University of Nevada, Reno

Predictors of Flexibility and the Participation of Referential Instructions in a Turn-Based Matching-to-Sample Procedure

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Psychology

By

William Andrew Fleming III

Dr. Linda J. Hayes/Thesis Advisor

December, 2021

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

WILLIAM ANDREW FLEM

entitled

Predictors of Flexibility and the Participation of Referential Instructions in a Turn-Based Matching-to-Sample Procedure

be accepted in partial fulfillment of the
requirements for the degree of

MASTERS OF SCIENCE

Linda J. Hayes, Ph.D.,
Advisor

Ramona A. Houmanfar, Ph.D.,
Committee Member

Kenneth Hunter, Jr. Sc.D.,
Graduate School Representative

David W. Zeh, Ph.D., Dean
Graduate School

December, 2021

Abstract

The contemporary culturo-behavioral enterprise concerned with studying cultural events is not compatible with a molar, interbehavioral orientation toward understanding cultural interbehavior, largely due to its behavior-analytic assumptions. Not only does the enterprise not have a definition of cultural behavior, but it directs attention away from factors that may participate in cultural change, such as individual proclivities to prefer immediate over delayed rewards and the referential properties of interactions. As turn-based matching-to-sample procedures (TBMTS) are capable of isolating cultural interbehavior and changes of such, the purpose of this thesis was to (1) determine if framing instructions referentially altered shared patterns of stimulus-response functions (SRFs) observed in TBMTS, (2) predict flexible and rigid patterns of SRFs occurring in TBMTS, and (3) address limitations from Fleming and colleagues' (2021) original TBMTS study. Results suggest that referentially framing of instructions for TBMTS is functionally related to the length of trials necessary for a shared pattern of SRFs to be established between dyad partners. While measures external to TBMTS were not predictive of flexible patterns of SRFs, delay discounting rates and typical numbers of hours allocated toward sleeping were found to be related to whether dyads established a shared pattern of SRFs, a requirement for establishing flexible patterns of SRFs in TBMTS. Major findings of Fleming and colleagues' (2021) study were replicated, but some limitations of their study were only marginally remediated.

Keywords: interbehaviorism, cultural interbehavior, delay discounting, reference

Dedication

For Rockne and JoAnne Flowers,

who have given me so much

Table of Contents

List of Tables

Figures

Predictors of Flexibility and the Participation of Referential Instructions in a Turn-Based Matching-to-Sample Procedure

In behavior analysis, the cultural and verbal sub-enterprises have evolved in parallel to one another without much overlap. In the cultural sub-enterprise, researchers have mainly been concerned with cultural selection occurring through metacontingencies (Glenn et al., 2016), whereas in the verbal sub-enterprise, researchers have focused on exploring different theories to explain the occurrence of verbal behavior (S. Hayes et al., 2001; Horne & Lowe, 1996; Sidman, 2000). While cultural selection is thought to be facilitated by verbal behavior (Glenn et al., 2016; Leite & de Souza, 2012) and verbal behavior is thought to be maintained through socially mediated contingencies of reinforcement (S. Hayes et al., 2001; Skinner, 1957), the establishment of conventional referential behavior has not itself been treated as the subject matter of cultural studies in behavior analysis. This is in part a consequence of the Skinnerian molecular orientation dominant within behavior analysis, and it may be addressed by orienting to cultural behavior from a molar and interbehavioral perspective.

The purpose of this thesis was to extend a line of research examining cultural behavior using turn-based matching-to-sample procedures (TBMTS). Specifically, given a molar and interbehavioral orientation, this study aimed to assess how referential characteristics of instructions provided to participants and other factors were related to flexible and rigid patterns of stimulus-response functions (SRFs) occurring in TBMTS. To explain why such factors were the focus of this analysis, this paper will begin with an outline of a molar interbehavioral perspective. In doing so, (a) an interbehavioral approach to studying the establishment of shared SRFs, (b) the utility of a molar orientation to studying the establishment of shared SRFs, and (c) an experiment conducted from an interbehavioral orientation using a turn-based matching-to-sample procedure (TBMTS) will be described. Second, the thesis experiment utilizing TBMTS conducted to orient to factors that contribute to flexible and rigid patterns of SRFs—in which responding either does or does not change across conditions, respectively—will be described. After doing so, the methods and results of the thesis experiment will be outlined, and its limitations and implications will be discussed.

Cultural Interbehavior

From an interbehavioral perspective, psychological events can be conceptualized as integrated fields comprising interrelated factors, including stimulus functions, response functions, interbehavioral history, and setting factors (L. Hayes & Fryling, 2018; Kantor; 1958). An integrated field is a fundamentally different unit of analysis than the operant, which is studied through analyzing response rates (Skinner, 1937) or probability of emissions (Skinner, 1950), because it is non-causal; an integrated field describes an interrelation of factors participating as a psychological event without ascribing causal properties to any one factor (L. Hayes et al., 1996). While Kantor often differentiates between stimulus and response functions to highlight that both stimuli and responses could have multiple functions (Kantor, 1958) and that either function—and not necessarily both—could be substitutional (Kantor, 1924), here stimulus-response functions (SRFs) will be discussed to highlight that neither function can be defined without respect to the other (Fleming et al., 2021).

Cultural psychological events are a specific class of psychological events (Kantor, 1982). Cultural psychological events are those involving shared SRFs, or arbitrary SRFs conventional across individuals and established under the auspice of a group (Fleming et al., 2021; Kantor, 1982). To say that such SRFs are arbitrary is to say that they do not refer to formal or natural properties of stimuli, but rather to attributed functions established through social interaction under the auspice of a group. These functions include a wide array of interactions, but most notably they include what is typically considered to be verbal behavior or social mediation. To say that such SRFs are shared is to make two claims. First, the same SRFs should be observed across multiple individuals in similar circumstances. Second, those SRFs must be established through interactions with others participating as both setting factors and auxiliary stimulus objects across whom such functions are shared. As such, examining cultural psychological events necessitates a dual analysis on the psychological level (i.e., in demonstrating the establishment of shared SRFs with respect to interactions with others) and the cultural level (i.e., in demonstrating that shared SRFs are shared across multiple individuals; Fleming et al., 2021), although such levels may be better characterized as dimensions given that analyses across both concern the same events.

While Kantor (1982) contended that an experimental analysis of cultural psychological events was possible, he did not construct experimental methods for analyzing such events. Others in behavior science have constructed methods to examine cultural events, particularly for events of cultural selection occurring through metacontingencies within the cultural sub-enterprise of behavior analysis. Their experimental effort has been in part a direct response to calls for demonstrating how conducive the metacontingency construct is for producing viable empirical data upon which the enterprise can be built (Branch, 2006; Marr, 2006; Mattaini, 2006). While metacontingency experiments are useful for studying factors that contribute to the production of aggregate products, they are not designed to investigate cultural psychological events. They are designed to study culturants, which orients their research away from observing shared patterns of SRFs.

Experimental Research on Cultural Events in Behavior Analysis

Several behavior-analytic oriented researchers have studied events similar to—but not the same as—cultural interbehavior. Following the trichotomy of selection processes outlined by Skinner (1971,1981), the cultural sub-enterprise of behavior analysis has largely been constructed on the assumption that behavioral selection is different from cultural selection. It is thought that, whereas reinforcement describes the selection of operants by reinforcing events, cultural selection describes the selection of culturants by selecting events, as depicted in Figure 1. Unlike operants, which refer to classes of environment-altering behavior upon which particular reinforcing events are contingent, culturants refer to aggregate-product producing interlocking behavioral contingencies, or IBCs, upon which particular selecting events are contingent (Baia & Sampaio, 2019; Glenn et al., 2016). Although the logic of cultural selection is distinct from that of reinforcement because it is concerned with a different unit of analysis (i.e., the culturant rather than the operant; Baia & Sampaio, 2019; Fryling et al., 2021), the contingency logic of cultural selection is derivative of or analogous to that of reinforcement (Fleming & L. Hayes, 2021; Glenn et al., 2016). Although the logics are similar, they are not the same, given that they concern different types of events (Fryling et al., 2021).

Given that experiments on culturants are concerned with cultural events rather than cultural interbehavior, experimental procedures constructed within the metacontingency enterprise are not suitable for studying cultural interbehavior. When reinforcement is assumed by cultural selection, it is not explicitly analyzed when iterations of culturants are shown to occur at a reliable rate or probability (Fleming & L. Hayes, 2021). When responses by individuals are analyzed, target responses investigated are those designated *a priori* by experimenters to produce larger rewards when they occur with respect to certain stimuli. For example, Soares and colleagues (2018) scheduled cultural consequences (i.e., donations of school supplies) to be contingent on all group members selecting even-numbered rows on a Wiggins' Matrix board. In another study, Sampaio and colleagues (2013) delivered the largest point bonuses when group members selected stimuli rotated in a direction relative to that of those selected by other members. Such preparations are coherent with the metacontingency construct because they involve delivering consequences contingent on environmental alterations produced by two or more individuals behaving with respect to each other, but they are ill-suited for studying the establishment of shared SRFs. This is because these types of arrangements allow for some variability to be observed between groups based on interactions occurring within them, but in each case the maximization of rewards was based on formal properties between stimuli determined by experimenters prior to interactions between experimental subjects rather than arbitrary properties that are the product of group interactions between participants given constraints imposed by experimenters. As Fleming and colleagues (2021) state, “Although such experiments [on culturants] may involve cultural interbehavior, cultural events—not behavioral events—are the focus of study,” (p. 3).

Even when operants, not culturants, are investigated, cultural interbehavior is typically not recognizable with high precision from data reported. For example, consider a series of experiments by Skinner and colleagues (Epstein et al., 1980; Epstein & Skinner, 1981; Lanza et al., 1982) on symbolic behavior among non-human organisms, as discussed by Fleming and colleagues (2021). In these experiments, the researchers individually established conditional discriminations for pairs of pigeons so that they would peck a certain key in the presence of a particular stimulus. After training, they were able to demonstrate “communication” through the production of discriminative stimuli. In the presence of a particular light, one pigeon would peck a key with a particular symbol on it. That key would then light up in view of the other pigeon, which would then peck a certain key of a color identical to that shown to the first pigeon for access to grain. While one may state that the first pigeon “told” the other what color they saw by pecking a symbolic key, the behavior is not cultural interbehavior with respect to groups of experimental subjects because (a) the pigeons did not participate in the establishment of each other’s behavior and (b) all behavior was established with respect to the formal, non-arbitrary properties of food. While their intertwined behavior may seem like, “a sustained and natural conversation” (Epstein et al., 1980, p. 544), it was not so if conversing is to be considered cultural interbehavior.

To some extent, experiments examining operant behavior are incompatible with examining cultural interbehavior because of their underlying molecular protopostulates. From a Skinnerian perspective, reinforcement is a process by which an operant is strengthened on the basis of contiguous relations between discrete events (Baum, 2020; Skinner, 1938). In this way, operant behavior is effective behavior, or behavior that is effective at producing reinforcers. But the establishment of arbitrary functions likely does not occur immediately in the natural environment. The very fact that Epstein and colleagues' (1980) pigeons needed to be trained substantially to meet performance criteria suggests that experimenters needed to maintain socially mediated contingencies of reinforcement for some time despite a lack of reinforcing consequences for doing so, for at least some time. A Skinnerian account of verbal behavior—which is definitively cultural in behavior analysis like it is in interbehavioral psychology as it is socially established—is only coherent when the behavior of listeners reinforces the behavior of speakers. In situations in which multiple verbal interactions are necessary to produce reinforcing events, a contiguity-based conceptualization of reinforcement is arguably more mentalistic (if appeals to reinforcement through private rule-governance are necessary to explain temporally-extended behavior) and less applicable in scope (if functional relations are thought to pertain to discrete events) than molar, correlation-based alternatives to understanding the organization of behavior.

A Molar Orientation to Cultural Interbehavior

A molar behavior-analytic approach to cultural interbehavior departs from a molecular approach in two fundamental ways: unit of analysis and types of relations between units. The units of molecular analyses are discrete events, like individual discriminated operant behavior and reinforcing events, whereas the units of molar analyses are patterns of events, extended through time. Given that patterns of events overlap and are intertwined with one another (Rachlin, 1992), relations between patterns are thought of as correlations or covariances, not contiguities (Rachlin, 2017). These two divergences have profound implications for behavioral science, just as they do for analyses of culturant interbehavior.

Consider Figure 2, which depicts two sets of coordinated operant contingencies in which patterns of behavior are reinforced. From a molar perspective, the contingency should not be thought to describe relations between discrete events but environmental regularity by which patterns of such events can become correlated (Fleming, Ardila-Sánchez, & L. Hayes, 2021). In the top panel, the contingency depicted is similar to that in Figure 1 where P1's responding functions as discriminative events for P2's responding, and both P1's and P2's responding are necessary to produce specific reinforcing events. When patterns of such events are considered, though, reinforcing events should be reconceptualized because (1) relevant patterns of events are correlated by the contingency and (2) conditional reinforcing events may only retain their reinforcing functions because of their correlation with other events (Baum, 1973). It is also important to note that, while reinforcing events may not be contingent on all responding, any responding that does occur can be correlated with reinforcing events and responding upon which such reinforcing events are contingent. Such a correlation-based approach orients research toward a wider array of circumstances, like that depicted in the bottom panel of Figure 2 in which reinforcing events are not only necessarily dependent on responding by P2 but only correlated with one another when P2 contacts verbal responding by P1.

A molar behavior-analytic approach has already produced experimental analyses concerning verbal contingencies like that depicted in the bottom panel of Figure 2, although mostly outside of behavior analysis. By adopting both the framework proposed by Lewis (1969) and an evolutionary perspective, Skyrms (2010) and other researchers have used Lewis signaling games to study the evolution of conventional systems in cooperative situations. Basic Lewis signaling games comprise a series of trials in which (1) given a state of the world, a sender sends a particular message (typically some symbol) to a receiver, then (2) given the message from the sender, a receiver acts in accordance with a possible state of the world, followed by (3) a shared payoff dictated by whether the receiver acted with respect to the state of the world contacted only to the sender. Each individual payoff is not dependent on what message is sent by the sender, but the receiver can only reliably behave with respect to the state of the world contacted only to the sender if states of the world and messages are correlated with one another. When there are an equal number of states of the world, messages the sender can send, and actions the receiver can take to produce shared outcomes, maximization of payoffs is produced in accordance with strict or separating Nash equilibria in which only one message is sent for each state of the world and one action is taken for each message. When such responding is shown to be consistent, a signaling system has been formed (Lewis, 1969).

Reinforcement has been incorporated in Lewis signaling game models to describe the establishment of signaling systems (Bruner et al., 2018) and to create learning algorithms for simulated agents (Catteeuw & Manderick, 2013), but these models are notably based on Herrnstein's (1970) matching law rather than Skinner's (1938) construction of reinforcement as a strengthening process. The matching law, which describes relations between relative rates of responding and relative rates of obtained reinforcement (Baum, 1974), is a distinctly molar description because it relates patterns of operant and reinforcing events rather than relations among their constituent elements. Simulations of Lewis signaling games often adjust response probabilities molecularly (i.e., trial by trial) through "urn learning," but the equilibria payoffs upon which such adjustments are based are molar variables. When maximization is assumed, such variables can be arranged as utility functions to predict and control behavioral patterns (Rachlin, 1992). Utility functions based on the matching law have accounted for a high proportion of variance in choice observed during Lewis signaling games with both humans and simulated agents (Huttegger et al., 2014).

While Lewis signaling games demonstrate the establishment of arbitrary SRFs, they are typically ill-suited to demonstrate shared SRFs. Sending a specific message given a particular state of the world and making a certain action given a particular message are arbitrary in Lewis signaling games because any message or action could be taken as long as they are coordinated. For such functions to be shared, individuals must participate in Lewis signaling games as both sender and receiver, and they must behave identically when participating in the same role. This is hardly—if ever—incorporated into Lewis signaling games with humans and is likely related to mentalistic orientations (i.e., possessed "beliefs" are thought to be updated based on empirical contacts; Lewis, 1969; Ventura, 2019). The lack of concern for accounting for processes relevant to relations between sender and receiver roles is most apparent by the lack of any parameter addressing the change in probability of a sender acting like a receiver or vice versa within their models. Certain features of Lewis signaling game experiments with humans also make it difficult to discern the participation of different factors in the establishment of functions, such as providing multiple points of feedback simultaneously (e.g., points as well as what all group members selected). Additionally, Lewis signaling game experiments only incorporate replications across groups of cooperating members rather than within each group, decreasing reliability that reliable signaling game systems were observed. These limitations are addressed by TBMTS, a procedure constructed to demonstrate the establishment of shared SRFs (Fleming et al., 2021).

Turn-Based Matching-to-Sample Procedures

Although TBMTS may be considered a special variant of a Lewis signaling game, it will be described here in behavioral terms more suited to the present analysis and theoretical orientation. In TBMTS, dyad partners take turns selecting comparison stimuli in the presence of sample stimuli across a series of trials in order to earn shared points. On each trial, (1) in the presence of a particular, randomly determined sample stimulus from Set B (either Sb_1, Sb_2, or Sb_3), one partner selects a comparison stimulus from Set A (either Sa_1, Sa_2, or Sa_3), then (2) in the presence of the comparison stimulus from Set A selected by the first partner, the other partner selects a comparison stimulus from Set B, after which (3) both partners receive a shared reward simultaneously. The amount of each reward is determined by whether correspondence or noncorrespondence occurred on the trial. Correspondence refers to the second partner to respond selects a comparison stimulus that is identical to that shown as a sample stimulus to the first, whereas noncorrespondence refers to the second partner to respond selects a comparison stimulus that is not identical to that shown as a sample stimulus to the first. Dyads switch between completing trials in a correspondence maximization (CM) condition or a noncorrespondence maximization (NCM) condition depending on stability criteria. In CM conditions, correspondence produces more points than noncorrespondence; in NCM conditions, noncorrespondence produces more points than correspondence. Every trial, partners switch roles so that, if one responds first on one trial, that participant responds second on the next. While partners are told that they will be working with one another to earn points and see one another before starting the experiment, partners cannot see or hear the other during TBMTS; their only contacts are through programmed dependencies. Because partners switch roles, shared SRFs in which both dyad partners select the same comparison stimuli in the presence of the same sample stimuli can be demonstrated. Preventing participants from responding with respect to one another except substitutionally through TBMTS prevents previously established communicative conventionality from permeating the task. Switching between CM and NCM conditions allows for differential patterns of SRFs to be established and reconstitute across conditions. As such, TBMTS satisfies requirements to demonstrate shared SRFs.

Fleming and colleagues (2021) demonstrated the establishment of shared SRFs using a version of TBMTS in which dyads began in a condition in which correspondence maximized points. In each analyzable case during stability trials in that condition across all dyads, dyad partners always selected an identical comparison stimulus in the presence of a particular sample stimulus. Not only were such functions shared, but they were symmetrical; if a participant selected a particular comparison stimulus in the presence of a particular sample stimulus, they selected that sample as a comparison in the presence of that comparison as a sample. In all subsequent like conditions where correspondence maximized points, participants generally responded in accordance with the pattern of SRFs established in the first condition. Replication of the establishment of shared SRFs, then, were observed in each dyad and across dyads. As would be expected if the functions were arbitrary, different patterns of shared SRFs were established across dyads, although some patterns were more likely than others.

In NCM conditions, one participant's responding in each dyad typically adjusted while their partner continued to respond as they did in CM conditions. Certain observations predicted flexible (i.e., of the individual whose responding changed more across conditions) and rigid (i.e., of the individual whose responding changed less across conditions) patterns of SRFs, including (1) the first deviation in the first NCM condition from the pattern of SRFs observed in the first CM condition, (2) consistency of responding in the first CM conditions, and (3) the number of initial responses (i.e., first responses with respect to a particular sample stimulus) that were consistent with the terminal pattern of SRFs established within the first

CM condition. These findings were related to discriminability across conditions, varied by whether dyads earned more, less, or the same number of points for correspondence in NCM conditions than they did in CM conditions. As flexible and rigid patterns of SRFs may have implications for propensities to engage in or deviate from established patterns of cultural conduct, understanding factors that contribute to flexible and rigid patterns of SRFs is a worthwhile endeavor. Of particular interest are factors that may be measured prior to and predictive of performance in TBMTS, as such measurements may also be able to predict cultural change.

From a molar perspective, few processes are as ubiquitous or relevant to understanding the organization of behavioral patterns as delay discounting. Delay discounting refers to the decrease in the value of a commodity or activity as delay to its receipt is increased (Odum, 2011). Said differently, delay discounting describes temporal horizons, or the temporal window by which operant events are organized by correlated consequences (Bickel et al., 2006). Since socially mediated consequences are contingent on the behavior of two or more individuals, temporal horizons may need to be expansive to accommodate delays to rewards imposed by the behavior of others. Explanations of this logic have been given to describe the lack of cooperation among pigeons and rats during prisoner's dilemma games (Green et al., 1995). Given that point maximization on any single trial reduces point maximization across trials in an iterated prisoner's dilemma game, the behavior of non-human organisms may not be sensitive enough to temporally-extended rewards to maintain cooperative responding. Said differently, outcomes of cooperation may be too remote to organize cooperative social interactions for some individuals that play prisoner's dilemma games, as well as whole species of organisms. A similar logic can be applied to dyads that fail to complete the initial condition within TBMTS; a consistent pattern of shared SRFs may not actualize in TBMTS if, in the initial CM condition, one participant's consistent responding does not persist long enough for their partner's responding to adjust with respect to each *Sb* sample stimulus. If this occurs, more frequent switching between different patterns of responding may be indicative of temporally-restricted or impulsive choice. In terms of delay discounting, impulsive choice is measured as a steep discounting rate (i.e., rewards steeply lose value when their receipt is increasingly delayed). Such rates are typically calculated from data collected using a monetary-choice questionnaire (Odum, 2011), which may also be used to predict performance in TBMTS if it is related to delay discounting.

Flexible and rigid patterns of SRFs observed during TBMTS may also be related to initial orientations toward the task, particularly contact with instructions that pertain to its referential characteristics. From an interbehavioral perspective, referential events are those in which a referor refers a referee to a referent (Kantor, 1977). Although rudimentary, TBMTS incorporates dependencies of this form of reference. Interactions on each trial can be seen as the first dyad partner to respond referring their partner to a particular stimulus shown to them through their selection of a comparison stimulus. Considering that it is possible, albeit unlikely, that shared SRFs could be established with participants not interacting verbally or referentially in the task (Fleming et al., 2021), constructing instructions that frame TBMTS in more or less referential terms may be alter patterns of SRFs observed. If it is shown to be related to flexible and rigid patterns of SRFs, such patterns may not only be predicted by measures of referential competency but controlled through framing circumstances in referential terms.

Method

The primary purposes of this thesis project were to identify factors relevant to the prediction of flexible and rigid patterns of SRFs observed during TBMTS and to examine how referential instructions participate in the establishment of shared SRFs. With respect to predicting flexible and rigid patterns of SRFs, (1) delay discounting, (2) the referential framing of instructions, (3) awareness of one's own and one's partner's responding across conditions in TBMTS, (4) demographic variables, including age, gender, and (5) typical time allocated to various activities were of particular interest. To this end, an experiment was performed in which dyads, assigned to groups that differed on the basis of how instructions emphasized referential characteristics of the task, completed TBMTS after each participant completed a delay discounting procedure similar to that constructed by Du and colleagues (2002). After TBMTS, participants were asked to report what comparison stimuli they chose in the presence of each sample stimulus across all conditions, as well as what they thought their partner was selecting, to determine the extent to which participants could reference their own behavior and that of their partner. They were also required to answer demographic questions and report how much of their time during a typical week was spent participating in various activities to assess relations between them and patterns of SRFs observed during TBMTS.

This project also sought to address limitations of Fleming and colleagues' (2021) TBMTS study. In their study, Fleming and colleagues found that the use of one set of stimuli composed on arbitrary symbols and one set of conventional shapes was likely related to an increased probability of observing a particular pattern of shared SRFs within CM conditions. The current study sought to balance the probability of observing any pattern of shared SRFs by using two sets of arbitrary symbols instead of just one. They also could not determine whether SRFs were shared with respect to responding the in presence of each sample stimulus for every dyad in every CM condition because not all dyad partners contacted the same sample stimuli during stability trials in CM conditions. To address this, the probability of presenting each *Sb* symbol as a sample stimulus was made more consistent (i.e., from three times for each block of nine trials to one time for each block of three trials) and the number of consecutive point-maximizing trials necessary to complete a condition was increased from nine to twelve. Additionally, Fleming and colleagues found that, when the number of points for correspondence decreased from CM to NCM conditions, dyads were more likely to complete the first NCM condition in less trials than the first CM condition. As this may have been related to decreasing the number of points for correspondence to the number of points for noncorrespondence in CM conditions specifically rather than decreasing points for correspondence generally, the thesis experiment sought to replicate this finding when points decreased to a value not previously contacted in TBMTS.

Subjects

Thirty-four undergraduate students from the University of Nevada, Reno through the Department of Psychology's SONA system. In order to participate in the study, participants were required to be students enrolled in at least one psychology course, over 18 years of age, and speak English. Participants received SONA credits for participating in the study and up to $16 for their performance in the study. All participants signed informed consent forms prior to participating. Participants were assigned to a dyad (i.e., pair of participants) based on whoever signed up to participate during the same timeslot on SONA.

Setting and Apparatus

Dyads completed experimental sessions in a small observation room outfitted with two desks separated by a divider and a table with two storage containers. One dyad partner used a computer at each desk to complete the study. Each participant wore headphones playing brown noise at 25% of their computer's total volume (or less if requested to a more comfortable volume) to minimize distractions and auditory stimulation from their partner. All participants were asked to silence or turn off their electronic devices and store their personal belongings in one of the two storage containers or under the table if their belongings did not fit. Experimental tasks, instructions, and questionnaires were programmed using Visual Studio 2019. The computers were networked together so that each computer could read and write to the same files stored in a shared folder.

Procedure

After gaining consent, each dyad completed a single experimental session. Each session comprised participants completing (1st) an adjusting-amount delay discounting procedure, (2nd) TBMTS, (3rd) a post-questionnaire about TBMTS, (4th) a demographic questionnaire, and (5th) a time allocation questionnaire. All questionnaires and tasks besides TBMTS were completed individually; participants only interacted with each other during TBMTS.

Adjusting-amount delay discounting procedure

The adjusting-amount delay discounting procedure used in this study was similar to that constructed by Du and colleagues (2002). Before beginning the procedure, participants were shown the following instructions:

> "Please read the following questions as they appear. You will be asked whether you would prefer a smaller amount of money today or a larger amount of money after a particular delay. Although these rewards are hypothetical, please answer as if you were actually receiving the rewards you choose. Choose your preference by clicking on the button with your preference when it appears. When you are ready to begin, click Start when it appears."

The procedure consisted of seven sets of six questions in which participants chose whether or not they preferred to receive an amount of money less than $200 today or $200 at a specific delay that remained the same in each set. All rewards were explicitly hypothetical and framed in US dollars. Across sets, the delays (i.e., 1 month, 3 months, 9 months, 2 years, 5 years, 10 years, and 20 years) progressively increased after completing each set. Within each set, the first question always asked participants whether they would prefer $100 today or $200 at the specified delay for the set. If a participant selected the smaller amount of money, the smaller amount of money decreased by $50 on the next question. If a participant selected $200, the smaller amount of money increased by $50 on the next question. On subsequent questions, each titration of the smaller amount of money was half of that of the previous titration (i.e., $25, $12.50, $6.25, and $3.13). Participants completed on set after the other; sets were only discriminable by changes in the smaller amount of money and the delay for the $200 option.

Each question was presented in the same way. Before each question was presented, the screen was blank. Then, different blocks of text appeared with 1 s intervals between them to increase the probability of participants attending to each part of each question. The first text to appear was the question, "What would you prefer?" Second, the smaller amount of money option would appear (i.e., $X Today, where X was the smaller amount of money rounded to the nearest cent) beneath the question. Third, the conjunction, "or" appeared beneath the smaller amount of money option. Fourth, the $200 option appeared (i.e., $200) beneath the "or" (i.e., $200). Fifth, on the same line as the $200 option, the delay for the $200 option (i.e., in X, where X was the specific delay in either months or years). Sixth, buttons the participant could click to choose either the smaller amount of money or the $200 appeared beneath the $200 option. The smaller amount of money option was always presented on the left side of the screen, and the $200 option appeared on the right. No time limit was imposed on responding to each question.

TBMTS

The version of TBMTS used in this study was like that constructed by Fleming and colleagues (2021). After completing the adjusting-amount delay discounting procedure, participants were brought directly to instructions for TBMTS. Instructions differed across groups, as shown in Table 1. Dyads were pseudo-randomly assigned to one of three groups—the D-NR, D-R, or the ND-NR group—so that, in every three consecutive dyads, one dyad was randomly assigned to each group. Each participant in dyads assigned to the D-NR group were presented instructions that described critical aspects of point-maximization contingencies in non-referential terms similar to instructions presented to participants in Fleming and colleagues' (2021) study. Each participant in dyads assigned to the D-R group was shown similarly descriptive instructions but in referential terms that highlighted the "messaging" aspect of TBMTS. Participants in dyads assigned to the ND-NR group was delivered instructions that neither described the point-maximization contingency nor were framed in referential terms. The differences in instructions across groups was the only programmed difference among them. Participants were shown instructions for TBMTS and not able to start it by clicking a "Continue" button until 15 s after both dyad partners completed the adjusting-amount delay discounting procedure. Once both participants clicked the "Continue" button, they started TBMTS.

For the sake of describing TBMTS and data analyses thereafter, the first partner in each dyad to respond first on a trial will be referred to as P1 and the first partner to respond second on a trial will be referred to as P2. As stated above, on each trial, (1st) one partner selected a comparison stimulus from set *Sa* (i.e., either Sa_1, Sa_2, or Sa_3) in the presence of a sample stimulus from set *Sb* (i.e., either Sb_1, Sb_2, or Sb_3), (2[nd]) the other partner selected a comparison stimulus from set *Sb* in the presence of a sample stimulus identical to that selected by the first partner to respond, and (3[rd]) both partners received a shared reward simultaneously. Table 2 shows stimuli included in each set. While one dyad partner was making a selection, their partner's screen only showed the text, "Please Wait" and their cursor was an inoperable red No signal (i.e., ∅). The presentation of stimuli was staggered so that comparison stimuli followed the presentation of the sample stimulus by a 1 s delay. After the comparison stimuli appeared, the partner's cursor switched from an inoperable red No signal to an operable, typical white cursor. Comparison stimuli appeared left to right in a randomized order for each selection. The sample stimulus presented to the first partner to respond on a trial was pseudo-randomly presented so that, in every set of three trials, all *Sb* stimuli were randomly presented once. The order of *Sb* sample stimuli was the same across all dyads. When dyad partners received their shared reward each trial, the text was staggered so that the text "Your team has earned" was followed by "X points" after a 1 s delay where X was the number of points earned on the trial. After another 1 s delay, partners could click a "Continue" button that started the next trial after each partner clicked it. Partners switched roles each trial so that if one selected a comparison stimulus first or second on a given trial, then they selected second or first, respectively, on the subsequent trial.

TBMTS consisted of two condition types: a correspondence-maximization (CM) and a noncorrespondence maximization (NCM) condition. In CM conditions, dyads received 25 points for correspondence (i.e., the second partner to respond on a trial selecting a comparison stimulus identical to that shown to the first partner to respond as a sample) and 5 points for noncorrespondence (i.e., the second partner to respond on a trial selecting a comparison stimulus that was not identical to that shown to the first partner to respond as a sample). In NCM conditions, dyads received 15 points for correspondence and 45 points for noncorrespondence. Participants completed trials in a given condition until they completed 12 consecutive point-maximizing trials, hereafter referred to as stability trials. Conditions alternated on a mixed schedule with the only discriminable differences across conditions being the number of points received for correspondence and noncorrespondence. All dyads began in the CM condition. Each CM condition was followed by a NCM condition, and each NCM condition as followed by a CM condition. There was no limit or minimum number of condition reversals; dyads simply continued to complete trials until forty-five minutes has elapsed.

TBMTS Post-Questionnaire

After completing TBMTS, all participants completed a post-questionnaire. The post-questionnaire consisted of two open response questions in which participants were asked:

> "Did you come up with a strategy to make the most amount of points in the fewest amount of responses? If so, what did you do?"

Participants could submit their answer by pressing the "Continue" button that appeared 10 s after they began to type their response. Afterwards, participants were asked to select which comparison stimuli they chose in the presence of each sample in each condition to maximize points. The general prompt for each question in this section was:

> "When the most amount of points you could earn on a trial was X, which Y symbols(s) did you choose in the presence of the Z symbol that will appear below to earn X points? Select all the Y symbols that apply by clicking on them. You can unselect symbols by clicking on them again. When the Continue button appears, click on it to proceed in the study"

where X was either 25 or 45 points (corresponding with CM or NCM conditions, respectively), Y was either black or blue, and Z was either blue if Y was black or black if Y was blue. Participants were able to select multiple stimuli for each response. With the exceptions of the prompt on the screen, the consistent placement of comparison stimuli, and text stating "Selected" appearing above selected comparison stimuli, how participants selected stimuli was identical to how participants selected stimuli during TBMTS to minimize differentiation between the post-questionnaire and TBMTS; the presentation of each sample stimulus and comparison stimuli was staggered and positioned on the screen as they were in TBMTS. After indicating they selected in the presence of each sample stimulus in both conditions, participants completed a nearly identical questionnaire in which they indicated what they thought their partner was selecting in the same circumstances. Participants in dyads that did not complete the first condition were only asked what they selected when the most number of points they could earn was 25.

Demographic Questionnaire

After the TBMTS post-questionnaire, participants completed a series of demographic questions pertaining to age, gender, languages spoken, semesters of college completed, and GPA. The details of each question are included in Table 3.

Time Allocation Questionnaire

After the demographic questionnaire, participants indicated how many hours was spent engaged in various activities during their typical week. Participants were required to allocate exactly 168 hours across eight activities, including (1) Eating (outside other activities), (2) Exercising, (3) Occupational Work, (4) School Work, (5) Sleeping, (6) Social Media, (7) Social Outings, and (8) Other. Participants allocated hours by clicking up and down on a numeric counter or typing their response for each activity. The total number of hours allocated was displayed on the screen. A message stating "Too High" or "Too Low" in red text was shown on the screen if the total number of hours allocated was over or under 168, respectively.

If participants indicated that they were 21 years of age or older on the demographic questionnaire, they were also asked how many days (0-7) they drank alcohol, used/consumed products with nicotine, and used/consumed products with THC in a typical week.

After both participants completed the time allocation questionnaire, the study was completed. Participants were each paid $2 for each condition completed. If participants had any questions about the study, they were answered by the researcher or the research assistant running the session.

Data Analyses

Raw data was organized in Excel prior to being subjected to refinement and formal statistical analyses using GraphPad Prism 8. Parametric statistical analyses were used unless data were determined statistically non-normally distributed by a Shapiro-Wilk test. Alpha was 0.05 for each statistical test.

Delay Discounting

Data from the adjusting-amount delay discounting procedure was used to calculate *k* values, or discounting rates according to the Mazur's (1987) hyperbolic discounting equation:

$$V = \frac{A}{1+kD} \qquad (1)$$

Equation 1 relates the value of the amount of money less than $200 available today to the $200 available after a specific delay. In this equation, *V* is the value of the amount of money less than $200, *A* is 200, as the value of the delayed amount of money was always $200, *D* is the delay, and *k* is the rate at which *V* decreases as a function of *D*. Values for *V,* which will hereafter be referred to as indifference points, were found for each delay by averaging the value of the amount of money less than $200 the last time the amount of money less than $200 was selected and the last time $200 was selected for a particular delay. If only the amount of money less than $200 was selected, the indifference point for that delay was $1.57 (the average of $0 and $3.13). If only $200 was selected, the indifference point for that delay was $198.44 (the average between $196.87 and $200). After an indifference point was calculated for each delay, a *k*-value was determined by fitting Equation 1 to a participant's set of indifference points across delays with *k* treated as the only free parameter.

As delay discounting data is sometimes unsystematic (Johnson & Bickel, 2008), delay discounting data was included in analyses according to particular criteria (Matthew Locey, personal communication). If no indifference points increased as delay increased, that participant's *k*-value was included within all relevant analyses. If no indifference points increased as delayed increased besides one, all indifference points besides that one were used to determine that participant's k-value. If more than one indifference points increased as delay increased or the indifference point for the longest delay was not less than that of the shortest delay, that participant's k-value was not included in any analysis. These inclusion criteria were used in lieu of Johnson and Bickel's (2008) criteria to retain as much completely systematic data as possible.

After determining a k-value for each participant with systematic data, k-values were normalized according to the following natural-log function:

$$k' = \ln(k+1) \tag{2}$$

where k was increased by 1 to keep all values positive. Individual and mean k' values for each group are shown in Table 4.

TBMTS

Data refinements for TBMTS data were similar to those conducted by Fleming and colleagues (2021). Refinements concerning SRFs occurring during TBMTS were calculated as relevant to demonstrating group differences and assessing the predictability of flexible and rigid patterns of SRFs. For each dyad, the number of trials completed in the first CM and NCM conditions, the number of different sample stimuli contacted in the first CM condition, the number of conditions completed, the proportion of CM trials completed in the first two conditions, and the number of trials into the first NCM condition necessary for responding with respect to *Sa* and *Sb* sample stimuli to deviate from the pattern established in the first CM condition were calculated. As was found in Fleming and colleagues' (2021) study, all SRFs observed in stability trials within CM conditions were shared (i.e., dyad partners selected the same comparison stimuli in the presence of the same sample stimuli) and symmetrical (i.e., participants selected the same comparison and sample stimuli in the presence of one another), allowing for each dyad to be characterized by the set of symmetrical SRFs observed during stability trials established within the first CM condition. For each participant, the total number of SRFs in non-stability trials in the first CM condition that were consistent with those observed during stability trials, the total number of SRFs consistent with those occurring in stability trials in the first CM condition occurring on the first opportunity to occur, and flexibility (i.e., the degree to which SRFs remained consistent across conditions) were calculated. Flexibility was determined by the average overall consistency scores between dyad partners using Fleming and colleagues' (2021) equation:

$$\frac{a}{b} \tag{3}$$

where *a* is the most times a particular comparison stimulus was selected with respect to a particular sample stimulus across stability trials in all conditions and *b* is the total number of comparison stimuli selected with respect to a particular sample stimulus across stability trials in all conditions. If the same comparison stimulus was selected more than once in stability trials in the same condition, those additional selections did not count toward *a* or *b*. Overall consistency scores for each sample stimulus were averaged together to produce an average consistency score for each participant. The pattern of SRFs for the partner in each dyad with the lower average overall consistency score was deemed flexible, whereas the patterns of

SRFs for the partner with the higher average overall consistency score was deemed rigid for the sake of statistical analyses.

TBMTS Post-Questionnaire

Across all participants, a correspondence score was calculated for each set of sample stimuli in each condition from the point of view of each dyad partner. Equation 3 was used to calculate each correspondence score, but in this case, *a* was the number of comparison stimuli selected in the post-questionnaire with respect to either *Sa* or *Sb* sample stimuli that corresponded with comparison stimuli that the participant actually selected in TBMTS to maximize points in a particular condition or what that participant thought their partner selected to maximize points in a particular condition, and *b* was the total number of comparison stimuli selected in the post-questionnaire for that particular set of sample stimuli in a particular condition by a particular dyad partner. As such, correspondence scores ranged from 0 to 1, in which a score of 0 indicates that a participant could not accurately describe what they or their partner selected in the presence of sample stimuli from a particular set to maximize points in a particular condition. Conversely, a score of 1 indicates perfect accuracy in describing such behavior.

With respect to typed responses to the open-ended question regarding each participant's strategy to earn points in TBMTS, responses were categorized as (1) incorporating one's partner or (2) not incorporating one's partner based on whether their response included any reference to their partner.

Other Measures

Responses to the demographic questionnaire were analyzed as reported, but responses to the time allocation questionnaire were transformed. Proportions of time allocation for each activity were calculated by dividing the reported hours spent engaged in that activity by 168, the total number of hours in a week. Due to the lack of participants that were 21 years of age or older (n = 4 of 34), data regarding alcohol use and consumption of products with nicotine or THC were not incorporated in data analyses. Demographic information for all participants is shown in Table 5.

Results

Data analyses were conducted to address three central aims: (1) addressing limitations from Fleming and colleagues' (2021) study; (2) how instructions that are more or less referential and more or less descriptive of relevant contingencies participate in patterns of SRFs observed during TBMTS; (3) whether or not external measures external to TBMTS (i.e., *k'*, awareness of one's own and one's partner's TBMTS selections, demographic factors, and time allocation) could predict flexible and rigid patterns of SRFs; and (4) whether or not the measures internal to TBMTS that Fleming and colleagues (2021) found to be predictive of flexible and rigid patterns of SRFs would predict such patterns again. The results reported below will thus address each of these aims separately.

Comparisons with Fleming and Colleagues' (2021) Findings

Table 6 shows findings pertinent to assessing how changes of TBMTS in the thesis experiment addressed limitations of Fleming and colleagues' (2021) study. In their study, Fleming and colleagues found that, when points for correspondence decreased from CM to

NCM conditions for dyads in the ↓ Reward group, dyads generally completed the first NCM condition in less trials than the first CM condition. As they point out, this finding may be specifically related to decreasing points to the number of points awarded for noncorrespondence in CM conditions rather than decreasing points in itself. This finding was replicated in the thesis experiment within the D-NR group, the group with instructions nearly identical to those delivered in Flemings and colleagues' study, despite points decreasing to a value not previously contacted in TBMTS. The proportion of trials completed in the first two conditions that were completed in the first CM condition did not statistically differ between comparable groups (i.e., the D-NR and ↓ Reward groups) these studies according to a Mann-Whitney U test (D-NR Median = .68, $n = 3$; ↓ Reward Median = .71, $n = 7$; $U = 8$, $p = .667$). With respect to the number of trials necessary to complete the first CM condition, a Mann-Whitney U-test did not find a significant difference between the D-NR (Median = 158, $n = 6$) and the ↓ Reward (Median = 32, $n = 8$) groups ($U = 10$; $p = .078$), although more dyads failed to complete the first CM condition in the thesis experiment (5) than in Fleming and colleagues' study (1).

Differences between the thesis experiment and Fleming and colleagues' (2021) related to the number of stability trials were also examined. While all SRFs in stability trials of Fleming and colleagues' study were shown to be shared (i.e., dyad partners selected identical comparison stimuli in the presence of identical sample stimuli) and symmetrical (i.e., pairs of stimuli were selected in the presence of one another), shared and symmetrical SRFs could only be analyzed in 110 of 138 cases (79.10%) in the first CM condition because not all participants contacted all sample stimuli during stability trial in that condition; on average, both dyad partners only contacted 5.39 of 6 possible sample stimuli. To address this, the number of stability trials in the thesis experiment was increased from nine to 12 consecutive point-maximizing trials. In the thesis experiment, both dyad partners contacted 5.5 of 6 possible sample stimuli, allowing for 60 of 70 cases (83.33%) to be analyzed.

Given that two sets of arbitrary stimuli were used in the thesis experiment instead of just one like in Fleming and colleagues' (2021) study, established patterns of shared SRFs were examined across studies. Only one set of stimuli, *Sb* (or *Sh* as it was referred to in Fleming and colleagues' study) differed across studies; in Fleming and colleagues' study, Sb_1 was a circle, Sb_2 was a square, and Sb_3 was a triangle. Fleming and colleagues found that the set [$Sa_1|Sb_1$ (selecting Sa_1 and Sb_1 in the presence of one another), $Sa_2|Sb_3$, $Sa_3|Sb_1$] was the most common set established, occurring across 10 of 23 dyads (43.48%). This set was also the most common in the thesis experiment, occurring in 5 of 12 dyads (41.67%), but the set [$Sa_1|Sb_2$, $Sa_2|Sb_1$, $Sa_3|Sb_3$] was nearly just as common (4 of 12 dyads, 33.33%). Unlike in Fleming and colleagues' study where five different sets were shown to be established across dyads, only four different sets were established in the thesis experiment. Proportions of sets observed across each study are illustrated by Figure 3.

Participation of Instructions in Patterns of SRFs

The differential participation of instructions in patterns of SRFs observed during TBMTS was assessed by measuring the number of conditions completed, the number of trials completed in the first CM and NCM conditions, and the consistency of flexible and rigid patterns of SRFs, as shown in Table 7. If dyads did not complete at least two conditions, analyses involving trials completed in the first NCM condition were not included. Brown-Forsythe ANOVAs did not find significant differences in mean conditions completed (F (2, 9.683) = 0.229, p = .800), mean trials completed in the first CM condition (F(2, 12.88) = 0.059, p = .943), or mean trials completed in the first NCM condition (F(2, 1.187) = 1.892, p = .428). However, a significant difference was found in mean proportions of trials completed in the first two conditions that were CM trials (i.e., trials completed in the first CM condition divided by the number of trials completed in the first CM and NCM conditions; F (2, 4.261) = 24.15, p = .005). Post-hoc Donnett's T3 multiple comparison tests found significant differences in the mean proportion of trials completed in the first two conditions that were CM trials between the D-NR (M = .650) and D-R groups (M = .245; p = .016) and between the ND-NR (M = .715) and D-R groups (p = .015) but not between the D-NR and ND-NR groups (p = .787). With respect to flexible and rigid patterns of SRFs across groups, the D-R group was the only group in which both partners within a dyad demonstrated a flexible pattern of SRFs. In the case of dyad 008, both partner's responding in the first NCM condition deviated from the pattern established in the first CM condition, an observation that was not seen in any other dyad.

Differences across groups were also examined in responses to the TBMTS post-questionnaire. Table 8 shows each participant's reported strategy—or lack thereof—for maximizing points during TBMTS. Groups differed in the number of dyads in which at least one partner within a dyad mentioned their partner in describing their strategy. In three of six dyads in the D-NR group, four of five dyads in the D-R group, and one of six dyads in the ND-NR group, one participant in a dyad referenced their partner in their reported strategy. A Kruskal-Wallis test did not find this difference to be significant (H (2) = 4.163, p = .160). Table 9 shows correspondence scores for participants across all dyads. Kruskal-Wallis tests did not find any significant differences in correspondence scores for one's own selection of stimuli or that of their partner's with respect to any set of stimuli in any condition among groups ($p \geq .222$).

External Predictors of Flexible and Rigid Patterns of SRFs

Given the limited number of dyads included in data analyses (due to COVID and the failure of eight dyads to complete at least two conditions) and the significant difference in mean proportions of trials completed in the first two conditions that were CM trials, differences correlated with flexible and rigid patterns of SRFs were assessed by combining dyads in the D-NR and ND-NR groups without including dyads in the D-R group. Differences in age, gender (self-identification as male or female), languages spoken (whether a participant reported only speaking English or English and at least one other language), semesters completed, GPA, hours allocated toward eating, exercising, school work, sleeping, social media, social outings, occupation work, or other activities in a typical week, *k*', self-awareness of one's responding in NCM conditions, and awareness of one's partner responding in NCM conditions between dyad partners with lower (flexible) and higher (rigid) average overall consistency scores were assessed, as shown in Table 10. If all tested data passed Shapiro-Wilk tests, a paired *t*-test was use; if not, a Wilcoxon matched-pairs signed rank test was used. Across all factors assessed, none were found to be significantly different across flexible and rigid patterns of SRFs ($p \geq .125$).

As determining flexible and rigid patterns of SRFs was conditional on dyads completing the first CM and NCM conditions, similar analyses conducted between partners participating within flexible and rigid patterns of SRFs were conducted for participants in dyads that did or did not complete two or more conditions, as detailed in Table 11. Since dyads were not paired across groups, unpaired *t*-tests and Mann-Whitney *U* tests were used to assess differences instead of their paired counterparts. Awareness of one's own responding and one's partner's responding in NCM conditions was not compared between these groups, as dyads in one group necessarily did not complete the first NCM condition. Participants in D-R dyads were also excluded from this analysis. While most differences were not shown to be significant ($p \geq .053$), the proportion of hours allocated toward sleeping in a typical week was found to be greater for participants in dyads that completed less than two conditions (Mean = 0.341) than those in dyads that completed two or more conditions (Mean = 0.265; t (22) = 2.421; p = .024). Additionally, *k*' was found to be greater for participants in dyads that completed less than two conditions (Mean = 0.805) than those in dyads that completed two or more conditions (Mean = 0.370; t (19) = 2.508; p = .021)), as shown in Figure 4. With respect to this finding, it is especially noteworthy that *k'*-values of dyad partners were found to be positively correlated (Pearson's r = .738; p = .006).

Internal Predictors of Flexible and Rigid Patterns of SRFs

Initial deviations from response patterns with respect to *Sa* and *Sb* sample stimuli were correlated with the establishment of flexible patterns of SRFs. Table 7 shows which trials initial deviation in the pattern of SRFs established occurred and the partner that participated in the deviation, as well as which partner participated in a flexible pattern of SRFs (i.e., their responding changed across different conditions; the partner with the lower average overall consistency score). Of the dyads that completed at least two conditions, the partner who participated in a flexible pattern of SRFs was also the partner whose responding first deviated in the first NCM condition from the pattern established in the first CM condition with respect to both *Sa* (8/9 dyads) and *Sb* (9/9 dyads) sample stimuli. In the case of dyad 008, deviations with respect to both *Sa* and *Sb* sample stimuli were correlated with flexible responding despite P1 first deviating with respect to *Sa* sample stimuli and P2 first deviating with respect to *Sb* sample stimuli because both participants demonstrated flexible responding in the first NCM condition. In subsequent conditions completed, only P2 demonstrated flexible responding in dyad 008.

Flexible patterns of SRFs were also generally associated with the dyad partner whose responding throughout the first CM condition was more consistent with the terminal pattern of SRFs established in the first CM condition. Table 6 details observations of responding for each participant and dyad in the first CM condition. In five of the eight dyads that completed at least two conditions (and in which only one dyad member engaged in flexible responding), the dyad partner participating in a flexible pattern of SRFs responded more consistently throughout all trials in the first CM with the pattern established in the first CM than their partner. In no case did a dyad partner who participated in a flexible pattern of SRFs respond less consistently in the first CM condition with the pattern established in the first CM condition on the first opportunity to respond to each sample stimulus than their partner.

Discussion

The results from the thesis experiment suggest that patterns of SRFs are functionally related to instructions delivered at the beginning of TBMTS. Not only were references to one's partner more likely in groups that received instructions in which instructions were framed referentially, but dyads that received such instruction—and were able to complete two or more conditions—were more likely to complete more trials in the first NCM condition than in the first CM condition. This was not observed in any other group, supporting the conclusion that framing instructions referentially can foster sensitivity to changes across conditions. Although conclusions are limited due to the lack of participants restricted by COVID-19, the thesis experiment suggests that pattens of SRFs observed in TBMTS are related to instructions. Specifically, referential instructions seem to participate in patterns of SRFs that are shared relatively soon into TBMTS or relatively later, as some dyads in the D-R failed to complete two conditions.

The differential speeds at which dyads in the D-R completed—or failed to complete—the first CM condition is likely related to referential patterns of SRFs. If patterns of SRFs that dyad partners were participating in at the beginning of TBMTS were more similar, fewer trials would be required for one participant to adjust to their partner's pattern of SRFs.

Participants in the D-R group may have more easily established consistent response patterns given their initial orientation to such response patterns through instructions, but this may interfere with the establishment of shared patterns of SRFs if their responding was rigid. More dyads are needed to demonstrate a "U" shaped function in which referential instructions either facilitates or impedes the establishment of shared SRFs, but current evidence supports the assertion that referential instructions may do so given how similar initial participations in patterns of SRFs are between dyad partners.

While no measure external to TBMTS predicted flexible or rigid patterns of SRFs during TBMTS, some measures predicted a dyad's ability to establish a shared pattern of SRFs in the first CM condition, a pattern that is seemingly necessarily to participate in the first NCM condition (as shared SRFs were probable, but not necessary, to complete the first CM condition). The relation between lower delay discounting rates (i.e., higher self-control) and a dyad's likelihood to complete more than two conditions was expected, given previous theorizing on dyad's inability to cooperate during prisoner's dilemmas games (Green et al., 1995). The increased probability of flexible patterns of SRFs for participants that responded more consistently throughout the first CM condition with the pattern of shared SRFs that was to be established in the first CM condition is different than that observed in Fleming and colleagues' (2021) study in which they saw an increased probability of flexibility among those who responded less consistently in the first CM condition for dyads in the No Disc. group (the group that did not experience a change in points for correspondence between conditions). This suggests that flexible patterns of SRFs are functionally related to historical circumstances; flexibility is not just a property of a particular dyad partner's history but that of the dyad in TBMTS.

Although most demographic factors were not shown to be related with dyad performances during TBMTS, participants in dyads who completed less than two conditions reported sleeping more during a typical week than those who completed two or more conditions. Some existing evidence suggests that this relation should be seriously considered, especially given deviations observed in the current findings. While researchers have found sleep deprivation to be related to decreased cooperation in Chicken games (Lin et al., 2020) and others have found patterns of short sleeping durations to be related to higher delay discounting rates or impulsive choice (Massar & Chee, 2019), participants who completed less than two conditions in the thesis experiment typically reported sleeping longer and had higher delay discounting rates than other participants. As sleep deprivation may potentially be more likely among participants who sleep a larger proportion of the time, more data on more immediate sleep patterns prior to TBMTS is needed to assert that sleeping more is related to poorer performance in TBMTS.

A rather surprising finding was that delay discounting rates of dyad partners were correlated. Given that participants could not see each other complete the adjusting-amount delay discounting procedure and completing that procedure was the first thing participants did, similarities in delay discounting rates cannot be attributed to experimental interactions. Further study is necessary to determine why partner's delay discounting rates were correlated, but it is clear with the current data that similarity in delay discounting rates between partners does not necessarily facilitate performance in TBMTS. Said differently, absolute values of delay discounting rates seem to be more important than relative rates between partners in relation to TBMTS performance.

It is important to state that the thesis experiment replicated major findings from Fleming and colleagues' (2021) study despite changes to TBMTS. When dyads completed the first CM condition, SRFs were always shared and symmetrical. Decreasing the number of points for correspondence from CM to NCM conditions was associated with completing more trials in the first CM condition than in the first NCM condition, unless the dyad was in the D-R group. First deviations in the first NCM condition from patterns of SRFs established in the first CM condition were highly predictive of flexible patterns of SRFs being established. These repeated findings suggest that TBMTS findings are consistent despite use of different stimuli sets, different randomizations of sample stimuli, different stability criteria, and requiring participants to complete other tasks prior to completing TBMTS.

While some differences may not have been statistically significant, they should be noted nonetheless. Less dyads in the thesis experiment completed the first CM condition than in Fleming and colleagues' (2021) study, suggesting that changing stability criteria and stimuli used made TBMTS more difficult for at least some dyads. It is also important to note, though, that some differences were only marginal. Like in Fleming and colleagues' study, not all participants contacted all sample stimuli during stability trials in the first CM condition. Furthermore, a similar distribution of sets of shared symmetrical SRF pairs was observed across both studies. Given that participants were selected from the same pool (i.e., undergraduates participating in psychology courses at UNR) and presumably participate in similar cultural auspices, similar patterns in the variability of patterns of SRFs established during TBMTS may always be expected. However, stricter stability criteria should be used to ensure that participants contact all sample stimuli, such as requiring that participants continue to complete trials in a particular condition until both dyad partners have contacted each sample stimulus a certain number of times. As doing so may reasonably extended the duration of conditions, additional criteria—such as forcing dyads to complete at least two conditions before concluding TBMTS—may be necessary to guarantee viable data for conducting relevant analyses after.

Additional studies may consider two lines of thought. The first concerns other possible predictors of flexible and rigid patterns of SRFs. It is possible that flexible and rigid patterns are properties of social interactions and, therefore, only loosely related to psychological indices. However, there are other measures of flexibility that were not examined in the current study, most notably the Acceptance and Action Questionnaire-II (Bond et al., 2011) and the Wisconsin Card Sorting Task (Heaton et al., 1993). If responding on such were shown to be related to performances during TBMTS, research would be oriented to exploring relations between clinical applications like Acceptance and Commitment Therapy and the establishment of patterns of shared SRFs and cultural change with respect to such patterns. The second line of thought concerns altering dependency relations within TBMTS. In the current variation of TBMTS, referential dependencies are forced, but in more natural settings, reference is never forced; verbal behavior is, by definition, ineffective with respect to its immediate relation to environmental outcomes—it bears indirect influence (Skinner, 1957). As such, modifying TBMTS so that referential responses may occur without points being necessarily dependent upon them may increase the social validity of the task as well as impact types of flexibility observed.

These types of research questions—as well as the relevance of verbal repertoires for TBMTS—are vitally important for cultural studies in behavioral science. Behavioral science certainly consists of many sub-enterprises that are related to studying conventionality, such as those interested in verbal behavior and cultural events, but not a specific enterprise dedicated to understanding conventionality itself. The construction of such an enterprise may be useful in several ways for behavioral science as a whole. Behavioral scientists interested in cultural events have long advocated for a comprehensive, interdisciplinary science (Mattaini & Cihon, 2020). Social psychologists, anthropologists, and linguists could certainly participate in such a science, but their involvement would depend on how coherent the science is with their own approach. Formalizing an account of conventionality could foster expanded participation within the science, especially considering the attention conventionality has received in other sciences concerned with similar events (Bartlett, 1932; Hanson, 1975; Saussure, 1959). It could also help unite behavior-analytic sub-enterprises that have long operated in parallel to one another, particularly those interested in cultural and verbal relations. At the very least, understanding conventionality may lend itself to systematizing behavioral science or, in other words, formalizing a conventional system in which behavioral scientists can operate in.

References

Baia, F. H., & Sampaio, A. A. S. (2019). Distinguishing units of analysis, procedures, and processes in cultural selection: Notes on metacontingency terminology. *Behavior and Social Issues, 28*(1), 204–220. https://doi.org/10.1007/s42822-019-00017-8.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge University Press.

Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior, 20*(1), 137–153. https://doi.org/10.1901/jeab.1973.20-137.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 32,* 269-281. https://doi.org/10.1901/jeab.1974.22-231.

Bickel, W. K., Kowal, B. P., & Gatchalian, K. M. (2006). Understanding addiction as a pathology of temporal horizon. *The Behavior Analyst Today,* 7(1), 32-47. http://dx.doi.org/10.1037/h0100148.

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

Branch, M. N. (2006). Reactions of a laboratory behavioral scientist to a "Think Tank" on metacontingencies and cultural analysis. *Behavior and Social Issues, 15,* 6-10. https://doi.org/10.5210/bsi.v15i1.343.

Bruner, J., O'Connor, C., Rubin, H., & Huttegger, S. M. (2018). David Lewis in the lab: Experimental results on the emergence of meaning. *Synthese, 195,* 603-621. https://doi.org/10.1007/s11229-014-0535-x.

Catteeuw, D, & Manderick, B. (2013). The limits of reinforcement learning in Lewis signaling games. In *Proceedings of the 13th adaptive and learning agents workshop* (pp. 22-30). Maastricht University.

Costa, D., Nogueira, C. P. V., & Vasconcelos, L. A. (2012). Effects of communication and cultural consequences on choices combinations in INPDG with four participants. *Revista Latinoamericana de Psicología, 44*(1), 121–131.

Du, W., Green, L., & Myerson, J. (2002). Cross-cultural comparisons of discounting delayed and probabilistic rewards. *The Psychological Record, 52*, 479-492. https://doi.org/10.1007/BF03395199.

Epstein, R., & Skinner, B. F. (1981). The spontaneous use of memoranda by pigeons. *Behavior Analysis Letters, 1*, 241–246.

Epstein, R., Lanza, R. P., & Skinner, B. F. (1980). Symbolic communication between two pigeons (Columba livia domestica). *Science*, *207*(4430), 543–545. https://doi.org/10.1126/science.207.4430.543

Fleming, W., Ardila-Sánchez, J. G., & Hayes, L. J. (2021). Culture and contingencies: Molar insights for the metacontingency enterprise. *Mexican Journal of Behavior Analysis. 47*(2). TBA.

Fleming, W., & Hayes, L. J. (2021). Relations between description and experimentation in the metacontingency enterprise: An interbehavioral analysis. *Perspectives on Behavior Science, 44,* 417-472. https://doi.org/10.1007/s40614-021-00286-y.

Fleming, W., Thomas, J., Lopez, O., Locey, M. L., & Hayes, L. J. (2021). Evolution of cultural interbehavior in a turn-based matching-to-sample procedure. *The Psychological Record.* https://doi.org/10.1007/s40732-021-00485-8.

Fryling, M. J., Hayes, L. J., & Fleming, W. (2021). Interbehavioral psychology and organizational behavior management. In R. Houmanfar, M. Fryling, and M. Alavosius (Eds.), *Applied behavior science in organizations: Consilience of historical and emerging trends in organizational behavior management.* Springer.

Glenn, S. S. (1986). Metacontingencies in Walden Two. *Behavior Analysis and Social Action, 5*(1–2), 2–8. https://doi.org/10.1007/BF03406059.

Glenn, S. S. (2004). Individual behavior, culture, and social change. *The Behavior Analyst, 27*(2), 133–151. https://doi.org/10.1007/BF03393175.

Glenn, S. S., Malott, M. E., Andery, M. A. P. A., Benvenuti, M., Houmanfar, R. A., Sandaker, I., Todorov, J. C., Tourinho, E. Z., & Vasconcelos, L. A. (2016). Toward consistent terminology in a behaviorist approach to cultural analysis. *Behavior and Social Issues, 25*(1), 11–27. https://doi.org/10.5210/bsi.v25i0.6634.

Green, L., Price, P. C., & Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. *Journal of the Experimental Analysis of Behavior, 64,* 1-17. https://doi.org/10.1901/jeab.1995.64-1.

Hanson, F. A. (1975). *Meaning in culture.* Routledge.

Hayes, L. J., Adams, M. A., & Dixon, M. R. (1996). Causal constructs and conceptual confusions. *The Psychological Record, 46,* 97-112. https://doi.org/10.1007/BF03395214.

Hayes, L. J., & Fryling, M. J. (2018). Psychological events as integrated fields. *The Psychological Record, 68*(2), 273–277. https://doi.org/10.1007/s40732-018-0274-3.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. Plenum Press.

Heaton RK, Chelune GJ, Talley JL, Kay GG, Curtiss G (1993) Wisconsin Card Sorting Test Manual: Revised and expanded. Psychological Assessment Resources.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*(2), 243-266. https://doi.org/10.1901/jeab.1970.13-243.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*(1), 185-241. https://doi.org/10.1901/jeab.1996.65-185.

Huttegger, S., Skyrms, B., Tarres, P., & Wagner, E. (2014). Some dynamics of signaling games. *PNAS, 111,* 10873-10880. https://doi.org/10.1073/pnas.1400838111.

Johnson, M. W., & Bickel, W. K. (2008). An algorithm for identifying nonsystematic delay-discounting data. *Experimental and Clinical Psychopharmacology, 16*(3), 264–274. https://doi.org/10.1037/1064-1297.16.3.264.

Kantor, J. R. (1924). *Principles of psychology (Vol. I).* Principia Press.

Kantor, J. R. (1958). *Interbehavioral psychology.* Principia Press.

Kantor, J. R. (1982). *Cultural psychology.* Principia Press.

Lanza, R. P., Starr, J., & Skinner, B. F. (1982). "Lying" in the pigeon. *Journal of the Experimental Analysis of Behavior, 38*(2), 201–203. https://doi.org/10.1901/jeab.1982.38-201.

Leite, F. L., & de Souza, C. B. A. (2012). Metacontingencies, cultural selection and social/verbal environment. *Revista Latinoamericana de Psicología, 44*(1), 35-42.

Lewis, D. (1969). *Convention: A philosophical study.* Oxford University Press.

Lin, Y., Hu, P., Mai, Z., Jiang, T., Mo, L., & Ma, N. (2020). Sleep deprivation impairs cooperative behavior selectively: Evidence from prisoner's and chicken dilemmas. *Nature and Science of Sleep, 12,* 29-37. https://doi.org/10.2147/NSS.S237402.

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior, 96*(3), 427-439. https://doi.org/10.1901/jeab.2011.96-423.

Marr, M. J. (2006). Behavior analysis and social dynamics: Some questions and concerns. *Behavior and Social Issues, 15,* 57-67. https://doi.org/10.5210/bsi.v15i1.345.

Massar, S. A. A., & Chee, M. W. L. (2019). Sleep and delay discounting: Is sufficient sleep a cause or a manifestation of short-sighted choice? *Sleep, 42*(4), 1-3. https://doi.org/10.1093/sleep/zsz005.

Mattaini, M. A. (2006). Will cultural analysis become a science? *Behavior and Social Issues, 15,* 68-80. https://doi.org/10.5210/bsi.v15i1.380.

Mattaini, M. A., & Cihon, T. M. (Eds.) (2020). *Behavior science perspectives on culture and community.* Association for Behavior Analysis International.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *The Effects of Delay and of Intervening Events on Reinforcement Value* (Vol. 5, pp. 55–73). New Jersey: Erlbaum.

Rachlin, H. (1992). Teleological behaviorism. *American Psychologist, 47*(11), 1371–1382. https://doi.org/10.1037/0003-066X.47.11.1371.

Rachlin, H. (2017). In defense of teleological behaviorism. *Journal of Theoretical and Philosophical Psychology, 37*(2), 65-76. http://dx.doi.org/10.1037/teo0000060.

Sampaio, A. A. S., Araújo, L. A. S., Gonçalo, M. E., Ferraz, J. C., Alves Filho, A. P., Brito, I. S., Barros, N. M., & Calado, J. I. F. (2013). Exploring the role of verbal behavior in a new experimental task for the study of metacontingencies. *Behavior and Social Issues, 22*(1), 87–101. https://doi.org/10.5210/bsi.v22i0.4180.

Saussure, F. (1959). *Course in general linguistics.* (C. Bally, A. Sechehaye, & A. Reiglinger, Eds.; W. Baskin, Trans.). Philosophical Library.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*(1), 127-146. https://doi.org/10.1901/jeab.2000.74-127.

Skinner, B. F. (1937). Two types of conditioned reflex: A reply to Konorski and Miller.

Journal of General Psychology, 16, 272–279.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review, 57*(4), 193–216.

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.

Skinner, B. F. (1971). *Beyond freedom and dignity.* Knopf.

Skinner, B. F. (1981). Selection by consequences. *Science, 213*(4507), 501-504. https://doi.org/10.1126/science.7244649.

Skyrms, B. (2010). Signals: Evolution, learning, and information. Oxford University Press, New York.

Soares, P. F. R., Rocha, A. P. M. C., Guimarães, T. M. M., Leite, F. L., Andery, M. A. P. A., & Tourinho, E. Z. (2018). Effects of verbal and non-verbal cultural consequences on culturants. *Behavior and Social Issues, 27*(1), 31-46. https://doi.org/10.5210/bsi.v27i0.8252.

Ventura, R. (2019). Ambiguous signals, partial beleifs, and propositional content. *Synthese, 196,* 2803–2820. https://doi.org/10.1007/s11229-017-1580-z.

Table 1
Instructions Given to Each Group

D-NR	**In this study, you and your partner will earn points together. You will take turns being the first player, who selects a black symbol, and the second player, who selects a blue symbol. On each trial, the amount of points you will earn depends on (1) what blue symbol is presented to the first player and (2) what blue symbol is selected by the second player.** Nothing can be clicked when the cursor is a red No Sign; the cursor must be a white Pointer to click anything. Earning the most amount of points in the fewest amount of responses will result in earning the most amount of money. Half of the money earned by you and your partner will be given to you at the end of the study; the other half will be given to your partner. When the Continue button appears, click it to progress in the study.
D-R	**In this study, you and your partner will earn points together. You will take turns being the first player, who tells the second player what blue symbol was presented to you by sending a black symbol, and the second player, who selects a blue symbol after being shown the black symbol sent by the first player. It is up to you and your partner how you will use the black symbols to tell each other which blue symbols are presented to you. On each trial, the amount of points you will earn depends on (1) what blue symbol is presented to the first player and (2) what blue symbol is selected by the second player.** Nothing can be clicked when the cursor is a red No Sign; the cursor must be a white Pointer to click anything. Earning the most amount of points in the fewest amount of responses will result in earning the most amount of money. Half of the money earned by you and your partner will be given to you at the end of the study; the other half will be given to your partner. When the Continue button appears, click it to progress in the study.
ND-NR	**In this study, you and your partner will earn points.** Nothing can be clicked when the cursor is a red No Sign; the cursor must be a white Pointer to click anything. Earning the most amount of points in the fewest amount of responses will result in earning the most amount of money. Half of the money earned by you and your partner will be given to you at the end of the study; the other half will be given to your partner. When the Continue button appears, click it to progress in the study.

Differences in instructions are bolded.

Table 2
Stimuli Sets

	Sa	
		
Sa_1	Sa_2	Sa_3
	Sb	
		
Sb_1	Sb_2	Sb_3

Table 3
Demographic Questions

Factor	Question	Response Type	Response Options
Age	How old are you? Either type your answer in the textbox below and press Enter or click on the arrows to select your answer.	Text Entry/Numeric Counter	0-99
Gender	What gender do you identify as? Check all that apply.	Checkbox (+ Text Entry if "Other" was selected)	Female; Male; Prefer not to answer; Other
Spoken Languages	What languages do you speak fluently? Check all that apply	Checkbox (+ Text Entry if "Other" was selected)	English; Spanish; French; German; Japanese; Chinese; Korean; Italian; Russian; Other
Semesters Completed	How many semesters of college have you completed? If this is your first semester, type or select "0".	Text Entry/Numeric Counter	0-99
GPA	What is your current GPA? Approximate if you are not sure of your answer.	Text Entry/Numeric Counter	0.00-99.00

Differences in instructions are bolded.

Table 4

Delay Discounting Rates

		k' (R^2)	
Group	Dyad	P1	P2
D-NR	002	0.235 (0.944)	0.406 (0.700)
	005	1.369 (0.963)	0.763 (0.990)
	009	0.391 (0.912)	0.260 (0.954)
	011	1.142 (0.457)	2.068 (0.416)
	013	0.039 (0.932)	0.107 (0.992)
	016	-[x]	0.034 (0.939)
Group Median		0.391 (0.932)	0.333 (0.946)
D-R	003	-[x]	0.073 (0.914)
	004	-[x]	-[x]
	008	0.014 (0.723)	0.111 (0.982)
	010	0.313 (0.871)	0.293 (0.924)
	014	0.186 (0.679)	0.134 (0.854)
Group Median		0.186 (0.723)	0.122 (0.919)
ND-NR	001	0.175 (0.858)	0.218 (0.804)
	006	0.401 (0.795)	0.687 (0.814)
	007	0.369 (0.977)	0.006 (0.938)
	012	-[x]	0.625 (0.888)
	015	0.558 (0.842)	-[x]
	017	0.058 (-0.122)	0.379 (0.546)
Group Median		0.447 (0.842)	0.379 (0.814)
Total Median		0.313 (0.858)	0.260 (0.914)

[x]Delay discounting data did not meet inclusion criteria

Table 5
Demographic Information

Group	Dyad	P	Age	Gender	Languages	Semesters	GPA	Alcohol	Nicotine	THC
D-NR	002	P1	21	Female	English	5	3.93	2 days	0 days	0 days
		P2	18	Female	English	1	4.00	-	-	-
	005	P1	19	Female	English	1	4.00	-	-	-
		P2	21	Female	English	5	3.9	0 days	0 days	0 days
	009	P1	18	Female	English	1	3.25	-	-	-
		P2	19	Female	English	1	2.1	-	-	-
	011	P1	18	Male	English	1	0.4	-	-	-
		P2	18	Male	English	1	0.7	-	-	-
	013	P1	18	Female	English, Russian	1	4.0	-	-	-
		P2	19	Male	English	3	3.7	-	-	-
	016	P1	18	Female	English	1	3.5	-	-	-
		P2	18	Female	English	0	3.5	-	-	-
D-R	003	P1	19	Male	English, Spanish	3	3.90	-	-	-
		P2	19	Female	English	3	3.98	-	-	-
	004	P1	18	Female	English	1	2.5	-	-	-
		P2	20	Male	English, Spanish	2	2.0	-	-	-
	008	P1	18	Female	English	1	3.0	-	-	-
		P2	18	Female	English	1	4.0	-	-	-
	010	P1	19	Female	Spanish	1	3.12	-	-	-
		P2	21	Male	English	0	3.7	0 days	7 days	0 days
	014	P1	19	Female	English	1	1.6	-	-	-
		P2	19	Female	English	3	3.7	-	-	-
ND-NR	001	P1	18	Male	English	1	3.8	-	-	-
		P2	19	Female	Vietnamese	0	2.8	-	-	-
	006	P1	19	Female	English	1	2.0	-	-	-
		P2	19	Female	English	1	3.5	-	-	-
	007	P1	19	Male	English	1	2.5	-	-	-
		P2	19	Female	English	3	3.5	-	-	-
	012	P1	22	Female	English	7	3.94	1 day	4 days	0 days
		P2	20	Female	English	3	3.3	-	-	-
	015	P1	18	Female	English, Spanish	1	3.7	-	-	-
		P2	19	Female	English, Spanish	1	3.5	-	-	-
	017	P1	18	Male	English	1	3.8	-	-	-
		P2	18	Female	English	1	3.05	-	-	-

P denotes partner, either P1 or P2. Participants without values for typical number of days in a week in which Alcohol, Nicotine, or THC is used were not asked because of their age

Table 6

Data Relevant to Comparisons between the Thesis Experiment and Fleming and Colleagues' (2021) Experiment

Group	Dyad	Trials Completed in First CM	Trial Prop. (1st CM Trials / [1st CM + NCM Trials])	Non-Stability SRFs Consistent with Stability SRFs		Stability SRFs Occurring on First Opportunity		Different Samples Contacted by Each Partner During Stability Trials	Symmetrical SRFs Pairs
				P1	P2	P1	P2		
D-NR	002	22	0.58	5/10	5/10	4/6	4/6	5	$Sa_1\|Sb_1, Sa_2\|Sb_3, Sa_3\|Sb_2$
	005	218	-	-	-	-	-	-	-
	009	85	0.69	45/73	39/73	5/6	3/6	5	$Sa_1\|Sb_1, Sa_2\|Sb_3, Sa_3\|Sb_2$
	011	267	-	-	-	-	-	-	-
	013	99	0.68	31/87	30/87	3/6	2/6	6	$Sa_1\|Sb_2, Sa_2\|Sb_1, Sa_3\|Sb_3$
	016	217	-	-	-	-	-	-	-
D-R	003	38	0.24	10/26	6/26	6/6[a]	4/6	4	$Sa_1\|Sb_2, Sa_2\|Sb_3, Sa_3\|Sb_1$
	004	166	-[x]	125/154	104/154	3/6	3/6	6	$Sa_1\|Sb_1, Sa_2\|Sb_3, Sa_3\|Sb_2$
	008	14	0.25	2/2	1/2	5/6	6/6	6	$Sa_1\|Sb_2, Sa_2\|Sb_3, Sa_3\|Sb_1$
	010	126	-	-	-	-	-	-	-
	014	200	-[x]	98/188	188/188	0/6	6/6	6	$Sa_1\|Sb_3, Sa_2\|Sb_2, Sa_3\|Sb_1$
ND-NR	001	43	0.74	23/31	24/31	1/6	5/6	5	$Sa_1\|Sb_1, Sa_2\|Sb_3, Sa_3\|Sb_2$
	006	31	0.53	8/19	7/19	3/6	2/6	6	$Sa_1\|Sb_1, Sa_2\|Sb_3, Sa_3\|Sb_2$
	007	105	0.86	38/93	66/93	2/6	2/6	6	$Sa_1\|Sb_2, Sa_2\|Sb_1, Sa_3\|Sb_3$
	012	154	-[x]	101/142	52/142	1/6	1/6	6	$Sa_1\|Sb_2, Sa_2\|Sb_1, Sa_3\|Sb_3$
	015	87	0.73	36/75	25/75	2/6	2/6	5	$Sa_1\|Sb_2, Sa_2\|Sb_1, Sa_3\|Sb_3$
	017	198	-	-	-	-	-	-	-

CM = correspondence maximization condition; NCM = noncorrespondence maximization condition;

Table 7

Conditions Completed, Trials Completed in the First Two Conditions, and Flexibility Across Conditions

						Trial in NCM of First Deviation from Set Established in CM by Sample Stimuli (Partner Who Deviated)		Average Overall Consistency Scores		
Group	Dyad	Conditions Completed	Trials in First CM	Trials in First NCM	Trial Prop. (CM/CM+NCM)	*Sa*	*Sb*	P1	P2	Flexible Partner
D-NR	002	14	22	16	0.58	5 (P2)	2 (P2)	1.00	0.43	P2
	005	0	218	-	-	N/A	N/A	-	-	-
	009	5	85	39	0.69	23 (P1)	26 (P1)	0.53	1.00	P1
	011	0	267	-	-	N/A	N/A	-	-	-
	013	7	99	46	0.68	35 (P1)	36 (P1)	0.59	1.00	P1
	016	0	217	-	-	N/A	N/A	-	-	-
D-R	003	6	38	119	0.24	66 (P1)	109 (P1)	0.55	1.00	P1
	004	1	166	11	-^{x}	N/A	9 (P1)	-	-	-
	008	9	14	43	0.25	2 (P1)	16 (P2)	0.98	0.61	P1/P2-P2
	010	0	126	-	-	N/A	N/A	-	-	-
	014	1	200	4	-^{x}	N/A	N/A	-	-	-
ND-NR	001	5	43	15	0.74	5 (P2)	4 (P2)	1.00	0.52	P2
	006	4	31	27	0.53	17 (P1)	16 (P1)	0.45	1.00	P1
	007	4	105	17	0.86	7 (P1)	8 (P1)	0.42	1.00	P1
	012	1	154	26	-^{x}	N/A	N/A	-	-	-
	015	3	87	33	0.73	3 (P1)	7 (P2)	1.00	0.64	P2
	017	0	198	-	-	N/A	N/A	-	-	-

CM = correspondence maximization condition; NCM = noncorrespondence maximization condition; N/A = no deviation occurred. A dash between partners indicates that flexibility switched between conditions. A forward slash between partners indicates that both partners demonstrated flexibility in the same NCM condition; - denotes that the value could not be calculated or identified because the dyad did not complete the first CM condition (Dyads 005, 010, 011, 016, 017) or the first NCM condition (Dyads 004, 012, 014).

Table 8
Point-Maximization Strategies Reported by Each Participant in the TBMTS Post-Questionnaire

Group	Dyad	P1	P2	Partner?
D-NR	002	I started to correlate one blue figure with one black figure that gave me the most amount of points each time.	At first I thought there was a certain amount of tries before it switched over to the "other" pattern so I would count how many times we were on one pattern before switching over. It worked for about two or three rounds, but then it stopped being the same amount of clicks per each pattern so there was no way for me to keep track anymore.	N
	005	NO	I tried to match the images together	N
	009	I began to associate symbols with other symbols, but when **we** began getting a smaller amount of points, I "broke the mold" and chose something completely different to get the most amount of points, it worked.	As time went on, I kept choosing the same symbols to match what I had been previously choosing, and that seemed to work in making the most points.	Y
	011	To be completely honest I had no idea what I was doing, I didnt understand the directions	I clicked the middle option everytime in hope **my partner** would catch on.	Y
	013	I figured out which paired shapes gave me the most points and selected those. Once the point value changed, I found another pair that gave me more points.	I tried to figure out what blue **she** was picking for each black and then I eventually figured it out and it went smoothly from there.	Y
	016	No. I just clicked randomly, I did not really find a pattern.	i tried seeing which shapes had lines crossing and loops in them to match to the other shape	N
D-R	003	When the amount increased I changed which blue symbols correlated to black symbols when I realized the changes I made were resulting in a higher amount for the changed response.	I could tell that **my partner** and I came up with a pattern of which blue shapes went with which black shapes, so in order to get the most points, I just followed the pattern.	Y
	004	Everytime I tried to match the blue symbol with the same black symbol with hopes that **my partner** would catch on	I did not do that, i simply attempted to make connections between the shapes	Y
	008	If it seemed like the points changed for the same answers, I tried changing the symbols I chose to see if that brought the points back up. If not I just went back to the 'tried and true' system.	My strategy was to focus on the points that were allocated for each response that I had made. The less the points were, the more I knew that I had to change my response to that specific image in order to gain more points.	N
	010	I did not use any strategy	Tried to guide the other to pick symbols that matched with the current order of the blue. and picked the black that corisponded with the last location of blue. Was still confused on how we were supposed to let **the other** know about the locations.	Y
	014	I just noticed which shapes matched which to maximize the points i could potentislly earn.	I correlated each blue pattern with one of the black patterns and stuck with those pairs the whole time. I thought that **my patner** should catch on to the pattern, and if she choses to change **her own pattern** to mine then I shouldn't change mine up.	Y

Bolded text denotes references to a participant's partner. Y = Yes, a participant did refer to their partner; N = No, a participant did not refer to their partner.

Table 8

Point-Maximization Strategies Reported by Each Participant in the TBMTS Post-Questionnaire (Cont.)

Group	Dyad	P1	P2	Partner?
ND-NR	001	All I did was continually click the same pictures that i associated with the earning of more points	Yes, I use to imagine in te spatial in my mind about the shape	N
	006	I used memory to remember which pieces would give me the highest points possible	I matched the same two shapes together eveytime the other one popped up	N
	007	Recognized a pattern which allowed for the answers worth the highest amount of points to be memorized.	The middle circle was the squiggly. The left circle was the shape that looks like a U. The big circle was the shape in zig zags.	N
	012	At first, if we matched and got 25 points then I would just do the same symbol the next time it came up. Eventually she caught on and we had 3 set pairs.	At first I tried to always only click the right one. But I wasnt sure if **the other persons** images were mixed up so I tried to match the shapes together. I only clicked the same thing for each image every time.	Y
	015	what I did was that I keep track of the pictures that matched.	I gave each shape a name to keep track of which one went with which.	N
	017	The only thing that I could fiugre out is that when the one image showed that was the answer to the following question.	I tried to memorize the symbols on each question and match them to where they previously were.	N

Bolded text denotes references to a participant's partner. Y = Yes, a participant did refer to their partner; N = No, a participant did not refer to their partner.

Table 9
Correspondence Scores for Each Participant

		P1								P2							
		Participant's Responding				Partner's Responding				Participant's Responding				Partner's Responding			
		CM		NCM		CM		NCM		CM		NCM		CM		NCM	
Group	Dyad	*Sa*	*Sb*	*Sa*	*Sb*	*Sa*	*Sb*	*Sa*	*Sb*	*Sa*	*Sb*	*Sa*	*Sb*	*Sa*	*Sb*	*Sa*	*Sb*
D-NR	002	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0
	005	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	009	1	1	1	1	1	1	.2	0	1	1	1	1	1	1	0	0
	011	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	013	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0
	016	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Median	1	1	1	1	1	1	.07	0	1	1	1	1	1	1	0	0
D-R	003	1	1	.5	.5	1	1	.3	0	1	1	1	1	1	1	0	0
	004	1	1	-	-	1	1	-	-	.6	.75	-	-	1	1	-	-
	008	1	1	1	1	1	.6	0	.4	1	1	1	1	1	1	0	0
	010	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	014	1	1	-	-	1	1	-	-	1	1	-	-	1	1	-	-
	Median	1	1	.75	.75	1	.9	.15	.2	.9	.94	1	1	1	1	0	0
ND-NR	001	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0
	006	.5	.5	.6	1	1	1	0	0	1	1	1	1	1	1	0	0
	007	1	1	1	1	1	1	0	0	1	1	1	1	1	1	.4	0
	012	1	1	-	-	1	1	-	-	1	1	-	-	1	1	-	-
	015	1	1	1	1	1	1	0	0	.6	.5	.6	.5	.3	.5	.4	.4
	017	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Median	.9	.9	.9	1	1	1	0	0	.73	.9	.9	.88	.86	.88	.2	.1

Table 10
Measuring Differences between Individuals by Flexible and Rigid Patterns of SRFs

Group	SRF Pattern Type	Pass Shapiro-Wilk?	Mean (SD)	Median (IQR)	*t*	*W*	*p*
Age	Flexible	No	18.57 (0.535)	19 (18-19)	-	5	.750
	Rigid	No	19.00 (1.000)	19 (18-19)			
Gender	Flexible	No	1.857 (0.378)	2 (2-2)	-	-2	>.999
	Rigid	No	1.714 (0.488)	2 (1-2)			
Languages Spoken	Flexible	No	1.429 (0.535)	1 (1-2)	-	-3	.500
	Rigid	No	1.143 (0.378)	1 (1-1)			
Semesters Completed	Flexible	No	0.857 (0.378)	1 (1-1)	-	10	.125
	Rigid	No	2.143 (1.574)	1 (1-3)			
GPA	Flexible	Yes	3.150 (0.758)	3.25 (2.5-4)	-	8	.547
	Rigid	No	3.461 (0.620)	3.7 (3.5-3.8)			
Eating	Flexible	Yes	0.104 (0.045)	0.089 (0.059-0.137)	0.634	-	.549
	Rigid	Yes	0.120 (0.038)	0.125 (0.083-0.161)			
Exercising	Flexible	Yes	0.067 (0.054)	0.083 (0.012-0.125)	1.052	-	.333
	Rigid	Yes	0.051 (0.023)	0.048 (0.030-0.077)			
School Work	Flexible	Yes	0.199 (0.071)	0.202 (0.125-0.250)	0.466	-	.658
	Rigid	Yes	0.218 (0.089)	0.208 (0.161-0.298)			
Sleeping	Flexible	Yes	0.253 (0.074)	0.268 (0.208-0.316)	0.454	-	.666
	Rigid	Yes	0.276 (0.113)	0.304 (0.238-0.333)			
Social Media	Flexible	Yes	0.144 (0.098)	0.125 (0.060-0.250)	.1338	-	.229
	Rigid	Yes	0.098 (0.052)	0.119 (0.042-0.143)			
Social Outings	Flexible	Yes	0.076 (0.060)	0.071 (0.030-0.119)	0.455	-	.665
	Rigid	Yes	0.065 (0.031)	0.071 (0.030-0.095)			
Occupational Work	Flexible	No	0.070 (0.106)	0.060 (0.000-0.071)	-	-4	.688
	Rigid	No	0.028 (0.044)	0.000 (0.000-0.048)			

Gender: 1 = Male; 2 = Female (no other options were selected for gender); Languages Spoken: 1 = English only; 2 = English and at least one other language.

Table 10
Measuring Differences between Individuals by Flexible and Rigid Patterns of SRFs (Cont.)

Group	SRF Pattern Type	Pass Shapiro-Wilk?	Mean (SD)	Median (IQR)	*t*	*W*	*p*
Other	Flexible	Yes	0.088 (0.077)	0.077 (0.000-0.167)	0.747	-	.483
	Rigid	Yes	0.144 (0.156)	0.113 (0.000-0.232)			
k'	Flexible	No	0.304 (0.147)	0.380 (0.173-0.402)	-	-7	.438
	Rigid	Yes	0.273 (0.252)	0.235 (0.090-0.474)			
Self-Awareness of NCM	Flexible	No	0.907 (0.174)	1.000 (0.800-1.000)	-	3	.500
	Rigid	Invalid	1.000 (0.000)	1.000 (1.000-1.000)			
Partner Awareness of NCM	Flexible	No	0.071 (0.150)	0.000 (0.000-0.100)		-2	.750
	Rigid	No	0.029 (0.076)	0.000 (0.000-0.000)			

Invalid denotes that the test could not be performed because all participants had the same value.

Table 11
Measuring Differences between Dyads that Completed Less than Two Conditions and Dyads that Completed Two or More Conditions

Group	Conditions Completed	Pass Shapiro-Wilk?	Mean (SD)	Median (IQR)	*t*	*W*	*p*
Age	< 2	No	18 (1.491)	18 (18-20.25)	-	64.5	.752
	2+	No	18.79 (0.802)	19 (18-19)			
Gender	< 2	No	1.7 (0.483)	2 (1-2)	-	64	>.999
	2+	No	1.786 (0.426)	2 (1.75-2)			
Languages Spoken	< 2	No	1 (0)	1 (1-1)	-	50	.114
	2+	No	1.286 (0.469)	1 (1-2)			
Semesters Completed	< 2	No	2.1 (2.234)	1 (1-3)	-	64	.726
	2+	No	1.5 (1.286)	1 (1-1.5)			
GPA	< 2	Yes	3.009 (1.332)	3.5 (2.463-3.910)	-	68.5	.943
	2+	No	3.306 (0.684)	3.5 (2.725-3.833)			
Eating	< 2	Yes	0.085 (0.050)	0.068 (0.040-.0127)	1.465	-	.157
	2+	Yes	0.112 (0.041)	0.107 (0.082-0.147)			
Exercising	< 2	Yes	0.043 (0.022)	0.042 (0.028-0.054)	1.100	-	.283
	2+	Yes	0.059 (0.041)	0.054 (0.022-0.083)			
School Work	< 2	Yes	0.145 (0.069)	0.140 (0.077-0.220)	2.044	-	.053
	2+	Yes	0.208 (0.078)	0.208 (0.152-0.262)			
Sleeping	< 2	Yes	0.341 (0.043)	0.333 (0.316-0.369)	2.421	-	**.024**
	2+	Yes	0.265 (0.093)	0.283 (0.222-0.320)			
Social Media	< 2	Yes	0.148 (0.077)	0.152 (0.086-0.188)	0.828	-	.416
	2+	Yes	0.121 (0.079)	0.122 (0.055-0.158)			
Social Outings	< 2	Yes	0.079 (0.037)	0.086 (0.061-0.095)	0.484	-	.633
	2+	Yes	0.071 (0.046)	0.071 (0.030-0.098)			
Occupational Work	< 2	No	0.054 (0.082)	0.000 (0.000-0.098)	-	69	.963
	2+	No	0.049 (0.081)	0.015 (0.000-0.063)			
Other	< 2	Yes	0.105 (0.092)	0.110 (0.000-0.165)	-	69.5	.989
	2+	No	0.116 (0.121)	0.101 (0.000-0.173)			
k'	< 2	Yes	0.805 (0.696)	0.694 (0.138-1.312)	2.508	-	**.021**
	2+	Yes	0.370 (0.281)	0.297 (0.152-0.497)			

Bolded p-values are significant (< .05).Only dyads in the D-NR and ND-MR groups are included; Participants that completed < 2 conditions: n = 10; those that completed 2+ conditions: n = 14

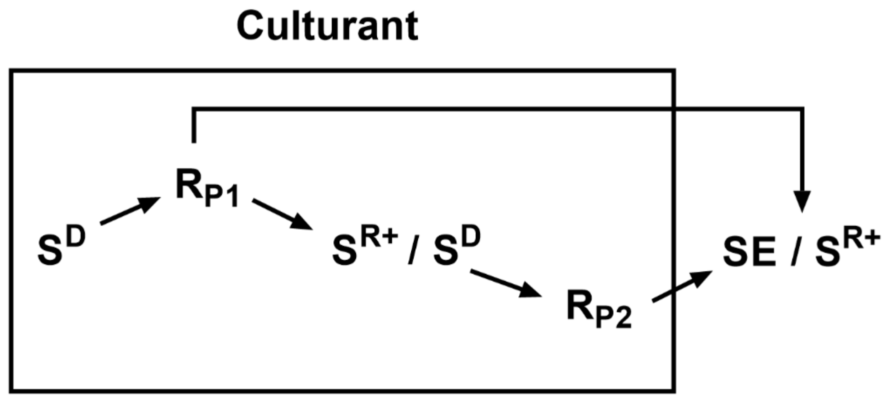

Figure 1. A metacontingency. S^D : discriminative event; R_{P1} : operant event of Person 1; R_{P2} : operant event of Person 2; S^{R+} : reinforcing event; SE : selecting event. The box denotes the boundary of the culturant iteration. Arrows denote dependency relations.

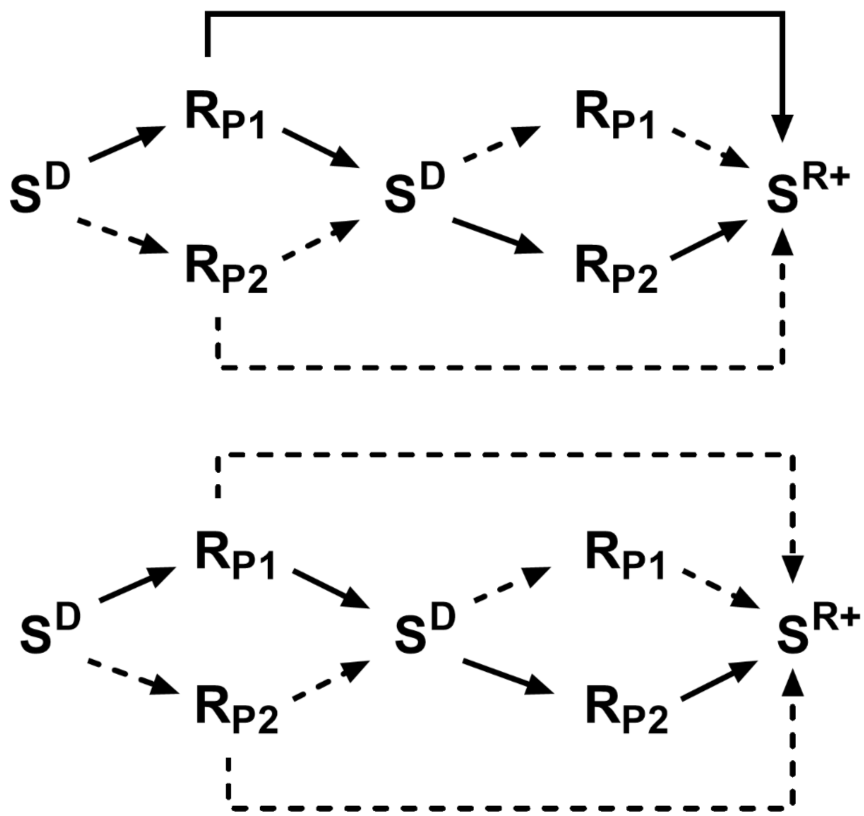

Figure 2. Coordinated operant contingencies. S^D : discriminative event; R_{P1} : operant event of Person 1; R_{P2} : operant event of Person 2; S^{R+} : reinforcing event. The box denotes the boundary of the culturant iteration. Solid arrows denote dependency relations. Dotted arrows denote correlated but not dependent relations between events.

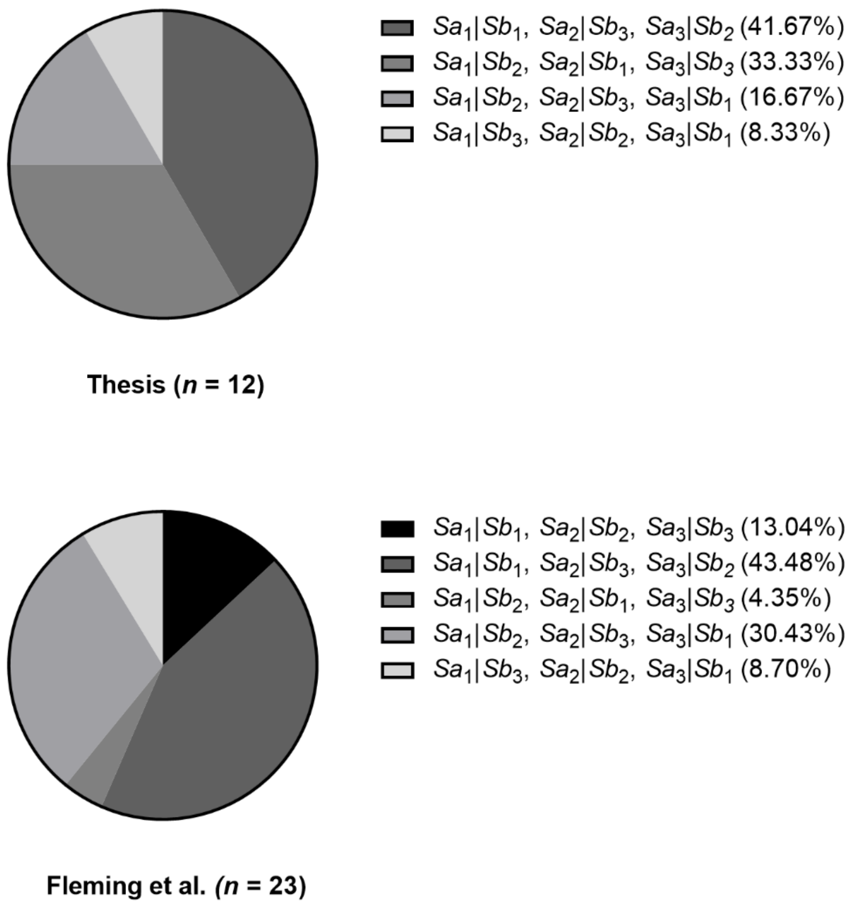

Figure 3. Proportions of sets of symmetrical and shared SRFs established in the first CM condition among dyads in the thesis experiment (*top*) and Fleming and colleagues' (2021) study (*bottom*).

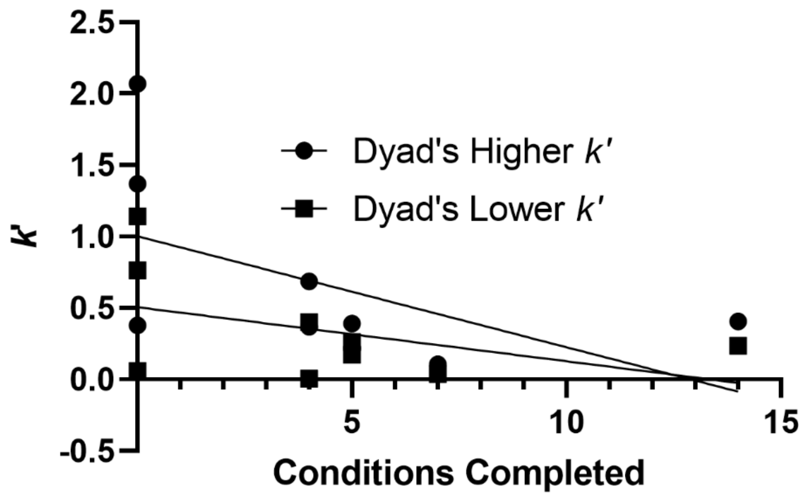

Figure 4. Values of *k*' as a function of conditions completed for participants in the D-NR and ND-NR groups. Circles denote the lower *k*'-value and squares denote the higher *k*'-value among partners in each dyad. Lines represent simple linear regressions, neither of which were significant ($p \geq .129$)