

# The use of mixed models in a modified Iowa Gambling Task and a prisoner's dilemma game

Jean Stockard\*

Department of Planning, Public Policy, and Management  
University of Oregon

Robert M. O'Brien

Department of Sociology  
University of Oregon

Ellen Peters

Decision Research  
Eugene, Oregon

## Abstract

Researchers in the decision making tradition usually analyze multiple decisions within experiments by aggregating choices across individuals and using the individual subject as the unit of analysis. This approach can mask important variations and patterns within the data. Specifically, it ignores variations in decisions across a task or game and possible influences of characteristics of the subject or the experiment on these variations. We demonstrate, by reanalyzing data from two previously published articles, how a mixed model analysis addresses these limitations. Our results, with a modified Iowa gambling task and a prisoner's dilemma game, illustrate the ways in which such an analysis can test hypotheses not possible with other techniques, is more parsimonious, and is more likely to be faithful to theoretical models.

Keywords: statistics, repeated measures, mixed models, Iowa Gambling Task, prisoner's dilemma.

## 1 Introduction

Experiments within the broad decision making tradition can involve subjects making multiple decisions, either with different partners or groups (e.g., Dawes & Messick, 2000) or as repeated decisions in a learning task (e.g., Damasio, 1994). Analyses of these data typically involve examination of how the rewards received are related to characteristics of the experiment and/or the individuals involved (e.g., Andreoni & Petrie, 2004; Bechara et al., 1996; Knutson et al., 2001; Parco et al., 2002). Complex multivariate analyses are relatively rare.<sup>1</sup>

This typical analytic approach has at least one major drawback. Most studies in this tradition involve multiple decisions for each individual. That is, each subject participates in several decisions. It is entirely possible

that individuals vary their decisions across the rounds of play in which they engage. Looking at the average of the results of these decisions could mask important variations and patterns within the data. Aggregated analyses do not allow the researcher to fully examine the ways in which characteristics of participants and the experimental design affect decisions. In other words, these analyses can provide descriptive results, but only hints of the decision-making dynamics within the experiments. We suggest that data from these experiments fit the classic form of multi-level, or nested, data sets in that they usually include multiple decisions from an actor, within one or more set conditions. Mixed models have been specifically developed to deal with multi-level units of observation and can help us understand the process or dynamics of decision making within our studies.

In this paper we describe the logic of mixed model analyses, contrasting them with the familiar within- and between- subject designs. We then demonstrate the use of these models with decision making data by reanalyzing data from two previously published articles. The first study (Peters & Slovic, 2000) involves a modification of the Iowa Gambling Task with multiple decisions by individual players in response to a variety of stimuli, and the second (Mulford, et al., 1998) involves prisoner dilemma games conducted within seven person groups. Finally, we briefly describe the wide range of possible applications

\*Address: Jean Stockard, Department of Planning, Public Policy, and Management, University of Oregon, Eugene, Oregon 97403. Emails: jeans@uoregon.edu bobrien@uoregon.edu, empeters@decisionresearch.org.

<sup>1</sup>An important exception is the recent work of Yechiam and Busemeyer (2005, see also Busemeyer & Stout 2002; Yechiam, et al., 2005; and Yechiam & Busemeyer, 2006) who explicitly analyzed models of choices in repeated play games. While these models are persuasive, we suggest that the techniques we describe are more general, are more flexible and can apply to more situations, and can incorporate important theoretical elements regarding repeated choices within decision making experiments by the way in which variables related to both the subjects and the experimental design can be considered.

Table 1: Examples of within subject, between subject, and mixed model designs.

1) The pure within subjects design: (AxBxS) design:								
	A <sub>1</sub>			A <sub>2</sub>				
	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>		
S <sub>1</sub>								
S <sub>2</sub>								
S <sub>3</sub>								
S <sub>4</sub>								
2) The pure between subjects design. Subjects are nested within A and B: AxB:								
	A <sub>1</sub>			A <sub>2</sub>				
	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>		
S <sub>1</sub>	S <sub>5</sub>	S <sub>9</sub>		S <sub>13</sub>	S <sub>17</sub>	S <sub>21</sub>		
S <sub>2</sub>	S <sub>6</sub>	S <sub>10</sub>		S <sub>14</sub>	S <sub>18</sub>	S <sub>22</sub>		
S <sub>3</sub>	S <sub>7</sub>	S <sub>11</sub>		S <sub>15</sub>	S <sub>19</sub>	S <sub>23</sub>		
S <sub>4</sub>	S <sub>8</sub>	S <sub>12</sub>		S <sub>16</sub>	S <sub>20</sub>	S <sub>24</sub>		
3) The mixed or split-plot design with (B) as the within subjects factor and (A) as the between subjects factor: AxB(S) design:								
	A <sub>1</sub>			A <sub>2</sub>				
	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>		
S <sub>1</sub>				S <sub>5</sub>				
S <sub>2</sub>				S <sub>6</sub>				
S <sub>3</sub>				S <sub>7</sub>				
S <sub>4</sub>				S <sub>8</sub>				

and the advantages of the approach, provide pragmatic advice for using hierarchical models, and encourage their broader use in decision making research. Although the major aim of this paper is to illustrate the utility of this method, we include an appendix with the SAS language used for our analyses to provide greater transparency and encourage use of the method.

### 1.1 The logic of mixed model analyses

A mixed model is one in which there are both between- and within-subject variables. As noted above, studies in the decision research literature often focus exclusively on between-subject variables in their analyses and ignore within-subject variables, those that affect the variations in decisions that one individual makes. Yet, variations in individuals' decisions are often theoretically and substantively important. Thus, we argue that it is more natural, fruitful, and parsimonious to analyze many decision mak-

ing studies using a mixed model approach. In addition, mixed models are very flexible and can adapt to a wide variety of experimental designs.

In a pure within-subjects design, all of the subjects serve in all of the treatment conditions. The first panel of Table 1 illustrates this design, assuming that there are two different design elements (A and B). In this design, each subject (S<sub>i</sub>) participates in each condition of design element A (A<sub>j</sub>) and design element B (B<sub>k</sub>). An example would be the classic Lichtenstein and Slovic (1973) study in which gamblers in a Las Vegas casino showed within-subject preference reversals based on response mode (bids and choices).

The pure between-subjects design is one in which each subject serves in only a single condition and there is no within-subject variation. This design is illustrated in the second panel of Table 1, where each subject (S<sub>i</sub>) is in only one condition of the two design conditions (A and B). In other words, in the pure between-subjects design, subjects are nested within A and B: AxB; each subject contributes a single observation on the dependent variable and thus appears in only a single treatment condition. Many decision making studies use this design (e.g., Tversky & Kahneman, 1974; Scharlemann et al., 2001) or simplify their data to mimic this design by summarizing across all of the decisions that an individual makes to form a summary score (Andreoni & Petrie, 2004; Parco et al., 2002). This summary number is then used as the dependent variable.

Researchers also use a design that involves repeated measures or conditions for one or more factors that are presented to each subject (within-subject variables such as a set of rewards in a neuroimaging study) and variables such as gender or culture or experimental conditions such as positive and negative frames that subjects do not share (between-subject variables). This design is frequently referred to by psychologists as a *mixed design* and is illustrated in the third part of Table 1, with B as the within-subjects factor and A as the between-subjects factor. The term mixed is particularly apt in that there are sources of variance that are produced by between-subjects differences and by within-subjects differences and both types of differences contribute to the statistical analysis.<sup>2</sup> This

<sup>2</sup>Sometimes researchers slip into mixed models when they use a paired t-test to see if a subject is more likely to cooperate when using Matrix A than when using Matrix B. We could use a paired t-test analysis that takes advantage of having two summary scores from each subject. But, as we will see, we can do this and much more by shifting our unit of analysis from the individual to the decision. Similarly, mixed models are more flexible than repeated measures analyses of variance, which are simply extensions of paired t-tests. The mixed model approach is most similar in logic to analysis of covariance, but is more flexible and elegant in its statistical properties. For instance, it allows us to include a measure of autocorrelation, which indicates the extent to which subjects do, or do not, make similar decisions from one decision to the next.

design allows us to assess within-subject differences at the same time as we study between-subject relationships. The design also allows us to model the dependency of successive decisions, what is technically called auto regression, or the tendency for individuals to behave in a similar manner from one decision to another.

Researchers have used mixed models in a variety of settings. Much of the original work on mixed models was conducted in agricultural settings and thus, statisticians also refer to such models as a *split-plot design*. In the agricultural setting, plots were often considered random with each plot exposed to the same set of fixed experimental conditions on one of the experimental variables, while the plots were nested within experimental conditions of a different experimental variable. These mixed designs contain at least one within-subject (plot) variable and at least one between-subject (plot) variable (Cochran & Cox, 1957). The panel models used by economists have both a between-units (often firms) component and a within-units (changes over time) component (Green, 2000). Sociologists have used Hierarchical Linear Models to examine the influence of characteristics of school classrooms and entire schools on individual student achievement (e.g. Bryk & Raudenbush, 1988). Developmental psychologists and educational researchers have used the models to examine "growth curves," changes in individuals' achievement or developmental characteristics (the within-subject variable) for students varying on different characteristics (the between-subject analysis) (e.g. Huttenlocher et al., 1991; see also Raudenbush & Bryk, 2002, pp. 5-14).

Thus, there are a wide variety of analytical methods that pay close attention to within and between individual (or group) variables. Two widely used general modeling strategies are hierarchical linear models and mixed models. Either of these allows for the modeling of random variation for individuals and of decisions nested within individuals. For the situations analyzed in this paper, either hierarchical linear modeling or a mixed modeling approach can be used.

We suggest that the general logic of this approach can be applied to the issues facing experimenters studying decision making and help researchers develop a more accurate understanding of the elements of the decision making process that lead to hypothesized outcomes. Within many such experiments, a given individual makes numerous choices, often across a variety of experimental settings. Because individuals make multiple decisions, these decisions are not statistically independent. The usual method of handling such data is simply to aggregate all decisions to the level of the individual and ignore any intra-individual variation. Such a practice, however, ignores potential variation and limits the extent to which researchers can assess the impact of experimental vari-

ables on individual choice. It also encourages researchers to conduct numerous separate bivariate analyses. Use of mixed models can provide both a more inclusive, yet more parsimonious and efficient, method of analysis. We now illustrate the utility of these models by reanalyzing data from two different studies.

## 2 Study 1: Reactivity to positive and negative events in a modified Iowa Gambling Task (Peters & Slovic, 2000)

Peters and Slovic (2000) examined the relationship of individual differences in affective information processing to choices in a game designed to mimic real life decisions involving gains, losses, and uncertainty. The task was a modified version of the Iowa Gambling Task. Peters and Slovic were especially interested in how individual differences in reactivity to positive and negative events were related to decisions. To examine this question, they used a card-selection task with four decks that varied in gains, losses, and expected value. Subjects were presented one card at a time and could either choose or reject the card. The four decks were labeled A, B, C, and D, and subjects initially knew nothing else about the decks but learned as they chose and received feedback about gains and losses. The task structure and characteristics of the decks are shown in Table 2.

Below we describe their original analysis and results. We then present our replication of their original analysis using mixed modeling and explore additional hypotheses. This additional exploration is only possible with the increased flexibility afforded by the use of mixed modeling. The additional hypotheses tested were not even developed in the original paper because it was not possible to test them with the analytic techniques used in that paper, yet, as explained below, the additional hypotheses address important theoretical issues.

### 2.1 The original analysis

The original work hypothesized that individuals who reported greater reactions to affectively charged events in life would make choices in this card task that were congruent with their self-reported tendencies for reactivity. Results consistent with this hypothesis would support the notion that experienced, anticipatory affective reactions are one of the mechanisms underlying choice. The study tested two hypotheses: 1) Greater reactivity to negative events would be related to fewer choices from high-loss decks; and 2) Greater reactivity to positive events would

Table 2: Payoff structure for the modified Iowa gambling card-selection task

Payoff Variables	Modified Schedule — Decks			
	A	B	C	D
Gain Amount	\$50–\$150	\$150–\$250	\$50–\$150	\$150–\$250
Average Gain Amount	\$100	\$200	\$100	\$200
Probability (Gain)	0.5	0.5	0.8	0.5
Loss Amount	\$100–\$200	\$200–\$300	\$200–\$300	\$100–\$200
Average Loss Amount	\$150	\$250	\$250	\$150
First loss appears at card	3 or 4	3 or 4	3 or 4	3 or 4
Expected value of a card	– \$25	– \$25	\$30	\$25
Coding of good decks	0	0	1	1
Coding of high-loss decks	0	1	1	0
Coding of high-gain decks	0	1	0	1

be related to more choices from high-gain decks (Peters & Slovic, 2000).

Peters and Slovic tested and found support for these hypotheses by using individual subjects as the unit of analysis and calculating three separate regression equations (Table 3, p. 1471). Each of the equations had the same independent variables, both measures of individual differences: BIS (a measure of reactivity to negative events) and Extraversion (a measure of positive reactivity).<sup>3</sup> The equations differed in the dependent variable: one was the number of selections from the “high-loss decks” (B and C), one was the number of selections from high gain decks (B and D), and the third was the number of selections from “good” decks (C and D). Hypothesis one predicted that those with high BIS scores would make fewer choices from the high-loss decks. Hypothesis two predicted that those with higher scores on “extraversion” would choose more cards from the high gain decks.

A separate analysis in the original article (Table 4) regressed the number of choices from high-loss decks and high-gain decks (separately) on the measures of reactivity (BIS for high-loss decks and extraversion for high-gain decks) and a measure of conscious knowledge about the decks. The analysis was conducted for the total group and for different sets of choices (the first set of 20 choices, the second, etc.), for a total of 12 regression equations. It was hypothesized that the relation of reactivity to choices would be greater in the early parts of the game before conscious knowledge influenced results. This hypothesis was supported for the analysis of high-loss decks, but

not for the analysis of high-gain decks. Finally, the original analysis used correlation coefficients to examine the possibility that subjects' current mood state influenced choices and found no support for this possibility.

## 2.2 Replicating and extending the analysis with mixed modeling

The original analysis, as is typical in this area, implicitly assumed that events occurring throughout the game, and the immediate influence of events within the game, “averaged out” across the entire experience of a subject. Yet, it is possible that subjects' experiences from play to play within the game could influence their individual decisions. For instance, the result of decision one could affect what the player does in decision two, etc. This hypothesis would be consistent with immediate experiences producing a mood state that influences subsequent decisions. In particular, based on the affect heuristic and affect's role as information in decisions, we might expect that a previous positive payoff would produce a positive mood and that this positive mood would be used as information in the evaluation of the subsequent choice, making it more likely that the subsequent deck would be chosen (Isen, 1997; Peters, 2006). Because their analysis used data aggregated across all decisions, Peters and Slovic could not test the hypothesis that irrelevant mood states produced from prior feedback would act as information in subsequent choices. As a result, estimates of the influence of individual differences might be inaccurately specified.

It would also be expected that subjects' experiences with a given deck would affect their response when pre-

<sup>3</sup> A third independent measure (REI-Rational) was also used in the analysis but was found to be unrelated to the dependent variables. We have omitted this variable from our reanalysis of the data.

sented with that deck again through some type of learning process. If subjects gained (or lost) money the previous time a card from a given deck was chosen, they could well expect to gain (or lose) when given that deck again, and this would affect their choice. By aggregating across all the decisions, as occurred in the original analysis, one cannot adequately test the impact of possible learning on subjects' choices.

With mixed modeling, we can control for the possible influence of the immediate situation and prior feedback and thus provide a more stringent test of the individual differences hypothesis.<sup>4</sup> In addition, gambling tasks, such as the one we analyze here, implicitly assume that subjects are linear and consistent in their behavior. Only through actually looking at variations in decisions, as is possible with mixed modeling, can we examine this implicit assumption.

### 2.3 Analysis plan

In the present analysis we focus not on the total number of choices from different types of decks but, instead, on each individual choice. Thus, instead of looking at the summary responses of 72 individuals, we look at the decisions that these individuals made: over 12,000 decisions. Instead of examining the total number of choices from a particular deck, our dependent variable is the actual decision that a respondent made – whether or not to accept a card (choosing a card is coded as “1;” rejecting a card is coded as “0”).

Our independent variables can be seen as either being “within subjects” or “between subjects.” The within-subject variables are those that relate to the individuals' decisions and allow us to examine the influence of the ongoing experience within the game: the type of deck presented,<sup>5</sup> whether the decision is made early in the experiment, and two variables that reflect the results of subjects' recent experiences in the game: specifically, the amount of money that they gained or lost in their previous play as well as how much money they gained or lost in their most recent play with the presented deck.

The between-subject variables involve the characteristics of the individual and interactions of these characteristics with the deck with which they are playing. To align with the description of the research in the original

article (p. 1467), our base model includes, as between-subject predictors, the BIS score (a measure of negative reactivity), Extraversion (a measure of positive reactivity), and the two pre-test measures of mood (positive and negative) as control variables. To replicate the analysis of results with different decks, we include four interaction terms: 1) between BIS and the high-loss decks, 2) between extraversion and the high-gain decks, 3) between negative mood and the high-loss decks, and 4) between positive mood and the high-gain decks. Based on the original analysis and its results, we expect only the first two of these interactions to be significant (although theoretical reasons exist for all four interaction terms to be significant). As in Peters and Slovic we use an additional variable indicating early choices from the decks, adding a dummy variable for the first 20 choices (we coded a variable “Early” as “1” if the choice was in the first 20 choices and “0” otherwise), and three-way interactions of BIS, choices from high-loss decks, and “Early” as well as extraversion, high-gain choices, and “Early.” Based on results from the original analysis, we expect a stronger influence of BIS on the early high-loss choices than later high-loss choices, but no significant influence of the interaction of extraversion scores and early high-gain choices.

The present analysis involves all decisions made in the game, using Deck A as the omitted category in the dummy variables for deck, as shown in the coding displayed in Table 2. The analysis includes the characteristics of the deck as independent variables, essentially looking at how these characteristics affect decisions both as main effects and in interaction with other variables. While the original analysis involved 15 separate regression equations and examined the influence of mood only through correlations, the present analysis incorporates all of the proposed independent variables into one model, including the measures of mood, thus providing a more parsimonious test of the hypotheses.

Finally, and perhaps most importantly, the present analysis controls for possible influences of the course of play – the immediate feedback that a player is receiving – on decisions. These situational variables could be seen as classic examples of “state” variables, in contrast to the “trait” variables captured by the BIS and Extraversion scores. This situational control is thus quite important in providing a more accurate estimate of the impact of individual differences (trait variables) on decisions. It also helps ensure that the models are properly specified because a broader range of appropriate control variables is included. In addition, this state variable is an important potential source of affect that may be used as information by decision makers, but that Peters and Slovic were unable to examine.

<sup>4</sup>Through including both of these variables in our model we could be seen as implicitly testing the expectancies and learning models identified by Yechiam and Busemeyer (2005). The variables related to the immediate situation may relate to expectancies, while those related to prior feedback relate to learning. To apply our technique to these learning models it would, of course, be necessary to create variables that represented memory of past pay-offs and to use a form of multinomial logistic regression to analyze the possibility of multiple choices.

<sup>5</sup>We include dummy variables to represent the three types of decks (high-loss, good, and high-gain decks). See Table 2 for the specific codes used.

Table 3: Regression (HLM) of decision to take a card on last pay, reactivity measures, mood, deck, time in game, and interactions

Fixed Effects	Model 1		Model 2		Model 3		Model 4	
	Estimate	t	Estimate	t	Estimate	t	Estimate	t
Intercept	0.47	5.77***	0.56	0.87	0.31	0.47	0.11	0.17
Level 1 (Within-Subject) Variables:								
Last Pay	-0.001	-8.52***	-0.001	-8.52***	-0.001	-8.47***	-0.001	-8.59***
Last Pay from Deck	0.002	20.12***	0.002	20.11***	0.002	19.86***	0.002	18.27***
Good Deck	0.23	5.87***	0.23	5.87***	0.23	5.87***	0.24	6.15***
High Reward Deck	-0.28	-7.24***	-0.28	-7.24***	-1.52	-7.68***	-1.62	-7.69***
High Loss Deck	0.28	7.14***	0.28	7.14***	2.13	8.40***	2.16	8.42***
Early	----	----			----	----	1.27	3.42***
Level 2 (Between-Subject) Variables:								
BIS	----	----	-0.28	-1.78	-0.03	-0.16	-0.01	-0.08
Extravert	----	----	-0.001	-0.01	-0.09	-1.31	-0.07	-1.09
Negative mood	----	----	0.23	1.17	0.32	1.55	0.31	1.49
Positive mood	----	----	0.17	1.57	0.10	0.89	0.10	0.90
BIS*High Loss	----	----	----	----	-0.53	-6.91***	-0.53	-6.32***
Extra * High Rew.	----	----	----	----	0.17	5.37***	0.18	5.30***
Neg. mood * Hi Loss	----	----	----	----	-0.19	-1.75	-0.19	-1.70
Pos. Mood * Hi Rew	----	----	----	----	0.14	2.59**	0.14	2.64**
Early * BIS	----	----	----	----	----	----	-0.05	-0.36
Early * Extravert	----	----	----	----	----	----	-0.10	-1.67
Early * Hi Loss	----	----	----	----	----	----	-0.11	-0.17
Early * Hi Rew	----	----	----	----	----	----	0.60	1.27
Early * Bis * High Loss	----	----	----	----	----	----	-0.09	-0.42
Early * Ex * Hi Rew	----	----	----	----	----	----	-0.08	-0.92
Autoregressive (1)	-0.06	-6.16***	-0.06	-6.18***	-0.06	-6.22***	-0.06	-6.64***

Note: For the simplest, intercept only, model the coefficient associated with the intercept is .50 (t = 6.80), whether or not the autocorrelation term is included. The Autoregressive (1) value for this model is -0.05, (z = 5.17).

## 2.4 Results

We began our analysis with the simple baseline, “intercept-only,” model. The dependent variable is the decision to choose a card, with a value of 1 indicating that the card was chosen and 0 indicating that it was not. Because the dependent variable is a dichotomy, we used generalized linear models (Raudenbush & Bryk, 2002), employing a Glimmix procedure within SAS. The Glimmix procedure is a subroutine designed to transform linear mixed models to formats that can incorporate dichotomous dependent variables as well as other data transformations such as logs (Littell et al., 1996).

The intercept-only model is equivalent to a logistic re-

gression testing the null hypothesis that individuals all have the same proportion of decisions to accept a card. Results indicate that subjects do differ significantly in this tendency ( $z = 5.39, p < .001$ )<sup>6</sup>

The models designed to replicate the material in Peters

<sup>6</sup>When the dependent variable is measured on an interval scale the results with this model provide additional information. Specifically, the estimate of the variance between subjects, the differences between the subjects in their decisions, can be compared to the estimate of the “residual variance,” the variance that is within subjects (the variance of scores around the average for each subject). These two values can be used to determine the proportion of variance in decisions (the dependent variable) that can be attributed to differences between individuals. This is called the intraclass correlation coefficient (rho) and is calculated by dividing the variance between individuals by the sum of this variance and that for the residual.

and Slovic (2000, Tables 3 and 4) using mixed modeling are shown in Table 3. The first model includes all the within-subject variables: the money received from the previous play, the money received from the previous play with the presented deck, and the dummy variables for the nature of the presented deck. Model 2 adds the two reactivity measures (BIS and extravert) and the two mood measures; Model 3 adds interactions of BIS with the high punishment deck, extraversion scores with the high-reward deck, negative mood with the high punishment deck, and positive mood with the high-reward deck. Finally, Model 4 adds a dummy variable indicating that the choices occurred in the first fifth of the game and interactions of these early choices with BIS and extraversion and with high loss and high-reward decks respectively.<sup>7</sup>

Each of the models in Table 3 includes an autoregressive term to control for the fact that the same individual makes multiple decisions. The autoregressive term can simply be seen as a correlation (varying from  $-1.00$  to  $+1.00$ ) between the choices that individuals make on successive decisions. Note that, for this data set, the autoregressive term is quite small in magnitude and negative for all models. (See the last row of figures, labeled Autoregressive (1) in Table 3.) This indicates that a subject who agrees to accept a card at one opportunity is slightly less likely to accept a card at the next opportunity (controlling for all of the other variables in a model). The autoregressive term is included to control for similarities within individuals in their decision-making patterns; and including this control helps ensure that the estimates of other effects are more accurate.

The regression coefficients associated with the variables in each of the models are given in Table 3. Because the dependent variable is a dichotomy, the regression coefficients are logs of the odds of choosing a card. A coefficient of zero indicates that an increase or decrease in the independent variable is not associated with subjects selecting additional cards from that deck or deck type, positive values indicate that a higher score is related to choosing a given card, and negative values indicate that a higher score is related to subjects rejecting a given card in

similar fashion. The t-values associated with each coefficient test the null hypothesis that a coefficient equals zero. The values may also be exponentiated to compute the associated odds.<sup>8</sup> For instance, with the base-line intercept-only model, the coefficient of .50 ( $t=6.80$ ) associated with the intercept indicates that subjects are significantly more likely to accept a card than to reject a card. To convert this log odds ratio to log odds we simply exponentiate the value of this coefficient ( $e^{.50} = 1.64$ ), indicating that the odds of a subject choosing rather than rejecting a card are 1.64 to 1.

The results in Table 3 indicate that all of the level one variables significantly influence subjects' choices. Subjects were significantly less likely to accept a card if their last decision to accept a card produced a higher payoff. This result runs counter to the affect-as-information (affect heuristic) hypothesis, but is consistent with previous research by Isen (1999) who finds that subjects in positive moods are more averse to risks in real choices than those in neutral moods, presumably in order to maintain their positive mood. Thus, subjects who received a higher payoff from a previous choice may have been in a more positive task-induced mood state and, when confronted with the next risky choice (and all of these choices are risky), were more likely to reject that deck. At the same time, subjects were significantly more likely to accept a card if their previous accept decision from that specific deck had produced a higher payoff. Finally, subjects' choices varied significantly between decks. These findings support the expectation that subjects learned about characteristics of the decks and were not influenced only by moods. Note that these influences could not be observed, or controlled, in the original analysis because all choices were aggregated and choices for each deck were examined in separate equations.

The coefficients associated with the between-subjects variables indicate that, net of the immediate influence of previous plays in the game, the measures of individual differences significantly influenced choice of cards. For the most part, these influences are parallel to the original hypotheses of Peters and Slovic. The BIS, extraversion, and mood variables do not affect choice by themselves, but only in combination or interaction with the characteristics of the cards. (Compare results in Model 2 with those in Models 3 and 4.) Specifically, subjects with greater reactivity to negative events (higher BIS scores) were significantly less likely to choose cards from the high-loss decks (significant two-way interaction of BIS and high-loss decks). Contrary to expectations and the results in the original analysis, however, this pattern was not stronger in the early stages of the game (as indicated

<sup>7</sup>When the dependent variable is measured on an interval scale it is possible to compare various models when one model is a subset of another. Three different statistics are commonly used in these comparisons: 1) comparing the between-subject variance from one model to another, using a proportionate reduction of error calculation; 2) comparing log likelihood ratios associated with each model using standard procedures of model fit; and 3) examining Schwarz's Bayesian Criterion (BIC) (Schwarz, 1978), a number calculated from the Likelihood Ratio Chi-square value, the number of variables in the model, and the number of cases (see Singer, 1998 and Raudenbush and Bryk 2002 for details). Unfortunately, such statistics are not yet available for analyses that use the glimmix procedure to produce logistic regressions or other procedures appropriate for non-intervally measured dependent variables (Phil Gibbs, SAS Corporation, personal communication, March 13, 2006).

<sup>8</sup>Probably the easiest way to do these calculations is through Excel, or another spreadsheet program, and the (exp) function; or with a calculator.

by the insignificant three-way interaction). Using simple arithmetic and the values for the dummy variables associated with time in the game and type of deck (shown in Table 2), the coefficients associated with the BIS measure for both high-loss decks and other types of decks can be easily calculated from the results in Model 3:  $-.56$  for high-loss decks and  $-.03$  for other decks.<sup>9</sup> In other words, the measure of negative reactivity (BIS) has an influence on only choices from the high-loss decks.

Similarly, as expected, there are significant interactions with the measure of extraversion, but again only the two-way interactions are significant. Subjects with higher extraversion scores are more likely to choose cards from high-gain decks and less likely to choose cards from low-gain decks. Calculations indicate that the coefficient associated with extraversion is  $.08$  for cards from the high-gain decks, but  $-.09$  for cards from the low-gain decks.

It is important to remember that these results are independent of subjects' pre-task mood. However, in contrast to what was expected from the original study, the interaction of positive mood and choices from the high-gain decks was significant. Subjects with a more positive pretest mood were more likely to choose cards from the high-gain decks even with the strong controls related to type of deck and previous payments. No significant effects were found for negative mood. This mood-congruency effect had been hypothesized but not found in the earlier study.

Finally, the results provide insight into the processes that guide subjects as they explore in this task. The significant coefficient associated with choices early in the game indicates that subjects were more likely to select rather than reject cards in the early stages of the game and that exploratory behavior declined with experience. However, as indicated by the insignificant interactions in Model 4, the effects of the measures of individual differences were consistent from early to later stages of the game and across the decks that were used.

To summarize, these results generally replicate the findings of the earlier analysis, but do so through the use of one equation rather than several, providing a much more parsimonious analysis and a more completely specified model. Because the analysis controls for the influence of immediate experiences within the game, estimates of the effects of measures of individual differences,

<sup>9</sup>These coefficients are calculated by summing the coefficients associated with the various component elements related to BIS in Table 3. For instance, the coefficient associated with BIS for high-loss decks ( $-.56$ ) is the sum of the coefficient for BIS ( $-.03$ ) and the interaction of BIS and high-loss decks ( $-.53$ ). The coefficient for BIS for non-high-loss decks later in the game (zero value for high-loss deck) is simply equal to the coefficient for BIS ( $-.03$ ). As noted in the text, these coefficients can be exponentiated to obtain odds for making a choice at various values of BIS. Coefficients from Model 3 were used for this calculation because the 3 way interactions added in Model 4 were insignificant.

such as BIS and extraversion, should be more accurate. In addition, the technique allows the exploration of hypotheses that could not have been examined with other methods. Specifically, we were able to examine the way in which information gained through the game affects decision making and how the process of decision making changes throughout the game. These findings could not have been obtained through the earlier analysis technique and are important both substantively and theoretically.<sup>10</sup>

### 3 Study 2: Using mixed models to analyze results from a prisoner's dilemma game

Our second example comes from an experimental design that used a group-level prisoner's dilemma game to examine the relationship of perceived attractiveness to outcomes. Subjects in this experiment (Mulford, et al., 1998) were divided into 32 seven-person groups. In these groups they participated in two rounds of six games, one with each person in the group. Subjects played with matrices with three different levels of difficulty, and the same matrix was used with a given partner in both rounds of play.<sup>11</sup>

Table 4 summarizes the design of the study. As listed in the table, the groups differed in the choices that were made, whether or not they had feedback, and whether or not the decision was the same from one round to the next. Half of the groups had a two-step process, first making a decision whether or not to enter play with the other person and then choosing whether or not to cooperate. Half of the groups made only the standard cooperate/defect choice; in other words they only had to decide whether or not to cooperate with the other person. Half of the groups had the same set of choices in both rounds, and half of the groups received feedback on the results of the first round before playing the second. The design resulted in eight possible combinations of conditions with two groups in each of these conditions. As shown in Table 4, slightly less than half of the subjects with the option to opt-out of the game did so. Of those who played the game, only about one-third chose to cooperate with their partner.

At the time of making each decision, subjects indicated their perception of the probability that the person they were playing with would cooperate. At the end of play they gave ratings of the attractiveness of all others in the

<sup>10</sup>The original analysis of Peters and Slovic only examined the first 100 choices subjects made. When our analysis was restricted to only these choices we obtained the same results as reported here.

<sup>11</sup>The payouts were as follows:

Matrix 1: cc = 2,2; cd = -7,5; dc = 5,-7; dd = -4,-4.

Matrix 2: cc = 1,1; cd = -4,4; dc = 4,-4; dd = -1,-1.

Matrix 3: cc = 1,1; cd = -2,2; dc = 2,-2; dd = -1,-1.



Table 4: Design of Mulford, et al., experiment and choices made by round of play

Condition	Trinary, Round 1	Trinary, Round 2	Feedback	Number of Subjects
1		X	X	21
2			X	19
3	X	X	X	22
4	X		X	21
5		X		27
6				26
7	X	X		26
8	X			23
Total				185

  

Choices	Round 1	Round 2	Total
Opt-Out	41%	40%	41%
Play	59%	60%	59%
Total	100%	100%	100%
N (choices)	552	576	1128

  

Defect	63%	68%	65%
Cooperate	37%	32%	35%
Total	100%	100%	100%
N (choices)	882	879	1761

Note: Subjects in the “trinary” condition had two sets of choices. The first was whether or not to enter play and the second was whether to cooperate or defect with the partner. In this condition only subjects who chose to enter play made the cooperate/defect decision. Subjects who did not give ratings on the attractiveness measure or who gave all others the same score were eliminated from the analysis.

group as well as themselves. Subjects that either gave all players the same attractiveness rating, or who refused to give the ratings, were omitted from the analysis, leaving 185 subjects (see Mulford et al., 1998, for details).

The published article focused on the influence of ratings of attractiveness on subjects' choices. The original analysis did not aggregate decisions across subjects but used decisions as the unit of analysis, a procedure that, as noted above, is not typical in the field. Decisions (play/not play and cooperate/defect) were regressed on the gender of the subject, gender of other, subjects' expectation that the other would cooperate, subjects' self-rating of attractiveness, and subjects' rating of the other's attractiveness. Findings indicated that subjects were more likely to enter play and to cooperate with others that they found attractive. Men who saw themselves as more attractive more often cooperated than other men, while women who saw themselves as more attractive cooperated less often than other women. In addition, there

was a significant interaction of self-ratings and ratings of others, with those who saw themselves as attractive especially likely to cooperate with others that they perceived to be attractive. The effect of perceived attractiveness on decisions was independent of expectations of others' cooperation.

To control for the possibility that individuals have certain patterns of decision-making that might persist across a game, a dummy variable was included for each individual in the analysis. This procedure could be seen as a gallant attempt to control for individual differences, but it consumed a large number of degrees of freedom. In addition, and, perhaps more important, this loss of degrees of freedom led the researchers to ignore the possible effect of feedback on decisions and any differences between the matrices used in the game. The original analysis only examined decisions in the first round of play and ignored the possible effects of different decision matrices.

Table 5: Regression (HLM) of decision to play (Mulford et al.) on gender of subject and other, matrix of play, ratings of attractiveness, and feedback.

Fixed Effects	Model 1		Model 2		Model 3	
	Estimate	t	Estimate	t	Estimate	t
Intercept	0.60	1.70	-0.20	-0.51	0.98	1.17
Level 1 (Within-Subject) Variables:						
Matrix 2	1.41	8.10***	1.39	7.64***	1.39	7.63***
Matrix 3	1.82	9.84***	1.81	9.41***	1.81	9.37***
Round 2	-0.16	-0.63	-0.15	-0.60	-0.16	-0.64
Male other	0.04	0.26	0.06	0.39	0.07	0.42
Subject's rating of other's attractiveness	0.08	2.51**	0.14	3.61***	0.14	3.66***
Subject's expectation other will defect	-0.03	-8.18***	-0.02	-5.05***	-0.02	-5.12***
Feedback	-0.54	-1.60	2.40	3.35***	2.41	3.35***
Feedback by other's attractiveness	----	----	-0.19	-2.85***	-0.19	-2.88**
Feedback by expectation of other	----	----	-0.03	-3.97***	-0.03	-3.91***
Level 2 (Between-Subject) Variables:						
Male subject	----	----	----	----	-0.49	-1.33
Subject's self rating of attractiveness	----	----	----	----	-0.13	-1.32
Autoregressive (1)	0.13	3.41***	0.13	3.04**	0.13	3.12**

Note: For the simplest, intercept only, model the coefficient associated with the intercept is .36 ( $t=2.80^{**}$ ). When the autocorrelation term is added the intercept value is .44 ( $t=2.92^{**}$ ). The autocorrelation term in this model is .19 ( $z = 4.76^{***}$ ). Intermediate models with only some of the Level 1 variables were examined and coefficients were remarkably similar until the interaction terms were added.

We suggest that a more appropriate analysis for this data set involves a mixed model, including an autoregressive term to model dependencies between individual decisions. Using a mixed model provides a more appropriate control for individual tendencies to react in given ways and also allows us to examine the effects of feedback and different matrices on subjects' decisions, tests that were not possible with the techniques that were previously used. Understanding how feedback affects the results is substantively important. Even though the results from the first round of play indicate that others' attractiveness induced a decision to play and cooperate, even when they thought the other would not cooperate, perhaps subjects learn from experiences. If such learning occurs we could expect a significant interaction between feedback and ratings of attractiveness. Including the matrix of play also provides an important control, helping to determine the extent to which choices reflect characteristics of the game (the presented matrix) or individual differences in cooperative behavior. Neither of these questions could be examined with the original analysis.

### 3.1 Decisions to play

Table 5 summarizes the models for the analysis of the decision to play. Predictor variables at the level of decisions (level one) include the matrix that was used in the play, whether the play occurred in round one or two, gender of the other person, the subject's rating of the other's attractiveness, the expectation that the other would cooperate, and whether or not the play occurred after the subject had received feedback. Predictor variables at level two, the level of the decision maker, are the gender of the subject and the subject's self-rating of attractiveness.

The results with the intercept-only model indicate that individual subjects vary in their choices to enter play ( $z = 5.85$ ,  $p < .001$ ). In addition, the autocorrelation term is both positive and significant (.19 in the intercept-only model) and, as shown in Table 5, slightly smaller, but still significant, in the more complex models. This indicates that subjects tend to make the same decision from one encounter to the next.

The results in Model 1 in Table 5, which includes only the level one, or within-subject variables, replicate

the original work with subjects more likely to enter play when they perceive that the other person in the interaction is more attractive and less likely to play when they believe that the other person will not cooperate with them. In addition, the results document the strong influence of the matrix of play on decisions, something that was not controlled in the original analysis. Subjects were much more likely to choose to play with others (even with expectations and perceived attractiveness of others controlled) with the two matrices that provide less risk of loss.

While simply receiving feedback about the previous decision of the other person did not affect subjects' overall tendency to play (as shown by the insignificant coefficient associated with feedback in Model 1), there were significant interactions of feedback with both the perceptions of others' attractiveness and expectations that the other would defect (Models 2 and 3). Subjects who did not receive feedback about the result of their first round of play continued to be more likely to play with others whom they saw as attractive ( $b$  for attractiveness = .14), no matter what their expectation of the other's behavior. In contrast, those who did receive feedback were actually somewhat less likely than other people to play with others they saw as attractive ( $b$  for attractiveness =  $-.05$ ). These differences are not trivial. For instance, if a subject rated the other person as a 10 (out of a range of 1–11) the odds that they would choose to play with him/her were 3.2 to 1.0 if they had not received feedback regarding their decision in round 1, but about even (1.1 to 1.0) if they had received feedback. If the other person received a rating of 5 (the middle to low part of the scale), the odds of the subject choosing to play were 1.6 to one without feedback and 1.4 to 1.0 if they had received feedback.<sup>12</sup>

In addition, the interaction of feedback and expectations of others was statistically significant; that is, feedback led subjects to be even more attuned to their expectations regarding others' defection. For those without feedback the  $b$  associated with expectations =  $-0.02$ ; for those with feedback the  $b$  associated with expectations =  $-0.05$ . The between-subject variables (subject's gender and self ratings of attractiveness) had no influence on subjects' decisions to play and the coefficients associated with the within-subject variables remained unchanged when these variables were added in Model 3.

### 3.2 Decisions to cooperate

Table 6 gives the coefficients associated with variables in the models predicting cooperative behavior. As with the decision to play, the results with the intercept-only model indicate that individual subjects vary in their choice to

cooperate ( $z = 6.73$ ,  $p < .001$ ). In addition, the autocorrelation term is both positive and significant (.11) in the intercept-only model and slightly smaller, but still significant, in the more complex models. This indicates that subjects tend to make the same decision from one encounter to the next, although this tendency is not as strong as with the decision to play.

The strongest influence on subjects' decision to cooperate is their expectation of the others' behavior, followed by the impact of Matrix 2. If subjects expect the other in the interaction to defect or if they played with Matrix 2, they were less likely to cooperate. At the same time, as found in the original study, when subjects played with others they perceived as being more attractive they were more likely to cooperate, independent of their expectations of the other's behavior.

In contrast to the analysis of decisions to play, the influence of perceptions of others' attractiveness persisted even when the presence of feedback was controlled. Also in contrast to the analysis of decisions to play, interactions involving subject-level variables also significantly influenced choices. Specifically, women who saw themselves as more attractive were less likely to cooperate, but self ratings of attractiveness had virtually no influence for men ( $b$  for attractiveness =  $-.25$  for women and  $+.03$  for men). In addition, there was a significant tendency for subjects who saw themselves as attractive to also cooperate with others whom they saw as attractive. Both of these results replicated the findings obtained in the original study.<sup>13</sup>

Again, the differences are not trivial. If subjects rated themselves as 10 on the 11 point scale of attractiveness, the odds of cooperating with others would be .85 to 1 for male subjects, but only .40 to 1 for female subjects (assuming all other characteristics were at the mean). The situation is reversed for those who rate themselves lower in attractiveness. Men who rated themselves as 5 would have an odds of cooperating of .42 to 1, while the odds for women are .80 to 1.<sup>14</sup>

To summarize, the results from our reanalysis of the data from Mulford and associates (1998) support their original findings, but provide important extensions and elaborations, none of which could have been accomplished with the original analytical technique. Specifically, the extended analyses confirm that subjects in prisoner dilemma games are more likely to enter play and to cooperate with others whom they perceive to be more at-

<sup>12</sup>These results were calculated in the manner explained in footnote 9.

<sup>13</sup>The three way interaction of gender, self ratings and ratings of others was insignificant, indicating that the result held for both men and women. In addition, neither of the significant interactions regarding subject level variables that appeared in the analysis of cooperation appeared in the analysis of play decisions.

<sup>14</sup>Again, these figures were calculated by substituting average values for all variables except those that vary in the interactions, calculating the expected value, and exponentiating the results.

Table 6: Regression (HLM) of decision to cooperate (Mulford et al.) on gender of subject and other, matrix of play, ratings of attractiveness, and feedback.

Fixed Effects	Model 1		Model 2		Model 3		Model 4	
	Estimate	t	Estimate	t	Estimate	t	Estimate	t
Intercept	0.44	1.58	0.49	1.66	1.14	1.88	2.45	3.10**
Level 1 (Within-Subject) Variables:								
Matrix 2	-0.65	-4.60***	-0.66	-4.61***	-0.66	-4.60***	-0.64	-4.52***
Matrix 3	0.22	1.63	0.22	1.62	0.22	1.63	0.23	1.68
Round 2	-0.12	-0.78	-0.12	-0.77	-0.11	-0.75	-0.12	-0.78
Male other	-0.21	-1.72	-0.21	-1.71	-0.21	-1.70	-0.22	-1.81
Other's attract.	0.10	3.52***	0.09	3.09**	0.09	3.17**	0.07	2.10*
Expectation of other	-0.03	-10.94***	-0.03	-9.88***	-0.03	-9.90***	-0.03	-9.89***
Feedback	-0.26	-1.18	-0.55	-0.93	-0.55	-0.94	-0.58	-0.99
Feedback X other attractiveness	----	----	0.03	0.46	0.03	0.47	0.04	0.54
Feedback X expectation of other	----	----	0.002	0.29	0.002	0.29	0.002	0.28
Level 2 (Between-Subject) Variables:								
Male subject	----	----	----	----	0.07	0.26	-2.04	-1.85
Self rating of attractiveness	----	----	----	----	-0.09	-1.27	-0.25	-2.58**
Male * self rating	----	----	----	----	----	----	0.28	1.98*
Both High Attract.	----	----	----	----	----	----	0.55	2.31*
Autoregressive (1)	0.09	2.89**	0.09	2.90**	0.09	2.89**	0.09	2.97**

Note: For the baseline intercept only model the coefficient associated with the intercept is  $-.85$  ( $t=6.66^{***}$ ). When the autocorrelation term is added, the coefficient associated with the intercept changes only slightly ( $-0.82$ ,  $t = 6.61^{***}$ ). The autocorrelation term in the base line model is  $.11$  ( $t=3.73^{***}$ ). Intermediate models with only some of the Level 1 variables were examined and coefficients were remarkably similar until the interaction terms were added.

tractive, even when they expect that the others will not cooperate with them. At the same time, the results suggest that receiving feedback about the results of the game can alter the impact of perceived attractiveness of others, at least with respect to the decision to enter play with others. Results with the analysis of cooperate decisions confirm the role of perceived attractiveness of others and expectations of others' behavior, but also indicate that receiving feedback had no effect on these influences. In other results that could not be obtained with the original analysis, the results with mixed modeling illustrate the important impact of the matrix of play on decisions and demonstrate (through the measure of autocorrelation) the extent to which individuals maintain consistent patterns of decision making from one decision to another.

## 4 Discussion and general applications

Using a general mixed model approach appears to have a number of advantages over the analyses used in the original articles and analyses common within the field. First, and perhaps most important, mixed models allow us to take advantage of the full richness of these types of decision-making data sets. These very flexible techniques allow researchers to explore changes in behavior throughout a game, include a range of control variables, and test hypotheses that otherwise could not be examined regarding the process by which decisions are made. The models also allow researchers to examine a wide range of possible interaction effects that could be theoretically and substantively important.

Second, while researchers in this field have used a variety of specific multivariate techniques in the past, we suggest that mixed models are more flexible, can be used with a wider variety of experimental designs, and can test hypotheses related to associations on varied levels of analysis (e.g. decisions, individuals, and their interactions). For instance, repeated measures analyses of variance are useful when the number of repeated trials is not overly large (e.g. Tenbrunsel and Messick, 1999), but are much harder to use and interpret when the number of decisions is larger, a situation that is very common within the field. Similarly, the models used by Busemeyer, Yechiam, and their associates (Busemeyer and Stout 2002; Yechiam and Busemeyer, 2005, 2006; Yechiam, et al., 2005) to examine repeated decisions in the Iowa Gambling Task are elegant and useful for that application, but more difficult to apply to a more general setting and generally only examine variables at the decision level. The mixed model approach applied to the modified Iowa Gambling Task in this study not only tests hypotheses regarding decision level variables, but also hypotheses regarding individual differences related to the subjects in the study and interactions between variables on the decision and individual levels.

Third, using mixed models can result in analytic models that more closely match theoretical views of how decisions occur. Our theories about decision making often address changing experiences throughout a game, which require that we examine individual decisions. We can examine these questions in the mixed model framework. Using these techniques can also help determine the impact of different characteristics of a design, such as the three different matrices used in the Mulford et al. prisoner dilemma game or the relative attractiveness of different card choices in the Peters and Slovic modified Iowa Gambling Task.

Fourth, correct specification of analytic models is more likely when we have the flexibility of the mixed-model approach. As one example, decisions may be related over time and we can correct for such dependencies by noting that decisions are nested within subjects and by including an autoregressive term. Using decisions as a unit of analysis also makes it possible to include a wider, and more appropriate, set of control variables, especially those that are directly related to actual decision making processes. This helps assure that estimates of individual differences are more accurate.

Fifth, the statistical characteristics of mixed models provide both appropriate controls and substantively important information. For instance, the autoregressive term provides not only an efficient method of controlling for dependencies, such as the tendency for individuals to behave in similar ways on successive opportunities to play a game, but also provides substantive information about

the magnitude and direction of such a tendency. In the analyses in this paper, the results with the card selection task suggested that there was little consistency and more exploratory behavior in choice from one selection to the next, while results with the prisoner dilemma data suggested a more consistent pattern.

Finally, when the dependent variable is measured on an interval scale (e.g., a task that involved receiving points or variable amounts of money), rather than as a dichotomy, as occurred in our analyses, additional statistical techniques are available. Specifically, it is possible to compare the relative fit of various models and to calculate the percentage of variation that can be attributed to the different levels of analysis. (See footnotes 6 and 7.)

While the computations involved in a mixed-model analysis are certainly more complex than those that are often used in analyses of decision making data sets, they are not overly onerous. The most important first step is making sure that data are recorded in a way that maintains decisions, rather than individuals, as the unit of analysis. Several different statistical programs provide the capacity for conducting these analyses. We have used SAS because it is widely available, allows the easy importation of data from other programs, and is very flexible. Both Singer (1998) and Raudenbush and Bryk (2002) provide very useful guides to the steps involved in using the statistical software; and in an appendix to this paper we include the language that we used for our analyses.

Even though mixed models provide many advantages when compared to the simpler techniques often used in the field, there are still limitations. One of the most important involves the number of independent variables that can be used. As with any analysis, researchers should be careful to limit the number of variables related to a given level (e.g. individuals or decisions) relative to the sample size on that level.<sup>15</sup>

In general, however, we believe that mixed models provide a more effective, efficient, and flexible way of ana-

<sup>15</sup>A standard rule of thumb is the ratio of no more than one variable for every five to ten subjects. For instance, with the decision to play in the Mulford et al. study, there were 1,128 choices and 96 people. Model 3 in this analysis used nine variables related to the decisions and two related to the subjects, well within these suggested limits. With the final model in Table 3, the reanalysis of the Peters and Slovic data, with over 12,000 decisions and 72 subjects, there are six variables related to the decisions, four related to the subjects and ten that involve interactions of subjects and decisions, a level that approaches the suggested limits. Note that the limits on subject level variables are no different than those that would occur in traditional analyses that aggregate data to the subject level. The method that we propose is, we believe, superior because it allows us to also examine variability in decisions within games. Even studies with clinical populations, which typically have much smaller samples, could conceivably use this approach. For instance, Bechara et al. (1996) had nineteen subjects (seven patients and twelve controls), with over 150 decisions per subject, which could allow the use of several variables on the decision level as well as the distinction between patients and controls on the subject level of analysis.

lyzing data from decision-making experiments than the traditional methods of aggregating data across decisions. Mixed models allow researchers to test complex hypotheses directly relevant to the way in which decisions are made, to obtain estimates of variance explained by both individual and game related characteristics, and to have more accurate estimates of effects. Using these more appropriate techniques is an important step in advancing our understanding of decision making.

## References

- Andreoni, J. & Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fundraising. *Journal of Public Economics*, 88, 1605–1623.
- Bechara, A., Tranel, D., Damasio, H. & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, 6, 215–225.
- Bryk, A. S. & Raudenbush, S. W. (1988). On heterogeneity of variance in experimental studies: A challenge to conventional interpretations. *Psychological Bulletin*, 104, 396–404.
- Busemeyer, J. R. & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment* 14, 253–262.
- Cochran, W. G. & Cox, G. M. (1957). *Experimental Designs*, 2<sup>nd</sup> edition. New York: John Wiley.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Avon.
- Dawes, R. M. & Messick, D. M. (2000). Social dilemmas. *International Journal of Psychology*, 35, 111–116.
- Greene, W. H. (2000). *Econometric Analysis (fourth edition)*. Upper Saddle River, New Jersey: Prentice Hall.
- Huttenlocher, J. E., Haight, W., Bryk, A. S., & Seltzer, M. (1991). Early Vocabulary Growth: Relation to Language Input and Gender. *Developmental Psychology*, 27, 236–249.
- Isen, A. M. (1997). Positive affect and decision making. In W. M. Goldstein & R. M. Hogarth (Eds.), *Research on judgment and decision making: Currents, connections, and controversies* (pp. 509–534). New York: Cambridge University.
- Isen, A. M. (1999). Positive affect. In T. Dalgleish and M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 521–539). Chichester, England, John Wiley & Sons Ltd.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, 21, 1–5.
- Lichtenstein, S. & Slovic, P. (1973). Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology*, 101, 16–20.
- Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. (1996). *SAS System of Mixed Models*. Cary, NC: SAS Institute.
- Mulford, M., Orbell, J., Shatto, C., & Stockard, J. (1998). Physical Attractiveness, Opportunity, and Success in Everyday Exchange. *American Journal of Sociology*, 103, 1565–1592.
- Parco, J. E., Rapoport, A., & Stein, W. E. (2002). Effects of financial incentives on the breakdown of mutual trust. *Psychological Science*, 13, 292–297.
- Peters, E. (2006). The functions of affect in the construction of preferences. In S. Lichtenstein & P. Slovic (Eds.), *The construction of preference*. (pp. 454–463) New York: Cambridge University Press.
- Peters, E., & Slovic, P. (2000). The springs of action: Affective and analytical information processing in choice. *Personality and Social Psychology Bulletin*, 26, 1465–1475.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage.
- Scharlemann, J. P. W., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, 22, 617–640.
- Singer, J. D. (1998). Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models. *Journal of Educational and Behavioral Statistics*, 24, 323–355.
- Tenbrunsel, A. E. & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly* 44, 684–707.
- Yechiam, E. & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin and Review* 12, 387–402.
- Yechiam, E. & Busemeyer, J. R. (2006). The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making* 19, 1–16.
- Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science* 16, 973–978.