The Psychological Record (2025) 75:41–55
https://doi.org/10.1007/s40732-024-00628-7

ORIGINAL ARTICLE

Relational Coherence, Speaker Preference, and Rule-Following: A Replication and Extension of Bianchi et al. (2021)

Paulo Henrique Bianchi[1,2] · **William Ferreira Perez**[2,3] · **Colin Harte**[2,4] · **Dermot Barnes-Holmes**[5] · **Antonio Teixeira e Silva**[1]

Accepted: 18 December 2024 / Published online: 22 January 2025

Abstract

The present study evaluated the influence of relational coherence upon speaker preference and subsequent rule-following behavior. Across two experiments, participants first learned a particular set of conditional relations via a stimulus-pairing procedure (e.g., A1B1, A2B2). Then, the same relations were presented in an MTS task to ensure that participants could correctly match these stimuli. Next, the participants were exposed to two animated characters or speakers that differed by the color of their t-shirt (green or purple). One of the speakers "stated" relations that cohered (e.g., A1B1, A2B2) with the participant's previous relational training and the other stated relations that were incoherent (e.g., A1B2, A2B1). The stimulus relations were presented inside a speech balloon, as if each speaker spoke them. Rule-following was assessed during a subsequent preference test in which participants were exposed to a simultaneous discrimination task with novel stimuli. On each trial, participants first chose which of the two speakers to provide an instruction to tell them how to respond on that trial. Both speakers presented rules that were consistent with the programed contingency. In Experiment 1, feedback during the rule-following task was delivered in CRF; In Experiment 2, the schedule was FR10. In both experiments and IRAP was implemented to assess the credibility of each speaker. Results showed that a history of relational coherence with each speaker had a reliable impact upon the preference test, rule-following behaviors and the IRAP performance, irrespective of the reinforcement schedule in place (CRF in Experiment 1 and FR10 in Experiment 2).

Keywords Relational frame theory · HDML · Coherence · Rule-following · Preference test · Adults

Rule-governed behavior was first defined by Skinner (1966) as behavior under the control of contingency specifying stimuli (i.e., a rule) that allowed a listener to interact with a reinforcement contingency without requiring their behavior to be directly shaped or reinforced by that contingency. This type of behavior was initially proposed by Skinner as being critical to a behavior-analytic account of problem solving. Decades of research has since explored the influence of rules (or instructions) on human performance on schedules of reinforcement (see Harte et al., 2020a, 2020b, for a recent overview of this work). A side effect of rule-following in this context has been termed the "insensitivity" effect. This term refers to a pattern of responding typically observed with verbal humans in which a particular rule or instruction continues to influence behavior despite changes in the reinforcement contingencies that do not support this pattern (see Hayes, 1989, for an early book-length treatment).

A vast literature has focused on evaluating the conditions that may influence rule-following and the persistence of instructed behavior when the contingencies specified by a rule no longer match the scheduled contingencies of reinforcement. For example, an early study showed that rule-following could come under discriminative control (Galizio, 1979); other research explored how sources of reinforcement may interact with rule-following (e.g., Hayes et al., 1985);

✉ Paulo Henrique Bianchi
phbianchi@gmail.com

1 Instituto de Pesquisas Energéticas e Nucleares IPEN-CNEN, São Paulo, Brazil

2 Paradigma—Centro de Ciências do Comportamento, São Paulo, Brazil

3 Instituto Nacional de Ciência e Tecnologia Sobre Comportamento, Cognição e Ensino (INCT-ECCE), São Paulo, Brazil

4 Universidade Federal de São Carlos, São Carlos, Brazil

5 School of Psychology, Ulster University, Northern Ireland, UK

some research investigated how different types of rules may affect upon performances on a schedule (Hayes et al., 1986a, 1986b); whereas other work investigated how different levels of rule-accuracy effected subsequent rule-following (DeGrandpre & Buskist, 1991; but see also, Baron & Galizio, 1983; Hayes et al., 1986a, 1986b; Hayes , 1989; Martinez-Sanchez & Ribes-Iñesta, 1996; Martinez & Tamayo, 2005; Paracampo & Albuquerque, 2005; Perez et al., 2010, for some more relevant experimental examples of the rich work in this broad area). This research agenda has been extended and explored through the lens of arbitrarily applicable relational responding (AARR), as defined by relational frame theory (RFT; Hayes et al., 2001; see also Barnes-Holmes & Harte, 2022a and Harte & Barnes-Holmes, 2021), a behavior-analytic account of human language and cognition.

In the original treatment (Hayes et al., 2001), the core analytic unit of the relational frame (the simplest form of relational network) was defined as comprising three properties: mutual entailment (e.g., if X is related to Y in some way, then Y is also related to X), combinatorial entailment (e.g., if X is related to Y and also to Z, then Y and Z are also related), and the transformation of stimulus functions (the nonrelational stimulus functions are transformed in accordance with the types of relations in which they participate). For example, if X was relationally greater than Z but less than Y, and an aversive function was established for X, then Y may acquire greater aversive functions than X, whereas Z may acquire less (e.g., Dougher et al., 2007). The concept of a relational frame intended, therefore, to capture the extent to which symbolic language, as AARR, alters human responding to the world around us.

From an RFT perspective, rules or instructions are made up of networks of relations involving the words in the rule and the stimuli and events to which they refer. In addition, contextual cues contained in the rule transform particular behavior functions. For example, the instruction "If the chicken meat is pink, then do not eat it" involves frames of coordination between the words "chicken," "meat," "pink," and "not eat" and actual chicken meat the color pink, and the act of not eating, respectively. In addition, the contextual cues "if" and "then" establish a conditional relation between the pink chicken and the act of not eating it. The functions of the pink uncooked chicken itself are thus transformed by the relational network such that it now controls the behavior of not eating this food in this context.

In recent research, a framework has been offered within RFT to provide a conceptual analysis to help systematize research in the area and highlight potentially important variables relevant to AARR. This framework is known as the hyperdimensional multilevel (HDML) framework (Barnes-Holmes et al., 2020, 2021) and conceptualizes AARR as varying along five levels and four dimensions. The five levels are based on analyses that have emerged from the RFT literature over the last 30 years and are viewed as increasingly complex patterns of relating from mutual entailment, combinatorial entailment, relational networks, relating relations, to the relating of relational networks. These five levels are conceptualized as varying along four dimensions; coherence, complexity, derivation, and flexibility (although brief explanations of the concepts involved in the HDML relevant to the current work will be provided below, the reader is referred to Barnes-Holmes et al., 2017, for a detailed explication of each of these concepts).

Researchers have begun to use this framework to explore the dynamics involved in persistent rule-following. In particular, this work has typically sought to assess the extent to which an experimenter-given rule, involving some aspect of an experimentally established derived relation, may affect persistent rule-following (in the face of reversed reinforcement contingencies). Furthermore, this work has also explored extent to which this persistence in rule-following may differ when the derived relation varies along a dimension of the HDML. For example, Harte et al., (2020a, 2020b, Experiment 2; Harte et al., 2021) targeted coherence. Within the HDML, coherence is defined as the extent to which a pattern of derived relating is consistent with or predicted by a previously established pattern. In these experiments, Harte et al. first established novel combinatorially entailed relations (A = B = C, C = A) in two groups of participants. For one group, performance feedback was provided on their derived test responses (C = A), in principle, increasing the coherence of such responses; the other group received no such feedback. The novel derived relation was subsequently inserted into a rule for responding on a matching-to-sample (MTS) task. That is, participants received the rule "Choose the comparison that is *BEDA* the sample stimulus" where *BEDA* had been derived to mean *least similar* in the previous training and testing phase (i.e., train Least similar = TTT = BEDA, test Beda = Least similar).

For the first 100 trials of the MTS task, reinforcement (i.e., gained points) was provided for choosing the comparison that was the least similar to the sample. However, on the 101st trial, an unsignaled contingency reversal occurred and now participants lost points for choosing the least similar comparison. The general finding of both experiments was that participants were more likely to persist with following the rule when the derived element of the rule had been reinforced with performance feedback. Or in other words, participants were more likely to persist with rule-following if they had been told that they had correctly derived the relationship between BEDA and Least Similar. It is important to note that the results of both of these studies suggested that this effect was moderated by another variable highlighted within the HDML, level of derivation (i.e., the extent to which a derived relation has been emitted previously). However, we will not consider this finding further in the current article because it is not directly relevant to the current research.

The above two studies explored how providing reinforcement in one condition (versus not in another condition) might produce relatively coherent patterns of responding. According to the HDML definition of relational coherence, both studies focused on manipulating reinforcement of previously established patterns and used rule-following to measure the effects on relational coherence. In contrast, other recent work has involved manipulating relational coherence directly to explore the extent to which relational coherence affects a choice or preference for following one rule over another, even when both rules produce the same degree of reinforcement when followed. A study by Bianchi et al. (2021) employed this latter strategy (see also Bern et al., 2020, and Harte et al., 2021, for other work that has used the HDML to explore the impact of coherence on rule persistence using alternative approaches to those described currently). In particular, participants first learned a particular set of conditional relations via a stimulus-pairing procedure (e.g., A1B1, A2B2). Then, the same relations were presented in an MTS task to ensure that participants could correctly pair these stimuli. Participants were first introduced to two animated characters, each identifiable by their distinct t-shirt colors—one green, the other purple. Each character presented different sets of relations. The first character presented relations that were consistent with the relational training the participants had previously undergone (examples being A1B1, A2B2), whereas the second character presented relations that did not align with the prior training (such as A1B2, A2B1). These relational statements were visually represented within speech balloons, intended to create the impression that they were being "spoken" by the characters. The experiment then progressed to a phase where the adherence to rules by the participants was evaluated. This was done through a preference test, which introduced a simultaneous discrimination task featuring new stimuli. Within each trial of this test, participants had to choose one of the two speakers to provide instructions about the appropriate response for that particular trial. It is important to note that both speakers offered guidance that adhered to the predetermined contingency of the experiment (e.g., "Click on the red circle to earn 10 points"). This allowed the researchers to assess the extent to which a history of "speaking" in a manner that was coherent or incoherent with the participant's relational training affected upon speaker preference and whether they then followed that rule. Results showed that all participants preferred the coherent speaker to provide instructions (at least initially) and that participants were more likely to follow the rules provided by that speaker throughout the test.

The results of the study by Bianchi et al. (2021) suggested that the extent to which an instruction was consistent (coherent) with previously established patterns of relational responding exerted control over both speaker preference and (at least initial) rule-following. However, given the small number of participants involved (four), the current study sought to replicate and extend these findings. In addition to systematically replicating the 2020 study, the current work also sought to assess the credibility of each speaker, using a reaction time-based procedure, following the preference test. In particular, participants were exposed to the implicit relational assessment procedure (IRAP; Barnes-Holmes et al., 2008) after they completed the preference test. The IRAP is a computer-based task designed to assess relational responding under time pressure. During some blocks of trials, participants are required to respond in a manner deemed coherent with a previous learning history (e.g., coherent speaker-reliable-True; incoherent speaker-reliable-False) and in other blocks of trials are required to responding in a manner assumed to be inconsistent with that history (e.g., coherent speaker-reliable-False; coherent speaker-unreliable-True). The general assumption is that participants should be faster when responding to relations that are consistent rather than inconsistent with previous learning histories (see Barnes-Holmes & Harte, 2022b, for a detailed explication of the IRAP and its use as a behavior-analytic tool; see also Hussey, 2023, for potential limitations and considerations when using the IRAP). We opted to include a latency-based procedure, such as the IRAP, because it would likely be less susceptible to experimenter demand effects. In particular, in the context of the current study, participants might choose to follow a particular rule and indicate a preference for a particular speaker because doing so was in line with what they believed were the experimenter's expectations. Inclusion of a latency-based procedure thus allowed for a less demand-sensitive experimental analytic test of reactions to the speakers.

The current study, therefore, sought to investigate whether the findings of Bianchi et al. would be reproduced with a greater number of participants and explore how patterns of relational responding produced on the IRAP did or did not reflect rule-following and speaker preference. In addition, a second experiment, planned from the outset as a separate study, was conducted, partially replicating Experiment 1, except modifying the schedule of reinforcement of the preference test to a Fixed Ratio (FR) 10 schedule. For the purposes of communication, this second experiment will be outlined in further detail following the presentation of Experiment 1.

Experiment 1

Method

Participants

Participants were 18 verbally competent adults (males = 14; age = 28–65 years, $M = 43.17$ $SD = 10.68$), recruited by convenience sampling and with no previous experience in any psychology-based research. Participants read and agreed to terms of consent (approved by the Brazilian platform for ethical committees, Plataforma Brasil, CAAE 19827719.0.0000.5493). At the end of the procedure, participants were fully debriefed and thanked. They did not receive any compensation or payment.

Setting, Equipment, and Stimuli

The research was conducted in a quiet 13 ft × 13 ft room equipped with a table, a chair for the participant, a chair for the experimenter positioned behind the participant and turned perpendicular to the participant's direction, and notebook computer. The custom-written software "Preferência Entre Falantes CRF" presented the rule-following task. The GO-IRAP software ran the latency-based IRAP credibility task and calculated D-IRAP scores for analysis from the final experimental phase. As depicted in Fig. 1, stimuli were two pictorial representations of "speakers," abstract black shapes presented on a white background, and abstract-colored shapes.

Procedure

Figure 1 provides an illustration of the five experimental phases involved in Experiment 1: (1) Relational training; (2) Relational testing; (3) Establishing coherent and incoherent speakers; (4) Preference test; and (5) IRAP credibility test.

Phase 1: Relational training. A stimulus-pairing training procedure (e.g., Delgado & Rodriguez, 2022; Leader & Barnes-Holmes, 2001) was implemented to teach four arbitrary relations: A1B1, A2B2, B1C1, and B2C2. As in Bianchi et al. (2021), this procedure was chosen in part because it seemed topographically similar to actual speaking behavior in which a word is paired with an object without the need for a matching response. In addition, using it in the training phase familiarized participants with the procedure that was used later in Phase 4 (see below). Before starting, the participants read the following instructions on the computer screen:

This is your first task. Pairs of symbols will be displayed on the computer screen. In each trial, a symbol will appear followed by a corresponding symbol. You must learn these pairs to answer a test in the next stage, so pay close attention (press spacebar to continue).

Each trial comprised the successive presentation of a pair of stimuli that were arbitrarily designated to be equivalent or coordinate with one another (e.g., A1B1). Trial onset presented the first stimulus of the pair (e.g., A1) in the center of the screen for 2 s followed by a 1-s interval (blank).

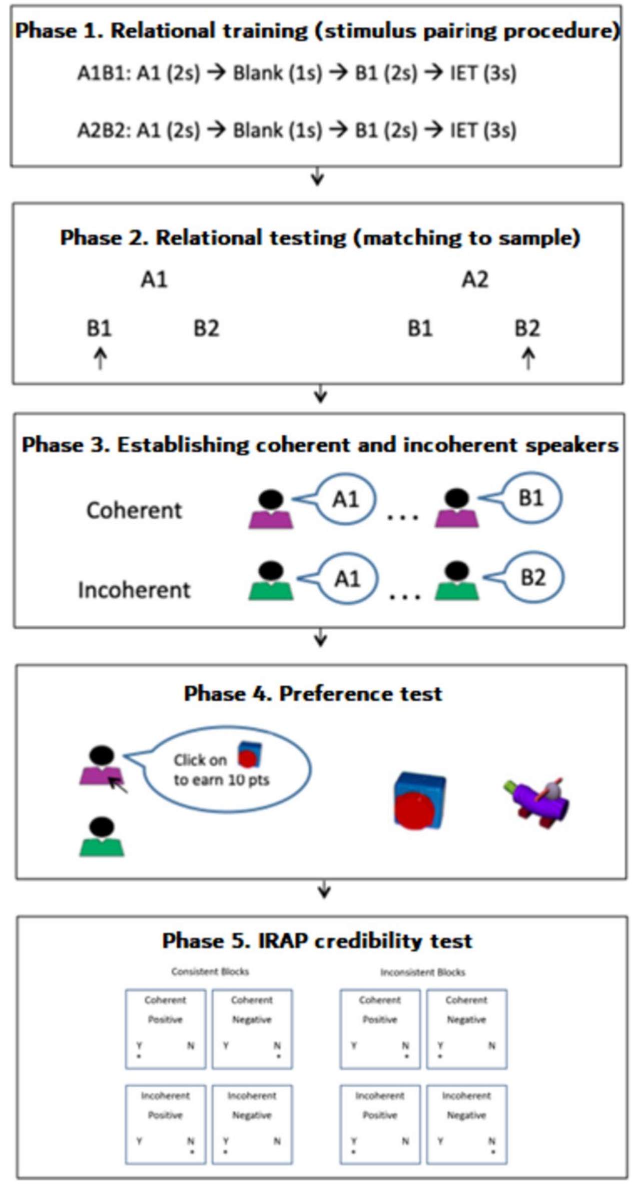

Fig. 1 Outline of the Experimental Phases. *Note.* In Phase 5, the asterisk below the answers represents the correct response for each trial in each block (consistent or inconsistent)

Once the interval ended the second stimulus of the pair was presented for 2-s followed by a 3-s intertrial interval (ITI, blank). After this ITI, another pair of stimuli were similarly presented. Training blocks comprised four trials in which the four stimulus pairs (A1B1; A2B2; B1C1; B2C2) were randomly presented. Each block was presented 13 times (52 trials in total).

Phase 2: Relational testing. After the relational training, participants were exposed to an MTS task that tested responding coherently with the relations established via stimulus pairing in Phase 1 (e.g., given A1 as a sample stimulus, choose B1 rather than B2 or a third novel stimulus as a comparison stimulus, and so on: A2B2; B1C1; B2C2). Before the MTS trials were presented, participants read the following instruction on the computer screen:

Now, let's test what you have learned. A symbol will appear at the top of the screen, followed by three symbols below. You will have to choose the symbol below that matches the symbol above. Choose it by clicking with the mouse cursor. Consider what you learned in the previous stage. The computer will record your hits and errors based on the previous stage, but will not show this information during the task (press spacebar to continue).

Each trial onset first presented a sample stimulus at the top of the screen. After 1 s, three comparison stimuli were simultaneously presented at the bottom of the screen. The first stimulus of each pair presented in Phase 1 was always presented as a sample stimulus (e.g., A1) whereas the second stimulus of each pair was always presented as one of the comparison stimuli (e.g., B1), along with the second stimulus from the other pair (e.g., B2), and a third novel stimulus (e.g., N1 or N2). The third comparison stimulus was not part of any programed relation and aimed to avoid responses based on reject control (see Sidman & Tailby, 1982; Perez et al., 2015). Mouse clicking on the comparison stimulus (e.g., B1) of the same pair as the sample (i.e., A1) was registered as a correct response; mouse clicking any of the two other comparisons was registered as an error. Participants responses had no programmed feedback and were followed by a 0.5-s ITI with a blank screen and the onset of the next trial. The position of the comparison stimuli was randomized. A total of 48 MTS test trials were presented in a single block that comprised the four relations established in Phase 1 (A1B1, A2B2, B1C1, and B2C2). Each relation was presented 12 times in a quasi-random order. The criteria to proceed to the next phase was 80% correct responses (i.e., 39/48 test trials). Although arbitrarily defined, the probability of a participant reaching this criterion when responding randomly is approximately 1.07×10^{-11}, as calculated by Binomial Trial method (Papoulis, 1965), which is significantly low. Participants that did not reach these criteria were thanked and dismissed.

Phase 3: Establishing coherent and incoherent speakers. This phase was largely similar to Phase 1. As can be seen from Fig. 1, however, the stimulus pairing on each trial was presented inside a speech balloon located above one of two speakers. The speakers were differentiated by the color of their t-shirts (green or purple). One of the speakers (the Coherent speaker) presented stimulus pairs that were *coherent* with the trained and tested stimulus relations from Phases 1 and 2 (i.e., A1B1, A2B2, B1C2, and C2B2); the other speaker (the Incoherent speaker) presented pairs that were *incoherent* with the previously established relations (i.e., A1B2, A2B1, B1C2 and B2C1). The color of the speaker designated as coherent and incoherent alternated between participants.

Before starting Phase 3, the participants read the following instruction on the screen:

Now you will meet two characters, one in a green t-shirt and one in a purple t-shirt. They will show you pairs of symbols, in a similar way to the first task. Later, you will have to choose one of them to help you solve a series of problems, so try to form an opinion about them by looking closely at the pairs of symbols they "speak" to you about (press spacebar to continue).

This level of instruction was designed to explain the task to the participants, but critically without directly influencing what kind of opinion they should form about the speakers, and neither what they should be specifically looking for as they engage in the task.

Each trial onset presented the speaker and the speech balloon (see Fig. 1) along with the successive presentation of a given stimulus pair inside the speech balloon as if they were spoken by one of the speakers. The coherent speaker presented the same pairs that were presented in Phase 1 (A1B1, A2B2, B1C1, and B2C2); the Incoherent speaker presented stimulus pairs that were different combinations of Phase 1 stimulus pairings (A1B2, A2B1, B1C2, and B2C1). The first stimulus of the pair (e.g., A1) was presented in the center of the speech balloon for 2 s followed by a 1-s interval (blank speech balloon) followed by the presentation of the second stimulus of the pair for 2 s. The presentation of the stimulus pair was then followed by a 0.5-s interval in which the speech balloon was blank. Next, all stimuli were withdrawn from the screen during a 2.5-s ITI until the next trial onset. No response was required from the participant but to observe the stimuli presented inside the speech balloon "spoken" by the speaker. Training trials were organized in eight-trial blocks. Each block presented all coherent (A1B1, A2B2, B1C1, B2C2) and incoherent pairs (A1B2, A2B1, B1C2, B2C1) once per block in a quasi-random order, to avoid any preference derived from order effect. Each block started with the coherent speaker and thus the presentation of one of the coherent pairs (which of the four coherent pairs presented was randomly selected). The remaining seven training trials within that block alternated between the coherent and incoherent speaker. A total of seven blocks were presented (i.e., a total of 56 trials).

Phase 4: Preference test. Before starting, the participants read the following instructions on the screen:

Ok, you advanced to the next stage! You will be presented with two images on the screen. You must choose one of them. Choosing the correct option (there is only one!) will give you points accumulated on a counter. In each trial, you must choose one of the characters from the previous phase to help you proceed and decide what image you should choose. Click on one of the characters to “ask for help.” After that, you must click on one of the images, to select it and proceed to the next trial onset. Try to accumulate as many points as possible (press spacebar to continue).

As depicted in Fig. 1, each test trial onset simultaneously presented the following elements on the screen: in the top right-hand side was a counter (to present points); on the left-hand side the two speakers appeared (i.e., purple and green), placed one above the other (the position of the green and purple speakers alternated across trials); on the center-right of the screen, two abstract colored images were displayed side-by-side. The novel colored stimuli were sorted from a pool that comprised 60 stimuli. The position of the correct stimulus was randomly determined by an algorithm (see Knuth, 2014).

Once the participant selecting one of the speakers using a mouse click, it immediately presented a “hint” (a rule) inside the speech balloon that would be displayed until the end of the trial. Only one speaker could be selected per trial. Clicking on the second speaker after having selected one of them produced no programed consequence. None of the two colored stimuli could be chosen before selecting one of the two speakers. In case that happened, a warning message would appear written on the screen: “You must request a hint before choosing an image!”.

During this phase, both speakers presented rules that were consistent with the programed contingency: “Click on [small version of the correct image for that trial] to earn 10 points.” This allowed assessment of the extent to which a history of “speaking” in a manner that was coherent or incoherent with the Phase 1 training (confirmed in Phase 2 testing), affected upon speaker preference. When the selection of the colored stimuli was enabled, correct responses (following the rule) immediately produced the written message “+10 points” and 10 points were added to the counter; incorrect responses produced "No points earned" and no points added to the counter. The feedback message was displayed on-screen for 1 s and was followed by a 1-s ITI. The preference test comprised a total of 30 trials.

Phase 5: IRAP credibility test. The IRAP aimed to evaluate the credibility of each of the speakers using positive and negative adjectives. The depiction of this phase in Fig. 1 presents an illustration of the IRAP trial-types. Each IRAP trial onset simultaneously presented one of the speakers (coherent or incoherent) on the top of the screen as a label stimulus, one word (a positive or a negative adjective) in the center as a target stimulus, and two response options in the lower left and right corners (“Yes” and “No”). The image of the speakers presented as label stimuli were the same as those from previous phases; the adjectives presented as target stimuli were polar adjectives that related to the context of credibility (positive adjectives: sincere, good, reliable, safe, accurate and true; negative adjectives: liar, bad, doubtful, insecure, inaccurate and false). Correct responses were followed by removal of all stimuli presented in that trial and a 400 ms ITI; incorrect responses were followed by a red “X” appearing in the center of the screen as feedback for error, while all stimuli remained on the screen; the withdrawal of all stimuli and the ITI only followed the emission of a correct response for that trial. IRAP blocks comprised 24 trials. During each block, four different trial types were presented pertaining to label-target stimulus combinations: Coherent–Positive, Coherent–Negative, Incoherent–Positive, Incoherent–Negative. The IRAP task always alternated between blocks that required responding assumed to be history-consistent or -inconsistent with the experimentally established credibility of the speaker. In particular, during history-consistent blocks, correct responding on the four trials types was (Speaker-Word/Correct Answer): Coherent–Positive/True, Coherent–Negative/False, Incoherent–Negative/True, Incoherent–Positive/False. During history-inconsistent blocks, the contingencies were reversed and correct responding on the four trial types was: Coherent–Negative/True, Coherent–Positive/False, Incoherent–Positive/True, Incoherent–Negative/False. The four trial types were randomly presented six times within each block. Half of the participants started the IRAP task with a history-consistent block of trials and the other half with a history-inconsistent block.

The IRAP started with practice (“warm up”) blocks to familiarize the participant with the task. Test blocks only started after the participants achieved set mastery criteria on those practice blocks. The practice phase started with a pair of consistent/inconsistent blocks, requiring 85% response accuracy in both blocks. After reaching criterion, participants are exposed to another pair of consistent/inconsistent blocks with an additional criterion of latency. Participants were also required to maintain a median response latency of 2,000 ms between stimuli onset and emission of a correct response. In case 2,000 ms had elapsed and no response was emitted, written feedback “Go faster next time!” appeared on the screen. Participants that failed to meet both accuracy and latency criteria after four pairs of practice blocks were thanked for their participation and dismissed from the experiment. Participants who met both accuracy and latency criteria across a successive pair of consistent/inconsistent practice blocks advanced to the testing phase.

The test phase comprised three pairs of consistent/inconsistent blocks programmed as described in the practice phase. Participants were not required to maintain accuracy or latency criteria during Test blocks in order to progress through the blocks. IRAP data analysis (DIRAP scores) were based only on the test block performances. Participant’s scores were excluded from data analysis if their accuracy fell below 85% in more than one test block, or average latency exceeded 2,000 ms in any test block. The series of six test blocks ended with a written message "Please, call the experimenter."

Results and Discussion

From the 18 participants that took part in Experiment 1, 8 were excluded from data analyses for not fulfilling at least one of the following criteria: for not achieving at least 80% accuracy at Stage 2 test (7 participants: E1, E2, E3, E4, E5, E6 and E7); for not reaching the criteria in the IRAP practice blocks (1 participant: E8); for not maintaining the necessary accuracy for at least five of the six IRAP test blocks (3 participants: E4, E5, and E6); or for not maintaining the necessary latency for at least five of the six IRAP test blocks (1 participant: E3; see Table 1). A total of 10 participants met all the criteria, with relevant demographic information available in Table 2.

Table 1 Inclusion/exclusion scores and criteria for Experiment 1

Participant	Phase 2 Score	IRAP Test blocks with accuracy < 85%	IRAP Test blocks with latency above 2000 ms
E1	33		
E2	22		
E3	25		5
E4	27	1	
E5	34	3	
E6	36	2	
E7	37		
E8	43	N/A	N/A
P1	47		
P2	47		
P3	42		
P4	48		
P5	48		
P6	45		
P7	43		
P8	48		
P9	48		
P10	41		

"E" refers to an excluded participant; excluded participant E8 did not reach the criteria in the IRAP training blocks and was dismissed before the test blocks

MTS, Preference Test, and Rule-Following

Data analyses are based on the 10 participants that maintained criteria across all the experimental phases. In Phase 2 (MTS test) participants scored between 41 and 48 correct responses. Table 3 presents results for Phase 4 (Preference test). Nine out of 10 participants chose to ask the coherent speaker for a rule for responding in the first trial. Five participants did not ask the incoherent speaker for a rule in any trial during the preference test. Instructional control (actually following the rule) for the coherent speaker was 100% (or close to it) for all participants. Among the five participants that chose the incoherent speaker in some of the trials, instructional control was marked by variability from 0 to 100% for following the rule provided by the incoherent speaker. In general, therefore, speaker preference appeared to be the main influence over subsequent rule-following.

One variable that may influence rule-following behavior is the accuracy of the initial rule (Fox & Kyonka, 2017). Given that both speakers were accurate in Phase 4, it may be useful to consider performance on the first trial before receiving any feedback. It is interesting that 9 out of 10 participants chose the coherent speaker on the first trial. To further test the extent to which participant rule-following was determined by previously established coherence versus incoherence of the two speakers, the design of Phase 4 in Experiment 2 only delivered feedback after 10 trials of rule following.

Table 2 Demographic Data for Experiment 1

Participant	Age	Sex	Race	Educational Level	Profession
P1	28	Male	Black	Bachelor's degree	IT professional
P2	54	Male	Black	Bachelor's degree	IT professional
P3	36	Female	Caucasian	Bachelor's degree	IT professional
P4	48	Male	Caucasian	Bachelor's degree	IT professional
P5	35	Male	Caucasian	Master's degree	Nuclear Security Advisor
P6	37	Male	Caucasian	Bachelor's degree	Administrative Assistant
P7	32	Male	Caucasian	Master's degree	Engineer
P8	41	Female	Caucasian	Bachelor's degree	Psychologist
P9	43	Male	Caucasian	Doctorate degree	Researcher
P10	34	Male	Black	Bachelor's degree	Lawyer

Table 3 Tips Asked and Followed of Both Speakers in Experiment 1

Participant	First tip	Coherent			Incoherent		
		Qty tips asked	Qty tips followed	Following index	Qty tips asked	Qty tips followed	Following index
P1	Coherent	30	30	1,00	0	0	N/A
P2	Coherent	30	30	1,00	0	0	N/A
P3	Coherent	30	30	1,00	0	0	N/A
P4	Coherent	16	16	1,00	14	13	0,93
P5	Coherent	11	11	1,00	19	17	0,89
P6	Incoherent	8	8	1,00	22	22	1,00
P7	Coherent	29	29	1,00	1	0	0,00
P8	Coherent	30	30	1,00	0	0	N/A
P9	Coherent	21	21	1,00	9	7	0,78
P10	Coherent	30	29	0,97	0	0	N/A
	Total	235	234		65	59	

Rule-following index was calculated by dividing tips followed by tips asked

IRAP Data

Response latency was the main datum recorded for analysis for the IRAP data. This variable was defined as time in milliseconds between trial onset and the emission of a correct response. All latency data were processed by the D-IRAP algorithm built into the IRAP software itself. In short, average latency scores for each of the four trial-types within each test block pair (i.e., two successive blocks of history-consistent and history inconsistent trials) were averaged and divided by the standard deviation calculated across that block pair (see Barnes-Holmes et al., 2010, for more detail). This produced a D-IRAP score for each of the four trial-types.

Given that the current experiment was focused on the difference in "credibility" between coherent versus incoherent speakers, the two coherent speaker trial-types (coherent–positive, coherent–negative) were collapsed, as were the two incoherent speaker trial-types (incoherent–positive, incoherent–negative). This produced a single "credibility" score for each speaker. Table 4 lists the collapsed D-IRAP scores for each speaker per participant. First, the IRAP literature (e.g., Barnes-Holmes et al., 2010) predicts that relationally coherent responses tend to be emitted more quickly than relationally incoherent responses. Thus, positive scores would be expected for both speakers; that is, a positive score for the coherent speaker indicates responding consistent with coherent–credible–true and coherent–not credible–false; whereas a positive score for the incoherent speaker indicates responding consistent with incoherent–credible–false, incoherent–not credible–true. Table 4 shows that only three of the 20 scores were negative. Based on a one-sided exact binomial test, the observed values ($N=20$, $K=17$) provided significant support for participants' responding on the IRAP consistent with the coherent speaker as credible and the incoherent as not credible ($p=0.001$). In addition, given the expectation that the coherent speaker would be more positively valenced than the incoherent speaker, the D-IRAP score for the former should be larger than for the latter (e.g., see Barnes-Holmes & Harte, 2022b). The data reveal that 8 out of the 10 participants produced a score that was more positive for the coherent speaker than the incoherent speaker. Based on a one-sided exact binomial test,[1] the observed values ($N=10$, $K=8$) approached statistical significance ($p=0.056$).

Table 4 IRAP Credibility Test Data for Experiment 1

Participant	Coherent Speaker (Trial-types 1 and 2 collapsed)	Incoherent Speaker (Trial-types 3 and 4 collapsed)
P01	.89	.41
P02	.51	-.16
P03	.70	.15
P04	-.44	.75
P05	.30	-.03
P06	.77	.02
P07	.49	.43
P08	.94	.81
P09	.48	.05
P10	.49	.56

[1] The exact binomial test was chosen because it does not require previous knowledge about the data distribution, which is particularly useful for small sample sizes. In addition, it is helpful when considering dichotomous outcomes (e.g., positive versus negative).

Overall, Experiment 1 replicated the findings reported by Bianchi et al. (2021). In particular, the results showed that relational coherence appeared to influence rule-following and speaker preference. In addition, relational responding on the IRAP suggested participants related the coherent speaker as more credible than the incoherent speaker overall.[2] Upon reflection, the preference test was run with a CRF feedback contingency, which ensured that participants had immediate feedback for following the rule presented by the speaker. In more naturalistic settings, however, such feedback may not be so immediate given that people do not necessarily provide feedback following every response in a given interaction (e.g., Malott, 2003; Maraccini et al., 2016). For example, a recent meta-analysis by Sleiman et al. (2020), in the context of organizational environments, reported that weekly feedback was the most commonly employed feedback frequency. Experiment 2, therefore, was designed to explore the extent to which coherence may or may not continue to influence rule-following and speaker preference under intermittent reinforcement contingencies.

Experiment 2

In Experiment 2, similar procedures to those of Experiment 1 were employed, except that an altered schedule of reinforcement was implemented during Phase 4. In particular, a Fixed Ratio (FR) 10 was employed so that feedback for rule-following was delivered only after 10 correct responses.

Method

Participants

Fifteen verbally competent adults (males = 8) ranging in age between 21 to 65 years ($M = 47.8$ $SD = 13.22$) participated in Experiment 2. Recruitment and consent were as described in Experiment 1.

Setting, Equipment, and Stimuli

The materials involved in the setting, equipment and stimuli for Experiment 2 were as described in Experiment 1.

Procedure

Experiment 2 had the same sequence of phases as described for Experiment 1 (see Fig. 1 above). However, during Phase 4 (Preference Test), the points were delivered according to an FR10 schedule of reinforcement (instead of a CRF); that is, points were only added to the counter after every 10 correct responses. In addition, the instructions given to the participant before the task began were modified slightly to reflect that change:

> Ok, you advanced to the next stage! You will be presented with two images on the screen that you must choose, and only one of them will give you 10 points. You should ask for tips from one of the two helpers from the previous stage once each attempt. After asking for the tip, click on one of the images to select it. Important: The computer will record your correct and wrong answers on every trial, but you will only receive your points at every 10 correct answers! Try to accumulate as many points as possible! (press spacebar to continue).

Results and Discussion

From the 15 participants that took part in Experiment 2, 5 were excluded from data analyses for not fulfilling at least one of the following criteria: for not achieving at least 80% accuracy during the Stage 2 test (4 participants: E2, E3, E4, E5); for not maintaining the required accuracy criteria for at least 5 of the 6 IRAP test blocks (2 participants: E1 and E2); or for not maintaining the required accuracy or latency criteria for at least 5 of the 6 IRAP test blocks (1 participant: E1; see Table 5). A total of 10 participants met all the criteria for Experiment 2, with relevant demographic information presented in Table 6.

MTS, Preference Test and Rule-Following

Data analyses are based on the 10 participants that maintained criteria across all experimental phases. In Phase 2 (MTS test) participants scored from 39 to 48 correct responses during tests. Table 7 presents the results of Phase 4 (preference test and rule-following). Eight of the 10 participants chose the coherent speaker to ask for advice on the first trial. Four participants did not ask the incoherent speaker on any trial during the preference test. Instructional control (following the rule) for the coherent speaker was 100% for all participants. Among the four participants who chose the incoherent speaker in some of the trials, instructional control was marked by variability from 0 to 100% of rule following (see Table 8). Once again, therefore, speaker preference appeared to be the main influence over subsequent rule-following.

In addition, P16 and P20 chose and followed the rule of the incoherent speaker on the first trial, but preferred the

[2] It should be noted that the labels "coherent" and "incoherent" for the speakers are used here to indicate if their rules were coherent (or incoherent) with the participants' experimental learning history (i.e., Phases 1, 2, and 3).

Table 5 Inclusion/Exclusion scores and criteria for Experiment 2

Participant	Phase 2 Score	IRAP Test blocks with accuracy < 85%	IRAP Test blocks with latency above 2000 ms
E1	46	2	1
E2	38	2	
E3	13		
E4	29		
E5	13		
P1	48		
P2	41		
P3	48		
P4	48		
P5	44		
P6	48		
P7	48		
P8	48		
P9	47		
P10	39		

coherent speaker on subsequent trials (see Table 9). This can be interpreted as additional evidence that the participants behavior in Phase 4 was not solely determined, at least initially, by the accuracy of the rules, but instead by their history with the previous phases.

The IRAP data was processed as in Experiment 1, and trial-types were similarly collapsed to generate a single credibility score for each speaker. Table 10 lists the collapsed D-IRAP score for each speaker per participant. As in Experiment 1, in only 3 of the 20 cases were the scores negative. Based on a one-sided exact binomial test, the observed values ($N = 20$, $K = 17$) provided significant support for participants responding on the IRAP consistent with the coherent speaker as credible and the incoherent as not credible ($p = 0.001$). With respect to whether the coherent speaker score was more positive than the incoherent speaker score, however, visual inspection of data reveal some variability. That is, 5 out of the 10 participants produced a score that was more positive for the coherent speaker than the

Table 6 Demographics Data for Experiment 2

Participant	Age	Sex	Race	Educational Level	Profession
P11	34	Male	Caucasian	Master's degree	Archivist
P12	56	Male	Asian	Bachelor's degree	IT professional
P13	55	Male	Caucasian	Doctorate degree	IT professional
P14	39	Male	Caucasian	Master's degree	IT professional
P15	46	Female	Caucasian	Bachelor's degree	IT professional
P16	34	Male	Caucasian	Bachelor's degree	IT professional
P17	38	Female	Caucasian	Bachelor's degree	Speech Therapist
P18	65	Female	Caucasian	Bachelor's degree	Sales executive
P19	21	Female	Caucasian	Bachelor's degree	Pharmacist
P20	39	Male	Caucasian	Bachelor's degree	IT professional

Table 7 Tips Asked and Followed of Both Speakers in Experiment 2

		Coherent			Incoherent		
Participant	First tip	Qty tips asked	Qty tips followed	Following index	Qty tips asked	Qty tips followed	Following index
P11	Coherent	30	30	1	0	0	N/A
P12	Coherent	30	30	1	0	0	N/A
P13	Coherent	26	26	1	4	0	0,00
P14	Coherent	30	30	1	0	0	N/A
P15	Coherent	25	25	1	5	0	0,00
P16	Incoherent	16	16	1	14	14	1,00
P17	Coherent	30	30	1	0	0	N/A
P18	Coherent	25	25	1	5	0	0,00
P19	Coherent	15	15	1	15	15	1,00
P20	Incoherent	28	28	1	2	2	1,00
	Total	255	255		45	31	

Rule-following index was calculated by dividing tips followed by tips asked

Table 8 Trial by Trial behavior of P11 to P20 in Phase 4 of Experiment 2

Trial	P11		P12		P13		P14		P15		P16		P17		P18		P19		P20	
	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F
1	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	C	Y	I	Y
2	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y
3	C	Y	C	Y	C	Y	C	Y	I	N	C	Y	C	Y	C	Y	C	Y	C	Y
4	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y
5	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y
6	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y
7	C	Y	C	Y	I	N	C	Y	I	N	C	Y	C	Y	C	Y	C	Y	C	Y
8	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y
9	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y
10	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	I	Y	C	Y
11	C	Y	C	Y	C	Y	C	Y	I	N	C	Y	C	Y	C	Y	I	Y	C	Y
12	C	Y	C	Y	I	N	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y
13	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	I	Y	C	Y
14	C	Y	C	Y	I	N	C	Y	C	Y	I	Y	C	Y	C	Y	C	Y	C	Y
15	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	I	Y	C	Y
16	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	C	Y	C	Y
17	C	Y	C	Y	C	Y	C	Y	I	N	I	Y	C	Y	I	N	I	Y	C	Y
18	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	I	N	C	Y	C	Y
19	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	I	Y	C	Y
20	C	Y	C	Y	I	N	C	Y	C	Y	I	Y	C	Y	C	Y	C	Y	C	Y
21	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y
22	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	C	Y	C	Y
23	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	I	N	I	Y	C	Y
24	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y
25	C	Y	C	Y	C	Y	C	Y	I	N	C	Y	C	Y	C	Y	I	Y	C	Y
26	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	N	C	Y	C	Y
27	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	I	N	I	Y	C	Y
28	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	C	Y	C	Y
29	C	Y	C	Y	C	Y	C	Y	C	Y	I	Y	C	Y	C	Y	I	Y	C	Y
30	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y	C	Y

T stands for "Tip"; F stands for "Followed"; C stands for "Coherent"; I stands for "Incoherent"; Y stands for "Yes/Followed"; N stands for "No/Not followed." For the first 10 trials (i.e., before receiving any feedback) tips were asked from the coherent speaker 88% of the time and was always followed

incoherent speaker, whereas the other five participants produced a score that was more positive for the incoherent than coherent speaker. Based on a one-sided exact binomial test, the observed values ($N = 10$, $K = 5$) did not provide support for participants relating to the coherent speaker as more credible than the incoherent speaker, $p = 0.5$.

Overall, Experiment 2 replicated the findings reported by Bianchi et al. (2021), and Experiment 1 of the current study, with respect to speaker preference and rule-following. That is, the results showed that relational coherence again appeared to influence rule-following and speaker preference. It is interesting that although relational responding on the IRAP suggested participants again related the coherent speaker as credible and the incoherent speaker as not credible overall, the extent to which responding to the coherent speaker generated more positive scores than the incoherent speaker was more varied than in Experiment 1. In particular, the results of Experiment 2 indicated a 50/50 split between the two speakers' scores on this comparison.

General Discussion

The current study was a systematic replication and extension of Bianchi et al. (2021), which evaluated the influence of relational coherence upon speaker preference and subsequent rule-following behavior. Results showed that a history of relational coherence with each speaker had a reliable impact upon the preference test and rule-following

Table 9 Trial by Trial behavior of P16 and P20 in Phase 4 of Experiment 2

Trial	P16		P20	
	Tip	Followed	Tip	Followed
1	Incoherent	Yes	Incoherent	Yes
2	Coherent	Yes	Coherent	Yes
3	Coherent	Yes	Coherent	Yes
4	Coherent	Yes	Coherent	Yes
5	Coherent	Yes	Incoherent	Yes
6	Coherent	Yes	Coherent	Yes
7	Coherent	Yes	Coherent	Yes
8	Coherent	Yes	Coherent	Yes
9	Coherent	Yes	Coherent	Yes
10	Incoherent	Yes	Coherent	Yes
11	Coherent	Yes	Coherent	Yes
12	Coherent	Yes	Coherent	Yes
13	Incoherent	Yes	Coherent	Yes
14	Incoherent	Yes	Coherent	Yes
15	Incoherent	Yes	Coherent	Yes
16	Incoherent	Yes	Coherent	Yes
17	Incoherent	Yes	Coherent	Yes
18	Incoherent	Yes	Coherent	Yes
19	Incoherent	Yes	Coherent	Yes
20	Incoherent	Yes	Coherent	Yes
21	Coherent	Yes	Coherent	Yes
22	Incoherent	Yes	Coherent	Yes
23	Incoherent	Yes	Coherent	Yes
24	Coherent	Yes	Coherent	Yes
25	Coherent	Yes	Coherent	Yes
26	Coherent	Yes	Coherent	Yes
27	Coherent	Yes	Coherent	Yes
28	Incoherent	Yes	Coherent	Yes
29	Incoherent	Yes	Coherent	Yes
30	Coherent	Yes	Coherent	Yes

In accordance with the FR10 schedule, points were first delivered on the 10th trial

Table 10 IRAP Credibility Test Data for Experiment 2

Participant	Coherent Speaker (Trial-types 1 and 2 collapsed)	Incoherent Speaker (Trial-types 3 and 4 collapsed)
P11	.51	.53
P12	.84	.27
P13	.48	.19
P14	.23	.28
P15	.12	.98
P16	.28	.30
P17	.72	.07
P18	.62	-.35
P19	.56	-.06
P20	-.09	.41

behaviors of participants, irrespective of the reinforcement schedule in place (CRF in Experiment 1 and FR10 in Experiment 2). That is, as in Bianchi et al., participants tended, overall, to prefer the coherent over the incoherent speaker, and respond in line with the rule provided by that (coherent) speaker. This suggests that the extent to which the rule was coherent with previously established patterns of responding seemed to exert at least some control over preference of a speaker and following rules provided by that speaker.

On balance, this control was not complete across all participants. In particular, five participants in Experiment 1 and six in Experiment 2 also chose the incoherent speaker on some trials, and in some cases followed the rule provided by that speaker. In having done so of course, the incoherent properties of that speaker would have been undermined because points were rewarded for following that speaker's rule. The tendency to choose and follow the rules of the incoherent speaker thereafter could thus be readily explained in terms of this history of reinforcement. However, it remains unclear why some participants chose the incoherent speaker and then followed their rule in the first instance (i.e., when there was no history of reinforcement for doing so). Although the use of points as reinforcers, when they are exchangeable for nothing, can be considered a valid approach in experimental settings (see Galizio & Buskist, 1988; see also Fox & Kyonka, 2017, for a commentary concerning the ubiquitous use of points in video and smartphone games), it is possible that in the current study that the points were relatively weak as reinforcers and thus other variables may have come in to play. For example, it may be that an inverted experimenter effect may have emerged in which these participants deliberately engaged in behavior that they perceived to be the opposite of what the experimenter "wanted." As an alternative, perhaps participants simply engaged in some element of exploratory behavior to test their own verbal rules about the nature of the experiment itself (perhaps future studies could also collect verbal reports from participants as a way of monitoring their self-generated rules). Irrespective of what variables may have been at play here, it may be useful in future studies to seek to increase the reinforcing value of the points (e.g., by rendering them exchangeable for money or course credit) and/or introduce the loss of points for incorrect responding.

Unlike Bianchi et al. (2021), the current study gave participants an IRAP after the rule-following task that sought to evaluate responding to the "credibility" of each speaker using positive (e.g., reliable) and negative (e.g., unreliable) words. Overall, the number of IRAP effects were significantly in the expected direction (i.e., in both experiments

17 out of the 20 scores were positive) and in no case did any of the 20 participants produce two negative IRAP scores (i.e., one for each speaker). It is interesting that the specific reinforcement schedule in place during the speaker preference and rule-following test seemed to differentially affect the relationship between the two IRAP scores. As pointed out previously, we argued that the coherent speaker may be more positively valenced than the incoherent speaker, and thus the D-IRAP score for the former should be larger than for the latter. In particular, in Experiment 1 (CRF schedule), 8 out of 10 participants clearly showed this pattern; but in Experiment 2 (FR10 schedule), only half did so. Of course, the difference between these two groups is only three participants showing the effect (i.e., 8 in Experiment 1 versus 5 in Experiment 2) and thus it would be unwise to draw any strong conclusions from this preliminary finding. In a conceptual sense, however, this finding suggests that that the delivery of more frequent reinforcement may increase the differential valence properties of the speakers as indicated by the IRAP. However, this effect would need to be replicated and developed before extrapolating further based on the current data and any future work employing the IRAP should be tied specifically to recently proposed functional-analytic models of IRAP effects (see Barnes-Holmes & Harte, 2022b).

In considering how to take the current research forward, the HDML framework described in the Introduction may continue to be of assistance. For example, the current experiments comprised a preparation whereby stimulus pairings were presented in the training, testing, and establishment of coherent and incoherent speakers. Future work could examine the impact of replacing these stimulus pairings with derived stimulus relations of increasing complexity (i.e., the levels of the HDML). One such experiment, for instance, could involve training baseline relations in Phase 1 (e.g., AB and AC), testing for mutually and combinatorially entailed relations in Phase 2 (BA, CA, BC, and CB), followed by presenting these derived relations in Phase 3 to establish coherent and incoherent speakers. Would the same speaker preference and rule-following patterns be observed in Phase 4 given derived rather than directly trained relations? How would these patterns potentially change with increasing relational complexity (e.g., invoking more complex derived relational networks instead of combinatorially entailed derived relations) or manipulations of other dimensions (e.g., providing many versus few opportunities to derive the novel relations)?

Furthermore, it may be interesting to manipulate the extent to which coherence affects upon rule-following and speaker preference when the speakers themselves participate in derived relations with other speakers. For example, imagine an experiment in which speakers A1, B1, and C1 are established as equivalent to one another, as are speakers A2, B2, and C2. Speakers A1 and A2 could then be established as coherent and incoherent, respectively, in a manner similar to the current preparation. Would presenting speakers C1 and C2 in the preference and rule-following tests produce similar results to those found currently even though these speakers had never been directly established as coherent and incoherent? Such explorations may contribute toward developing a more complete model of rule-following in the natural environment and may speak more directly to socially relevant areas such as social prejudice. For example, individuals may be more likely to follow the advice or rules provided by a stranger (C1) if that stranger belongs to the individual's ingroup (A1 and B1) rather than their outgroup (A2 and B2).

Given that the current work was largely a systematic replication of the study by Bianchi et al. (2021), the same limitations associated with that experiment still apply here and should be considered in any future work. In particular, Bianchi et al. suggested that the order in which speakers were established as coherent and incoherent in Phase 3 may have presented a confound. In this Phase, the coherent speaker was always presented to participants first and thus the resulting patterns of preference and rule-following may have been subject to an order effect. That is, it is possible that such a procedure may have biased participants in favor of the coherent speaker. Would the same results be found if this had been counterbalanced across participants? In addition, given the number of participants who failed to meet criteria in during relational testing (Phase 2) following the stimulus pairing procedure, future studies might incorporate interspersed probes during training (e.g., see Delgado & Rodriguez, 2022).

In closing, it should be emphasized once again that the current preliminary work should be seen as part of a broader attempt to build an experimental model of rule-following as symbolic relations, using relatively precise technical terms, guided by the HDML framework (see Harte & Barnes-Holmes, 2022, for an overview of other relevant work from this general research agenda). The current experiment involved exploring the extent to which relational coherence affected upon preference for speakers who provide "rules" that differ in the extent to which they cohere with previously established patterns of relational responding. As suggested above, further exploration of this line of research could have interesting real-world implications in terms of ingroup/outgroup effects and social prejudice, for example. In addition, it seems important to continue to develop multiple experimental paradigms for assessing coherence and rule-following different from those already in the literature (e.g., Harte et al., 2020a, 2020b). Doing so will be particularly important in developing a more complete understanding of the impact of relational/verbal coherence in various contexts, thus working towards the behavior-analytic goals of prediction-and-influence, with precision, scope and depth. The current study provides a successful preliminary step in this regard.

Author Contributions All authors contributed to the study conception and design. Material preparation and data collection were performed by Paulo H. Bianchi and William F. Perez. Data analyses were performed by all authors. All authors elaborated the first draft and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The present research was supported by the São Paulo Research Foundation (FAPESP; Grant numbers 2019/24210-0, 2017/10037-0, 2016/05935-6, and 2008/57705-8) and by the Brazilian National Research Council (CNPq, Grant numbers 573972/2008-7 and 465686/2014-1); the present study is also part of the research program of the Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE).

Data Availability All data supporting the findings of this study are available within the article and its tables. Any additional data can be obtained by directly contacting the corresponding author.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The research is approved by the Brazilian platform for ethical committees (Plataforma Brasil, CAAE 19827719.0.0000.5493).

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

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

References

Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEnteggart, C. (2017). From IRAP and REC model to a multi-dimensional multi-level framework for analyzing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science, 6*(4), 473–483. https://doi.org/10.1016/j.jcbs.2017.08.001

Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2020). Updating RFT (more field than frame) and its implications for process-based therapy. *The Psychological Record, 70*, 605–624. https://doi.org/10.1007/s40732-019-00372-3

Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., & Harte, C. (2021). Back to the future with an up-dated version of RFT: More field than frame? *Perspectivas em Análise do Comportamento, 12*(1), 33–51. https://doi.org/10.18761/PAC.2021.v12.RFT.03

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *The Psychological Record, 60*, 527–542. https://doi.org/10.1007/BF03395726

Barnes-Holmes, D., & Harte, C. (2022a). Relational frame theory 20 years on: The Odysseus voyage and beyond. *Journal of the Experimental Analysis of Behaviour, 117*(2), 240–266. https://doi.org/10.1002/jeab.733

Barnes-Holmes, D., & Harte, C. (2022b). The IRAP as a measure of implicit cognition: A case of Frankenstein's monster. *Perspectives on Behaviour Science, 45*, 559–578. https://doi.org/10.1007/s40614-022-00352-z

Baron, A., & Galizio, M. (1983). Instructional control of human operant behavior. *The Psychological Record, 33*(4), 495–520.

Bern, R., Persdotter, T., Harte, C., & Barnes-Holmes, D. (2020). Relational coherence and persistent rule-following: The impact of targeting coherence in a "non-critical" component of a relational network. *The Psychological Record, 71*(2), 279–290. https://doi.org/10.1007/s40732-020-00414-1

Barnes-Holmes, D., Hayden, E., Barnes-Holmes, Y., & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study. *The Psychological Record, 58*, 497–516. https://doi.org/10.1007/BF03395634

Bianchi, P. H, Perez, W. F, Harte, C., & Barnes-Holmes, D. (2021). Effects of coherence and speaker preference on rule-following. *Perspectivas em Análise do Comportamento, 12*(1), 214–227. https://doi.org/10.18761/PAC.2021.v12.RFT.07

DeGrandpre, R. J., & Buskist, W. F. (1991). Effects of accuracy of instructions on human behavior: Correspondence with reinforcement contingencies matters. *The Psychological Record, 41*(3), 371–384. https://doi.org/10.1007/BF03395119

Delgado, D., & Rodríguez, A. (2022). Stimulus equivalence using a respondent matching-to-sample procedure with verification trials. *The Psychological Record, 72*, 1–9. https://doi.org/10.1007/s40732-020-00438-7

Dougher, M. J., Hamilton, D., Fink, B., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior, 88*, 179–197. https://doi.org/10.1901/jeab.2007.45-05

Fox, A. E., & Kyonka, E. G. (2017). Searching for the variables that control human rule-governed "insensitivity." *Journal of the Experimental Analysis of Behavior, 108*(2), 236–254.

Galizio, M. (1979). Contingency-shaped and rule-governed behavior: Instructional control of human loss avoidance. *Journal of the Experimental Analysis of Behavior, 31*(1), 53–70. 10. 1901/jeab.1979.31–53

Galizio, M., & Buskist, W. (1988). Laboratory lore and research practices in the experimental analysis of human behavior: Selecting reinforcers and arranging contingencies. *The Behavior Analyst, 11*(1), 65.

Harte, C., & Barnes-Holmes, D. (2021). The status of rule-governed behavior as pliance, tracking and augmenting within relational frame theory: Middle-level rather than technical terms. *The Psychological Record, 72*, 145–158. https://doi.org/10.1007/s40732-021-00458-x

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Kissi, A. (2020a). The study of rule-governed behavior and derived stimulus relations: Bridging the gap. *Perspectives on Behavior Science, 43*, 361–385. https://doi.org/10.1007/s40614-020-00256-w

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2021). Exploring the impact of coherence (through the presence versus absence of feedback) and levels of derivation on persistent rule-following. *Learning & Behavior, 49*, 222–239. https://doi.org/10.3758/s13420-020-00438-1

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., Gys, J., & Hassler, C. (2020b). Exploring the potential impact of relational coherence on persistent rule-following: The first study. *Learning & Behavior, 48*, 373–391. https://doi.org/10.3758/s13420-019-00399-0

Hayes, S. C. (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control*. Plenum.

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

Hayes, S. C., Brownstein, A. J., Haas, J. R., & Greenway, D. E. (1986a). Instructions, multiple schedules, and extinction: Distinguishing rule-governed from schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior, 46*, 137–147. https://doi.org/10.1901/jeab.1986.46-137

Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986b). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior, 45*(3), 237–256.

Hayes, S. C., Rosenfarb, I., Wulfert, E., Munt, E. D., Korn, Z., & Zettle, R. D. (1985). Self-reinforcement effects: An artifact of social standard setting? *Journal of Applied Behavior Analysis, 18*(3), 201–214. https://doi.org/10.1901/jaba.1985.18-201

Hussey, I. (2023). A systematic review of null hypothesis significance testing, sample sizes, and statistical power in research using the implicit relational assessment procedure. *Journal of Contextual Behavioral Science, 29*, 86–97. https://doi.org/10.1016/j.jcbs.2023.06.008

Knuth, D. E. (2014). *The art of computer programming: seminumerical algorithms*, Vol. 2. Addison-Wesley Professional.

Leader, G., & Barnes-Holmes, D. (2001). Matching-to-sample and respondent-type training as methods for producing equivalence relations: Isolating the critical variable. *The Psychological Record, 51*, 429–444. https://doi.org/10.1007/BF03395407.

Malott, R. W. (2003). What OBM needs is more Jewish mothers. *Journal of Organizational Behavior Management, 22*(2), 71–87. https://doi.org/10.1300/J075v22n02_07

Maraccini, A. M., Houmanfar, R. A., & Szarko, A. J. (2016). Motivation and complex verbal phenomena: Implications for organizational research and practice. *Journal of Organizational Behavior Management, 36*(4), 282–300. https://doi.org/10.1080/01608061.2016.1211062

Martinez, H., & Tamayo, R. (2005). Interactions of contingencies, instructional accuracy, and instructional history in conditional discrimination. *The Psychological Record, 55*, 633–646.

Martinez-Sanchez, H., & Ribes-Inesta, E. (1996). Interactions of contingencies and instructional history on conditional discrimination. *The Psychological Record, 46*(2), 301–307.

Papoulis, A. (1965). *Random variables and stochastic processes*. McGraw-Hill.

Paracampo, C. C. P., & de Albuquerque, L. C. (2005). Comportamento controlado por regras: revisão crítica de proposições conceituais e resultados experimentais. *Interação em Psicologia*, *9*(2),

Perez, W. F., dos Reis, M. D. J. D., & de Souza, D. D. G. (2010). Efeitos de história experimental com diferentes instruções e do controle por contingências sobre o seguimento de instruções. *Acta Comportamentalia, 18*(1), 55–85.

Perez, W. F., Endemann, P., Pessôa, C. V., & Tomanari, G. Y. (2015). Assessing stimulus control in a discrimination task with compound stimuli: Evaluating testing procedures and tracking eye fixations. *The Psychological Record, 65*, 83–88. https://doi.org/10.1007/s40732-014-0092-1 .

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. (1966). An operant analysis of problem solving. In B. Keinmuntz (Ed.), *Problem-solving: Research, method, and therapy* (pp. 225–257). Wiley.

Sleiman, A. A., Sigurjonsdottir, S., Elnes, A., Gage, N. A., & Gravina, N. E. (2020). A quantitative review of performance feedback in organizational settings (1998–2018). *Journal of Organizational Behavior Management, 40*(3–4), 303–332. https://doi.org/10.1080/01608061.2019.1632242