Effects of Coherence on Speaker Preference and Rule-Following

Paulo H. Bianchi[1, 2], William F. Perez[1, 3], Colin Harte[4] and Dermot Barnes-Holmes[5]

1. Paradigma – Centro de Ciências e Tecnologia do Comportamento, Brazil
2. IPEN - Instituto de Pesquisas Energéticas e Nucleares, Brazil
3. Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE), Brazil
4. National College of Ireland, Dublin, Ireland
5. Ulster University, Northern Ireland, UK

Author Note: This manuscript is part of the PHB's Master Thesis. Preparation of the manuscript was supported by the Brazilian National Research Council (CNPq, Grants 573972/2008-7,465686/2014-1) and the São Paulo Research Foundation (FAPESP, Grant 08/57705-8), both for Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE), chaired by Deisy G. de Souza (UFSCar).

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

Abstract

Rule-following is affected by multiple variables. A relevant aspect of rules regards whether they "make sense", that is, the extent to which the instruction coheres with previously reinforced patterns of relational responding. The present study aimed to evaluate the influence of relational coherence upon rule-following. After mastering a particular set of conditional relations (e.g., A1B1, A2B2), the participants were exposed to two speakers, one of which would “state” relations that cohered (e.g., A1B1, A2B2) with the participant’s previous relational training and the other that would present relations that were incoherent (e.g., A1B2, A2B1). Then, rule-following was measured in a preference test in which the participant would have to choose which of the two speakers would provide instructions in each test trial. Results show that the participants preferred the coherent speaker to provide instructions and followed the rules presented by that speaker throughout the test. Coherence is discussed as a critical aspect of rule following and preference for particular narratives.

Key-words: relational frame theory, rule-governed behavior, instructional control, speaker preference, coherence.

In the last number of years, a programme of research (summaries of which can be found in Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020, and Harte & Barnes-Holmes, 2021) has sought to bridge the gap between two disparate areas of research within the behavior-analytic literature: rule-governed behavior and derived stimulus relations. Rule-governed behavior was first described by B.F. Skinner (1966) in the context of problem solving. Rules were defined as contingency specifying stimuli that allowed a listener to solve problems without having to directly contact contingencies in the environment. A wealth of research in the decades that followed sought to explore the impact of rules (or instructions) on human performances on schedules of reinforcement. One key finding that emerged from this work was that instructed behavior often led to varying degrees of 'insensitivity' to the scheduled contingencies, at least for verbally-able humans (e.g., Hayes et al., 1986; Shimoff et al., 1981). That is, verbal humans would often produce patterns of responding that did *not* reflect changes in the scheduled reinforcement contingencies. For example, when instructed how to earn reinforcers on a schedule of reinforcement, human participants tended to adapt less readily to un-cued changes in contingencies than participants who were not initially instructed (see Hayes, 1989, for an early book-length review).

The second area of research, derived stimulus relations, first emerged with the work of Murray Sidman and colleagues (e.g., Sidman, 1971; Sidman & Tailby, 1982), the basic phenomenon of which came to be known as stimulus equivalence (see Sidman, 1994 for a book length treatment). The key finding was that after training a small number of matching responses (e.g., A choose B; and A choose C), unreinforced or untrained matching responses often emerged spontaneously (e.g., B choose C; and C choose B). When such novel matching responses emerged, the three stimuli (A, B, & C) were said to be participating in a derived equivalence relation. Furthermore, other untrained responses also often emerged when a particular function was trained to a stimulus participating in this equivalence relation. For example, if stimuli A, B and C participated in an equivalence relation, and stimulus A was paired with an aversive stimulus (e.g., presentation of mild electric shock), then stimulus C may also acquire aversive functions, all in the absence of direct reinforcement. This latter effect has often been referred to as a derived transfer of stimulus functions. While derived relational responding, including transfer of functions, appears to occur with relative ease in verbally-able humans, it has not been readily observed in nonhuman animals (e.g. Sidman et al., 1982; Dugdale and Lowe, 2000).

A link between the research on rule-governed behavior and derived equivalence relations was made initially when it was argued that rules may control behavior because the words contained within the rule participate in equivalence relations. Thus, the instruction or rule "When the light turns green, then go" controls appropriate behavior because the word "green" is in an equivalence relation with the actual color green. As a result, rules or instructions may come to control behavior in the absence of direct reinforcement because equivalence relations themselves do not require direct reinforcement for all of the defining relations. In fact, Sidman (1994) suggested that when we say that rules "specify" or "refer" to contingencies, these terms (i.e., specify and refer) simply indicate that the events "specified" in the rule participate in equivalence relations with the words in the rule.

Despite considerable conceptual overlap between the study of rule-governed behavior and derived stimulus relations (e.g., Hayes, 1989; Hayes et al., 2001; Sidman, 1994), research has only recently sought to integrate these areas empirically (see Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020 for a recent review). The basic approach has involved providing participants with a rule that contains some level of novel, within-experiment derivation and exploring the extent to which participants persist with rule-following in the face of reversed reinforcement contingencies. For example, in a number of studies (e.g., Harte et al., 2018; Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart, Gys et al., 2020), participants were first trained that the phrase 'least similar' was equivalent to a nonsense symbol 'XXX', before being trained that 'XXX' was equivalent to a nonsense word 'Beda'. Thus, the equivalence relation between the phrase 'least similar' and the nonsense word 'Beda' would emerge. 'Beda' was then inserted into a rule for responding on a subsequent matching-to-sample (MTS) task (e.g., to earn points "choose the image that is 'beda' to the sample image"). Each MTS trial involved presenting participants with a sample shape at the top of the screen with three comparison shapes along the bottom of the screen. Each comparison shape varied to degree in terms of their similarity to the sample shape. That is, one shape was clearly very similar to the sample shape, one shape was quite similar to the sample but with more variations, and one shape was clearly completely different to the sample with little or no overlapping features. Within the MTS task, responding in accordance with the partially derived rule was first reinforced but subsequently punished following a contingency reversal. Specifically, for the first 100 MTS trials, participants received one point per trial upon choosing the 'beda' (i.e., least similar) comparison, and lost one point for choosing either of the other two comparisons. Participants were required to get a minimum of eight out of the first 10 trials correct, and a minimum of 80 out of the first 100 trials correct to ensure that they were responding in accordance with the derived rule rather than simply learning how to respond in accordance with the task contingencies independent of the rule. On the 101st trial, the MTS task contingencies reversed for a further 50 trials, uncued to participants, such that points were now awarded for choosing the most similar comparison, and lost for choosing either of the other two options. Points visibly accrued on screen to participants throughout the task. The main focus of this research was to determine the extent to which participants would persist in following a rule (choosing the comparison that was least like the sample) when it contained a relation that had been previously derived in the experiment. The critical test involved determining the level of persistence in rule-following when the contingencies changed and the rule, therefore, ceased to specify the contingencies.

In a number of recent studies that have explored the impact of deriving part of a rule on persistent rule-following, the role of relational coherence has been manipulated. Coherence, in this context, is used to refer to the extent to which a particular pattern of derived relational responding overlaps functionally with a specific previously observed pattern of such responding[1]. One approach to exploring the impact of coherence may involve providing reinforcement in one condition, versus no reinforcement in another condition, for producing a coherent pattern of responding. Two recent studies adopted this strategy (Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart, Gys et al., 2020; Harte, Barnes-Holmes, Barnes-Holmes, & McEnteggart, 2020). Specifically, the same paradigm described above was used to train novel derived relations within the experiment (Least similar = XXX = Beda), and then manipulate coherence through the presence versus absence of performance feedback for deriving the relations between the nonsense word 'beda' and key phrase 'least similar'. Next, the novel relation was inserted into the rule required for responding on a subsequent MTS procedure. For the first 100 trials of the MTS procedure, the scheduled contingencies supported the derived rule (i.e., participants gained points for responding in accordance with the rule). On the 101st trial, however, these contingencies reversed so that the scheduled contingencies were now in opposition with the derived rule (i.e., participants lost points for responding in accordance with the rule). The general finding was that participants showed greater persistence in rule following when the derived part of the rule

[1] The reader should note that we are using the term coherence here in a relatively technical way, in that it is restricted to the functional overlap (or lack thereof) between patterns of derived relational responding. For example, the simple statement, 'if A is bigger the B, then B is bigger than A' would typically be seen as lacking in coherence with the way in which the verbal community employs the term 'bigger than' (i.e., in most contexts, the coherent derivation would be that 'B is smaller than A'). Coherence has sometimes been used in behavior analysis in a similar but broader and perhaps less precise way when discussing "sense making" (e.g., Wray et al., 2017). In this case, lack of coherence in a derived relational response might be described as "not making sense." Following on from the previous example, a listener might accuse a speaker of not making sense if "B is bigger than A" was derived from "A is bigger than B."

had been reinforced with appropriate feedback. Or more informally, participants persisted with rule-following to a greater extent when they were informed that they had previously derived the “correct” relation between ‘beda’ and ‘least like’. As an aside, this effect was moderated by at least one other variable, but that finding is not directly relevant to the current research and thus will not be discussed here.

While the foregoing studies explored the impact of coherence through providing reinforcement versus no reinforcement for producing a coherent pattern of responding, another study assessed coherence through providing reinforcement versus punishment for coherent responding (Bern et al., 2020). In this sense, coherence was *undermined* in one experimental condition, as opposed to simply involving the absence of reinforcement. Additionally, Bern and colleagues sought to explore the impact of undermining a ‘non-critical’ component of a derived relational network as opposed to a critical component as in the studies described above. As with the research described previously, the experiment involved first training participants on novel relations within the experiment before inserting one of these relations into a rule for responding on a contingency-switching MTS task. Specifically, participants were trained on a six member relational network (i.e., A=B=C=D=E=F). In one condition, the researchers reinforced the derived F=D relational response, making it maximally coherent, while in a second condition the researchers introduced an element of incoherence by punishing the derived F=D relational response. Critically, this part of the network (i.e., D=E=F) was not necessary for deriving the rule required for responding on the MTS task, which was restricted to the A=B=C part of the network. Thus, the experiment involved undermining an element of a novel relational network that was not necessarily critical for derived rule following.

Results showed that undermining a non-critical part of the network significantly impacted upon persistence in rule-following. Specifically, participants in the condition in which coherence of a non-critical part of the network was undermined persisted with rule-following for significantly more trials following the contingency reversal than the maximal coherence group. Interestingly, this result appears to contradict the earlier finding that increased coherence (in the derived rule) produced increased persistence in rule-following. On balance, the researchers suggested that the use of punishment in this latter study (rather than simply the absence of reinforcement) may have undermined the coherence functions of the feedback itself. In other words, when computer-generated feedback was used to undermine the coherence of part of a derived relational network, the behavior-controlling properties of the feedback were reduced in the MTS task. Consequently, the feedback itself was deemed unreliable and participants were more likely to ignore the feedback when the contingencies switched in the MTS task (i.e., they persisted with following the rule).

The studies described above all employed group designs, but the most recent study in this line of research has begun to explore coherence and persistent derived rule-following using single-case experimental designs (Harte et al., 2021). Specifically, this recent research began to explore the impact of flexibility in reversing derived relations on persistent rule-following; the fact that the study involved reversing previously derived relations also made it relevant to the issue of coherence (defined as functional overlap in distinct patterns of relational responding).

In the first of three experiments, researchers first sought to assess flexibility in reversing derived relations. Specifically, participants were initially trained on a relational network comprising two, three member equivalence relations (i.e., A1=B1=C1 and A2=B2=C2), before testing their ability to derive A1=C1 and A2=C2. Next, participants were trained and tested in a similar network but that now involved reversing the B and C relations (i.e., train A1=B1=C2 and A2=B2=C1; test for A1=C2 and A2=C1). Three participants were required to complete these training and testing reversal sequences three times, and in each case they successfully produced test performances in accordance with the most recently trained relations, thus demonstrating highly flexible relational responding.

In a second experiment, 3 additional participants were trained and tested on the same derived relational networks as above. In this experiment, however, each training and testing sequence was followed by an MTS task in which a rule was presented that employed a derived relation that had been trained and tested immediately before the MTS task (i.e., either the derived A1=C1 or A1=C2 relation was inserted into the rule for responding). In the first instance (i.e., the A1=C1 relation was inserted into the rule), the MTS task contingencies cohered with the derived rule. Following training and testing of the reversed network, however (i.e., the A1=C2 relation was inserted into the rule), the MTS feedback contingencies were now in opposition to the derived rule, and thus responding in accordance with the previously trained and tested pattern of relational responding was punished (i.e., undermining coherence between the rule and the MTS contingencies). The main aim was to assess the impact of flexibility in deriving a relational network on persistent rule-following. Results showed that all 3 participants generally responded in accordance with the MTS-feedback contingencies when the derived rule no longer cohered with the MTS feedback (immediately after deriving A1=C2 and A2=C1). More informally, participants readily reversed derived relations, but then ignored those relations when they were included in a rule that did not cohere with a current MTS task.

A final experiment partially replicated the foregoing procedure with another 3 participants, but in this case the MTS task also involved a reversal in task contingencies. Specifically, following the training and testing of the relational networks as in the previous two experiments, the derived rule and MTS contingencies cohered throughout the first task (i.e., when the rule contained the A1=C1 and A2=C2 relations). However, when the derived relations reversed (i.e., A1=C2 and A2=C1), the MTS contingencies for the MTS task also reversed and thus cohered with the derived rule; after 15 MTS trials the feedback contingencies reversed and thus coherence between the rule and the MTS feedback was undermined. In this experiment, all 3 participants produced evidence of persistent rule-following (i.e., they continued to follow the rule, at least initially after the MTS feedback contingencies were reversed). This experiment suggested that increasing coherence between the derived rules and the MTS task, by reversing the feedback contingencies for both, the behavior-controlling properties of the derived rules increased. Or more informally, participants were less likely to "ignore" the rules if they had experienced reversals in deriving the relations in those rules *and* in the MTS feedback contingencies.

The studies outlined above suggest that relational coherence within a rule, and coherence between a rule and the feedback contingencies for following the rule, may impact upon the extent to which participants show persistent rule-following (in the face of reversed MTS feedback contingencies). In general, it appears that reducing coherence in some way (for the rule, for the feedback, or for the relationship between the rule and the MTS feedback) reduced behavioral control (either for the rule or for the MTS feedback). In pursuing the potential impact of coherence on rule-following it seems important to explore a range of different methods for assessing its impact beyond rule-persistence *per se.* For example, one potential approach could involve exploring the extent to which relational coherence impacts upon a choice or preference for following one instruction or rule over another, even when both instructions yield the same levels of reinforcement when they are followed.

In line with this general strategy, the current study sought to explore the extent to which manipulating coherence would impact upon the extent to which a listener would follow the advice of a speaker and would show a preference for one speaker over another. Specifically, after mastering a particular set of conditional relations (e.g., A1B1, A2B2), the participants were exposed to two speakers, one of which would "state" relations that cohered (e.g., A1B1, A2B2) with the participant's previous relational training and the other that would present relations that were incoherent (e.g., A1B2, A2B1). Then, rule-following was measured in a preference test in which the participant would have to choose which of the two speakers would provide instructions for a different task. Although both speakers would provide accurate information in how to complete this second task, based on the studies described previously, a differential preference may be observed between the two speakers. Specifically, participants may prefer a speaker who possesses increased coherence functions (i.e., increased behavioral control properties, assuming that such control is of benefit to the listener).

Method

Participants

Four verbally competent adults (Male = 3, Female = 1) ranging in age from 28 to 54 years (*M* = 42.50, *SD* = 12.58) participated. Participants were recruited through personal contact with the first author (sample of convenience) and none had previous experience with similar psychology experiments. Before the experiment began, 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 experimental procedures, they were fully debriefed and thanked. Participants did not receive any compensation for participation.

Equipment and Setting

The experiment took place in a quiet room with a table, chair, and notebook computer. The custom-written software "Preferência Entre Falantes CRF" presented the rule-following task. Two pictorial representations of "speakers" were presented throughout the phases (see Figure 1). Stimuli from Phases 1-3 were abstract black shapes presented on a white background. Stimuli from the preference test were abstract colored shapes.

Procedure

Figure 1 presents an outline of the experimental phases divided into 4 stages: (1) Relational training, (2) Relational testing, (3) Establishing coherent and incoherent speakers, and (4) Preference test.

Phase 1: Relational training. A respondent-type training procedure (Leader & Barnes-Holmes, 2001) aimed to teach four arbitrary relations: A1B1, A2B2, B1C1 and B2C2. Before starting, the participants read the following instruction: "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 training trial comprised the successive presentation of a given pair of stimuli arbitrarily designated to relate to each other (e.g., A1B1). Each trial onset started with the presentation of the first stimulus of the pair (e.g., A1) in the center of the screen for 2 s followed by a 1 s interval in which no stimulus was presented. Once the interval ended the second stimulus of the pair was presented for 2 s followed by a 3 s intertrial interval with no stimulus on the screen. Training trials were presented in blocks of four, such that each block comprised the random presentation of the four stimulus pairs (A1B1; A2B2; B1C1; B2C2). Each block was presented 13 times, thus involving a total of 52 trials.

Phase 2: Relational testing. Immediately following relational training, participants were exposed to a matching-to-sample (MTS) task that sought to test for relational responding based on the (respondent-type) relational training phase (e.g., given A1 as a sample stimulus, choose B1 rather than B2 as a comparison stimulus, and so on: A2B2; B1C1; B2C2). Before starting, the 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 presented a sample stimulus on the top of the screen. Following a 1-s interval, three comparison stimuli appeared at the bottom, aligned horizontally presented in random order across trials (see below). The first stimulus of each pair presented in Phase 1 were always presented as sample stimuli (e.g., A1). The second stimulus of each pair was always presented as one of the comparison stimuli (e.g., B1), with the second stimulus from the other pair (e.g., B2), and a third novel stimuli (e.g., N1 or N2). The third comparison stimulus was presented to control for rejection responses (see Sidman, 1982; Perez, Tomanari, & Vaidya, 2015). Selecting the comparison (e.g., B1) stimulus that was paired with the sample (i.e., A1) was considered a correct response, whereas selecting either of the two other comparisons was registered as an error. The position of the three stimuli, including the correct one, varied based on an analysis of all possible combinations for 3 symbols, in such way that for each relation six possibilities were presented in random order, using a combination of the Fisher–Yates shuffle algorithm with the subtractive random number generator algorithm (Knuth, 2014). No differential feedback was provided for participants' responses. Thus, the comparison selection was followed by removal of all four stimuli from the screen, with a 0.5 s intertrial interval, during which the screen remained blank, followed by onset of the next trial (i.e., presentation of a sample stimulus). Each pair (A1B1, A2B2, B1C1, and B2C2) was presented 12 times in a quasi-random sequence, comprising a 48-trial

MTS test block. To proceed to the next phase, participants had to produce a minimum of 80% correct responses (i.e., 39 correct test trials).

Phase 3: Establishing coherent and incoherent speakers. The procedure implemented in this phase was similar to Phase 1, except that the stimulus pairings on each trial were presented inside a speech balloon next to one of two speakers (see Figure 1). The speakers were two characters differentiated by the color of their t-shirts: green or purple. One of the speakers 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, however, presented pairs that were *incoherent* with the previously established relations (i.e, A1B2, A2B1, B1C2 and B2C1). The t-shirt color assigned for the coherent and incoherent speakers alternated between participants.

Phase 3 started with the presentation of 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)".

Each training trial comprised of the successive presentation of a given stimulus pair. Each pair was graphically displayed inside a speech balloon spoken by one of the speakers. For one of the speakers the stimulus pairs were always coherent with the trained and tested relations from Phases 1 and 2 (e.g., A1B1), but for the other speaker the stimulus pairs did not cohere with the previous training and testing (e.g., A1B2). The speaker and the speech balloon remained on the screen until the end of the trial. 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 in which the speech balloon was empty; after that, the second stimulus of the pair was displayed for 2 s followed by a 0.5 s interval in which the speech balloon was empty. Next, all stimuli were withdrawn from the screen during a 2.5 s intertrial interval. No action was required from participants to advance to the next trial (i.e. they were expected only to observe the screen). Training trials were presented in eight-trial blocks. Each block presented all coherent (A1B1, A2B2, B1C1, B2C2) and incoherent pairs (A1B2, A2B1, B1C2, B2C1), once each per block in a quasi-random order. 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. This phase started with the presentation of 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 in 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)”.

Each test trial simultaneously presented the following elements on the screen: in the top right-hand side was a counter accumulating points; on the left-hand side the two speakers appeared with different t-shirts (i.e., purple and green), placed one above the other (the position of the green and purple speakers alternated across trials); on the centre-right of the screen, two abstract colored images were displayed side-by-side. These novel colored stimuli were selected from a 60-stimulus pool. The position of the correct stimulus was randomly assigned using the subtractive random number generator algorithm (Knuth, 2014) to generate an integer number between 0 and 1, and assigning the correct stimulus to the left if the result was 0 and to the right if it was 1.

Clicking on one of the speakers immediately displayed a “hint” (a rule) inside a speech balloon located in the center of the screen to the right of that character. No image could be chosen before clicking on one of the speakers. If the participant tried to select one of the images without first selecting one of the speakers, the following warning message appeared: “You must request a hint before choosing an image!” along with an OK button to return to the previous screen and proceed with that trial. Once the participant clicked on one of the speakers, the rule inside the speech balloon was available until the end of the trial. Only one speaker could be selected per trial. Thus, clicking on the second speaker after having selected one of them produced no programed consequence.

During the preference test, the rule stated by both speakers inside the speech balloon was always consistent with the programed contingency: “Click on [small version of the correct image for that trial] to earn 10 points.” Making the rules/hints produced by both speakers consistent with the task contingencies in Phase 4 allowed for an assessment of the extent to which *a history* of “speaking” in a manner that was coherent or incoherent with the Phase 1 training, and performance in Phase 2 testing, impacted upon speaker preference. After having selected one of the speakers, image selection was enabled. The message “+10 points” followed correct responses, while "No points earned" followed incorrect responses. The feedback message was displayed on-screen for 1 s. Correct responses were always in accordance with the rule provided inside the speech balloon. The delivery of consequences initiated a 1 s intertrial interval. The preference test comprised a total of 30 trials.

Results

During Phase 2, all participants scored from 43-48 on the MTS test, indicating that the respondent-type training presented in Phase 1 had established the predicted relational responding. Table 1 presents the results from the preference test (Phase 4). All four participants selected the coherent speaker in the first test trial. Three participants (P1, P2 and P3) always selected the coherent speaker, and followed the rule provided by that speaker, throughout the 30 test trials (i.e., a "speaker coherence preference" index of 1.00). As presented on Figure 2, P4 selected the coherent speaker on each of the first 9 trials of the test and followed the rule. The incoherent speaker was selected on trials 10 to 16, and on trials 18, 22, 23, 25, 26, 29, and 30. For each of these trials, except for trial 26, the participant followed the rule provided by the speaker (thus obtaining 10 points on each of these trials). On trial 26, however, the participant did not follow the speaker's rule and thus failed to obtain any points for that trial (yielding a rule-following idex of .93).

Discussion

The current study sought to extend research exploring the impact of relational coherence on rule-following by investigating the extent to which manipulating this variable would influence whether a listener would follow the advice of a speaker and show a preference for one speaker over another. The results showed that all four participants initially showed a differential preference for the speaker who provided information coherent with previous relational training. In addition, three out of four participants continued to show an exclusive preference for that speaker, and followed the rule provided by the speaker, for the entirety of the task. Participant 4 demonstrated a more variable performance when compared to the other three participants, although responding on the first 9 trials indicated a preference for the coherent speaker. Thereafter, the participant alternated their preference response across the two speakers and on all but one trial followed the hint/rule provided by the speaker. Overall, therefore, the extent to which the rule was coherent (i.e., consistent) with previously established patterns of relational responding appeared to control both speaker preference and rule-following (at least initially).

At the current time it remains unclear why P4 chose the ‘incoherent’ speaker on the 10th trial. It could have simply been a “genuine” error through lack of attention, etc. or alternatively the participant may have chosen to “test” the incoherent speaker. Having done so and thus “discovering” that the speaker’s hint/rule cohered with the current task contingencies, the participant then alternated back and forth across the two speakers. Interestingly, the only trial in which they failed to follow the speaker’s hint/rule was on a trial in which they had chosen the incoherent speaker. Although highly tentative, this could suggest that the reduced coherence for this speaker established in Phase 3 led the participant to “test” the accuracy of this speaker’s hint/rule (but only on one trial). Irrespective of the reason why the participant chose the incoherent speaker on the 10th trial, doing so appeared to undermine the incoherence functions for this speaker because the participant failed to show a strong preference for the coherent speaker thereafter. In other words, having been exposed to coherence between the speaker’s hint/rule and the reinforcement contingency for obtaining points, the previous incoherence functions for that speaker appeared to be much reduced.

The present study is, of course, exploratory and was designed largely to develop an experimental paradigm for systematically examining the impact of relational coherence on subsequent preferences for speakers who produce relational responses that are coherent versus incoherent with previously established stimulus relations. In reflecting on the aims of the current study and the results found, a number of issues seem worth considering. First, as mentioned in the Introduction, developing different methods of assessing coherence beyond those employed in the literature currently (e.g., the use of feedback and reversed reinforcement contingencies in Harte et al., 2020) seems important. Indeed, pursuing this research agenda will be essential in order to explore the potential impact of relational coherence in multiple contexts, and the current study could be seen as successful in this regard. Specifically, in the current study, the speakers in Phase 3 provided information that was consistent (i.e., coherent) or inconsistent with the relational training and testing from Phases 1 and 2, and the impact of coherence/incoherence was then tested in Phase 4.

In developing the foregoing strategy, it seems important to note that a distinction between coherence as an operation versus coherence as a process may be drawn. Specifically, relational training and testing followed by exposure to two different “speakers”, one of whom produced relational responses that cohered with the prior training/testing and one who did not, involved defining coherence as an operation. Coherence as a process, however, was then inferred based on the relative preference responses observed in Phase 4. This distinction between behavioral operation and process is similar to the distinction that applies to the concept of reinforcement; that is establishing a contingency between responding and consequences (reinforcement as an operation) and then inferring the process when response rate, for example, changes as a result of the operation (Catania, 1984).

In moving forward with the current research program numerous questions seem to emerge. For example, the current study involved presenting stimulus pairings in the same sequence across Phases 1, 2, and 3 (i.e., AB and BC relations). A future study could attempt to (partially) replicate this procedure but train AB and BC relations in Phase 1 but test for derived BA, CB, AC and CA relations in Phase 2 and present these derived relations through the speakers in Phase 3. One speaker could produce relations that cohered with Phases 1 and 2 (e.g., B1A1, C1A1) and the other speaker could produce relations that did not (B1A2, C1A2). Would we again observe a preference for the coherent speaker during Phase 4 when the training and testing involved tests for derived relations rather than directly trained relations?

Other studies could explore the extent to which different levels of coherence might impact on speaker preference. In the current study coherence was dichotomous, in that one speaker always produced a stimulus relation that cohered with the previous training and testing whereas the other speaker always produced a stimulus relation that did not. It remains to be seen if preference for one speaker over another is sensitive to relative levels of coherence in which one speaker produces a high level of coherent relations (e.g., 70%) versus the other speaker who produces a low level of coherent relations (e.g., 30%). Indeed, it would be interesting to determine if there is some mid-point of indifference (e.g., 51% versus 49%?).

A related line of inquiry may also further explore the types of complex relational networking that are involved when a participant spontaneously switches from choosing a coherent to an incoherent speaker (similar to P4 in the current study). For example, is there a difference in subsequent responding when this switch is due to a genuine error on behalf of the participant versus a type of ‘testing’ of the speaker’s reliability? Perhaps, research of this nature could incorporate a ‘think aloud’ procedure to investigate the “private” relational networking occuring during the task. In any case, these types of experimental analyses would allow us to more fully explore the concept of relational coherence in the context of derived relational responding than has been possible so far.

One possible limitation to the current experiment that should be addressed in future studies occurred in Phase 3. Specifically, in this phase one speaker was established as coherent and the other as incoherent by presenting them alongside stimulus pairings that were either coherent or incoherent with relational training and testing in Phases 1 and 2. A potential order effect could be involved here because the coherent speaker was always presented to participants first. It is possible, therefore, that always being exposed to the coherent speaker first biased participants in favour of this speaker. Thus we can ask, would the same results have been observed if this feature was counterbalanced across participants? If the results differed across participants based on this order effect, it would indicate that coherence may be better defined not just in terms of the functional overlap in the stimulus relations established across Phases 1 and 2, but also in terms of the order in which the blocks are presented in Phase 3.

In closing it may be useful to consider if only briefly the potential contribution the current lab-based research may have in the applied domain. Or more precisely, the current research may begin to provide an experimental approach that allows us to explore the behavioral variables that increase the probability of one speaker being preferred over another in terms of following the instructions they present to relevant listeners (Pennypacker & Hench, 1997). In organizations, for instance, leadership could benefit from understanding how to create a narrative that engages employees to follow the organizational mission. This same rationale could also apply to education, considering that teachers need students to follow instructions to perform tasks during and outside classes. In a broader sense, the understanding of social dynamics might also benefit eventually from the type of research reported in the current article, insofar as studying speaker preference may be considered as relevant to phenomena such as persuasion (Biglan, 2016; Galbraith, 1983). One recent example is the use of social media platforms to influence “real world” behaviors, ranging from shopping to illicit actions (Johnson et al., 2019; Matz et al., 2017). In general, a behavioral model clarifying how a particular speaker (or narrative) becomes more prefered compared to another could potentially contribute towards mitigating the impact of fake news, stigma, and political polarization generally.

References

Biglan, A. (2016). A functional contextualist approach to cultural evolution. In Zettle, R. D., Hayes, S. C., Barnes-Holmes, D. & Biglan, A. (Eds.), The Wiley Handbook of Contextual Behavioral Science (1st ed., pp. 385-397). John Wiley & Sons, Ltd. http://hdl.handle.net/1854/LU-8510038

Dugdale, N. & Lowe, C.F. (2000). Testing for symmetry in the conditional discriminations of language trained chimpanzees. *Journal of the Experimental Analysis of Behaviour, 73,* 5-22. https://doi.org/10.1901/jeab.2000.73-5.

Galbraith, J. K. (1983). The anatomy of power. *Houghton Mifflin Harcourt, 26*(3), 26-33.

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.* Advanced Online Publicaiton. https://doi.org/10.1007/s40732-021-00458-x

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

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2018). The impact of high versus low levels of derivation for mutually and combinatorially entailed relations on persistent rule-following. *Behavioural Processes, 157,* 36-46. https//:doi.org/10.1016/j.beproc.2018.08.005.

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

Harte, C., Barnes-Holmes, D., Moreira, M., de Almeida, J.H., Aparecida-Passarelli, D., & de Rose, J.C. (2021). Exploring a Training IRAP as a single participant context for analyzing reversed derived relations and persistent rule-following. Journal of the Experimental Analysis of Behavior. Advanced online publication. https://doi.org/10.1002/jeab.671

Hayes, S.C. (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 153-190). 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. (1986). Instructions, multiple schedules, and extinction: Distinguishing rule-governed from schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior, 46*(2), 137-147. https://doi.org/10.1901/jeab.1986.46-137

Johnson, N. F., Leahy, R., Restrepo, N. J., Velasquez, N., Zheng, M., Manrique, P. & Wuchty, S. (2019). Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature, 573*(7773), 261-265.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the national academy of sciences, 114*(48), 12714-12719.

Pennypacker, H. S., & Hench, L. L. (1997). Making behavioral technology transferable. *The Behavior Analyst, 20*(2), 97-108.

Shimoff, E., Catania, A.C., & Matthews, B.A. (1981). Uninstructed human responding: Sensitivity of low-rate performance to schedule contingencies. *Journal of the Experimental Analysis of Behavior, 36*(2), 207-220. https://doi.org/10.1901/jeab.1981.36-207

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Language, and Hearing Research, 14*, 5-13. https//:doi.org/10.1044/jshr.1401.05

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Authors Cooperative.

Sidman, M. & Tailby, W. (1982). Conditional discrimination versus matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behaviour, 37,* 5-22. https://doi.org/10.1901/jeab.1982.37-5

Sidman M, Rauzin R, Lazar R, Cunningham S, Tailby W, & Carrigan P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behaviour, 37*, 23–44. https://doi.org/10.1901/jeab.1982.37-23

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

Wray, A.M., Dougher, M.J., Hamilton, D.A., & Guinther, P.M. (2017). Examining the reinforcing properties of making sense: A preliminary investigation. *The Psychological Record, 62*, 599-622. https://doi.org/10.1007/BF03395823

Table 1

Results From the Preference Test.

		Coherent			Incoherent		
Part	Preference 1st trial	#selections	#rule-following	Following index	#selections	#rule-following	Following index
P1	Coherent	30	30	1,00	0	0	-
P2	Coherent	30	30	1,00	0	0	-
P3	Coherent	30	30	1,00	0	0	-
P4	Coherent	16	16	1,00	14	13	0,93

Note: Following index was calculated dividing #selections/#rule-following.

Figure 1

Outline of the Experimental Phases

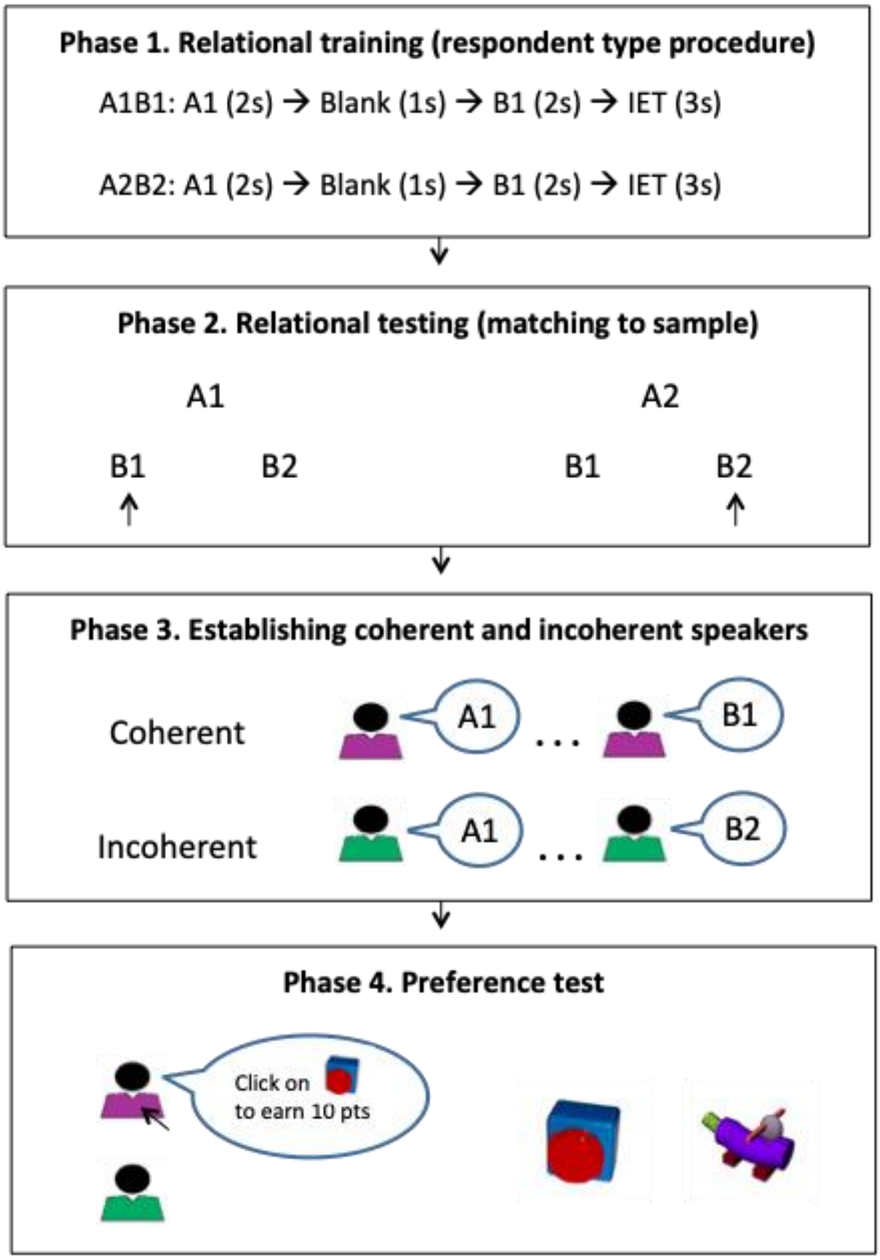
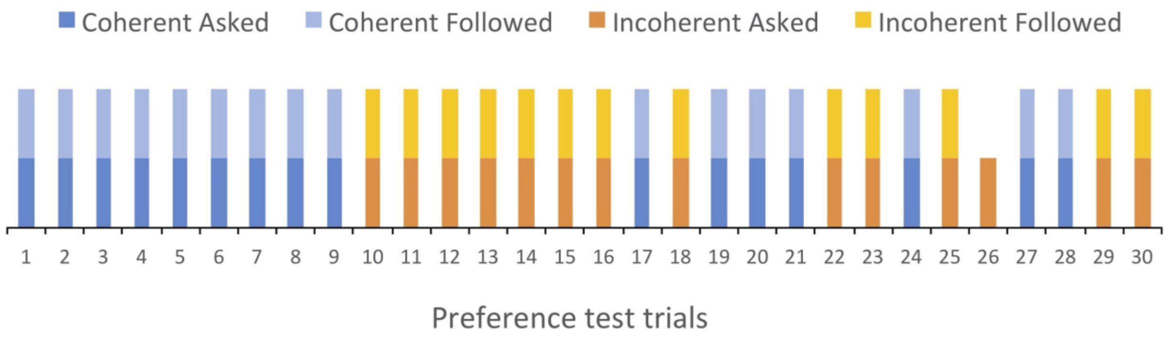


Figure 2

P4 Results From Phase 4.



Note. On each test trial, the speaker selected by this participant (Asked) and the subsequent occurrence of instructional control (Followed) are presented in the colored bars.