

COMPARING TWO METHODS OF QUANTIFYING RESPONSE-STIMULUS
RELATIONS: CONTINGENCY SPACE ANALYSIS AND YULE'S Q

By

Blair Lloyd

Thesis

Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

in

Special Education

December, 2011

Nashville, TN

Approved:

Professor Craig H. Kennedy

Professor Paul J. Yoder

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	iii
LIST OF FIGURES.....	iv
Chapter	
I. INTRODUCTION.....	1
II. METHOD.....	12
Subjects.....	12
Apparatus.....	12
Procedure.....	13
Sequential analysis.....	15
Correlational analysis.....	16
III. RESULTS.....	18
IV. DISCUSSION.....	38
REFERENCES.....	45

LIST OF TABLES

Table	Page
1. Construction of 2 x 2 Contingency Tables Using Lag-1 Event Sequential Analysis.....	7
2. Calculations of CSA Probabilities and YQ Statistics from the 2 x 2 Contingency Tables	8
3. Calculations of Relative Risk Ratios and Difference Scores from the CSA Probabilities	9
4. Correlations Among Indices of Contingency	27
5. Correlation Coefficients Between Each Index of Contingency and Total Frequency Target Responses	29

LIST OF FIGURES

Figure	Page
1. Graphic Representation of Contingency Space	5
2. Intended and Observed Mean Points in Contingency Space per Condition across Subjects	19
3. Examples of Undifferentiated Indices of Contingency and Undifferentiated Responding Between Experimental Conditions	21
4. Mean Points in Contingency Space for Subjects 2, 6, 10, and 14	23
5. Yule's Q and Relative Risk Ratios for Subjects 2, 6, 10, and 14	24
6. Relative Risk Ratios and Difference Scores for Subjects 2, 6, 10, and 14	25
7. Target Responses per Min for Subjects 2, 6, 10, and 14	26
8. Examples of Positive Linear Relations between Indices of Contingency and Total Frequency Target Responses	30
9. Examples of Negative Linear Relations between Indices of Contingency and Total Frequency Target Responses	33
10. Examples of Quadratic Relations between Indices of Contingency and Total Frequency Target Responding	34
11. Example of Within-Subject Variability Among Relations between Indices of Contingency and Total Frequency Target Responding	36

CHAPTER I

INTRODUCTION

Analogue functional analysis (AFA) methodology (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994) is the current standard for experimentally identifying variables maintaining problem behavior for individuals with disabilities. However, there are certain circumstances in which highly controlled experimental analyses may not be appropriate. For cases of severe self-injury or aggression, temporarily increasing these behaviors to confirm reinforcing variables may not be ethical (Durand, 1993). In other cases, experimental analyses may not be feasible given the setting and available resources (Horner, 1994).

Even when AFAs are feasible and ethically sound, the outcomes may not accurately reflect the function of problem behavior as it occurs in the natural environment. Some individuals may not engage in problem behavior during contrived test conditions, resulting in false negative findings (Wacker, Berg, Harding, & Cooper-Brown, 2004). False negatives may be due to isolating single antecedents and consequences in test conditions when in fact a combination of variables more reliably evoke or maintain problem behavior (Call, Wacker, Ringdahl, & Boelter, 2005; Dolezal & Kurtz, 2010). Idiosyncratic variables in the individual's natural environment may also contribute to false negative AFA outcomes if these variables are not included in test conditions (Carr, Yarbrough, & Langdon, 1997; Lang et al., 2009; McComas, Hoch, Paone, & El-Roy, 2000; Ringdahl & Sellers, 2000).

False positive outcomes are also possible for cases in which a consequence is not provided contingent on problem behavior in the natural environment, yet is identified as a reinforcer in an experimental analysis (Shirley, Iwata, & Kahng, 1999). Both false negative and false positive AFA outcomes seem to result from a mismatch between AFA test conditions and an individual's environment, yet the same set of test conditions (i.e., social positive reinforcement in the form of adult attention, social negative reinforcement in the form of escape from demands, automatic reinforcement, and a control condition) are included in the majority of AFAs (Hanley, Iwata, & McCord, 2003).

For cases in which AFAs may be inappropriate, infeasible, or do not provide clear results, descriptive assessments may be the best alternative. Descriptive assessments consist of direct observations in the environment in which problem behavior occurs during which data are collected on behavior-environment interactions (Bijou, Peterson, & Ault, 1968; Mace & Lalli, 1991). Data collected during descriptive assessments may be used to quantify response-stimulus relations and identify contingencies between problem behavior and consequences in the natural environment. These descriptive data may be used to form hypotheses concerning which variables are functioning as reinforcers.

The term contingency has been defined broadly as a relation between two or more events with varying probabilities (Vollmer, Borrero, Wright, Van Camp, & Lalli, 2001). Within reinforcement theory, contingencies are defined by two transitional probabilities: the probability of a stimulus given a response and the probability of a stimulus given no response (Catania, 1971; 2007). Contingent relations, in which a stimulus may occur both in the presence and absence of a response, may be contrasted with dependent relations, in which a stimulus occurs only in the presence of a response and never in its absence.

Contingent relations are more likely to be identified in natural settings, whereas dependent relations are characteristic of experimental analyses (e.g., AFA test conditions). Positive contingencies are identified when a stimulus occurs more often in the presence of a response than in its absence. Negative contingencies are identified when a stimulus occurs more often in the absence of a response than in its presence. Although a positive contingency may be necessary for reinforcement (Hammond, 1980; Martens, DiGennaro, Reed, Szczech, & Rosenthal, 2008), the precise conditions sufficient to produce a reinforcement effect remain unknown (Vollmer et al., 2001). Therefore, calculating contingencies via descriptive assessments may be used to suggest, but not confirm, the function(s) of problem behavior.

Several strategies for quantifying response-stimulus relations have been proposed and applied within the literature on descriptive assessments via sequential analysis. Methods of sequential analysis are used to determine whether the occurrence of one observed event is associated with the subsequent occurrence of another observed event within one or more observation sessions (Yoder & Symons, 2010). Sequential analysis methods may be applied to any pair of observed events. For the purpose of identifying potential reinforcers for problem behavior, the primary events of interest include a target response (i.e., problem behavior) and a change in the status of a stimulus (e.g., adult attention).

McComas et al. (2009) reviewed several methods of calculating contingencies in natural environments, including simple probabilities, transitional probabilities, and Yule's Q (YQ; Yule & Kendall, 1957). The authors identified simple probabilities as an insufficient and imprecise method of calculating contingency. Simple probabilities do not

control for the base rates of either the response or consequent stimulus and are therefore uninterpretable. Transitional probabilities in isolation are also insufficient, because transitional probabilities only control for the base rate of the response, but not for the consequent stimulus. YQ, a linear transformation of the odds ratio, is the recommended method for statistically quantifying sequential associations between events (Bakeman, McArthur, & Quera, 1996; Yoder, Short-Meyerson, & Tapp, 2004). YQ is bounded by -1.0 and 1.0 and is therefore more interpretable than the unbounded odds ratio. The YQ statistic controls not only for the base rates of each event, but for the chance occurrence of the two events occurring in sequence (Yoder & Symons, 2010). Positive YQ values indicate an observed sequential frequency greater than expected by chance. Negative YQ values indicate an observed sequential frequency less than expected by chance. Zero YQ values indicate an observed sequential frequency equal to that expected by chance (Yoder & Symons, 2010).

Contingency space analysis (CSA; Gibbon, Berryman, & Thompson, 1974; Matthews, Shimoff, & Catania, 1987; Martens et al., 2008) is a method of quantifying response-stimulus relations consistent with Catania's (1971; 2007) definition of contingency. CSA involves a comparison between two transitional probabilities: the probability of a stimulus given a target response ($p(S/R)$) and the probability of a stimulus in the absence of a target response ($p(S/\bar{R})$). The CSA probabilities are plotted as points in contingency space, with the y -axis representing the $p(S/R)$ and the x -axis representing the $p(S/\bar{R})$. Points located on, or near, the diagonal line between the x - and y -axes represent stimuli independent of the target response (i.e., the probability of the stimulus is the same whether or not the target response occurs; see Figure 1). The

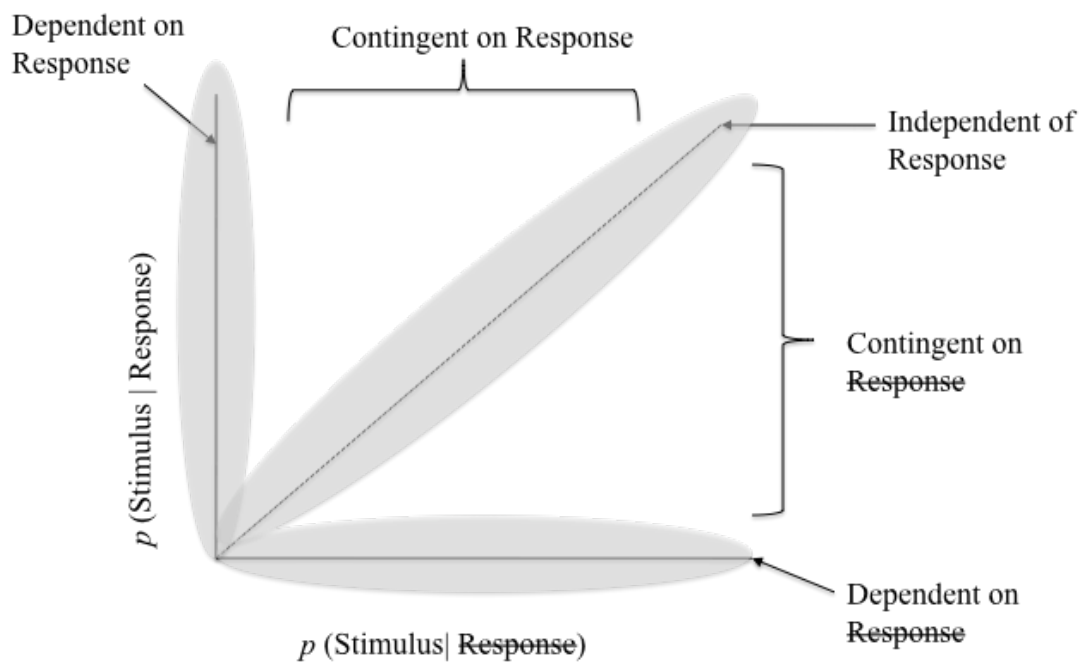


Figure 1

Graphic Representation of Contingency Space

location of points in contingency space allows a visualization of contingent and noncontingent relations between target responses and environmental stimuli.

Both YQ and CSA are calculated using the same 2 x 2 contingency table values (see Table 1), with each cell representing one of four possible sequences (i.e., response occurred and was followed by stimulus [Cell A], response occurred and was not followed by stimulus [Cell B], response did not occur but stimulus did occur [Cell C], and neither response nor stimulus occurred [Cell D]). By virtue of the formulae used to calculate YQ and CSA (see Table 2), these measures always correspond in terms of the direction of the association or contingency. A positive YQ will always correspond to a point in contingency space above the diagonal (i.e., the $p(S/R)$ is greater than the $p(S/\bar{R})$), indicating a positive contingency. A negative YQ will always correspond to a point in contingency space below the diagonal (i.e., the $p(S/R)$ is less than the $p(S/\bar{R})$), indicating a negative contingency. However, the extent to which the strength of the sequential association (YQ) corresponds with the strength of the contingency (CSA) is less clear.

In order to compare YQ and CSA in terms of strength, an effect size metric is needed for CSA. Currently, there is no recommended index of contingency strength for CSA aside from the difference between the two transitional probabilities (Martens et al., 2008). This difference score (DS; $p(S/R) - p(S/\bar{R})$) represents the distance from the diagonal of independence in contingency space, thus presumably, the farther the point lies from the diagonal, the stronger the contingency. Another potential measure of contingency strength is a ratio of the transitional probabilities ($p(S/R) / p(S/\bar{R})$). This alternative effect size metric, which has not been used in studies on CSA, is known as a relative risk ratio (RR) in epidemiology research (Merrill & Timmreck, 2006; see Table 3

Table 1

Construction of 2 x 2 Contingency Tables Using Lag-1 Event Sequential Analysis

	Reinforcer Contact	No Reinforcer Contact
Target Response	A	B
No Target Response	C	D

Cell	Contents
A	Number of event lag-1 pairs with first event as target response and second event as contact with reinforcer
B	Number of event lag-1 pairs with first event as target response and second event as any coded event except contact with reinforcer
C	Number of event lag-1 pairs with first event as any coded event except target response and second event as contact with reinforcer
D	Number of event lag-1 pairs with first event as any coded event except target response and second event as any coded event except contact with reinforcer

Table 2

Calculations of CSA Probabilities and YQ Statistics from the 2 x 2 Contingency Tables

Contingency Table		CSA $p(S/R)$	CSA $p(S/\mathbb{R})$	YQ
A	B	$\frac{A}{A+B}$	$\frac{C}{C+D}$	$\frac{(AD-BC)}{(AD+BC)}$
C	D			

Table 3

Calculations of Relative Risk Ratios and Difference Scores from the CSA Probabilities

RR	DS
$\frac{p(S/R)}{p(S/\bar{R})}$ or $\frac{\frac{A}{A+B}}{\frac{C}{C+D}}$	$p(S/R) - p(S/\bar{R})$ or $\frac{A}{A+B} - \frac{C}{C+D}$

for effect size formulae). The RR takes into account the location of the transitional probabilities along the x - and y -axes, in addition to the distance from the graphic diagonal. To illustrate the potential distinction between a difference in probabilities and a ratio of probabilities, consider the following two points in contingency space: (1) $p(S/R) = .95$, $p(S/\bar{R}) = .85$ and (2) $p(S/R) = .15$, $p(S/\bar{R}) = .05$. The difference in probabilities would be the same for each point (.10) whereas the ratio of probabilities would be greater for the second point (3.0), than the first point (1.1). The RR may better reflect the strength of contingency than the DS, as this index seems to account for the overall probability of the stimulus.

Because YQ is the current recommended approach for statistically quantifying sequential associations, we sought to compare this method to one uniquely suited to operant theory (CSA). For the purpose of identifying potential response-reinforcer relations, indices of contingency may be more informative than an index of sequential association (Yoder & Symons, 2010). We chose an animal model to compare CSA and YQ for three reasons. First, an animal model allows a simplified environment in which a single target response and a single stimulus may be isolated. Second, animal models allow for confirmation of the identified stimulus as a reinforcer. Third, laboratory environments allow increased control over contingent and noncontingent reinforcement schedules, not readily available in natural environments.

The purpose of the current study was to compare CSA and YQ measures in terms of the effects on response patterns in mice. We also sought to compare two possible effect size metrics for CSA (i.e., DS and RR) in terms of correspondence with overall responding. The decision to use frequency of responding as the outcome measure on

which to compare the three indices was based on an assumption that higher indices of contingency should correspond to higher rates of responding (e.g., Hammond, 1980). We selected two distinct points in contingency space for which the DS and RR differed but the YQ values were the same. These points in contingency space were approximated using concurrent schedules of reinforcement. We hypothesized that potential differences in response patterns between the approximated points in contingency space would reveal the most useful measure in predicting the strength of the response-reinforcer relation. Higher rates of responding in the condition for which the RR was greater, for example, may suggest RR as the most appropriate indicator of contingency strength.

CHAPTER II

METHOD

Subjects

Thirteen male C57BL/6J mice were included in the study and were approximately four months of age at the start of test conditions. A 12:12 hr light/dark cycle was in effect throughout the experiment and all experimental sessions occurred during the light cycle. Animals had ad libitum access to water outside of experimental sessions and had access to food for 4 hr each day beginning 1 hr after experimental sessions were completed. An 18-hr food deprivation period preceded all experimental conditions.

Apparatus

Operant conditioning chambers (OFA-510, Med AssociatesTM, Inc) were used for all experimental sessions. The chambers were controlled by modified programs included in the MedLab 8 Package (SOF-700LA-1, Med AssociatesTM, Inc) for MED-PC® IV software. Each chamber included an ENV-313M illuminated infrared nose poke sensor and an ENV-302M-S liquid dipper, which dispensed .01 cc of a 1:1 mixture of Ensure and water. An ENV-302HD infrared head entry detector for the dipper receptacle allowed recording of when the animals contacted the Ensure mixture. Two additional nose poke sensors were located to the left of the dipper and in the center of the wall opposite the dipper. Only nose pokes to the illuminated sensor at the right of the dipper constituted target responses and resulted in food delivery. Nose pokes to the two remaining sensors

were recorded but had no effect on food delivery. An ENV-315M house light remained illuminated for the duration of all sessions.

Procedure

Shaping. Shaping sessions were initially conducted to increase the occurrence of the target response of right sensor nose pokes and to establish contact with food as a positive reinforcer. All animals were initially exposed to a continuous schedule of reinforcement (CRF) following right sensor nose pokes for 11 consecutive sessions, each lasting 30 min. At the start of each session, the house light came on and remained illuminated until the session ended. The right nose poke sensor was illuminated until the sensor was activated, at which point the light turned off and the dipper became available. The dipper remained available until 10 s following a head entry to the dipper. The right nose poke sensor became illuminated again once the dipper was no longer available. A minimum of 5 consecutive sessions with target response (i.e., right sensor nose poke) rates above 1 per min were required before moving to the comparison schedules of reinforcement.

Comparison schedules of reinforcement. A concurrent schedule of reinforcement was selected to approximate each of the targeted points in contingency space (i.e., the probability of reinforcement given a target response versus the probability of reinforcement given no target response as .50 and .33, respectively and .20 and .11, respectively). To approximate the probabilities of reinforcement given a target response, random ratio schedules of reinforcement were selected. The random ratio schedule is a type of variable ratio reinforcement schedule in which the probability of reinforcement

remains constant across aggregations of responses (Bancroft & Bourret, 2008). Thus, a random ratio 2 was selected to approximate the .50 probability of reinforcement following a target response and a random ratio 5 was selected to approximate the .20 probability of reinforcement following a target response.

To approximate the probabilities of reinforcement given no target response, variable time (VT) schedules were added to each of the random ratio schedules. A “non-response” was arbitrarily defined as 5 s elapsing without a right sensor nose poke. To approximate the .33 probability of reinforcement given no target response (i.e., reinforcement following every third “non-response”), we multiplied 5 s by three and selected a VT 15-s schedule. The VT 15-s schedule was in effect concurrently with the random ratio 2 schedule to approximate the .50 probability of reinforcement following a target response and the .33 probability of reinforcement following no target response.

To approximate the .11 probability of reinforcement given no target response (i.e., reinforcement following every ninth “non-response”), we multiplied 5 s by nine and selected a VT 45-s schedule. The VT 45-s schedule was in effect concurrently with the random ratio 5 schedule to approximate the .20 probability of reinforcement following a target response and the .11 probability of reinforcement following no target response. These VT schedule components resulted in food delivery every 15 s (on average) during one comparison schedule and every 45 s (on average) during the other comparison schedule independent of responding. We selected random ratio + VT schedules to approximate an environment in which reinforcement is sometimes delivered following target responses and sometimes delivered in the absence of target responses.

During each of these test conditions, the right nose poke sensor remained on throughout the session. When either the random ratio response requirement was met or the VT interval elapsed, the dipper became available for 5 s. The VT schedule component was independent of the random ratio schedule component, thus it was possible for the reinforcement intervals to overlap, resulting in food availability for greater than 5 s.

Experimental design. Each animal was exposed to each concurrent schedule of reinforcement via an A-B-A-B ($n = 3$) or B-A-B-A ($n = 9$) reversal design. An A-B-A design was used for one animal that ceased responding during the third phase. The A condition corresponded to the random ratio 2 + VT 15-s schedule; the B condition corresponded to the random ratio 5 + VT 45-s schedule. Animals were exposed to each concurrent schedule for a minimum of 10 sessions. Conditions were changed after response rates in four of five consecutive sessions were within 20% of the mean of the previous five sessions. The final session had to be one of the four within this 20% range. Response rates beyond 30% of this mean could not be included in these five consecutive sessions. This criterion was added to prevent any extreme data points from significantly altering the mean stability range. Dependent variables included right, left, and rear sensor nose pokes per min during shaping and right sensor nose pokes in addition to the three indices of contingency (i.e., YQ, RR, DS) throughout the comparison schedules.

Sequential Analysis

A lag-1 event method of sequential analysis was used to construct the 2 x 2 contingency tables (see Table 1). Although event-based sequential analyses can be problematic due to the challenge of identifying an exhaustive set of relevant events

(Yoder & Symons, 2010), this method was feasible given the highly controlled setting and the low number of relevant events (i.e., right sensor nose poke, left or rear sensor nose poke, contact with Ensure mixture). The lag-1 event method was selected over time window analysis due to the rapid sequences of coded events displayed in the raw data streams. Because it was possible for several coded events to occur within the same second, an event-based method allowed for only the responses that triggered food delivery to be counted in the A cells of the contingency tables.

In the lag-1 event sequential analysis, pairs of coded events (i.e., events 1 and 2, events 2 and 3, events 3 and 4, and so on) were the coded units tallied into the 2 x 2 contingency tables. This approach to contingency table construction, in which events overlap between pairs, has been justified empirically (Bakeman & Dorval, 1989). In fact, the resulting doubling of the number of coded units in the contingency table is considered beneficial, as it facilitates interpretability of sequential associations (Yoder & Symons, 2010). The first event or target response included the right sensor nose poke and the second event or consequence included contacting food. The total number of event pairs or coded units (i.e., A+B+C+D) varied across sessions, as the number of coded events varied across sessions. CSA probabilities and YQ indices were calculated from the 2 x 2 contingency table values according to the formulae in Table 2. CSA probabilities were plotted in contingency space with $p(S/R)$ along the y -axis and $p(S/\mathbb{R})$ along the x -axis.

Correlational Analysis

Due to unexpected variability in indices of contingency and target response rates from one session to another (as opposed to systematic differentiation between

conditions), correlational analyses were conducted to explore whether one or more indices was more closely associated with total responding. Product moment correlation coefficients were calculated to measure the strength of the association between each pair of indices of contingency (i.e., $r[DS, RR]$, $r[DS, YQ]$, $r[RR, YQ]$). Correlation coefficients were also calculated to measure the strength of the association between each index of contingency and the total frequency of target responses. All correlations were calculated per subject and across conditions.

Based on variable and unexpectedly low correlation coefficients, quadratic terms (i.e., the squared values of each index) were added to determine whether any of the relations between indices and responding were curvilinear. Only quadratic terms that were statistically significant were included and the resulting R values (i.e., measures of association between each index and total responding) were reported. Three paired sample t -tests were conducted to determine whether there was a statistically significant difference between the correlations of each index and total responses (e.g., $r[YQ, \text{Total Responses}]$ vs. $r[RR, \text{Total Responses}]$). The t -tests were performed on z -transformed r values, as z -scores are normally distributed and thus more likely to relate linearly to other variables whereas Pearson correlation coefficients are not (Cohen, Cohen, West, & Aiken, 2003). Statistical t -tests and resulting p values are therefore more accurate when z (rather than r) is used as the dependent measure. Three additional paired sample t -tests were performed on z -transformed r values of correlations between each pair of indices to test for significant differences among these pairs.

CHAPTER III

RESULTS

After 11 sessions of shaping, target response rates were at or above 1 per min for at least five consecutive sessions across all subjects. Clear differentiation between target response rates (right sensor nose poke $M = 2.31$; $SD = 0.67$) and non-target response rates (left sensor nose poke $M = 0.07$; $SD = 0.06$; rear sensor nose poke $M = 0.17$; $SD = 0.11$) were observed for all subjects during the final five sessions. Because the frequency of the target response increased following the contingent presentation of food, contact with food (i.e., head entry to dipper when dipper was available) was identified as a positive reinforcer for right sensor nose pokes.

The concurrent schedules of reinforcement for each comparison condition did not produce the intended points in contingency space (see Figure 2). In addition, considerable variability in the indices of contingency was observed both within conditions and among subjects. The variability was, at least in part, due to variability in overall target response rates, as the rate of responding affected the probabilities of reinforcement in the presence versus absence of target responses. Some subjects, for example, emitted the target response rapidly throughout the session such that very few response-independent (i.e., VT-based) reinforcers were contacted.

Experimental control was not demonstrated between condition (concurrent schedule of reinforcement) and target response rate. For nine of the 13 subjects, there was no systematic distinction in probabilities of reinforcement given a target response

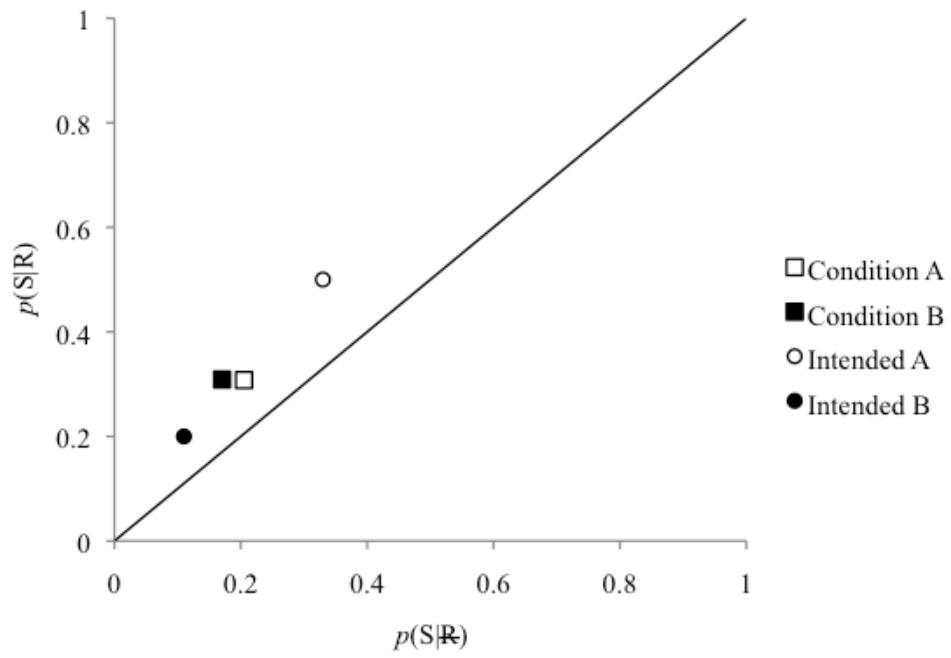


Figure 2

Intended and Observed Mean Points in Contingency Space per Condition across Subjects

(response-dependent reinforcers) and probabilities of reinforcement in the absence of a target response (response-independent reinforcers) between experimental conditions. The indices of contingency were therefore undifferentiated between conditions for these animals, as were the target response rates. See Figure 3 for example subjects for which no differentiation was observed between conditions.

For the remaining four subjects (2, 6, 10, and 14), experimental control was demonstrated between the experimental condition (concurrent schedule of reinforcement) and the probabilities of reinforcement following a target response and in the absence of a target response. For these subjects, all three indices of contingency were higher in the B condition (random ratio 5 + VT 45-s) than the A condition (random ratio 2 + VT 15-s; see Figure 4 for mean CSA points and Figures 5 and 6 for reversal graphs). Responding by Subject 14 extinguished in the third phase before criteria were met to implement the final reversal. The resulting A-B-A design permitted only two demonstrations of effect for this subject as opposed to three demonstrations for Subjects 2, 6, and 10. Despite the observed differences in contingencies between conditions, there were still no systematic differences in response rates between conditions for these four subjects (see Figure 7).

For all subjects, across all sessions, the three indices of contingency (DS, RR, YQ) indicated the same direction of association (positive versus negative; by virtue of the formulae for each index). In addition to the correspondence between indices shown in the reversal graphs (Figures 5 and 6), the correlational analysis revealed strong positive associations between each pair of contingency indices (see Table 4). The correlations between the DS and YQ ($M = .97$, $SD = .04$) were slightly yet consistently higher than the correlations between DS and RR ($M = .88$, $SD = .12$; $t(12) = 12.09$, $p < .0001$) and

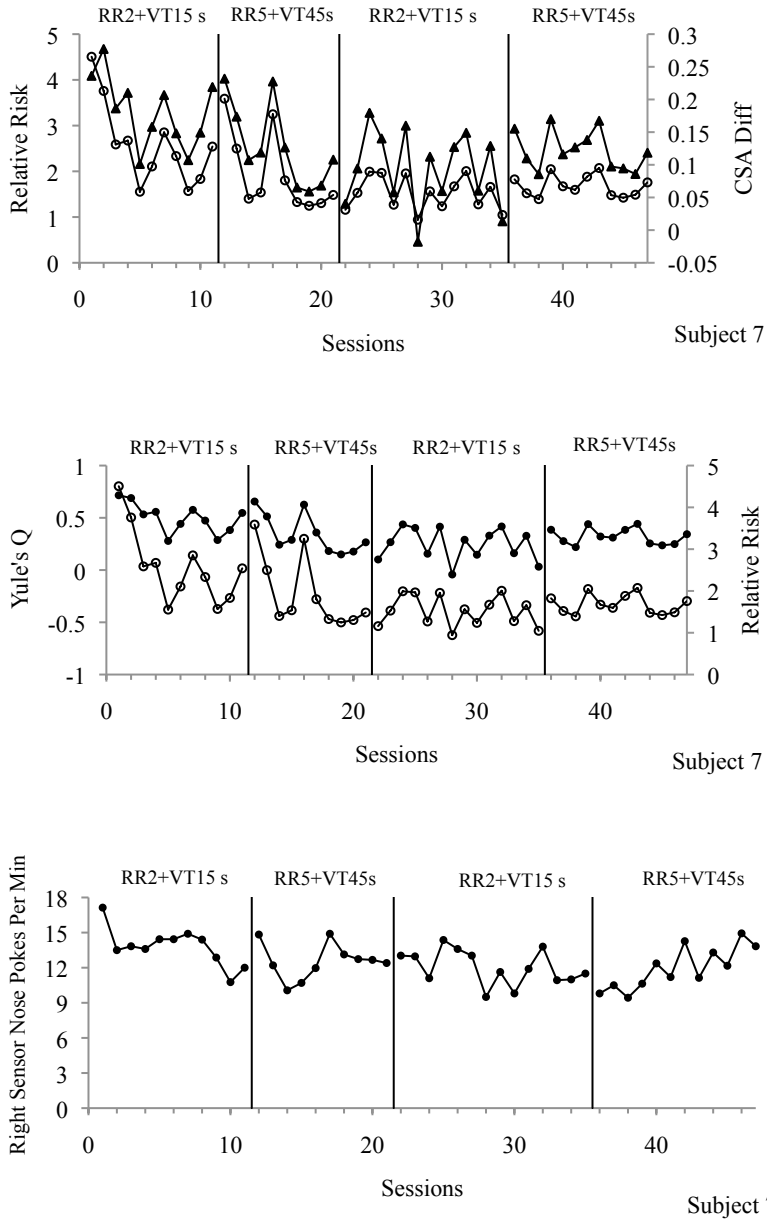


Figure 3

Examples of Undifferentiated Indices of Contingency and Undifferentiated Responding Between Experimental Conditions

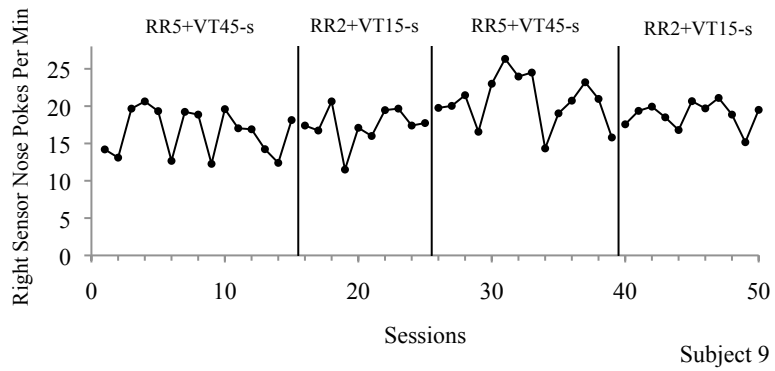
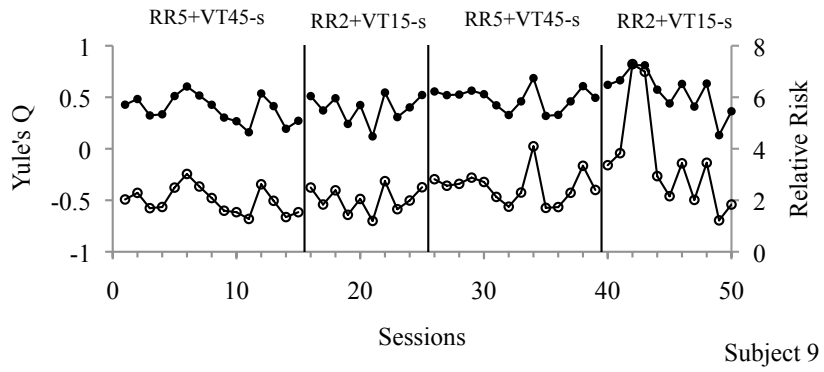
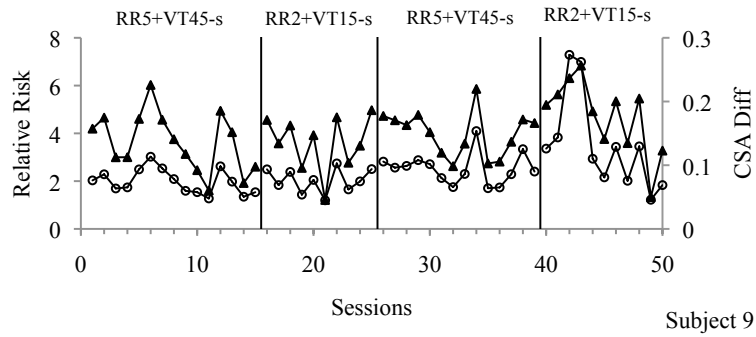


Figure 3, continued

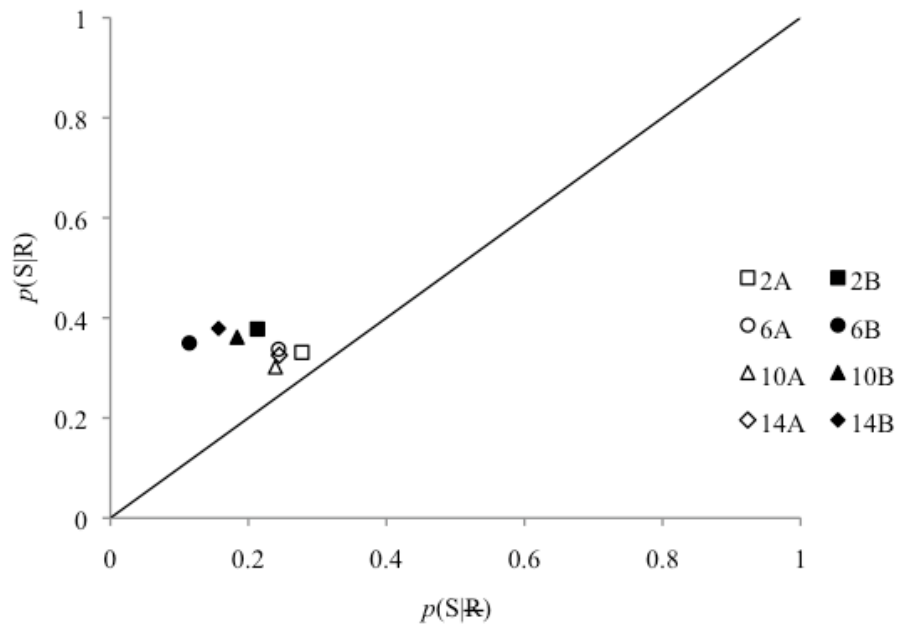


Figure 4

Mean Points in Contingency Space for Subjects 2, 6, 10, and 14

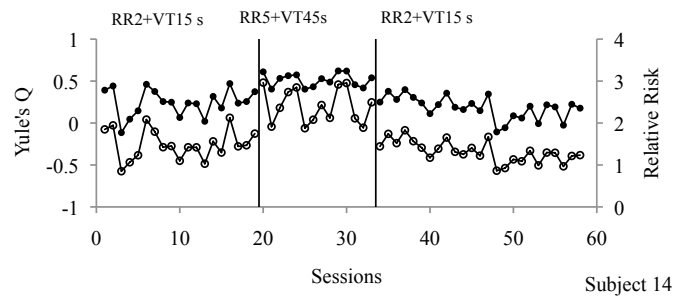
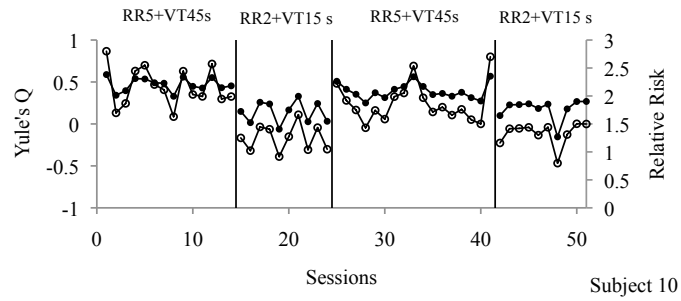
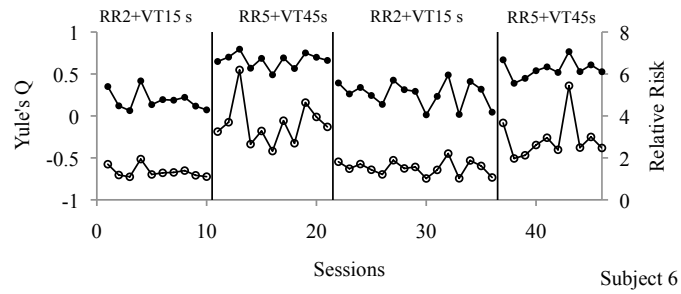
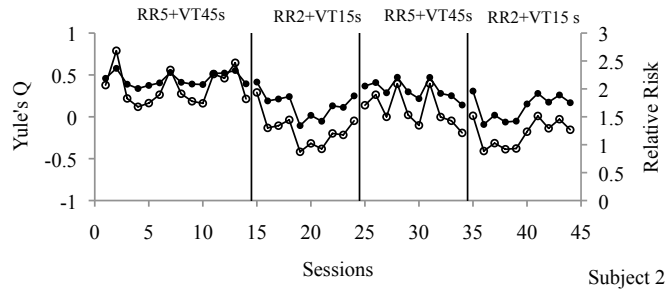


Figure 5

Yule's Q and Relative Risk Ratios for Subjects 2, 6, 10, and 14

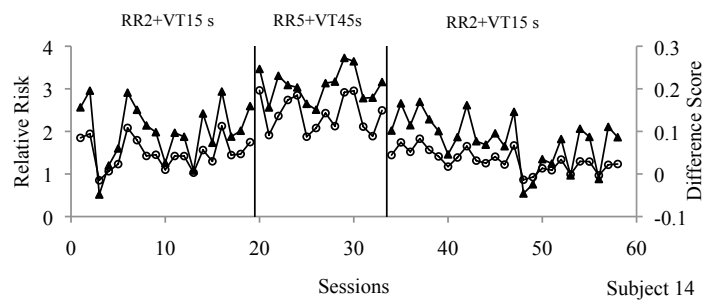
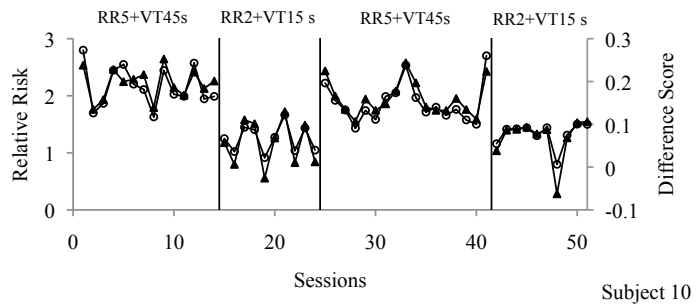
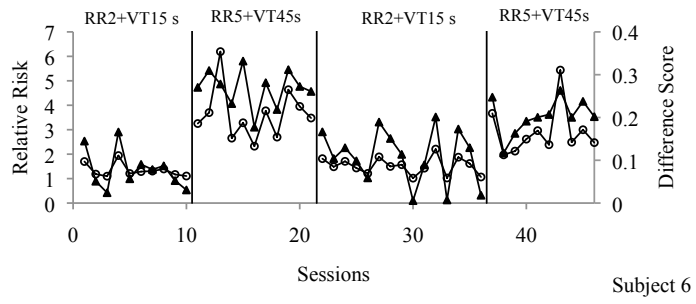
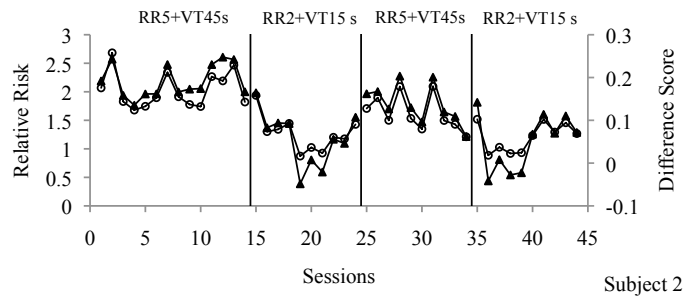


Figure 6

Relative Risk Ratios and Difference Scores for Subjects 2, 6, 10, and 14

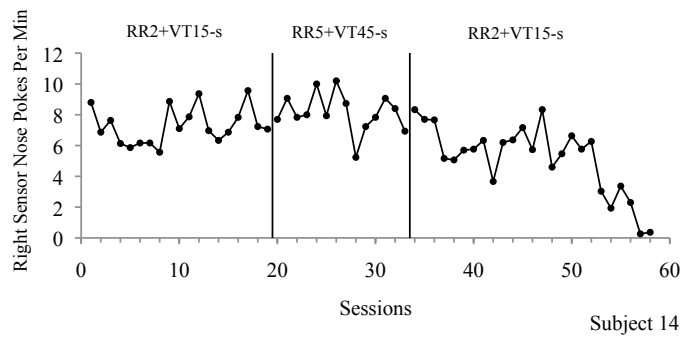
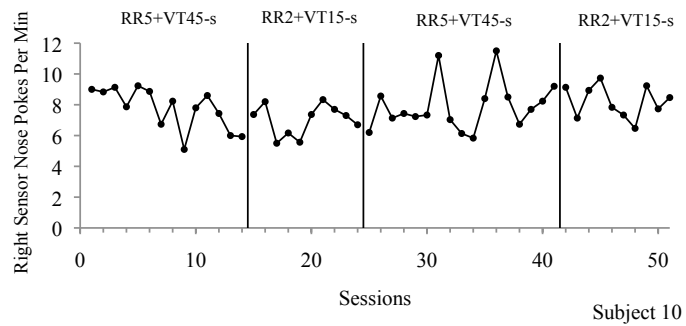
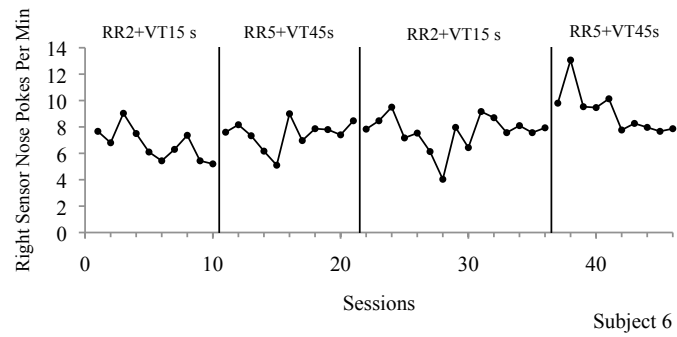
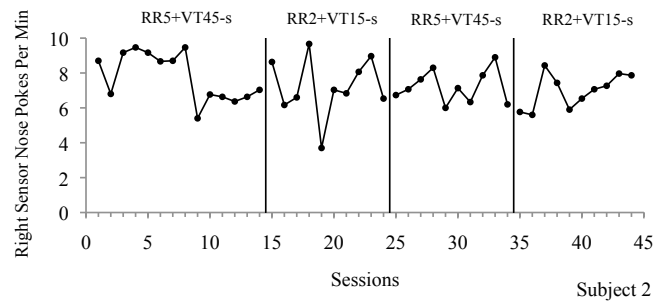


Figure 7

Target Responses per Min for Subjects 2, 6, 10, and 14

Table 4

Correlations Among Indices of Contingency

Subject	r (DS,RR)	r (DS, YQ)	r (RR, YQ)
2	0.965	0.997	0.97
3	0.931	0.996	0.949
4	0.985	0.999	0.992
5	0.915	0.99	0.958
6	0.863	0.981	0.909
7	0.903	0.986	0.935
8	0.836	0.966	0.875
9	0.824	0.975	0.873
10	0.951	0.993	0.963
11	0.967	0.991	0.98
12	0.886	0.944	0.955
13	0.521	0.856	0.746
14	0.937	0.993	0.956

between RR and YQ ($M = .93$, $SD = .07$; $t(12) = 7.24$, $p < .0001$). This difference in correlations among pairs of indices was likely due to a distinction in the range of DS and YQ (-1, 1) and RR (0, ∞). Both YQ and the DS are bounded whereas RR is unbounded. Although the ranges of values for each index in the current data set were restricted, the asymmetrical distribution of RR values compared to the symmetrical distribution of YQ and DS values may explain why correlations including RR were slightly lower than correlations between the DS and YQ.

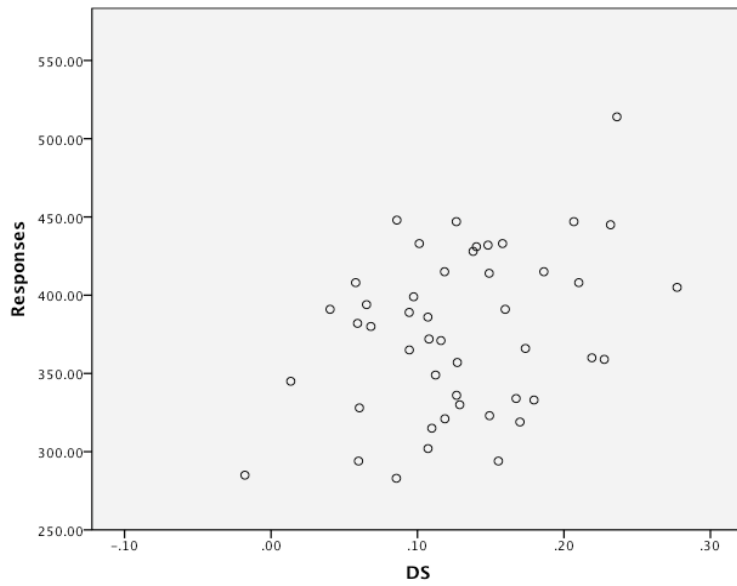
The correlation coefficients describing the association between each index of contingency and total responding were variable (see Table 5). Eight of 36 regression coefficients (3 indices per 13 subjects) included a statistically significant quadratic term. Five of these significant quadratic terms included regressions of response frequency on the DS. Although positive linear relations between indices of sequential association or contingency and total responding were identified for a subset of subjects (e.g., Subjects 7 and 14; see Figure 8), other regression analyses revealed negative linear relations (see Figure 9) or quadratic relations (see Figure 10). In fact, some distinctions were identified among the relations between each index and total responding within the same subject (e.g., Subject 13; see Figure 11). The paired sample t -tests did not reveal statistically significant differences between the r values and total responding among the three indices.

Table 5

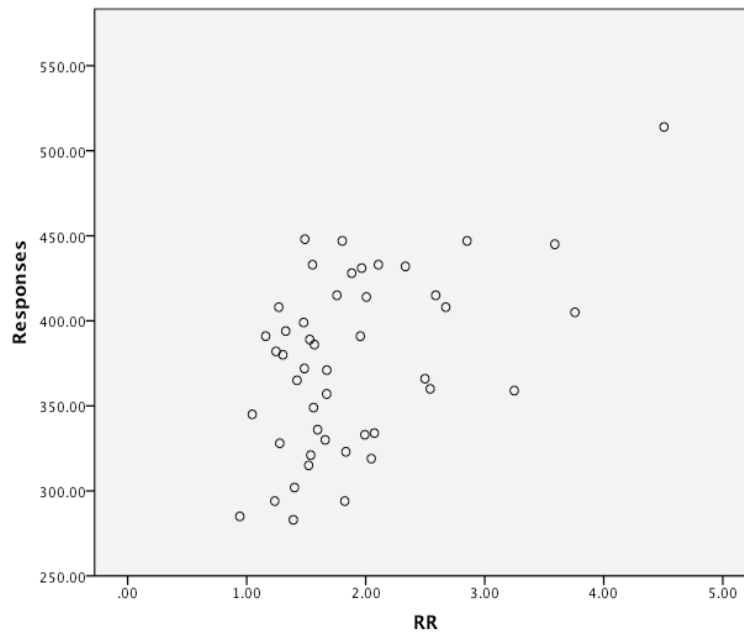
Correlation Coefficients Between Each Index of Contingency and Total Frequency Target Responses

Subject	r (DS, Resp)	r (RR, Resp)	r (YQ, Resp)
2	0.43*	0.39*	0.41*
3	-0.17	-0.12	-0.13
4	-0.25	-0.27	-0.26
5	0.23	0.4	0.41*
6	0.31*	0.16	0.23
7	0.35	0.5	0.45
8	-.47*	-0.18	-0.2
9	.28*	0.11	0.17
10	.36*	0.09	0.08
11	-0.16	-0.08	-0.09
12	-0.54	-0.25	-0.32
13	-0.11	0.33	0.3
14	0.39	0.46	0.46

*Includes quadratic term



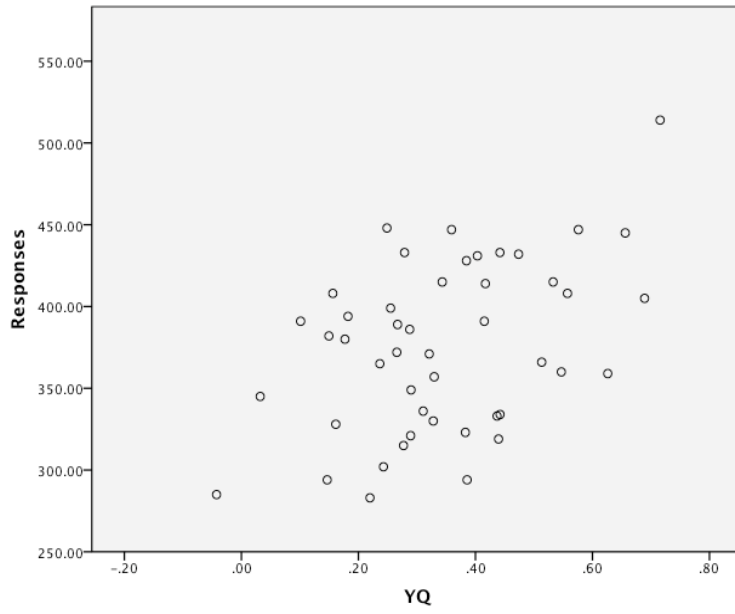
Subject 7



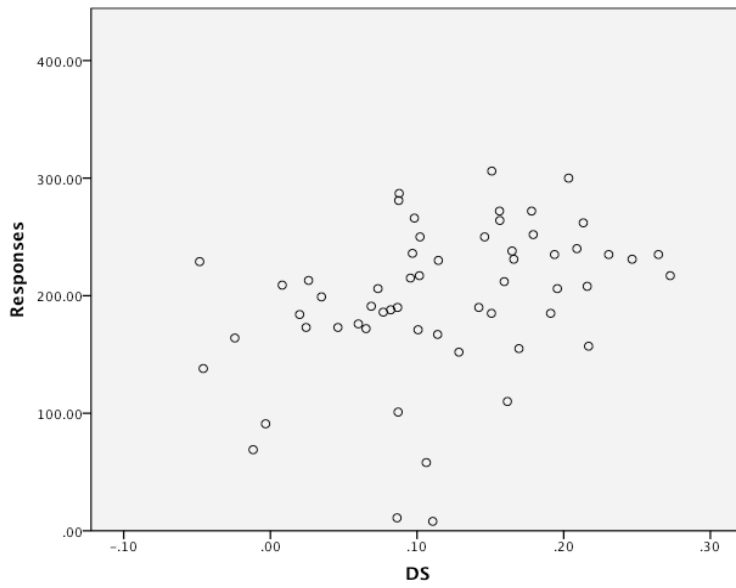
Subject 7

Figure 8

Examples of Positive Linear Relations between Indices of Contingency and Total Frequency Target Responses

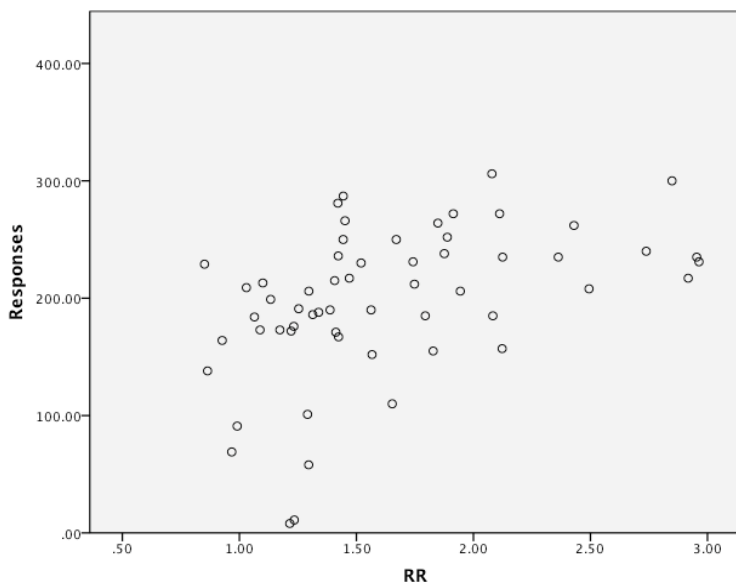


Subject 7



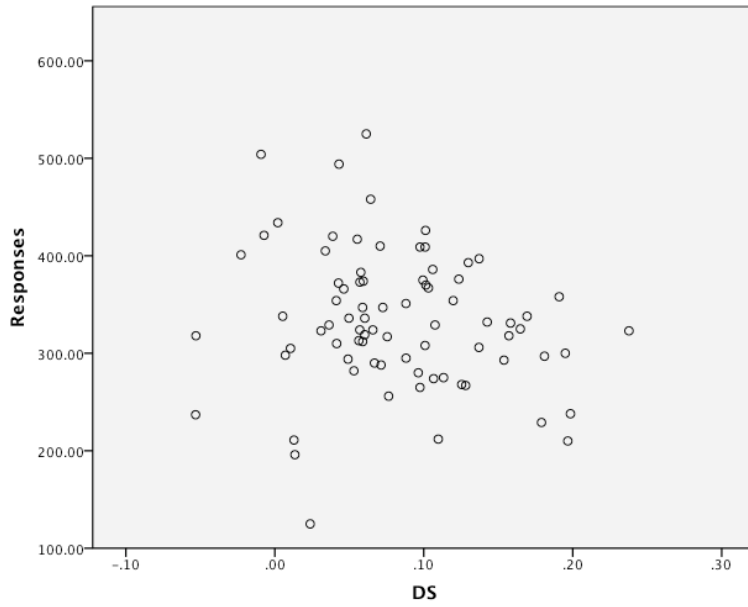
Subject 14

Figure 8 continued

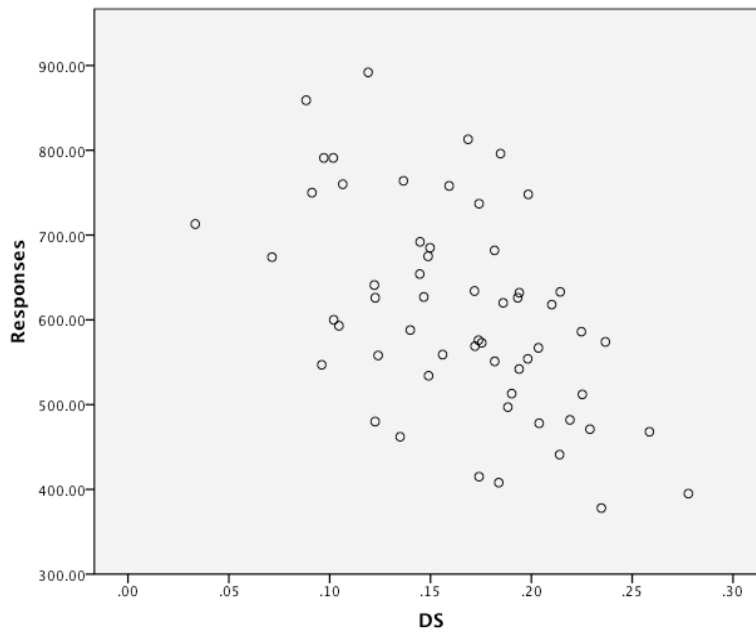


Subject 14

Figure 8 continued



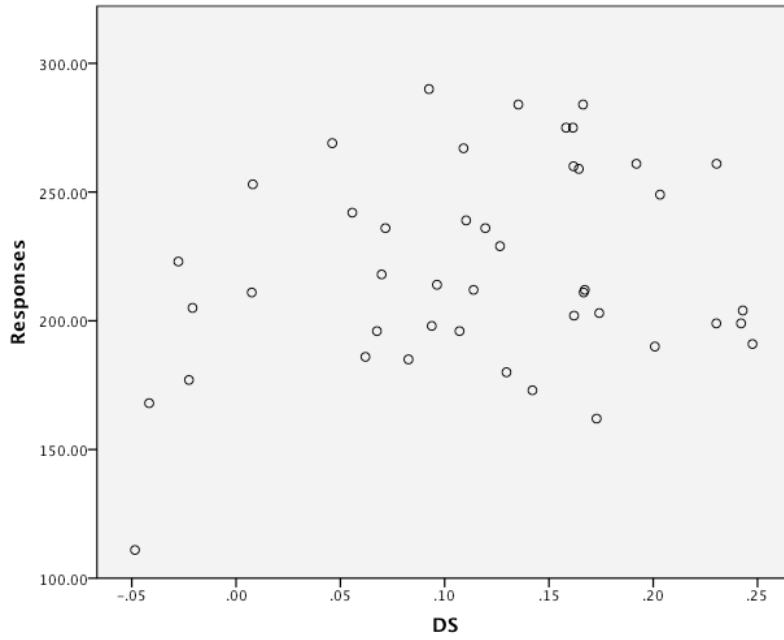
Subject 3



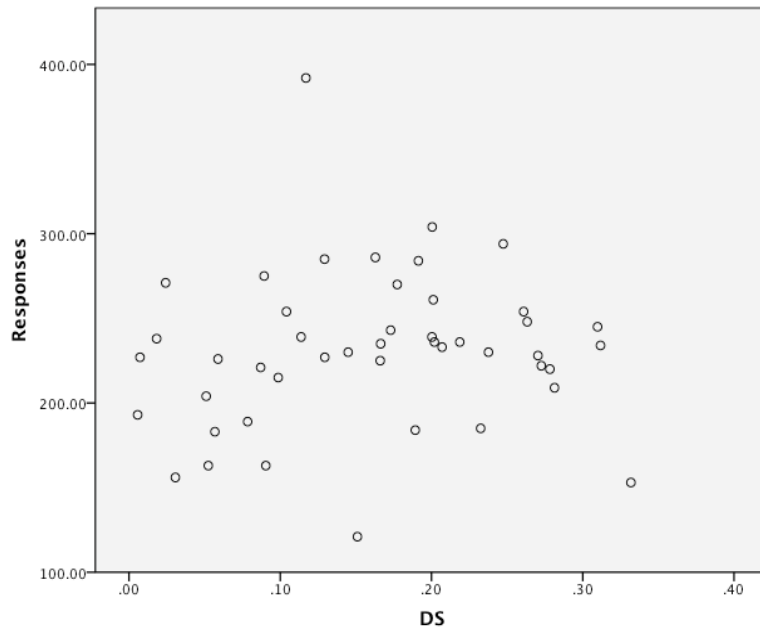
Subject 12

Figure 9

Examples of Negative Linear Relations between Indices of Contingency and Total Frequency Target Responses



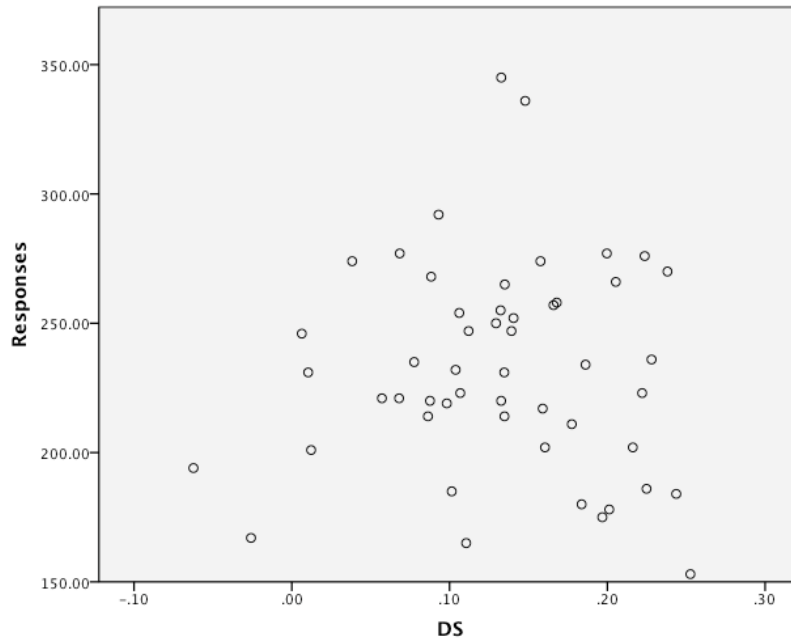
Subject 2



Subject 6

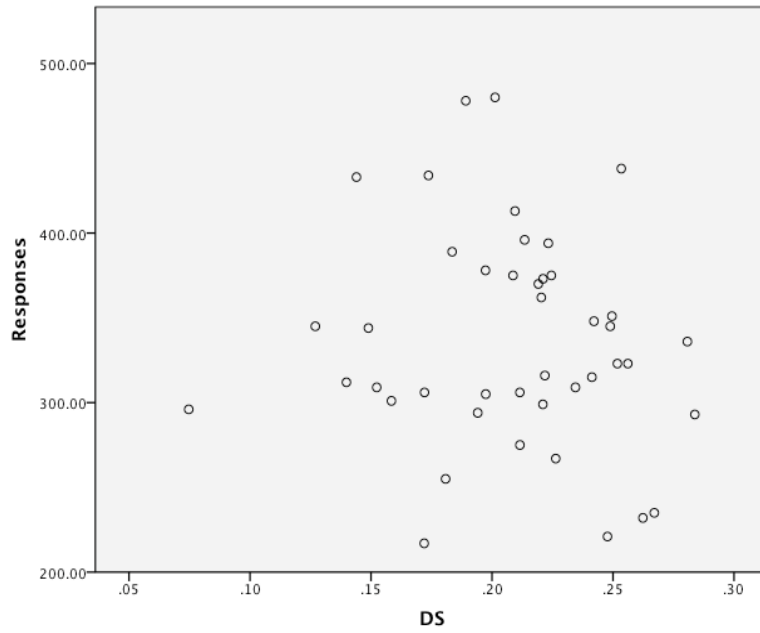
Figure 10

Examples of Quadratic Relations between Indices of Contingency and Total Frequency Target Responding

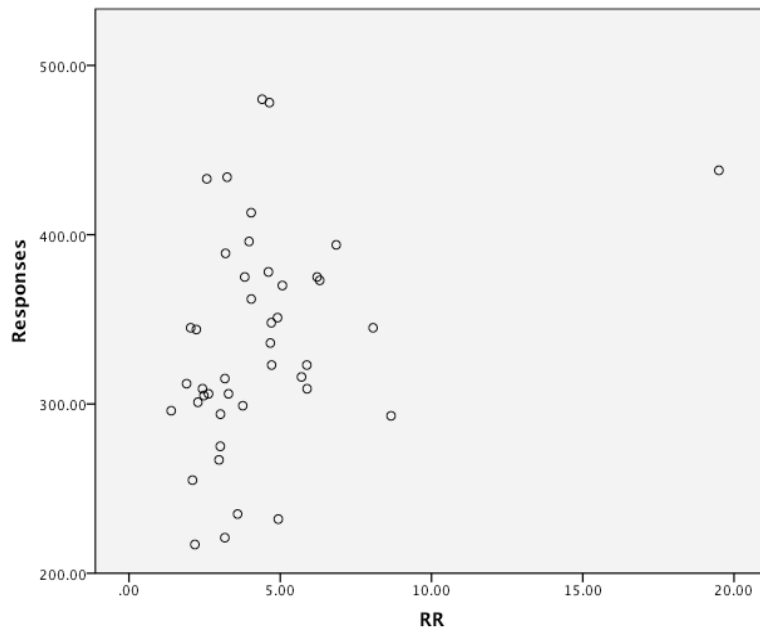


Subject 10

Figure 10 continued



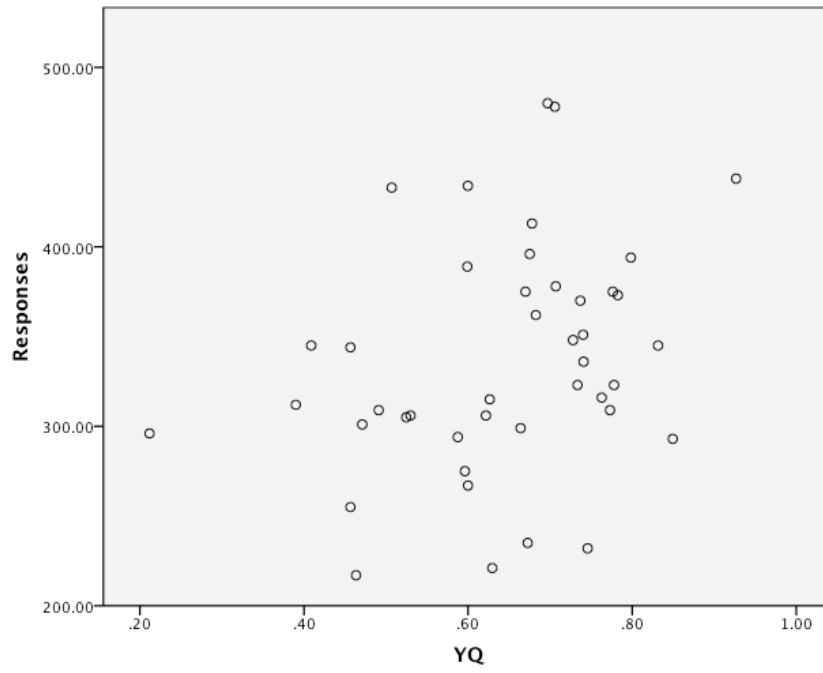
Subject 13



Subject 13

Figure 11

Example of Within-Subject Variability Among Relations between Indices of Contingency and Total Frequency Target Responding



Subject 13

Figure 11 continued

CHAPTER IV

DISCUSSION

The aim of the current study was to compare CSA and YQ as two methods of quantifying response-stimulus relations. To compare these methods in terms of strength, two effect size metrics were evaluated for CSA. In addition to a difference in transitional probabilities (DS), a ratio of transitional probabilities (RR) was included to better account for the base rate of the consequent stimulus. Across all subjects, each pair of indices was highly correlated. This finding is not surprising considering each of the indices is calculated using the same contingency table values. The degree of correspondence among indices, however, suggests a low likelihood of any two indices leading to different conclusions of whether a consequence is identified as a potential reinforcer.

The primary limitation of the current study was the discrepancy between the intended points in contingency space and the transitional probabilities resulting from the concurrent schedules of reinforcement. Despite careful approximations of reinforcement probabilities via the concurrent schedules, the ultimate probabilities were still partially dependent on the subjects' responding and thus were not entirely under experimenter control. Some subjects, for example, emitted target responses rapidly such that reinforcers were rarely contacted in the absence of a response. Other subjects did not contact the VT-based (i.e., response independent) reinforcers when they did become available. These scenarios resulted in decreased probabilities of reinforcement in the

absence of target responses, which increased the overall index of contingency or sequential association.

Although variable, the transitional probabilities resulting from the concurrent schedules of reinforcement were less distinct (i.e., closer in contingency space) than the intended probabilities. This outcome likely contributed to the lack of differentiated responding between the two experimental conditions. Even though there was a systematic distinction in indices of contingency between conditions for a subset of subjects, response rates for these subjects were still undifferentiated. The differences in transitional probabilities may not have been large enough to produce differentiated response rates. The lack of discriminative stimuli associated with each condition also may have contributed to the lack of response differentiation.

Our primary research question was whether the DS, RR, or YQ more accurately predicted response-reinforcer relations for cases in which contingency strength differed among indices. The variability in transitional probabilities across sessions and the lack of differentiated responding between experimental conditions precluded addressing the research question based on the experimental A-B-A-B design. Instead, we addressed this research question indirectly by comparing correlations between each index and total responding across subjects. The lack of statistically significant differences among correlations between each index and responding prevented the identification of any single index as a better predictor of responding.

The regression analyses did, however, provide insight on the relations between indices of contingency and frequency of target responses. The design of the current study was based on an assumption that greater indices of contingency (as measured by DS, RR,

or YQ) would generally correspond with higher rates of responding. Results of the regression analyses allowed an empirical evaluation of this assumption. Variability in the form of the relation between contingency indices and responding was identified among (and in some cases, within) subjects. Only a subset of positive linear relations between contingency strength and total responding was identified. In fact, a subset of negative linear relations was also identified. Other relations were curvilinear, including five statistically significant quadratic functions among the correlations between the DS and responding.

Curvilinear relations between contingency strength and response frequencies may, in fact, better approximate relations between indices of contingency and stimuli functioning as reinforcers. When considering the $p(S/R)$ in isolation, the highest response rates likely would be produced by probabilities between 0 and 1 (i.e., intermittent reinforcement), whereas lower response rates likely would be produced by probabilities at the two extremes (i.e., extinction and continuous reinforcement, respectively; Catania, 2007). Although the interaction between the $p(S/R)$ and the $p(S/R)$ is not fully understood, a curvilinear relation in which lower response rates correspond with high and low extremes of contingency indices (i.e., a quadratic relation) seems possible and remains consistent with operant theory. The lack of consistency in the relations between contingency indices and responding identified in the current data set may be related to the restricted ranges of contingency indices. If a general quadratic function does explain the relation between the full range of values of contingency strength and response frequency, restricted ranges of contingency strength may reveal positive linear, negative linear, curvilinear, or no clear relations depending on the location of the restricted range.

The method of contingency table construction (i.e., lag-1 event) also may have contributed to unexpected relations between contingency indices and responding (e.g., negative linear relations). Subject 12, for example, engaged in the highest rates of responding across conditions ($M = 20.32$ per min, $SD = 4.03$). For this subject, it appears that strings of rapid responding led to decreased probabilities of reinforcement given a target response, as only the response coded immediately prior to contacting reinforcement was counted as reinforced (Cell A). All previous responses in the string were not counted as reinforced (Cell B). High values in Cell B relative to Cell A result in a decreased $p(S/R)$. For subjects engaging in strings of rapid responses, it was possible for lower indices of contingency (due to a decreased $p(S/R)$) to be associated with higher overall response rates, producing a negative correlation. This pattern was apparent for Subject 12 (see Figure 9).

The potential influence of the method of contingency table construction on the relation between contingency indices and responding emphasizes the importance of selecting a method that accurately reflects sequences of responses and stimuli as they occur in the relevant environment. In cases of high rate ‘bursts’ of target responding, the lag-1 event method may result in an underestimation of contingency strength due to high B cell counts. A partial interval method, in which intervals of time would represent the coded units used to construct the contingency table, would likely produce more balanced cell counts. Although an event-based method seemed most appropriate to measure contingency for a highly-controlled animal model in which multiple events occurred per second, partial interval methods of data collection would be more appropriate for collecting data on behaviors occurring under free operant conditions within an applied

setting (Ayres & Gast, 2010). Indeed, partial interval methods of data collection and contingency table construction (as opposed to real-time event coding) are commonly used in the literature on descriptive assessments (e.g., Anderson & Long, 2002; Martens et al., 2008; Martens et al., 2010).

We hypothesized RR would better predict responding than the DS because RR appeared to better reflect the impact of the base rate of the stimulus on the contingency index. This prediction was based on the possibility of equal DSs for stimuli with very different base rates (e.g., $p(S/R) = .95$, $p(S/\bar{R}) = .85$ and $p(S/R) = .15$ and the $p(S/\bar{R}) = .05$). All else being equal, stimuli with lower base rates increase the RR value compared to stimuli with higher base rates. This adjustment may be considered consistent with a behavior analytic perspective, as stimuli presented frequently regardless of the occurrence of the target response are less likely to function as reinforcers than stimuli rarely presented in the absence of the target response. The comparisons of correlations between contingency indices and responding, however, did not reveal a significant difference between RR and DS in terms of predicting response frequency. Aside from predicting response-reinforcer relations, however, an advantage of using RR is its simplicity in terms of what is conveyed to consumers. For example, if the RR for attention given problem behavior is 2.0, the index indicates attention was twice as likely to occur when the individual engaged in problem behavior than when the individual did not engage in problem behavior.

Although neither the DS nor RR was identified to better predict response-reinforcer relations than YQ, the single subject graphs of contingency indices in addition to the correlations between pairs of contingency indices suggest no loss of information

from using CSA. Because YQ is a well-established statistic, guidelines exist for the amount of data necessary to yield interpretable YQ values (Yoder & Feurer, 2000; Yoder & Symons, 2010) as well as benchmarks for small, medium, and large effect sizes (derived from those identified for the odds ratio; Rosenthal, 1996). In contrast, no such guidelines exist for CSA, revealing clear avenues for future research. Another critical avenue for future research is evaluating the accuracy of CSA outcomes in natural settings. Accuracy may be measured as the extent to which CSA outcomes match FA outcomes (e.g., Martens, Gertz, de Lacy Werder, & Rymanowski, 2010), or the extent to which behavioral interventions based on CSA outcomes effectively reduce problem behavior within that setting (e.g., Dolezal & Kurtz, 2010).

Despite its limitations, the present study supports the use of CSA in identifying environmental contingencies, as this method is conceptually systematic with operant theory and the measures of contingency strength appear to correspond closely with the YQ measure of sequential association. This research extends previous research on evaluating sequential analysis methods for descriptive assessments in several ways. First, CSA and YQ have not been compared previously as methods for identifying potential reinforcement effects. Second, the consequent stimulus in the current study was shown to function as a positive reinforcer, which allowed confirmation that the indices of contingency and sequential association reflected reinforcement effects. Third, the utility of each index in predicting response-reinforcer relations was evaluated based on animal responding. Regression analyses allowed the evaluation of an assumption of a positive linear relation between contingency strength and response frequencies. Finally, the findings demonstrate preliminary support for two potential effect size measures for CSA,

which would allow comparisons of contingency strength among multiple consequent stimuli.

REFERENCES

- Anderson, C. M., & Long, E. S. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavior Analysis, 35*, 137-154.
- Ayres, K., & Gast, D. L. (2010). Dependent measures and measurement procedures. In D. L. Gast (Ed.), *Single subject research methodology in behavioral sciences* (pp. 129-165). New York: Routledge.
- Bakeman, R., & Dorval, B. (1989). The distinction between sampling independence and empirical independence in sequential analysis. *Behavioral Assessment, 11*(1), 31-37.
- Bakeman, R., McArthur, D., & Quera, V. (1996). Detecting group differences in sequential association using sampled permutations: Log odds, kappa, and phi compared. *Behavior Research Methods, Instruments and Computers, 28*, 446-457.
- Bancroft, S. L., & Bourret, J. C. (2008). Generating variable and random schedules of reinforcement using Microsoft Excel macros. *Journal of Applied Behavior Analysis, 41*, 227-235.
- Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1*, 175-191.
- Call, N. A., Wacker, D. P., Ringdahl, J. E., & Boelter, E. W. (2005). Combined antecedent variables as motivating operations within functional analyses. *Journal of Applied Behavior Analysis, 38*, 385-389.
- Carr, E. G., Yarbrough, S. C., & Langdon, N. A. (1997). Effects of idiosyncratic stimulus variables on functional analysis outcomes. *Journal of Applied Behavior Analysis, 30*, 673-686.
- Catania, A. C. (1971). Elicitation, reinforcement and stimulus control. In R. Glaser (Ed.), *The nature of reinforcement* (pp. 196-220). New York: Academic Press.
- Catania, A. C. (2007). *Learning* (4th Ed.). Cornwall-on-Hudson, NY: Sloan Publishing.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

- Dolezal, D. N., & Kurtz, P. F. (2010). Evaluation of combined-antecedent variables on functional analysis results and treatment of problem behavior in a school setting. *Journal of Applied Behavior Analysis, 43*, 309-314.
- Durand, V. M. (1993). Causes of behavior: Functional assessment and functional analysis. In M. D. Smith (Ed.), *Behavior modification for exceptional children and youth* (pp. 38-60). Stoneham, MA: Andover Medical Press.
- Gibbon, J., Berryman, R., & Thompson, R. L. (1974). Contingency spaces and measures in classical and instrumental conditioning. *Journal of the Experimental Analysis of Behavior, 21*, 585-605.
- Hammond, L. J. (1980). The effect of contingency upon the appetitive conditioning of free-operant behavior. *Journal of the Experimental Analysis of Behavior, 34*, 297-304.
- Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*, 147-185.
- Horner, R. H. (1994). Functional assessment: Contributions and future directions. *Journal of Applied Behavior Analysis, 27*, 401-404.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982/1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 215-240 (reprinted from *Analysis and Intervention in Developmental Disabilities, 2*, 1-20).
- Lang, R., O'Reilly, M., Lancioni, G., Rispoli, M., Machalicek, W., Chan, J. M., et al. (2009). Discrepancy in functional analysis results across two settings: Implications for intervention design. *Journal of Applied Behavior Analysis, 42*, 393-397.
- Mace, F. C., & Lalli, J. S. (1991). Linking descriptive and experimental analyses in the treatment of bizarre speech. *Journal of Applied Behavior Analysis, 24*, 553-562.
- Martens, B. K., DiGennaro, F. D., Reed, D. D., Szczech, F. M., & Rosenthal, B. D. (2008). Contingency space analysis: An alternative method for identifying contingent relations from observational data. *Journal of Applied Behavior Analysis, 41*, 69-81.
- Martens, B. K., Gertz, L. E., de Lacy Werder, C. S., & Rymanowski, J. L. (2010). Agreement between descriptive and experimental analyses of behavior under naturalistic test conditions. *Journal of Behavioral Education, 19*, 205-221.
- Matthews, B. A., Shimoff, E., & Catania, A. C. (1987). Saying and doing: A contingency-space analysis. *Journal of Applied Behavior Analysis, 20*, 69-74.

- McComas, J., Hoch, H., Paone, D., & El-Roy, D. (2000). Escape behavior during academic tasks: A preliminary analysis of idiosyncratic establishing operations. *Journal of Applied Behavior Analysis, 33*, 479-493.
- McComas, J. J., Moore, T., Dahl, N., Hartman, E., Hoch, J., & Symons, F. (2009). Calculating contingencies in natural environments: Issues in the application of sequential analysis. *Journal of Applied Behavior Analysis, 42*, 413-423.
- Merrill, R. M., & Timmreck, T. C. (2006). *An introduction to epidemiology* (4th ed.). Sudbury, MA: Jones and Bartlett Publishers.
- Ringdahl, J. E., & Sellers, J. A. (2000). The effects of different adults as therapists during functional analyses. *Journal of Applied Behavior Analysis, 33*, 247-250.
- Rosenthal, J. A. (1996). Qualitative descriptors of strength association and effect size. *Journal of Social Science Research, 21*(4), 37-59.
- Shirley, M. J., Iwata, B. A., & Kahng, S. (1999). False-positive maintenance of self-injurious behavior by access to tangible reinforcers. *Journal of Applied Behavior Analysis, 32*, 201-204.
- Vollmer, T. R., Borrero, J. C., Wright, C. S., Van Camp, C., & Lalli, J. S. (2001). Identifying possible contingencies during descriptive analyses of severe behavior disorders. *Journal of Applied Behavior Analysis, 34*, 269-287.
- Wacker, D., Berg, W., Harding, J., & Cooper-Brown, L. (2004). Use of brief experimental analyses in outpatient clinic and home settings. *Journal of Behavioral Education, 13*, 213-226.
- Yoder, P. J., & Feurer, I. D. (2000). Quantifying the magnitude of sequential association between events or behaviors. In T. Thompson & D. Felce (Eds.), *Behavioral observation: Technology and application in developmental disabilities* (pp. 317-333). Baltimore: Brookes.
- Yoder, P. J., Short-Meyerson, K., & Tapp, J. (2004). Measurement of behavior with a special emphasis on sequential analysis of behavior. In E. Emerson, C. Hatton, T. Thompson & T. R. Parmenter (Eds.), *The international handbook of research methods in intellectual disabilities* (pp. 179-199). New York: Wiley.
- Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York: Springer.
- Yule, G. V., & Kendall, M. G. (1957). *An introduction to the theory of statistics*. London: Griffin.