill in answers to each of the 60 questions. The test took between 15 and 40 min to complete.

After taking the test, the experimenter informed the participants about the online training that they were to complete before returning for the Day 2 session. They were also informed that they could take the online training at any time and from anywhere with internet access using a device capable of running a Google Chrome or Firefox browser, but that all the self-learning training had to be completed within 3 days. To ensure the participants’ understanding of the online training performance requirement, the experimenter showed the participants three example questions from the BH protocol (one of each trial type: simple, reversed, and double reversed), on the monitor. The experimenter asked the participant to answer each question and then provided immediate feedback (correct or incorrect) for each response. For group BH+, the experimenter introduced the deictic protocol (see Appendix A) used by McHugh et al. (2004a). For group BH–, the experimenter introduced a nondeictic protocol that we adapted from the original protocol (see Appendix B). For the control group, the experimenter first made sure that the participants understood the meaning of “deictic” by providing the Oxford dictionary definition of the word and answering any questions from the participants about it. Then, the experimenter introduced the original deictic protocol with all the problem-solving components removed, and trained the participants to perform a simple identification task, encouraging them to point to all the deictic expressions used in a sentence (see Appendix C). The participants were asked to complete a total of four tasks of 15

questions during the 3-day period. They had to repeat each training task if they did not meet the criteria of more than 90% correct within 5 min for the 15 questions. The participants’ answers were automatically evaluated, and feedback was provided. Their completion of the training was monitored by the experimenter. Lastly, the participant booked another appointment for Day 2 within 5 days of the Day 1 session.

On Day 2, upon the participant’s arrival in the same room used in the Day 1 session, they first completed the IRAP task. The experimenter asked the participant to put a blue Post-it note on their arm or chest (where the participant could see it) and the experimenter put a pink Post-it note on her own chest. The experimenter then positioned herself near the participant, who was sitting in front of the desktop computer in the room. The experimenter ensured that participant could see the experimenter wearing the pink Post-it and held a pen and a notebook in her hand while the participant worked on the IRAP task. The order of the IRAP was alternated to check for order bias: the consistent-trials-first group ($n = 45$) was presented with “if you were you and Tokiko were Tokiko” first. The inconsistent-trials-first group ($n = 45$) was presented with “if you were Tokiko and Tokiko were you” first. Within each of the three groups, the participants were randomly assigned to the two IRAP orders.

After completing the IRAP task, we asked the participant to move over to the 4-x-4 cupboard, which had a drape covering it completely from view. The participant was seated on the side where none of the cells were occluded from their view. The experimenter explained that she would play a director’s role, providing a total of 16 different instructions to move an object to a different cell in the cupboard and the participant was told to follow the instruction provided from the director (Table 1). The instructions followed a pattern of “Move the...” + the target object noun (e.g., ball, shoe, truck) + a direction (up, down, left, or right), based on the instructions used by Keysar et al. (2000).

The experimenter provided the following verbal instruction:

Now, we are moving onto the next task, the visuospatial experiment. For this visuospatial task, I am going to be a director who will simply give you some directions to move an object around in the cupboard, which is covered from our view at this point. This is because my research assistant has already set this up, so I have no knowledge on the details of the set-up. I am just going to give you some directions from my perspective (pointing to where I stand at the other side of the cupboard, which is still draped) and all you need to do in this experiment is to follow what I say.

After receiving acknowledgement from the participant, the experimenter proceeded to the practical demonstration as described below. The experimenter started giving directions as follows:

Now, you can remove the drape and put it under the desk. As you can see, different objects are placed randomly in the cells. Please notice that some cells are occluded from my perspective, so I can see all the objects placed in the open cells, but not the ones in the occluded cells.

Next, the experimenter asked the participant to come to her side to experience the experimenter's (the director's) view, then the participant returned to their original seat on the non-occluded side. Then, the experimenter (the director) gave an instruction for the purpose of giving the participant an idea of how the task worked. The instruction was as follows:

As I said earlier, all I am going to do is to give you some directions and you need to follow what I say. So, now we can try some practical trials, "move the Rubik's cube to the next cell on your right." [After the participant moved the object correctly] "Good, now move the bunny one cell down." [After the participant responded], "Good, the directions will be something like that, asking you to move an object to a different cell on the cupboard."

After the practical trial, the participant was asked to open an envelope titled "Picture 1" and to place all the items in the cupboard exactly as shown in the picture. This ensured their knowledge of the objects and their locations (i.e., placed in an occluded or open cell). During this process, the experimenter wore a blindfold to convince the participant that she had no knowledge of the locations of the objects. This process was repeated for the remaining three blocks of test trials. Once the participants finished placing all the objects in the cupboard, the first block of trials began. Instructions were administered in a similar manner as demonstrated in the practical

demonstration (a full list of the objects used and sequences of the test trials can be found in Table 1). Upon completion of all the trials, the participant finished the experiment, and the experimenter thanked them for their participation.

Results

Performance during BH+ and BH- Online Training

Training data showing each participant's performance for both BH+ and BH- groups on number of correct responses, numbers of sessions taken to reach the mastery criteria (i.e., 90% correct responses within 5 min), and total duration taken were summarized in two tables (see Appendix D Tables 2 and 3). These data were analyzed to evaluate participants' performance as they completed the two types of online training over three days. Participants in both groups showed progressive improvement in terms of reducing the numbers of repetitions needed to reach mastery, and becoming faster in completing each task. Major differences in training performance between the two groups were not observed.

The mean number of repeated sessions for the BH+ group was 2.21 (SD = .99), 1.63 (SD = .7), 1.45 (SD = .67), and 1.24 (SD = .44) from Tasks 1 to 4. For the BH- group, they were 1.84 (SD = .76), 1.33 (SD = .69), 1.24 (SD = .66), and 1.24 (SD = .61). For the analysis, a mixed ANOVA used the number of repeated sessions taken to achieve mastery from Task 1 to 4 as a within-subjects factor and the two groups (BH+ and BH-) as a between-subjects factor. All dependent data met the assumption of homogeneity; however, the assumption of sphericity was not met, $\chi^2(5) = 28.32$, $p < .001$. Huynh-Feldt corrections showed a significant change in

Table 1 Objects Placed in an Array of Cells with either Occluded or Open Views during Critical or Baseline Trials

Critical trials with ambiguous objects			Baseline trials with unambiguous objects		
Trial order	Occluded cell	Open cell	Trial order	Occluded cell	Open cell
1	Pillar candle	Glass jar candle	5	Yoyo	Glass jar candle
2	Glue bottle	Glue stick	6	Toy furniture (a rocking chair)	Glue stick
3	Large measuring cup	Middle and small sized measuring cup	7	Notebook	Middle and small sized measuring cup
4	A paper cup	An egg cup	8	Toy furniture (a bed)	An egg cup
9	Safety pins	Hair pins	13	Candies	Hair pins
10	Hairbrush	Paintbrush	14	Toy furniture (a piano)	Paintbrush
11	Water glasses	A pair of glasses	15	Toy furniture (a table)	A pair of glasses
12	Wood stick	Lip balm (Chap Stick)	16	A Lego figurine	Lip balm (Chap Stick)

the numbers of the repeated sessions, $F(2.43, 155.32) = 20.94$, $p < .001$, $\eta^2 = .25$; however, no significant interaction effect between the numbers of the repeated sessions and the groups was observed, $F(2.43, 155.32) = 1.08$, $p = .35$, $\eta^2 = .02$. In addition, there was no significant difference between the two groups, $F(1, 64) = 3.86$, $p = .054$, $\eta^2 = .06$.

The mean durations (in seconds) for the BH+ group were 273.33 (SD = 29.70), 238.6 (SD = 42.73), 211.93 (SD = 40.47), and 208.7 (SD = 36.57) for Tasks 1 to 4; for BH- group, they were 275.97 (SD = 40.18), 227.6 (SD = 44.44), 215.64 (SD = 41.10), and 199.04 (SD = 38.72). A mixed ANOVA was performed on the participants' average durations (in seconds) from Tasks 1 to 4 as a within-subjects factor and the two groups (BH+ and BH-) as a between-subjects factor. All dependent data met the assumption of homogeneity; however, the assumption of sphericity was not met, $\chi^2(5) = 13.66$, $p = .018$. Greenhouse-Geisser corrections showed a significant change in the mean duration, $F(2.67, 170.64) = 53.23$, $p < .001$, $\eta^2 = .45$, but no significant interaction effect was observed between the duration and the two groups, $F(2.67, 170.64) = .83$, $p = .47$, $\eta^2 = .0$. In addition, there was no significant difference between the two groups, $F(1, 64) = .33$, $p = .57$, $\eta^2 = .005$.

For the analysis of the performance results from the cupboard visuospatial perspective-taking task, we conducted a mixed ANOVA on the number of correct trials with critical items or baseline items as a within-subjects factor and the three groups (BH+, BH-, and Control) as a between-subjects factor. SPSS (<https://www.ibm.com/analytics/spss-statistics-software>) was used for all analyses. The assumption for homogeneity of variance was met, $F(2, 95) = 1.42$, $p = .26$. As the within-subject factors had only two levels, corrections were not needed. There was a significant effect of trial type, $F(1, 95) = 883.24$, $p < .001$, $\eta^2 = .9$, indicating that participants performed better on baseline trials with unambiguous items than on the critical trials with ambiguous items (Fig. 2). There was no significant interaction between the performance on the trial types for the three groups, $F(2, 95) = 1.76$, $p = .18$, $\eta^2 = .04$. However, there was a significant main effect of group, $F(1, 95) = 3.62$, $p = .03$, $\eta^2 = .07$. Bonferroni-corrected pairwise comparisons for the main group effect indicated that there were significant differences ($p = .04$) between the BH+ and BH- groups, but not between the Control group and the BH+ ($p = 1$) or the BH- groups ($p = .12$) for the critical trials.

Implicit Relational Assessment Procedure Perspective-taking Experiment

An analysis of the IRAP block-sequence order effect between the D-IRAP scores of the group of 45 participants who were exposed to the I versus I & Other versus Other trial first (i.e., the participants who were required to answer as if you were you and Tokiko were Tokiko), $M = .60$, 95% CI [.49, .71], and the group of remaining 45 participants who

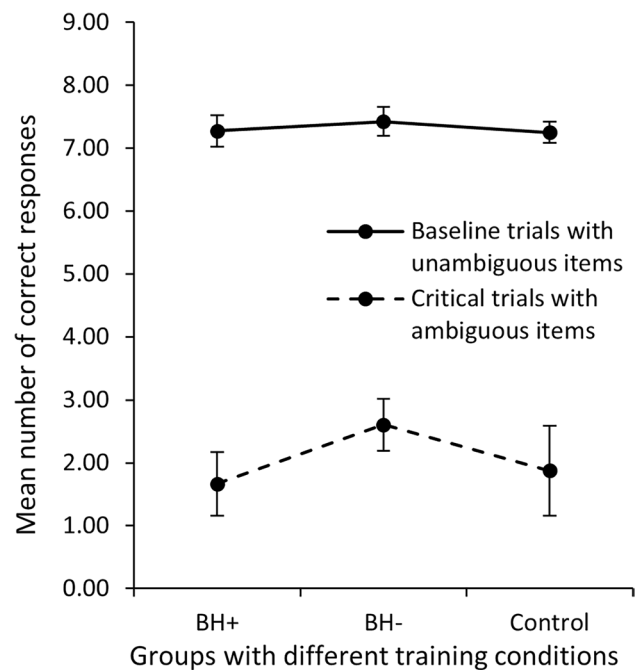


Fig. 2 Mean Number of Correct Responses across Three Groups in the Cupboard Task. Note. The maximum number of correct responses that a participant could obtain was eight for each trial type. There was a total of eight trials per block of testing with either ambiguous or unambiguous items. Error bars indicate the 95% confidence intervals

were exposed to inconsistent trials first (i.e., answering as if you were Tokiko and Tokiko were you), $M = .66$, 95% CI [.56, .75], showed no significant effect of order on D-IRAP scores, $t(88) = -.74$, $p = .45$. Levene's test indicated equal variances ($F = 1.07$, $p = .3$).

As indicated in Figure 3, the mean D-IRAP scores from the three groups showed a strong pro-self IRAP effect in all four trial types because the D-IRAP scores were all positive (i.e., participants gave faster and more accurate responses to the consistent trials with I versus I's action and Other versus Others' action, compared to the inconsistent trials with I vs. Other's action and Other vs. I's action). We conducted a mixed ANOVA to evaluate the influence of the different training conditions assigned to each group on D-IRAP scores in all four trial types. There was no significant difference in D-IRAP scores among the three groups, $F(1,2) = 1.98$, $p = .14$, $\eta^2 = .04$. The assumption of sphericity was met, based on Mauchly's Test, $X^2(5) = 4.78$, $p = .45$.

In terms of the RPM test, the scores were equivalent among the three groups. The mean RPM score for BH+ group was 51.70 (SD = 5.72), for BH- group it was 51.55 (SD = 5.23), and for the control group, 52.10 (SD = 6.79), $F(2, 95) = .07$, $p = .92$. The RPM scores and number of correct responses in the visual cupboard perspective-taking task were moderately and significantly related, $r(98) = .34$, $p < .001$. There was no significant relationship between responses on the cupboard task or the four trial types of

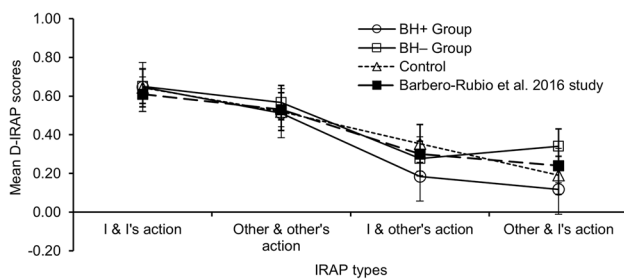


Fig. 3 D-IRAP Scores in Each Group Compared to Barbero-Rubio et al.'s (2016) Results. Note. Mean D-IRAP scores above zero indicate the participants' attitude inclined towards a pro-self-perspective (egocentric attitude), and scores below zero indicate a pro-other-perspective (faster in thinking about other's points of view rather than one's own). A score close to zero indicates that the individual is neutral and unbiased towards either of the two attitudes indicated by the pro-self and pro-other trials. The error bars indicate the 95% confidence interval

IRAP measures, $r(97) = -.007$, $p = .95$ (I vs. I's action), $r(97) = -.06$, $p = .59$ (I vs. Other's action), $r(97) = -.07$, $p = .51$ (Other vs. I's action), and $r(97) = -.09$, $p = .36$ (Other vs. Other's action). There were also no significant correlations between the RPM scores and each of the four trial types of IRAP measures; $r(97) = -.04$, $p = .68$ (I vs. I's action), $r(97) = -.04$, $p = .73$ (I vs. Other's action), $r(97) = -.01$, $p = .89$ (Other vs. I's action), and $r(97) = -.09$, $p = .4$ (Other vs. Other's action).

Discussion

To test the hypothesis of whether or not deictic expressions are effective in producing general perspective-taking improvements, we examined whether accuracy in responding to the BH protocol had any influence on participants' perspective-taking ability. Effectiveness of the deictic expressions would have been evidenced in a higher response accuracy relative to the control group on the visuo-spatial cupboard task or the IRAP. To support the idea that deictic relational framing serves as a core component of perspective taking, the BH+ group alone (or both BH+ and BH- groups, if the proposition that deictic words can be replaced with specific nouns is correct) should have performed better than the control group in the two experimental perspective-taking tasks. Overall, no significant difference in perspective-taking ability, compared to the control group, was observed in the results of the two experimental groups in the two tasks.

In terms of the cupboard visuospatial task, the main finding was that the BH+ group did not show any difference in response accuracy compared to the control group's performance on critical trials with the ambiguous items. We therefore failed to reject the null hypothesis. However, there was a significant difference between the BH+ and BH- groups' performance on

the critical-item trials, with participants in the BH- group performing better than those in the BH+ group. This indicates that it is worth conducting additional experiments to reexamine the criticality of the inclusion of deictic expressions in the protocol and further clarify which component (e.g., any other relational framing that was part of the BH protocol such as "if . . . then" causal relational framing) is responsible for improved performance on the other types of perspective-taking tasks when improvements are observed. As noted previously, however, there is limited evidence that the BH protocol leads to improved performance on other types of perspective-taking tasks.

We concluded that there was insufficient evidence to support the alternative hypothesis that neither the deictic expressions, specific nouns, nor any other relational framing that was part of the BH protocol (e.g., "if . . . then" causal relational framing) had an effect on performance in the perspective-taking tasks. It is interesting that our findings aligned with the outcome of previous studies demonstrating that verbally competent adults (i.e., university students) made substantial errors during the critical trials in the cupboard task testing the participants' accuracy in predicting what object could be seen from another's point of view (De Lillo & Ferguson, 2022; Epley & Caruso, 2009; Keysar et al., 2003; Samson et al., 2010). Despite the fact that the adult participants knew that the director could not see some of the objects on the cupboard, some of the participants selected the object that was not mutually visible. To do well in the cupboard visual perspective-taking task, the response should be controlled by a discriminative stimulus, which is often referenced as a perceptual common ground between an interlocutor and a listener (Clark & Marshall, 1981; Keysar, 1997), and is manipulated by the presence or absence of the stimuli that block one's view of an object or an event. The ability to discriminate the perceptual common ground between self and others does not appear to be facilitated by the BH protocol.

The main finding from the IRAP experiment was that there was no evidence to support the hypothesis that the 3-day period of training with the BH protocol had an effect on the IRAP performance results. Proficiency in perspective taking was measured by the participant's ability to demonstrate their indifference to the arranged dichotomy between the pro-self-perspective (i.e., responding faster in consistent trials, when answering as if "I am me, others are others") and pro-other-perspective (i.e., responding faster in inconsistent trials when answering as if "I am another, another is me"). If the IRAP score is at or close to zero, it is indicative of indifference to the two biases. However, the D-IRAP scores of our participants were consistent with those obtained by Barbero-Rubio et al. (2016), demonstrating a strong "pro-self" bias. Barbero-Rubio et al. suggested that this "pro-self" bias is the result of people being more often exposed to situations that require self-centered perspective in daily life, and also due to the immediacy of private events that are only available to individuals.

Deictic framing may lack the universality required to engage in successful perspective taking in different contexts, but it may generalize to situations with a similar format to the original BH protocol. Indeed, experiments have shown successful generalization of deictic relational responding when the test items were similar to the training items (e.g., participants were provided with tasks involving if–then framing, in the form of verbal questions in both training and testing; Davlin et al., 2011; Heagle & Rehfeldt, 2006; Gilroy et al., 2015; Montoya-Rodríguez et al., 2017). Skills gained from the BH protocol, however, do not appear to generalize to other perspective-taking tasks. Deictic framing may be one part of the broad and diverse set of skills that contribute to perspective taking; however, given the concerning theoretical issues that we and Guinther (2017) have identified, we should proceed cautiously with the assumption that deictic framing is the core operant involved in perspective taking. Rather, there are multiple ways that people can generalize or accurately respond when engaging in perspective-taking tasks. Identifying factors that increase accuracy in predicting others' behavior is likely a better approach to understanding this type of behavior than searching for a single “core” component of perspective-taking.

One of the most important requirements of the BH protocol is the requirement to produce a response that corresponds with either an initial statement or the reverse of that statement, dependent upon the wording of a conditional statement. For example, following the initial statement, “I am in a green chair, and you are in a red chair,” and the conditional statement, “If I were you and you were me, where would I be?” a response that corresponds with the reverse of the original statement is reinforced (i.e., “in a red chair”). With the same initial statement but the absence of a conditional statement (only “where am I?”), participants can simply respond to what is expressed on the initial statement (see Figure 4). Likewise, with questions involving a double conditional statement such as “if you were me and I were you, and if here was there and there was here” a response that corresponds with the first part of the conditional statement is sufficient to provide a correct response. The only conditional statement that requires a reversed response is the single-reversal statement. In fact, the demonstration of faster acquisition of tasks involving double conditional statements compared to single conditional ones has been observed and discussed in the literature (Heagle & Rehfeldt, 2006; Rehfeldt et al., 2007; Weil et al., 2011). This might be due to a training effect, because the most complex version of the question (double reversed) is usually provided last. However, it may also be indicative of the participants' reliance upon simple discriminative stimuli, such as the length of the conditional statement (which is the longest in the case of double-reversals) to complete the tasks accurately (Taylor & Edwards, 2021). If participants do come to rely on such strategies to solve the BH protocol tasks then, at face value, this type of behavior appears to have very little to do with what can meaningfully be described as perspective taking.

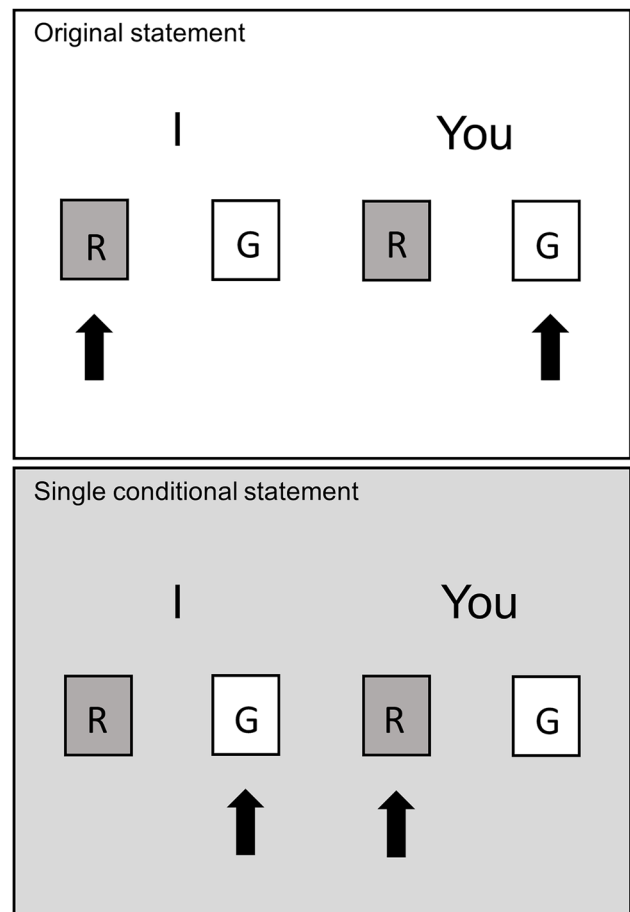


Fig. 4 Arrays of Both Sample and Comparison Stimuli in the BH Protocol in a Conditional Discrimination Task. Note. The top figure represents a simple relation task such as “I have a red brick and you have a green brick. Which brick do I have? Which brick do you have?” The correct responses are indicated by the black arrows. “R” indicates the color red, and “G” indicates green. The bottom figure represents a reversed relation task such as “I have a red brick and you have a green brick. If I were you and you were me, which brick would I have? Which brick would you have?” The background color can be arranged to serve as a conditional stimulus, as an alternative to the absence or presence of the phrase “If I were you and you were me.” The grey background color indicates the single conditional statement would signal reinforcement for a response that corresponds with the reverse of the original statement. A response to the original condition would be reinforced in the presence of the white background

There are some limitations to our study. We did not conduct a preliminary investigation to determine an appropriate duration of training for adult participants that would have allowed them to gain the optimum benefit from their exposure to the BH protocol. Each participant was exposed to the protocol for less than 15 min per day for 3 days (the daily requirement was to complete at least one task containing 15 question items). Even though the mastery criterion was set at 90% correct responses in less than 5 min and participants had to repeat until they passed the criterion, this type of arbitrarily constructed criterion may

not have resulted in participants achieving maximum proficiency with the protocol. Another limitation was that the instructions provided during the cupboard visual perspective-taking task were originally designed to measure eye movement at the onset of the verbalization of the target object name (for example, the “g” sound in the target object name “glasses”), rather than solely overt behavioral measures (i.e., the number of correct responses per participant, of those whose hand reached out to the target object in the cupboard). Because we did not use an eye-tracking device, we did not collect duration records to compare how long it took for participants to identify the target object, as was done in previous studies (Keysar et al., 2000; Keysar et al., 2003). In addition, the specific instructions provided may have influenced the performance outcome. For instance, in the cupboard task we could have used instructions similar to the BH protocol, such as, “consider what you would see if you were me,” or directions to move the target object could have been omitted from the instructions, simply asking, “can you give me the glasses please?” Future studies may modify the instructions (e.g., by enhancing the saliency of relevant stimuli) to better measure the ability of participants to discriminate the perceptual common ground between self and others. Moreover, our participants’ general use of various object names was not subject to preliminary examination. A preliminary session of one-to-one matching of the names and the objects used in the experiment may be necessary, especially when the participants are university students who have different cultural backgrounds. For example, one of the pairs of ambiguous nouns, hair pin and safety pin, was confusing for some of our participants as they were more familiar with calling a hair pin a bobby pin, so they took longer to find the intended object. Regarding IRAP, it was originally designed to gauge people’s attitude or preference through a latency measure that captures their strong preestablished learning (Watt et al., 1991). We did not conduct preliminary investigations (e.g., word usage frequency measure) to control such historical valences preestablished with the experimental stimuli used in the current study.

Conclusion

In light of the current findings, we question not only the validity of the theory of deictic framing and the utility of the BH protocol, but also suggest that previous approaches for investigating perspective taking may need to be reexamined. We showed that participants’ fluent performance with the BH protocol task was not associated with enhanced accuracy on the other types of perspective-taking tasks: the cupboard visual perspective taking task and a perspective-taking IRAP task. It appears that changes in behavior brought about by BH protocol training can only be observed under conditions that are extremely similar to the BH protocol itself, which casts doubt upon the proposition that the protocol is relevant to behavioral processes that are

fundamental to perspective taking more broadly. Going beyond these potential issues with the BH protocol, there also appear to be construct validity issues associated with the tasks designed to measure perspective-taking skills, due in part to perspective taking being a poorly defined social construct. Researchers need to develop a more systematic approach to identifying the specific stimulus functions of relevant stimuli in instances of successful perspective taking, and work to understand how these functions can be established. Such an approach, in our view, would be more effective than passive “train and hope” approaches (Stokes & Baer, 1977, p. 351). Stokes and Baer (1977) warn that the absence of a program of generalization for various behavioral interventions is a problem. This concern remains significant in relation to our approaches to examining the generalization of acquired perspective-taking abilities from various training and testing schemes.

Appendix A

Deictic Framing Protocol (adapted from McHugh et al., 2004a study) 60 Questions total

Simple Relations

Simple I-You

I have a red brick and you have a green brick.

Which brick do I have? [Answer: red brick]

Which brick do YOU have? [green brick]

I have a green brick and you have a red brick.

Which brick do YOU have? [red brick]

Which brick do I have? [green brick]

Simple HERE-THERE

I am sitting here on the blue chair and you are sitting there on the black chair.

Where am I sitting? [blue chair]

Where are YOU sitting? [black chair]

I am sitting here on the black chair and you are sitting there on the blue chair.

Where are YOU sitting? [blue chair]

Where am I sitting? [black chair]

Simple NOW-THEN

Yesterday I was watching television, today I am reading.

What am I doing now? [reading]

What was I doing then? [watching television]

Yesterday I was reading, today I am watching television.

What was I doing then? [reading]

What am I doing now? [watching television]

Yesterday you were reading, today you are watching television.

What are YOU doing now? [watching television]

What were YOU doing then? [reading]

Yesterday you were watching television, today you are reading.

What were YOU doing then? [watching television]

What are YOU doing now? [reading]

REVERSED RELATIONS

Reversed I-YOU
 I have a red brick and you have a green brick. If I was you and you were me,
 Which brick would I have? [green brick]
 Which brick would YOU have? [red brick]
 I have a green brick and you have a red brick. If I was you and you were me,
 Which brick would YOU have? [green brick]
 Which brick would I have? [red brick]
 I have a red brick and you have a green brick. If I was you and you were me,
 Which brick would YOU have? [red brick]
 Which brick would I have? [green brick]
 I have a green brick and you have a red brick. If I was you and you were me,
 Which brick would I have? [red brick]
 Which brick would YOU have? [green brick]
 I am sitting here on the black chair and you are sitting there on the blue chair. If I was you and you were me,
 Where would YOU be sitting? [black chair]
 Where would I be sitting? [blue chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If I was you and you were me,
 Where would I be sitting? [blue chair]
 Where would YOU be sitting? [black chair]
 I am sitting here on the blue chair and you are sitting there on the black chair. If I was you and you were me,
 Where would I be sitting? [black chair]
 Where would YOU be sitting? [blue chair]
 I am sitting here on the blue chair and you are sitting there on the black chair. If I was you and you were me,
 Where would YOU be sitting? [blue chair]
 Where would I be sitting? [black chair]
 Reversed HERE-THERE
 I am sitting here on the blue chair and you are sitting there on the black chair. If here was there and there was here,
 Where would YOU be sitting? [blue chair]
 Where would I be sitting? [black chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If here was there and there was here,
 Where would I be sitting? [blue chair]
 Where would YOU be sitting? [black chair]
 I am sitting here on the blue chair and you are sitting there on the black chair. If here was there and there was here,
 Where would I be sitting? [black chair]
 Where would YOU be sitting? [blue chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If here was there and there was here,
 Where would YOU be sitting? [black chair]
 Where would I be sitting? [blue chair]
 Yesterday, I was sitting there on the blue chair, today I am sitting here on the black chair. If here was there and there was here,
 Where would I be sitting now? [blue chair]

Where was I sitting then? [black chair]
 Yesterday, I was sitting there on the black chair, today I am sitting here on the blue chair. If here was there and there was here,
 Where was I sitting then? [blue chair]
 Where would I be sitting now? [black chair]
 Yesterday, I was sitting there on the blue chair, today I am sitting here on the black chair. If here was there and there was here,
 Where was I sitting then? [black chair]
 Where would I be sitting now? [blue chair]
 Yesterday, I was sitting there on the black chair, today I am sitting here on the blue chair. If here was there and there was here,
 Where would I be sitting now? [black chair]
 Where was I sitting then? [blue chair]
 Yesterday, you were sitting there on the blue chair, today You are sitting here on the black chair. If here was there and there was here,
 Where would you be sitting now? [blue chair]
 Where were you sitting then? [black chair]
 Yesterday, you were sitting there on the blue chair, today You are sitting here on the black chair. If here was there and there was here,
 Where were you sitting then? [black chair]
 Where would you be sitting now? [blue chair]
 Yesterday, you were sitting there on the black chair, today You are sitting here on the blue chair. If here was there and there was here,
 Where would you be sitting now? [black chair]
 Where were you sitting then? [blue chair]
 Yesterday, you were sitting there on the black chair, today You are sitting here on the blue chair. If here was there and there was here,
 Where would you be sitting now? [black chair]
 Where were you sitting then? [blue chair]
 Yesterday, you were sitting there on the black chair, today You are sitting here on the blue chair. If here was there and there was here,
 Where were you sitting then? [bleu chair]
 Where would you be sitting now? [black chair]
 Reversed NOW-THEN:
 Yesterday I was watching television, today I am reading.
 If now was then and then was now,
 What was I doing then? [reading]
 What would I be doing now? [watching television]
 Yesterday I was watching television, today I am reading.
 If now was then and then was now,
 What would I be doing now? [watching television]
 What was I doing then? [reading]
 Yesterday I was reading, today I am watching television.
 If now was then and then was now,
 What would I be doing now? [reading]
 What was I doing then? [watching television]
 Yesterday I was watching television, today I am reading.
 If now was then and then was now,
 What would I be doing now? [watching television]
 What was I doing then? [reading]
 Yesterday I was reading, today I am watching television.
 If now was then and then was now,
 What was I doing then? [watching television]

What would I be doing now? [reading]
 Yesterday you were watching television, today you are reading. If now was then and then was now,
 What were you doing then? [reading]
 What would you be doing now? [watching television]
 Yesterday you were reading, today you are watching television. If now was then and then was now,
 What were you be doing then? [watching television]
 What would you be doing now? [reading]
 Yesterday you were watching television, today you are reading. If now was then and then was now,
 What would you be doing now? [watching television]
 What were you doing then? [reading]
 Yesterday you were reading, today you are watching television. If now was then and then was now,
 What would you be doing now? [reading]
 What were you doing then? [watching television]
 Yesterday I was sitting there on the blue chair, today I am sitting here on the black chair. If now was then and then was now.
 Where would I be sitting now? [blue chair]
 Where was I sitting then? [black chair]
 Yesterday I was sitting there on the blue chair, today I am sitting here on the black chair. If now was then and then was now.
 Where was I sitting then? [black chair]
 Where would I be sitting now? [blue chair]
 Yesterday I was sitting there on the black chair, today I am sitting here on the blue chair. If now was then and then was now.
 Where would I be sitting now? [black chair]
 Where was I sitting then? [blue chair]
 Yesterday I was sitting there on the black chair, today I am sitting here on the blue chair. If now was then and then was now.
 Where was I sitting then? [blue chair]
 Where would I be sitting now? [black chair]
 Yesterday you were sitting there on the blue chair, today you are sitting here on the black chair. If now was then and then was now.
 Where were you sitting then? [black chair]
 Where would you be sitting now? [blue chair]
 Yesterday you were sitting there on the blue chair, today you are sitting here on the black chair. If now was then and then was now.
 Where would you be sitting now? [blue chair]
 Where were you sitting then? [black chair]
 Yesterday you were sitting there on the black chair, today you are sitting here on the blue chair. If now was then and then was now.
 Where were you sitting then? [blue chair]
 Where would you be sitting now? [black chair]
 Yesterday you were sitting there on the black chair, today you are sitting here on the blue chair. If now was then and then was now.
 Where would you be sitting now? [black chair]
 Where were you sitting then? [blue chair]

DOUBLE REVERSED RELATIONS

I-YOU/HERE-THERE:

I am sitting here on the blue chair and you are sitting there on the black chair. If I was you and you were me and if here was there and there was here,
 Where would I be sitting? [blue chair]
 Where would YOU be sitting? [black chair]
 I am sitting here on the blue chair and you are sitting there on the black chair. If I was you and you were me and if here was there and there was here,
 Where would YOU be sitting? [black chair]
 Where would I be sitting? [blue chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If I was you and you were me and if here was there and there was here,
 Where would I be sitting? [black chair]
 Where would YOU be sitting? [blue chair]
 I am sitting here on the blue chair and you are sitting there on the black chair. If I was you and you were me and if here was there and there was here,
 Where would YOU be sitting? [black chair]
 Where would I be sitting? [blue chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If I was you and you were me and if here was there and there was here,
 Where would YOU be sitting? [blue chair]
 Where would I be sitting? [black chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If I was you and you were me and if here was there and there was here,
 Where would YOU be sitting? [black chair]
 Where would I be sitting? [blue chair]
 I am sitting here on the black chair and you are sitting there on the blue chair. If I was you and you were me and if here was there and there was here,
 Where would I be sitting? [black chair]
 Where would YOU be sitting? [blue chair]
 HERE-THERE / NOW-THEN
 Yesterday I was sitting there on the blue chair, today I am sitting here on the black chair. If here was there and there was here and if now was then and then was now,
 Where would I be sitting then? [blue chair]
 Where would I be sitting now? [black chair]
 Yesterday I was sitting there on the blue chair, today I am sitting here on the black chair. If here was there and there was here and if now was then and then was now,
 Where would I be sitting now? [black chair]
 Where would I be sitting then? [blue chair]
 Yesterday I was sitting there on the black chair, today I am sitting here on the blue chair. If here was there and there was here and if now was then and then was now,
 Where would I be sitting then? [black chair]
 Where would I be sitting now? [blue chair]
 Yesterday I was sitting there on the black chair, today I am sitting here on the blue chair. If here was there and there was here and if now was then and then was now,
 Where would I be sitting now? [blue chair]
 Where would I be sitting then? [black chair]

Yesterday you were sitting there on the blue chair, today you are sitting here on the black chair. If here was there and there was here and if now was then and then was now,

Where would YOU be sitting then? [blue chair]

Where would YOU be sitting now? [black chair]

Yesterday you were sitting there on the blue chair, today you are sitting here on the black chair. If here was there and there was here and if now was then and then was now,

Where would YOU be sitting now? [black chair]

Where would YOU be sitting then? [blue chair]

Yesterday you were sitting there on the black chair, today you are sitting here on the blue chair. If here was there and there was here and if now was then and then was now,

Where would YOU be sitting then? [black chair]

Where would YOU be sitting now? [blue chair]

Yesterday you were sitting there on the black chair, today you are sitting here on the blue chair. If here was there and there was here and if now was then and then was now,

Where would YOU be sitting now? [blue chair]

Where would YOU be sitting then? [black chair]

Appendix B

Non-Deictic Component Protocol 60 questions Total

SIMPLE RELATIONS

Simple X - Y

X has a red brick and Y has a green brick.

Which brick does X have? [Answer: red brick]

Which brick does Y have? [green brick]

X has a green brick and Y has a red brick.

Which brick does Y have? [red brick]

Which brick does X have? [green brick]

Simple location

X is sitting on the blue chair and Y is sitting on the black chair.

Where is X sitting? [blue chair]

Where is Y sitting? [black chair]

X is sitting on the black chair and Y is sitting on the blue chair.

Where is Y sitting? [blue chair]

Where is X sitting? [black chair]

Simple Monday and Wednesday

Monday X is watching television, Wednesday X is reading.

What is X doing on Wednesday? [reading]

What is X doing on Monday? [watching television]

Monday X is reading, Wednesday X is watching television.

What is X doing on Monday? [reading]

What is X doing on Wednesday? [watching television]

Monday Y is reading, Wednesday Y is watching television.

What is Y doing on Wednesday? [watching television]

What is Y doing on Monday? [reading]

Monday Y is watching television, Wednesday Y is reading.

What is Y doing on Monday? [watching television]

What is Y doing on Wednesday? [reading]

REVERSED RELATIONS

Reversed X - Y

X has a red brick and Y has a green brick. If X was Y and Y was X,

Which brick would X have? [green brick]

Which brick would Y have? [red brick]

X has a green brick and Y has a red brick. If X was Y and Y was X,

Which brick would Y have? [green brick]

Which brick would X have? [red brick]

X has a red brick and Y has a green brick. If X was Y and Y was X,

Which brick would Y have? [red brick]

Which brick would X have? [green brick]

X has a green brick and Y has a red brick. If X was Y and Y was X,

Which brick would X have? [red brick]

Which brick would Y have? [green brick]

X is sitting on the black chair and Y is sitting there on the blue chair. If X was Y and Y was X,

Where would Y be sitting? [black chair]

Where would X be sitting? [blue chair]

X is sitting on the black chair and Y is sitting on the blue chair. If X was Y and Y was X,

Where would X be sitting? [blue chair]

Where would Y be sitting? [black chair]

X am sitting on blue chair and Y is sitting on the black chair. If X was Y and Y was X,

Where would X be sitting? [black chair]

Where would Y be sitting? [blue chair]

X is sitting on the blue chair and Y is sitting on the black chair. If X was Y and Y was X,

Where would Y be sitting? [blue chair]

Where would X be sitting? [black chair]

Reversed location

X is sitting on the blue chair and Y is sitting on the black chair. If the blue chair was the black chair and the black chair was the blue chair,

Where would Y be sitting? [blue chair]

Where would X be sitting? [black chair]

X is sitting on the black chair and Y is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair,

Where would X be sitting? [blue chair]

Where would Y be sitting? [black chair]

X is sitting on the blue chair and Y is sitting on the black chair. If the blue chair was the black chair and the black chair was the blue chair,

Where would X be sitting? [black chair]

Where would Y be sitting? [blue chair]

X is sitting on the black chair and Y is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair,

Where is Y sitting on Monday? [black chair]
 Where is Y sitting on Wednesday? [blue chair]
 Monday Y is sitting on the blue chair, Wednesday Y is sitting on the black chair. If Wednesday was Monday and Monday was Wednesday,
 Where would Y be sitting on Wednesday? [blue chair]
 Where was Y sitting on Monday? [black chair]
 Monday Y was sitting on the black chair, Wednesday Y is sitting on the blue chair. If Wednesday was Monday and Monday was Wednesday,
 Where is Y sitting on Monday? [blue chair]
 Where is Y sitting on Wednesday? [black chair]
 Monday Y is sitting on the black chair, Wednesday Y is sitting on the blue chair. If Wednesday was Monday and Monday was Wednesday,
 Where is Y sitting on Wednesday? [black chair]
 Where is Y sitting on Monday? [blue chair]
DOUBLE REVERSED RELATIONS
 X – Y / location
 X is sitting on the blue chair and Y is sitting on the black chair. If X was Y and Y was X and if the blue chair was the black chair and the black chair was the blue chair,
 Where would X be sitting? [blue chair]
 Where would Y be sitting? [black chair]
 X is sitting on the blue chair and Y is sitting on the black chair. If X was Y and Y was X and if the blue chair was the black chair and the black chair was the blue chair,
 Where would Y be sitting? [black chair]
 Where would X be sitting? [blue chair]
 X is sitting on the black chair and Y is sitting on the blue chair. If X was Y and Y was X and if the blue chair was the black chair and the black chair was the blue chair,
 Where would X be sitting? [black chair]
 Where would Y be sitting? [blue chair]
 X is sitting on the blue chair and Y is sitting on the black chair. If X was Y and Y was X and if the blue chair was the black chair and the black chair was the blue chair,
 Where would Y be sitting? [black chair]
 Where would X be sitting? [blue chair]
 X is sitting on the black chair and Y is sitting on the blue chair. If X was Y and Y was X and if the blue chair was the black chair and the black chair was the blue chair,
 Where would Y be sitting? [blue chair]
 Where would X be sitting? [black chair]
 X is sitting on the black chair and Y is sitting on the blue chair. If X was Y and Y was X and if the blue chair was the black chair and the black chair was the blue chair,
 Where would Y be sitting? [blue chair]
 Where would X be sitting? [black chair]
 Location / Monday - Wednesday
 Monday X is sitting on the blue chair, Wednesday X is sitting on the black chair. If the blue chair was the black chair

and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would X be sitting on Monday? [blue chair]
 Where would X be sitting on Wednesday? [black chair]
 Monday X is sitting on the blue chair, Wednesday X is sitting on the black chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would X be sitting on Wednesday? [black chair]
 Where would X be sitting on Monday? [blue chair]
 Monday X is sitting on the black chair, Wednesday X is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would X be sitting on Monday? [black chair]
 Where would X be sitting on Wednesday? [blue chair]
 Monday X is sitting on the black chair, Wednesday X is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would X be sitting on Monday? [black chair]
 Where would X be sitting on Wednesday? [blue chair]
 Monday X is sitting on the black chair, Wednesday X is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would X be sitting on Monday? [black chair]
 Where would X be sitting on Wednesday? [blue chair]
 Monday Y is sitting on the blue chair, Wednesday Y is sitting on the black chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would Y be sitting on Monday? [blue chair]
 Where would Y be sitting on Wednesday? [black chair]
 Monday Y is sitting on the blue chair, Wednesday Y is sitting on the black chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would Y be sitting on Monday? [blue chair]
 Where would Y be sitting on Wednesday? [black chair]
 Monday Y is sitting on the black chair, Wednesday Y is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would Y be sitting on Monday? [black chair]
 Where would Y be sitting on Wednesday? [blue chair]
 Monday Y is sitting on the black chair, Wednesday Y is sitting on the blue chair. If the blue chair was the black chair and the black chair was the blue chair and if Wednesday was Monday and Monday was Wednesday,
 Where would Y be sitting on Monday? [black chair]
 Where would Y be sitting on Wednesday? [blue chair]

Appendix C

Deictic Reading Comprehension Tasks 60 Items

I have a red brick and you have a green brick. [Answer: I, you]
 I have a green brick and you have a red brick. [I, you]

I am sitting here on the blue chair and you are sitting there on the black chair. [I, here, you, there]

I am sitting here on the black chair and you are sitting there on the blue chair. [I, here, you, there]

Yesterday I was watching television, today I am reading. [Yesterday, I, today, I]

Yesterday I was reading, today I am watching television. [Yesterday, I, today, I]

Yesterday you were reading, today you are watching television. [Yesterday, you, today, you]

Yesterday you were watching television, today you are reading. [Yesterday, you, today, you]

I have a red brick and you have a green brick. [I, you]

I have a green brick and you have a red brick. [I, you]

I am sitting here on the black chair and you are sitting there on the blue chair. [I, here, you, there]

I am sitting here on the blue chair and you are sitting there on the black chair. [I, here, you, there]

Yesterday, I was sitting there on the blue chair, today I am sitting here on the black chair. [Yesterday, I, there, today, I, here]

Yesterday, I was sitting there on the black chair, today I am sitting here on the blue chair. [Yesterday, I, there, today, I, here]

Yesterday I was watching television, today I am reading. [Yesterday, I, today, I]

Yesterday I was reading, today I am watching television. [Yesterday, I, today, I]

Yesterday you were watching television, today you are reading. [Yesterday, you, today, you]

Yesterday you were reading, today you are watching television. [Yesterday, you, today, you]

I have an orange pencil and you have a yellow pencil. [I, you]

I have a yellow pencil and you have an orange pencil. [I, you]

I am sitting here on the pink couch and you are sitting there on the purple couch. [I, here, you, there]

I am sitting here on the purple couch and you are sitting there on the pink couch. [I, here, you, there]

Yesterday, I was playing video games, today I am listening to music. [Yesterday, I, today I]

Yesterday I was listening to music, today I am playing video games. [Yesterday, I, today, I]

Yesterday you were listening to music, today you are playing video games. [Yesterday, you, today, you]

Yesterday you were playing video games, today you are listening to music. [Yesterday, you, today, you]

Yesterday I was sitting there on the pink couch, today I am sitting here on the purple couch. [Yesterday, I, there, today, I, here]

Yesterday I was sitting there on the pink couch, and today I am sitting here on the pink couch. [Yesterday, I, there, today, I, here]

Yesterday you were sitting there on the pink couch, today you are sitting here on the purple couch. [Yesterday, you, there, today, you, here]

Yesterday you were sitting there on the pink couch, and today you are sitting here on the pink couch. [Yesterday, you, there, today, you, here]

I have the hamburger and you have the grilled cheese. [I, you]

You have the hamburger and I have the grilled cheese. [You, I]

I am standing in the classroom and you are standing on the playground. [I, you]

You are standing in the classroom, and I am standing on the playground. [You, I]

Yesterday I was playing X-Box, today I am watching "The Incredible." [Yesterday, I, today I]

Yesterday you were watching "The Incredible," today I am playing X-box. [Yesterday, you, today, I]

Yesterday you were reading comic books, today you are talking on the phone. [Yesterday, you, today, you]

Today I am talking on the phone, yesterday I was reading comic books. [Today, I, yesterday, I]

I am holding the puppy and you are holding the kitten. [I, you]

You are holding the puppy and I am holding the kitten. [You, I]

Yesterday I was swimming there in the pool, today I am swimming here in the lake. [Yesterday, I, there, today, I, here]

Today you are swimming here in the lake, yesterday you were swimming there in the pool. [Today, you, here, yesterday, you, there]

Yesterday I was doing my homework; today I am taking a nap. [Yesterday, I, today, I]

Yesterday you were playing soccer, today you are playing basketball. [Yesterday, you, today, you]

Today I am playing soccer, yesterday I was playing basketball. [Today, I, yesterday, I]

I am sleeping here in the bedroom and you are sleeping there in the living room. [I, here, you, there]

You are sleeping here in the living room, and I am sleeping there in the bedroom. [You, here, I, there]

I am here eating McDonalds and you are eating there at Wendy's. You are eating here at Wendy's and I am eating there at McDonalds. [I, here, you, there, You, here, I there]

Yesterday I was shopping there at the mall; today I am shopping there at the grocery store. [Yesterday, I, there, today, I, there]

Yesterday you were running there in the park; today you are running here in gym class. [Yesterday, you, there, today, you here]

Today I am running here in the park; yesterday I was running there in gym class. [Today, I, here, yesterday, I, there]

I have a book and you have a comic book. [I, you]

You have a book and I have a comic book. [You, I]

I am driving Honda and you are driving Toyota. [I, you]

You are driving Honda, and I am driving Toyota. [You, I]

Yesterday I was cooking rice, today I am cooking pasta. [Yesterday, I, today I]

Yesterday you were cooking pasta, today I am cooking rice. [Yesterday, you, today, I]

Yesterday you were skiing, today you are skating. [Yesterday, you, today, you]
 Today I am skating, yesterday I was skiing. [Today, I, yesterday, I]

I own a hotel here in Hamilton and you own a house there in Queenstown. [I, here, you, there]

Appendix D

Table 2 A summary of online training data showing correct responses and duration required for each participant assigned to BH+ group, who were trained using the existing BH protocol (see Appendix A) until meeting the criteria to pass each task

Participants	Task 1		Task 2		Task 3		Task 4	
	% ^a	Duration ^b	%	Duration	%	Duration	%	Duration
1	81.5 100	4:20 1:23	93.75	3:28	62.5 100	2:30 1:27	100	3:58
2	56.25 93.75	5:00 3:31	68.75 100	5:00 3:38	37.5 100	5:00 3:07	75 93.75	5:00 2:34
3	56.25 93.75	5:00 5:00	87.5 100	4:04 3:36	100	4:04	100	3:24
4	100	5:00	87.5 50.0 87.5 93.75	4:50 5:00 5:00 4:39	100	3:34	100	3:27
5	43.75 75 100	5:00 5:00 3:36	75 100	5:00 2:50	37.5 100	2:43 3:18	62.5 100	3:30 2:49
6	87.5 93.75	4:43 3:40	87.5 100	2:27 2:59	100	3:36	100	2:46
7	25 43.75 100	5:00 5:00 4:54	56.25 100	5:00 3:22	6.25 ^c 100	- 3:45	6.25 ^c 100	- 3:15
8	100	4:59	93.75	3:48	93.75	4:00	100	3:26
9	56.25 93.75	5:00 4:22	74 100	5:00 4:53	25 87.5 100	5:00 4:09 1:52	62.5 100	2:59 2:41
10	81.25 100	5:00 3:53	93.75	3:36	100	3:48	100	3:49
11	81.25 62.5 93.75	5:00 5:00 3:23	93.75	3:22	100	3:02	50 100	5:00 3:19
12	68.75 75 87.5 93.75	5:00 4:53 4:09 3:19	93.75	3:10	56.25 87.5 93.75	5:00 2:59 3:02	100	2:44
13	81.25 62.5 93.75	5:00 3:33 1:50	75 100	5:00 1:57	50 100	4:44 1:49	75 93.75	4:48 2:15
14	68.75 93.75	5:00 3:48	93.75	2:38	87.5 93.75	2:52 2:46	100	2:41
15	50 81.25 100	5:00 4:52 4:32	75 81.25 93.75	4:41 4:11 3:08	81.25 87.5 100	4:18 3:42 3:00	93.75	3:23
16	68.75 81.25 81.25 100	5:00 5:00 5:00 2:18	62.5 93.75	2:30 2:57	75 100	3:41 2:04	37.5 100	5:00 4:02

Table 2 (continued)

Participants	Task 1		Task 2		Task 3		Task 4	
	% ^a	Duration ^b	%	Duration	%	Duration	%	Duration
17	56.25	5:00	81.25	5:00	100	4:43	100	4:52
	100	4:48	100	4:35				
18	43.75	5:00	75	5:00	56.25	2:04	100	3:57
	62.5	5:00	100	1:28	6.25 ^c	-		
	100	2:53			100	1:29		
19	81.25	5:00	56.25	2:44	62.5	4:41	56.25	1:58
	100	2:13	100	2:02	100	2:06	100	3:20
20	56.25	5:00	100	4:29	100	3:55	93.75	3:42
	62.5	4:03						
	56.25	5:00						
	75	5:00						
	93.75	4:30						
21	68.75	5:00	81.25	4:30	87.5	4:16	93.75	3:20
	50	5:00	93.75	4:18	100	3:57		
	93.75	3:58						
22	100	5:00	81.25	4:15	100	4:11	100	4:12
			93.75	3:33				
23	100	4:06	81.25	4:59	100	2:45	100	2:41
			93.75	3:45				
24	56.25	5:00	93.75	4:16	93.75	4:44	100	3:41
	93.75	5:00						
25	62.5	5:00	93.75	4:44	100	4:15	87.25	4:52
	100	4:42					100	4:21
26	93.75	5:00	100	4:19	100	3:30	100	3:04
27	75	4:58	100	4:50	100	3:22	100	2:58
	100	4:15						
28	81.25	4:55	93.75	4:02	100	3:30	100	3:18
	93.75	3:58						
29	81.25	5:00	75	4:58	100	4:23	100	4:20
	100	4:35	93.75	4:36				
30	93.75	5:00	100	4:35	100	3:32	100	3:28
31	62.5	5:00	81.25	4:35	100	3:56	100	3:33
	81.25	5:00	100	4:20				
	100	3:57						
32	93.75	5:00	93.75	4:36	100	3:15	100	2:46
33	100	4:56	100	3:20	100	2:45	100	2:46
Average score across all individuals	83.89	4:33	90.44	3:59	90.34	3:32	93.37	3:29

The participants were asked to complete a total of four tasks of 15 questions, during a 3-day period. They had to repeat each training task if they did not meet the criteria of more than 90% correct within 5 min for the 15 questions. The repeated numbers of each task were presented as a stack of percentages of correct responses for each participant.

^a% of correct responses (6.25% per each question)

^bDuration [mm:ss] (minutes:seconds)

^cTask incomplete

Table 3 A summary of online training data showing correct responses and duration required for each participant assigned to BH- group, who were trained using a modified version of the BH protocol with specific nouns instead of deictic expressions (see Appendix B) until meeting the criteria to pass each task

Participants	Task 1		Task 2		Task 3		Task 4	
	% ^a	Duration ^b	%	Duration	%	Duration	%	Duration
1	100	4:59	100	5:00	100	3:40	100	3:21
2	93.75	5:00	100	3:03	100	2:22	100	2:19
3	93.75	4:02	93.75	3:12	93.75	3:22	100	3:45
4	100	3:59	93.75	4:29	100	3:38	100	3:24
5	62.5 100	5:00 5:00	100	4:50	100	4:40	100	4:06
6	87.5 100	4:29 3:59	93.75	3:20	100	3:24	100	2:39
7	100	5:00	100	3:36	100	3:30	100	3:15
8	31.25 75 100	5:00 5:00 5:00	81.25 87.5 100	5:00 4:44 3:14	93.75	5:00	87.5 100	5:00 2:41
9	81.25 93.75	5:00 4:40	75 93.75	5:00 5:00	100	4:08	100	3:26
10	93.75	4:44	68.75 37.5 93.75 100	5:00 5:00 4:34 4:23	93.75	4:47	68.75 75 25c 100	5:00 4:37 - 5:00
11	87.5 100	5:00 4:57	93.75	3:04	100	3:20	100	3:48
12	68.5 75 100	5:00 5:00 3:40	25 100	3:40 4:01	18.75 ^c 37.5 ^c 100	- - 4:19	100	3:15
13	100	3:40	100	3:11	100	2:52	93.75	2:50
14	68.75 100	4:59 5:00	100	4:03	100	3:44	100	3:43
15	68.75 87.5 93.75	5:00 5:00 3:16	100	3:49	100	3:57	93.75	3:14
16	75 100	5:00 3:50	100	4:05	93.75	3:37	100	3:09
17	100	4:46	100	3:49	100	3:51	100	2:48
18	81.25 100	4:59 3:16	100	3:36	100	2:52	100	3:01
19	56.25 100	4:56 1:52	50 100	3:35 1:30	68.75 87.5 93.75	4:46 1:57 1:15	43.75 100	3:37 2:51
20	81.25 87.5 100	5:00 2:23 1:11	100	4:40	100	3:50	100	2:51
21	100	4:50	100	3:23	100	2:26	100	2:50
22	68.75 100	5:00 5:00	93.75	4:07	100	2:57	100	2:31
23	87.5 100	5:00 3:56	93.75	4:03	100	3:45	93.75	4:11
24	93.75	4:37	87.5 100	3:58 3:21	100	3:44	87.5 93.75	2:57 2:41
25	87.5 6.25c 93.75	4:59 - 4:32	100	3:38	100	3:14	50 100	2:44 3:07
26	50 100	5:00 2:01	56.25 100	2:53 1:16	62.5 87.5 93.75	2:08 3:33 1:34	62.5 93.75	2:56 1:28

Table 3 (continued)

Participants	Task 1		Task 2		Task 3		Task 4	
	% ^a	Duration ^b	%	Duration	%	Duration	%	Duration
27	62.5 87.5 100	5:00 5:00 3:52	93.75	3:41	100	3:43	100	3:18
28	75 100	5:00 4:38	100	3:01	100	2:45	100	2:47
29	25 93.75	4:28 3:07	56.25 100	5:00 5:00	12.5 87.5 93.75	5:00 3:05 3:51	100	3:50
30	100	4:41	100	2:44	93.75	3:56	93.75	4:13
31	93.75	4:33	100	3:39	100	3:26	93.75	3:10
32	75 81.25 100	5:00 5:00 4:44	93.75	4:35	100	4:43	100	4:51
33	68.75 93.75	5:00 5:00	100	3:32	93.75	4:01	100	3:00
Average score across all individuals	87.91	4:36	95.53	3:48	95.27	3:36	95.31	3:19

The same criteria and experimental settings described above applied to BH- group.

^a% of correct responses (6.25% per each question)

^bDuration [mm:ss] (minutes:seconds)

^cTask incomplete

Code Availability Not Applicable.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data Availability Data and materials are available from the first author upon request.

Compliance with Ethical Standards

Ethics Approval Division of Arts, Law, Psychology & Social Sciences Human Research Ethics Committee of the University of Waikato [FS2020-62]

Consent to Participate Participants consented to their participation and signed written consent forms that were approved by the ethics committee specified above.

Consent for Publication All authors consent to publication

Conflicts of Interest/Competing Interests None

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdel-Khalek, A. M. (2005). Reliability and factorial validity of the standard progressive matrices among Kuwaiti children ages 8 to 15 years. *Perceptual and Motor Skills, 101*(2), 409–412. <https://doi.org/10.2466/pms.101.2.409-412>
- Barbero-Rubio, A., López-López, J. C., Luciano, C., & Eisenbeck, N. (2016). Perspective-taking measured by implicit relational assessment procedure (IRAP). *The Psychological Record, 66*(2), 243–252. <https://doi.org/10.1007/s40732-016-0166-3>
- Barnes-Holmes, D., & Harte, C. (2022). Relational frame theory 20 years on: The Odysseus voyage and beyond. *Journal of the Experimental Analysis of Behavior, 117*(2), 240–266. <https://doi.org/10.1002/jeab.733>
- Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*(7), 169–177.
- Barnes-Holmes, Y., McHugh, L., & Barnes-Holmes, D. (2004). Perspective-taking and theory of mind: A relational frame account. *The Behavior Analyst Today, 5*(1), 15–25. <https://doi.org/10.1037/h0100133>
- Barron, B. F., Verkuylen, L., Belisle, J., Paliliunas, D., & Dixon, M. R. (2018). Teaching "then-later" and "here-there" relations to children with autism: An evaluation of single reversals and transformation of stimulus function. *Behavior Analysis in Practice, 12*(1), 167–175. <https://doi.org/10.1007/s40617-018-0216-1>

- Belisle, J., Dixon, M. R., Stanley, C. R., Munoz, B., & Daar, J. H. (2016). Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: Single-reversal “I-you” deictic frames. *Journal of Applied Behavior Analysis*, 49(4), 965–969. <https://doi.org/10.1002/jaba.324>
- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 11–63). Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.
- Cotton, S. M., Kiely, P. M., Crewther, D. P., Thomson, B., Laycock, R., & Crewther, S. G. (2005). A normative and reliability study for the Raven’s Coloured progressive matrices for primary school aged children from Victoria, Australia. *Personality & Individual Differences*, 39(3), 647–659. <https://doi.org/10.1016/j.paid.2005.02.015>
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality & Social Psychology*, 44(1), 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>
- Davlin, N. L., Anne Rehfeldt, R., & Lovett, S. (2011). A relational frame theory approach to understanding perspective-taking using children’s stories in typically developing children. *European Journal of Behavior Analysis*, 12(2), 403–430. <https://doi.org/10.1080/15021149.2011.11434392>
- De Lillo, M., & Ferguson, H. (2022). Perspective-taking and social inferences in adolescents, young adults and older adults. *Psychology Archives*, 1–57. <https://doi.org/10.31234/osf.io/8z2tf>
- Dymond, S., & Barnes, D. (1997). Behavior-analytic approaches to self-awareness. *The Psychological Record*, 47(2), 181–200. <https://doi.org/10.1007/BF03395219>
- Epley, N., & Caruso, E. M. (2009). Perspective taking: Misstepping into other’s shoes. In K. D. Markman, W. M. Klein, & J. A. Suhr (Eds.), *Handbook of imagination and mental simulation* (pp. 297–311). Psychology Press.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/BF03193146>
- Ferguson, H. J., & Cane, J. (2017). Tracking the impact of depression in a perspective-taking task. *Scientific Reports*, 7(1), 14821–14829. <https://doi.org/10.1038/s41598-017-13922-y>
- Gilroy, S. P., Lorah, E. R., Dodge, J., & Fiorello, C. (2015). Establishing deictic repertoires in autism. *Research in Autism Spectrum Disorders*, 19, 82–92. <https://doi.org/10.1016/j.rasd.2015.04.004>
- Guinther, P. M. (2017). Contextual influence over deriving others’ true beliefs using a relational triangulation perspective-taking protocol (RT-PTP-M1). *Journal of the Experimental Analysis of Behavior*, 108(3), 433–456. <https://doi.org/10.1002/jeab.291>
- Harman, I. P. (1990). Teaching indirect speech: Deixis points the way. *ELT Journal*, 44(3), 230–238. <https://doi.org/10.1093/elt/44.3.230>
- Hayes, S. C. (1984). Making sense of spirituality. *Behaviorism*, 12(2), 99–110.
- Hayes, S. C., & Barnes-Holmes, D. (2004). Relational operants: Processes and implications: A response to Palmer’s review of relational frame theory. *Journal of the Experimental Analysis of Behavior*, 82(2), 213–224. <https://doi.org/10.1901/jeab.2004.82-213>
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Kluwer Academic/Plenum.
- Heagle, A., & Rehfeldt, R. (2006). Teaching perspective-taking skills to typically developing children through derived relational responding. *Journal of Early & Intensive Behavior Intervention*, 3(1), 1–34. <https://doi.org/10.1037/h0100321>
- Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior*, 74(2), 207–227. <https://doi.org/10.1901/jeab.2000.74-207>
- Howlin, P., Baron-Cohen, S., & Hadwin, J. (1999). *Teaching children with autism to mind-read: A practical guide for teachers and parents*. John Wiley & Sons.
- Jackson, M. L., Mendoza, D. R., & Adams, A. N. (2014). Teaching a deictic relational repertoire to children with autism. *The Psychological Record*, 64(4), 791–802. <https://doi.org/10.1007/s40732-014-0078-z>
- Janssen, G., De Mey, H., Hendriks, A., Koppers, A., Kaarsemaker, M., Wittman, C., & Egger, J. (2014). Assessing deictic relational responding in individuals with social anxiety disorder: Evidence of perspective-taking difficulties. *The Psychological Record*, 64(1), 21–29. <https://doi.org/10.1007/s40732-014-0013-3>
- Kavanagh, D., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C., & Finn, M. (2018). Exploring differential trial-type effects and the impact of a read-aloud procedure on deictic relational responding on the IRAP. *The Psychological Record*, 68(2), 163–176. <https://doi.org/10.1007/s40732-018-0276-1>
- Keysar, B. (1997). Unconfounding common ground. *Discourse Processes*, 24(2–3), 253–270. <https://doi.org/10.1080/01638539709545015>
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), 32–38. <https://doi.org/10.1111/1467-9280.00211>
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41. [https://doi.org/10.1016/S0010-0277\(03\)00064-7](https://doi.org/10.1016/S0010-0277(03)00064-7)
- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fisher, J. S. (2004). *Neuropsychological assessment* (4th ed.). Oxford University Press.
- Lovett, S., & Rehfeldt, R. A. (2014). An evaluation of multiple exemplar instruction to teach perspective-taking skills to adolescents with Asperger syndrome. *Behavioral Development Bulletin*, 19(2), 22–36. <https://doi.org/10.1037/h0100575>
- McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004a). Perspective-taking as relational responding: A developmental profile. *The Psychological Record*, 54(1), 115–144. <https://doi.org/10.1007/BF03395465>
- McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004b). A relational frame account of the development of complex cognitive phenomena: Perspective-taking, false belief understanding, and deception. *International Journal of Psychology & Psychological Therapy*, 4(2), 303–324.
- Montoya-Rodríguez, M. M., & Cobos, F. J. M. (2016). Relationship between deictic relational responding and theory of mind tasks in children: A pilot study. *The Psychological Record*, 66(4), 573–587. <https://doi.org/10.1007/s40732-016-0193-0>
- Montoya-Rodríguez, M. M., McHugh, L., & Cobos, F. (2017). Teaching perspective-taking skills to an adult with down syndrome: A case study. *Journal of Contextual Behavioral Science*, 6(3), 293–297. <https://doi.org/10.1016/j.jcbs.2017.04.012>
- O’Neill, J., & Weil, T. (2014). Training deictic relational responding in people diagnosed with schizophrenia. *The Psychological Record*, 64(2), 301–310. <https://doi.org/10.1007/s40732-014-0005-3>
- Palmer, D. C. (2004a). Data in search of a principle: A review of relational frame theory: A post-Skinnerian account of human language and cognition. *Journal of the Experimental Analysis of Behavior*, 81(2), 189–204. <https://doi.org/10.1901/jeab.2004.81-189>
- Palmer, D. C. (2004b). Generic response classes and relational frame theory: Response to Hayes and Barnes-Holmes. *Journal of the Experimental Analysis of Behavior*, 82(2):225–34. <https://doi.org/10.1901/jeab.2004.82-225>
- Premack, D., & Woodruff, G. (1978). Chimpanzee problem-solving: A test for comprehension. *Science*, 202(4367), 532–535. <https://doi.org/10.1126/science.705342>

- Raven, J. C. (1984). Manual for the Raven colored matrices (rev.). NFER-Nelson.
- Rehfeldt, R., Dillen, A., Ziomek, J., & Kowalchuk, E. (2007). Assessing relational learning deficits in perspective-taking in children with high-functioning autism spectrum disorder. *The Psychological Record*, *57*(1), 23–47. <https://doi.org/10.1007/BF03395563>
- Samson, A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception & Performance*, *36*(5), 1255–1266. <https://doi.org/10.1037/a0018729>
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*(1), 5–22. <https://doi.org/10.1901/jeab.1982.37-5>
- Skinner, B. F. (1957). *Verbal behavior*. (The Century psychology series). Appleton-Century-Crofts.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*(2), 349–367. <https://doi.org/10.1901/jaba.1977.10-349>
- Taylor, T., & Edwards, T. L. (2021). What can we learn by treating perspective taking as problem solving? *Perspectives on Behavior Science*, *44*(2–3), 359–387. <https://doi.org/10.1007/s40614-021-00307-w>
- Tibbetts, P. A., & Rehfeldt, R. A. (2005). Assessing relational learning deficits in perspective-taking in children with high-functioning autism. *Behavioral Development Bulletin (Philadelphia, PA)*, *12*(1), 62–68. <https://doi.org/10.1037/h0100562>
- Thirus, J., Starbrink, M., & Jansson, B. (2016). Relational frame theory, mathematical and logical skills: A multiple exemplar training intervention to enhance intellectual performance. *International Journal of Psychology & Psychological Therapy*, *16*(2), 141–155.
- Villatte, M., Monestès, J. L., McHugh, L., i Baqué, E. F., & Loas, G. (2010). Adopting the perspective of another in belief attribution: Contribution of relational frame theory to the understanding of impairments in schizophrenia. *Journal of Behavior Therapy & Experimental Psychiatry*, *41*(2), 125–134. <https://doi.org/10.1016/j.jbtep.2009.11.004>
- Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record*, *41*(1), 33–50. <https://doi.org/10.1007/BF03395092>
- Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *The Psychological Record*, *61*(3), 371–390. <https://doi.org/10.1007/BF03395767>

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