Indiana University Cognitive Science Tech Report #227

Exploring the Relations between Categorization and Decision Making with Regard to Realistic Face Stimuli

James T. Townsend, Kam M. Silva, Jesse Spencer-Smith, Indiana University

and Michael J. Wenger, University of California, Santa Cruz

Abstract

Categorization and decision making are combined in a task with photorealistic faces. Two different types of face stimuli were assigned probabilistically into one of two fictitious groups; based on the category, faces were further probabilistically assigned to be hostile or friendly. In Part I, participants are asked to categorize a face into one of two categories, and to make a decision concerning interaction. A Markov model of categorization followed by decision making provides reasonable fits to Part I data. A Markov model predicting decision making followed by categorization is rejected. In Part II, a no-parameter model predicts decisions using categorization and decision responses collected in separate trials, suggesting that Part 1 results are not an artifact of the presentation of categorization and decision questions within a single trial. Decisions concerning interaction (defensive/friendly) appear to be based on information from the category decision, and not from the face stimuli alone.

Exploring the Relations between Categorization and Decision Making with Regard to Realistic Face Stimuli

Every day we make decisions about how we are going to interact with strangers. Perhaps in many cases, we opt for a strategy based on some categorization made on the basis of appearance. For example, convenience store clerks may decide to act in a friendly or defensive manner toward their customers based on the clerk's perceived threat of robbery or possible physical harm from their customers as opposed to a more benign mode of dress, race (in the case of stereotype), style of grooming, posture, language and so on. In fact, categorization might contain the necessary and/or sufficient information to form and execute an interaction decision. Such a decision involves executing a particular type of response toward another individual. The focus of this investigation is on initial perception of two classes of fairly realistic faces, created from photographs, differentiated according to physical characteristics, on the intervening categorization and on the final interaction decision. Actually, the categorization and final decision are under direct experimental manipulation, whereas the initial perception lies in the background of this particular investigation.

Individuals might employ their categorization label to determine their decision in a relatively strong form, in which case a type of decision might follow deterministically from the categorization. A weaker form would allow the decision to depend probabilistically on the categorization: given a categorization label, the probabilities of the various possible decisions would be determined, and no further direct dependence on the stimulus (e.g., the appearance of the "customer" in the shop) would appear, but any of the various decisions would be possible in principle. We can call this "conditional independence of stimulus" since once the categorization process has concluded, there is no further information associated with the stimulus, relative to the succeeding decision response. (Of course, a deterministic strategy could form an extreme form of conditional independence.) Alternatively, it might be that the categorization applied in, say a perceptual/social situation has little or no bearing on a subsequent decision. For instance, the original appearance of an individual may control the ensuing decision regardless of the intervening categorization.

Little, if any, research has investigated the relationship between categorization and decision-making, though a massive quantity of research has been devoted to investigating these concepts as distinct processes both in psychology per se as well as in various spheres of cognitive science. An appropriate key to unlocking the relationship of categorization to decision making is the so-called Markov property, named after the famous Russian mathematician.

The Markov Property

Under the Markov property, the future state of a system relies only on the present state, say that of time "n"; any of the past states of a system, at times n-1, n-2, etc., do not influence the future states. In the present instance, that "present" state, at time n=2, is taken to be the categorization made on that trial. Here, time n=1 is the time at which the face is presented and perception occurs, and we are interested in whether the decision, to be made at time n=3, is dependent on the categorization but not the state, that is the perception at time n=1. Satisfaction of the Markov property in this sense immediately implies and is implied by, the conditional independence stated above.

Models based on the Markov property have been of considerable value in many areas of science including physics, economics, sociology and psychology[1]. In psychology, Markov models have been studied extensively in the areas of learning (e.g., Atkinson, Bower & Crothers, 1965; Brainerd, 1979; Brainerd & Howe, 1980; Brainerd, Howe & Desrochers, 1982; Greeno, 1974), memory (Healy, 1978; Morganstein, 1974; Robertson, 1982), perception (Hacisalihzade, Stark & Allen, 1992), problem solving (Farley, 1983), signal detection and signal recognition (Sandusky, 1971; Ward, 1973).

In an analysis somewhat propaedeutic to the present one and exemplary of our method, the Markov property was investigated in a study of feature processing in letter identification by Townsend and Ashby (1982). Their study of letter recognition supported the Markov concept with regard to the determination of the frequency of letter responses dependent (in the Markov sense) on the set of extracted features. It was feasible to assay these mechanisms directly, as in the present study, because the observers were required to state what features they believed they saw, as well as a final letter response. Viewing the hypothesized Markov process in the letter experiment as a three step process, the first stage involved presentation of a letter to an observer, the second step consisted of probabilistic extraction and report of a feature set (i.e., there was a degree of randomness involved in the extraction process), and the last step involved the participant making a letter identification. The analysis could not reject the Markov assumption, and a strong relation between the features and letter reports was demonstrated. However, in that study, the reverse dependency, Letter → Feature, could not be decisively falsified.

It is true that the letter experiment and the present, three-stage implementations of the Markov question are rather trivial versions of the latter, and that the important critical question is whether conditional independence of the stimulus holds or not. However, it may be that future investigations of sequential relationships among various cognitive mechanisms will expand to include a number of successive stages of processing, much as Sternberg's additive factors method (1969) and later generalizations (e.g., Townsend & Schweickert, 1989; Schweickert & Townsend, 1989) may include a number of potential mechanistic contributions to overall response times.

With regard to our example of a store-keeper who is cognitively and behaviorally reacting to the entrance of an unknown person into her/his store then, would their cognitive activity be at least partially Markovian? Observe that there is nothing in the situation that forces such behavior and also note that maximal information is carried by the original unique stimulus. However, in some sense, the employment of conditional independence is more efficient since the decision maker need not retain excessive detail or irrelevant information about the original stimulus pattern, only the "pigeon-hole" into which the stranger has been placed. In the three-step process (perception of a face, categorization of the face into an invariantly defined group, and then decision regarding a particular action), the Markov hypothesis, as in the letter identification experiment, is that the decision will be dependent only on the categorization "state", and not at all influenced by other features of the face pattern itself, once a categorization has been made. The alternative hypothesis is that the individuals' interaction decisions will still be influenced by other aspects of the face such as certain facial features. In other words, the decision-making process may be dependent on the individual face pattern throughout the psychological chain of processes, hence violating the Markov property. It should also be noted that the fact that the facial characteristics will generally affect the categorization response (in fact is expected to!), does not in any way compromise the Markov question with regard to the final decision response.

The answer to the Markov question, as posed here, may hold implications for a number of important issues in several areas of cognitive science, including social/personality psychology. In some of the stereotyping literature, researchers have found that observers increase efficiency in information processing once categorization features have been identified (Andersen, Klatzky & Murray, 1990; Macrae, Stangor & Milne, 1994). As noted above, if observers follow a Markov property, it would allow greater efficiency in formulating an interaction decision since observers would not have to consider the actual facial pattern again. On the other hand, a strict adherence to conditional independence of the original stimulus might be correlated with rigidity of stereotypy.

During the experiment, participants were shown pictures of faces and asked to categorize them into one of two fictitious groups. Since learning was not the theme of the present study, they were told the defining features of the two groups. These two groups were defined by face shape and lip thickness. In the main procedure, after participants categorized a face, they were asked to make a decision about how they would interact with the face, knowing that one population tended to be friendly and the other population tended to be hostile (exact procedure given below). Once a decision was made, participants received feedback as to whether they made the correct categorization and interaction decision. A correct categorization and decision equates to observer's response corresponding to the face assignments generated by computer, based on probabilities.

Perhaps as in real life, the categories were related only probabilistically, rather than deterministically, to the featural structure of the faces. Thus, a round-like face with thin lips tended to be, but was not always, an "Adok" (see Figure 1), whereas a narrow face with fat lips tended to be, but not invariably, a "Lork" (see Figure 2). Similarly, also in accord with real-life categories (e.g., stereotypes), Adoks were more likely to be "friendly" rather than "hostile" and the Lorks were more likely to be hostile than friendly, but either could, on any given trial, be either friendly or hostile. Rewards and punishments were arranged to symbolically be in concordance with what might occur in a real-life situation. Often we will simply refer to "round faces" or "narrow faces" for brevity, even though the faces always differed both on the overall shape, as well as on lip width, dimensions. Our Markov model for this situation is given in Equation (1) as follows:

Prob(DEC | FACE TYPE) = Prob(DEC | CAT LORK)*Prob(CAT LORK | FACE TYPE) +
Prob(DEC | CAT ADOK)*Prob(CAT ADOK | FACE TYPE). {1}

There are two other important facets of our experimental design. Whether or not the data departed significantly from a Markov assumption can be statistically probed and was. However, the question of a statistically significant difference is, after all, somewhat arbitrary, since in principle, the study can be made as statistically powerful or as weak as the experimenter desires, simply by adjusting the number of trials run. From this vantage point it appeared interesting to compare the Markov fit to the sequence: Perception→Categorization→Decision with one where observers were requested to respond with the decision first, then categorization, that is, in the sequence: Perception→Decision→Categorization. If it is irrelevant which comes first, then there should be no difference in the appropriateness of the Markov property: the formulas Prob(DEC | CAT) in the standard sequence, should be no better (nor worse) than Prob(CAT | DEC) in the reversed response order. Therefore, we mingled the reverse order conditions in blocks along with blocks of the standard order. In addition, we can analyze the reverse order data as if the observers are actually, and covertly, making a categorization and then their decision and then simply reporting them in the reversed order, to comply with the experimenter's instructions. If there is something psychologically natural about the categorization-then-decision sequence vs the decision-then-categorization sequence, then we predict that our analyses in the forward order for the reversed report condition, should perform poorly, whereas analyzing the data in the "covert" order, should predict the data about as well as the forward order of analysis in the standard report order condition[2]. Such a finding would indicate that good fits from the Perception→Categorization→Decision Markov model were likely not a simple case of insufficient power.

Next, we were interested in the possibility of whether overt report of categorization and decision might distort what is really going on, psychologically. To partially assay this possibility, we composed a second part of the experiment where the participants only had to make either the categorization, or the decision response and never both in any one trial. Then we tested the "non-distortion" hypothesis in the following manner. We estimated the "true" categorization probability from the "categorization-alone" condition in Part II, Prob(CAT | FACETYPE), and the conditional probability of making a certain decision, given a certain categorization, Prob(DEC | CAT), from Part I of the experiment where both responses were give by participants. We then formed a prediction for Prob (DEC | FACETYPE) composing these probabilities, and compared the prediction to the observed quantitiy for Prob(DEC | FACETYPE) from Part II of the experiment.

Let $\text{Prob}_2(\text{ CAT } | \text{ FACETYPE })$ and $\text{Prob}_2(\text{ DEC } | \text{ FACETYPE })$ be the two appropriate quantities from Part II of the study, and $\text{Prob}_1(\text{ DEC } | \text{ CAT })$ be the probability of that decision given the particular categorization from Part I of the study. Then the prediction was composed of the foregoing quantities in Equation 1 as follows:

$$\text{Prob}_2(\text{ DEC } | \text{ FACETYPE }) = \text{Prob}_2(\text{ CAT } | \text{ FACETYPE })*\text{Prob}_1(\text{ DEC } | \text{ CAT}) \qquad \{2\}$$

It may be observed that no new parameters need to be estimated for the prediction.

Method

Participants

One hundred and fifty-four undergraduate students from Indiana University were recruited for the experiment. The students participated in the experiment for two sessions lasting one hour per day over a two-day period. They received credit for their introductory psychology courses as well as any monetary bonus they earned.

Materials

Six IBM compatible computers were used to present the programs and record participants' data. The faces were digitally altered versions of photographs from the book Heads (Kayser, 1985). These faces were kindly provided to us in computer-coded form by Dr. Thomas Busey. The faces used in the study were chosen based on two features: the shape of the face and the thickness of the lips, and were manipulated to enhance those features relevant for categorization. There were a total of 34 faces; half of the faces were round with thin lips and the other half were narrow faces with thick lips. Some of the features were made more extreme, using Aldus PhotoStyler version 2.0, to ensure a significant difference between the average sizes of the features defining each group. The face widths ranged from 5.1 to 5.7 cm for the round faces and 2.8 to 3.7 cm for the narrow faces. The lip thickness' ranged from 0.3 to 0.7 cm for the round faces and 0.5 to 1.0 cm for the narrow faces. All faces were 8.3 cm tall from the chin to the top of the head. All of the photographs shown were headshots of bald Caucasian men. Ancillary MDS analyses using the ALSCAL procedure with a different set of observers confirmed that people could and did readily classify our round face + thin lips vs. narrow face + fat lips into two separate categories.

Design

The two different types of stimuli, defined by the shape of their face and thickness of their lips, were reassigned by the computers into one of two fictitious groups. The round faces with thin lips had a 60% chance of being assigned as an "Adok" and a 40% chance of being called a "Lork." The narrow faces with thick lips had a 60% chance of being a "Lork" and a 40% chance of being an "Adok." The computers also assigned each member of the groups to have a "friendly" or "hostile" nature. The Adoks had a 70% chance of being friendly and a 30% chance of being hostile. The Lorks had a 70% chance of being hostile and a 30% chance of being friendly. Thus, the chance that a narrow-faced individual (with thick lips) was both an Adok as well as friendly was .40 x .70 = .28. Obviously, the overall probabilities of being confronted with a narrow vs. round face, a Lork vs. Adok, or a friendly vs. hostile demeanor was in each case, equal to .50.

Over all blocks of the experiment, the participants were required to answer two questions for each face, the categorization question “Is this face an Adok or a Lork?” and the interaction question “Would you be friendly or defensive towards this face?” The participants were told that the round faces with thin lips tend to be Adoks and Adoks tend to be friendly while the narrow faces with thick lips tend to be Lorks and that Lorks tend to be hostile. They were not given exact probabilities.

Part I of the experiment was divided into six blocks of 34 trials each and presented over a two day period. Four block were presented on the first day and two were presented on the beginning of the second day. The first 17 trials of each block were grouped by which question appeared first. Half of the participants answered the categorization question first while the other participants answered the interaction decision first. The order of the questions was reversed for the last 17 trials. During the experiment, participants viewed a face, answered the two questions and then only after both responses were made did the feedback appear on the computer screen.

Part II of the experiment tested the same participants on the second day with the same faces as in Part I. Part II consisted of two blocks of 34 trials each. In the first 17 trials of each block, half the participants were asked the categorization questions only, while the other half were asked decision questions only. In the last 17 trials, participants were asked the alternative questions. The rationale for this phase of the experiment is explained just before the method section.

In Part I of the experiment, the faces were presented one at a time and the participants viewed the faces and the first question for, at most, eight seconds. Participants were allowed to respond within the eight seconds. Their responses to the first question caused the next question to appear for up to eight seconds and their responses to the second question triggered the feedback to appear on the screen for two seconds. While each question was asked and the feedback was presented, the faces remained on the computer screen unless the participants failed to respond within the eight seconds. If the participants did not respond within the eight seconds, the faces disappeared and the computer continued to wait for a response. Once a response was made, the face reappeared with either the next question or the feedback. In Part II of the experiment, the faces were also presented one at a time for eight seconds. Only a single question was asked during each presentation. Feedback was given after each response.

In regard to the categorization portion of the experiment, the feedback informed the participants whether they made a correct categorization. For the decision-making feedback, participants earned five cents if they made the correct interaction decision. The participants were (a) told that they were either friendly to a friendly face and the person with that face handed them twenty dollars or (b) told that they were defensive to a hostile face and therefore escaped with their lives. The participants could also lose five cents if they made an incorrect interaction decision. The participants were (a) told that they were friendly to a hostile face and mugged as a result or (b) told that participants were defensive to a friendly face, and hence received a dirty look. At the end of the experiment, participants received any bonus money earned. Any debt earned by participants was forgiven.

Procedure

All participants were placed in a well-lit room containing the six IBM compatible computers. For Part I of the experiment, the computer instructions stated the following:

> You have been chosen by NASA to travel to the planet Meboo to categorize two colonies, the Adoks and the Lorks.
>
> As you interact with the two colonies, you will be first asked to categorize each face as either an 'Adok' or a 'Lork'. The Adoks tend to have round faces and thin lips, and the Lorks tend to have long faces with thick lips. But, this is not absolute! As in any culture, there is cross-over. A face with the features of an Adok may actually be a Lork. You have up to 8 seconds to view each face (you may answer before the 8 seconds are up). You should press 'a' for an 'Adok' or 'l' for a 'Lork'.
>
> Then, you have a choice to make, you can be friendly or defensive to the face. Adoks have the tendency to be friendly while Lorks tend to be hostile. This is not absolute! Since you do not know how the individual will, act towards you, make your decision carefully. You should press 'f' for friendly or 'd' for defensive.

For the trials in which the participants first answered the interaction decision, the second paragraph referring to the categorization question and the third paragraph referring to the interaction decision in the instructions were interchanged (with appropriate rewording). In Part II of the experiment, only the categorization or the decision response was called for on the presentation of each stimulus face.

The participants completed the first four blocks of Part I of the experiment on the first day. On the second day of testing, the participants finished the last two blocks of Part I, then completed two blocks of trials comprising Part II. In all blocks, each of the 34 faces (17 round, 17 narrow) was presented once, and were randomly assigned (using the percentages described above) to be Adoks or Lorks, friendly or unfriendly. The participants were paid any bonus money earned at the end of the second day.

Analyses and Results

Part I

Chi-square (χ^2) statistical tests were developed to test the Markov prediction. In Part I and the standard order of responses, the model is constructed as follows. The major “categorical information” lies in whether the stimulus was a narrow face with wide lips vs. round face with narrow lips. Even though the Adoks tend to be friendly and the Lorks hostile, if the Markov assumption holds and given a particular categorization response (even if a round face is categorized as a Lork and vice versa for a narrow face categorized as an Adok) a participant’s decision should no longer depend on whether the presented face was round or narrow; only on the categorization. By dividing the set of faces into narrow vs. round, it was feasible to test the Markov prediction for each individual participant separately.

Now, we assume that for each round face, there is a probability P_1 that the observer classifies that face as an Adok (refer to Figure 3) and a probability of $1-P_1$ that he or she categorized this face as a Lork. Note that the observer does not, and cannot know whether the face was actually a Lork or Adok until feedback is given at the end of a trial. Similarly, if the face is narrow, there is a probability P_2 that the observer classifies the face as a Lork and $1-P_2$, as an Adok. In general, we expect that $P_1>P_2$, since the probability of assignment as an Adok with a round face should be greater than the same assignment in the presence of a narrow face.

The Markov assumption enters in our next assumption that there is a probability Q_1 that, given categorization as an Adok, the decision will be to be friendly toward the face, and $1-Q_1$, to be defensive toward it. Naturally, it is further assumed that there is a probability Q_2 that, given categorization as a Lork, the decision will be to be friendly toward that face. The expectation is that $Q_1>Q_2$. Under the Markov property (implying that the interaction decisions were conditionally independent of face patterns), the probabilities, P_i and Q_j (i and j = 1,2), can be multiplied together to determine the expected proportions for the eight possible combinations of narrow vs. round, Adok vs. Lork categorization, and Friendly vs. Defensive decision. These probabilities were calculated by breaking down the observed values into two 2x2 tables, corresponding to when each face was round and when it was narrow, and recording the number of responses (categorizations and interaction decisions) to each face in either case (see Figure 4), as indicated in Equation (2) below, which is simply the realization of Equation (1) in terms of the proper indexes and in the special case of presentation of a narrow face.

Prob(DEC FRIENDLY | NARROW FACE) =

Prob(CAT ADOK | NARROW FACE) * Prob(DEC FRIENDLY | CAT ADOK) +

Prob(CAT LORK | NARROW FACE) * Prob(DEC FRIENDLY | CAT LORK) = P2 * Q1 + (1-P2) * Q2 {2}

Figure 4 shows how the expected (i.e., predicted) proportions are computed from the data tables for round and narrow faces. The results are estimates of the P_i and Q_j theoretical probabilities. Note that these are accomplished for each participant, thus permitting individual differences to appear. The frequencies required for the χ^2 analyses can, of course, either be gathered directly from the tables as shown or by multiplying the estimated probabilities by the appropriate marginal frequencies. Once these expected values were calculated, a χ^2 test was performed for each individual collapsed across the six blocks. The χ^2 tests had two degrees of freedom, the eight cells of the two 2x2 tables minus the four estimated parameters (probabilities P_1, P_2, Q_1, Q_2) minus the two overall frequencies (M and N).

χ^2 statistics were calculated for each participant for trials in which the categorization question was asked first, collapsing across blocks and face type (round/narrow). Out of 121 individual participants, 28 had χ^2 values that exceeded the .10 α level. Since approximately 12 significant values would be expected by chance, it is certainly not the case that everyone is strictly following the Markov assumption. Note also that 13 participants were excluded from analysis since they gave only "Adok" responses for round faces, or gave only "Lork" responses for narrow faces. These participants were responding in an near-optimal fashion: round faces were categorized as Adoks only, or narrow faces were categorized as Lorks. Only one participant responded in an optimal fashion: round faces were categorized strictly as Adoks, and narrow faces were categorized strictly as Lorks. Responding in an optimal or near-optimal fashion resulted in expected observations of 0, precluding χ^2 analysis. Data from most participants were well-accounted for by the Markov model.

Another set of χ^2 statistical tests, with similar calculations for expected frequency (see Figures 5 & 6), were used to test for independence between the face pattern and the participants' categorizations when the decision interaction question was asked first. For these trials, the categorization was significantly dependent on the face pattern, after conditioning on the stated decision. The "decision then categorization" Markov model was rejected in all but 4 of 144 analyses at the .10 α level, reflecting terrible fits of the Markov model under these conditions.

These data were then subjected to the same formulas as in the first analysis to test if the interaction decision (step 3) was independent of the face pattern (step 1) but dependent on the categorization (step 2). That is, could the participants have been covertly performing the categorization first, then the interaction decision, but reporting in the reverse direction, as requested by the experimenter? The results indicated that this was the case. Out of 121 individual participants, 24 had χ^2 values that exceeded the .10 α level. Again, it is certainly not the case that everyone is completely following a "categorization first, then decision" Markov process. But, the results suggest that participants were much more likely to be performing the “natural” conditional sequences covertly, than to be conditionally basing their categorization on their decisions, even when forced to respond in that order. Note that 13 participants were excluded from analysis as well, since they responded near optimally, giving only "Adok" responses for round faces, or giving only "Lork" responses for narrow faces, precluding χ^2 analysis. Only one participant responded perfectly optimally on both round and narrow faces.

In sum, even though the first analyses indicated that the Markov assumption was not entirely correct, it serves considerably better than does a “decision first, then categorization” with a Markov property in that direction, even for the trials when participants had to report in that order.

Part II

It may be recalled that Part II of the experiment required participants to give either the categorization response to the presented face, or the interaction decision, but not both within a single trial. The goal was to attempt to predict the interaction decision alone, given presentation of a round or narrow face, in a completely non-parametric way. Recall that the interaction decision from Part II is predicted by combining the conditional probabilities of the decision Part I, given either of the particular categorizations (Adok or Lork), and the categorization probabilities (given a round or narrow face as appropriate) from the Part II data. For a narrow face and a decision to be defensive, the prediction is derived by inserting the proper quantities into Equation (2):

$$\begin{aligned} &\text{Prob}_2\text{ (DEC DEFENSIVE | NARROW FACE) =} \\ &\text{Prob}_2\text{ (CAT ADOK | NARROW FACE) * Prob}_1\text{ (DEC DEFENSIVE | CAT ADOK) +} \\ &\text{Prob}_2\text{ (CAT LORK | NARROW FACE) * Prob}_1\text{ (DEC DEFENSIVE | CAT LORK),} \qquad \{3\} \end{aligned}$$

where Prob$_2$ (DEC DEFENSIVE | NARROW FACE), Prob$_2$ (CAT ADOK | NARROW FACE), and Prob$_2$ (CAT LORK | NARROW FACE) are obtained during Part II, and Prob$_1$ (

DEC DEFENSIVE | CAT ADOK) and Prob_1 (DEC FRIENDLY | CAT LORK) are derived from data from Part I.

For narrow faces, the non-parametric prediction was rejected at a .05 α level in 38 out of 138 cases. For round faces, the prediction was rejected in 34 out of 138 cases. It appears that the non-parametric fits to the Part II decision data are in the same ballpark with regard to tendency to adhere or violate the prediction found in Part I analyses. Therefore, we conclude on the basis of these analyses, that the moderate tendency to base one's interaction decisions on one's categorizations are carried on in more or less the same way regardless of whether categorizations and decisions must be made public or not.

The use of varying α levels in the analysis of model fits bears comment. In Part I of the experiment, we use α=.10 since we wish to expose participants who definitively do not follow the Markov model. In Part II, the aim is to falsify the non-parametric model, and a tighter α=.05 is used.

Discussion

The purpose of this experiment was to investigate the relationship between perception, categorization and decision-making when participants were required to form an interaction decision toward face stimuli. The hypothesis was that participants would follow a Markov property, in that their interaction decision would be based on the categorization information used and not on the original face pattern itself. In other words, the probability of a friendly or defensive interaction decision would be dependent only on the categorization and no longer on the physical characteristics of the original face. The results of this experiment suggest that most individuals, although not all, are at least moderately influenced by their categorization. Furthermore, our analyses demonstrated that it is important which response is categorization and which is decision: There was virtually no tendency to evince Markov tendencies when the interaction decision was made first and the categorization second, and conditional independence of the categorization relative to the interaction decision was tested. But if we still conditionalized on the categorization, even though it was actually made second, then the Markov property held pretty well and as good as it did when the responses were made in the natural order. This strongly suggests that the observers tended to make their categorization first, covertly, then the decision, but report them in the order requested by the experimenter.[3]

One possible confound raised by a reviewer involves the repeated use of faces between blocks. If participants determined that the same faces were appearing with different attributions (e.g. a face was friendly in block 1 but defensive in block 2), participants might “give up” and decide to rely solely on category information. An analysis of changes in response patterns between the first three blocks and the last three blocks of trials in Part I revealed no pattern indicating a shift in response probabilities that would accompany such a change in strategy.

It must be admitted that our experimental paradigm certainly encouraged the cognitive conjoining of categorization and decision making. In fact, it is difficult to conceive of an experimental design that completely unravels them without some type of experimental juxtaposition. We leave this conundrum open for further thought and implementation of potential solutions.

Next, it is interesting that it is easily shown that an optimal strategy would have been to classify any round face as an Adok and any narrow face as a Lork. Then the further optimal strategy would have been to always formulate a friendly interaction, given categorization as an Adok and a defensive interaction, given categorization as a Lork. Participants did not do this: with few exceptions, they responded probabilistically. In only two instances did participants reply in a completely optimal fashion. In this sense, most of the participants acted like the legions of human participants in probabilistic reinforcement regimes (Bush & Mosteller, 1955; Healy, 1981; Erev, 1998)

It is true that the actual set of faces employed in our study may have partly encouraged such behavior. That is, due to the arduous procedure involved in producing our face stimuli, they were sufficiently few that participants saw the same face more than once, often with different feedback as to proper categorization on the distinct occasions. Among further possible research topics, it would be interesting to use schematic faces, with which the experimenter would have more flexible and immediate control and no face need be viewed more than once by any individual. That situation would be more realistic in that sense, if less so in the use of schematic faces.

So why would an individual follow a Markov property in formulating an interaction decision? Some possibilities are based on efficiency in decision making. For instance, it has been shown in several stereotype studies that once a categorization characteristic has been identified, related information can be retrieved and processed more efficiently (Andersen et al., 1990; Macrae et al., 1994).[4] In our experiment, once the categorization characteristics of a face were identified, perhaps it was more efficient for the participants to rely on that stereotypic information than to refer back to the facial pattern in order to formulate an interaction decision.

The way in which categorization and decision making interact together and with facial (or other) information, holds implications for social interaction in a number of spheres, perhaps most patently in stereotyping behavior. For instance, in relearning programs intended to break stereotyping of individuals with certain types of physiognomy, for instance, belong to a particular group, such as race, there are several ways such learning might progress. The education process might emphasize unlearning stereotyped categories (e.g., categorization with regard to race), or learning new more enriched categories (e.g., “Asian and educated and…”, rather than simply “Asian”). It might focus on new interaction decision strategies to learn new more amicable behavior patterns, even if the old stereotyped categories still exist. What would be the relative independence and efficiency of reeducation on the various fronts, relative to societal effort and expense?

Needless to say, there are implications, not unencumbered by ethical questions, that will arise in potential application of information concerning perception, categorization and behavior propensities within social engineering contexts. For some years, security people have monitored closed circuit TV networks in order to anticipate or apprehend thieves. Even now, engineers and computer scientists are devising pattern recognition systems that will scan customers faces and other characteristics in order to assess the likelihood of shoplifting or other unacceptable behaviors.

Quantitative theories of categorization (e.g., Ashby & Maddox (1993); Nosofsky & Palmeri (1997); Anderson (1991); Estes (1986)) might be brought to bear on the present kind of phenomena. Theories of learning in decision making per se (e.g., Busemeyer & Myung (1997); Busemeyer & Townsend (1993); Groen & Atkinson (1966)) could be integrated with those of categorization in order to produce an overall theory of learning in situations that call for, or allow, both decision making as well as categorization. Furthermore, the ways in which features are employed in classification (e.g., Townsend & Ashby (1982); Goldstone (1994); Tversky (1977)) could be investigated in our dual task environment and in particular, in situations that do, vs. do not, involve emotionally charged categorization (e.g., stereotyping).

Finally, we noted that participants rarely responded deterministically, even though one version of that strategy would have been optimal, and even though that tack would have been easier to carry out by your average Introduction-to-Psychology participant. In the classical probability learning literature, it was discovered that the more extreme the probabilities (e.g., probability of being rewarding by choosing a left turn rather than a right turn) or the positive and negative reinforcements, the more human and animal participants deviated from probability matching (i.e., going to the left with the same frequency with which that side was rewarded) in the direction of optimality (Dusoir, 1974; Creelman & Donaldson, 1968). It would be interesting to learn if extreme reinforcements or probabilities lead to a greater tendency toward deterministic behavior, and if categorization strategies vs. decision strategies can be affected separately by reinforcement structures.

We hope that the present study is propaedeutic to further research in categorization and decision making that includes richer perceptual sources and dynamic situations; such experiments are beyond the scope of this first foray into the area. Of particular interest are designs in which categorization and decision making are more subtly related. We hope to motivate and carry out such research.

References

Andersen, S. M., Klatzky, R. L., & Murray, J. (1990). Traits and Social Stereotypes: Efficiency Differences in Social Information Processing. Journal of Personality and Social Psychology, 59, 192-201.

Anderson, J. R. (1991). The adaptive nature of human categorization. Psychological Review, 98(3), 409-429

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. Journal of Mathematical Psychology, 37, 372-400.

Atkinson, R. C., Bower, G. H. & Crother, E. J. (1965). An introduction to mathematical learning theory. New York: Wiley.

Brainerd, C. J. (1979). Markovian interpretations of conservation learning. Psychological Review, 86, 181-213.

Brainerd, C. J., & Howe, M. L. (1980). Developmental invariance in a mathematical model of associative learning. Child Development, 51, 349-363.

Brainerd, C. J., Howe, M. L., & Desrochers, A. (1982). The general theory of two-stage learning: A mathematical review with illustrations from memory development. Psychological Bulletin, 91, 634-665.

Busemeyer, J. R. & Myung, I. J. (1997). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. Journal of Experimental Psychology General, Vol 121(2), 177-194

Busemeyer, J. R. & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100(3),* 432-459.

Bush, R. R., & Mosteller, F. (1955). Stochastic models for learning. New York: Wiley.

Creelman, C. D., & Donaldson, W. (1968). ROC curves for discrimination of linear extent. Journal of Experimental Psychology, 77, (3, PT. 1). 514-516.

Duffie, D., Geanakoplos, J., Mas-Colell, A., & McLennan, A. (1994). Stationary Markov equilibria. Econometrica, 62, 745-782.

Dusoir,A.E. (1974). Thomas and Legge's matching hypothesis for detection and recognition tasks: Two tests. Perception and Psychophysics 16(3), 466-470.

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. Psychological Review, 105(2), 280-298.

Estes, W. K. (1986). Array models for category learning. Cognitive Psychology, 18(4), 500-549

Farley, A. M. (1983). A probabilistic model for uncertain problem solving. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 568-579.

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. Cognition, 52(2), 125-157.

Gordon, S. (1996). Using mixtures of flexible functional forms to estimate factor demand elasticities. Canadian Journal of Economics, 29, 717-737.

Greeno, J. G. (1974). Representation of learning as discrete transition in a finite state space. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), Contemporary developments in mathematical psychology: I. Learning, memory and thinking . San Francisco, CA: WH Freeman.

Groen, G. J. & Atkinson, R. C. (1966) Models for optimizing the learning process. Psychological Bulletin, 66(4), 309-320

Hacisalihzade, S. S., Stark, L. W., & Allen, J. S. (1992). Visual perception and sequences of eye movement fixations: A stochastic modeling approach. IEEE Transactions on Systems, Man, and Cybernetics, 22, 474-481.

Healy, A. (1978). A Markov model for the short-term retention of spatial location information. Journal of Verbal Learning and Verbal Behavior, 17, 295-308.

Healy, A. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. Journal of Experimental Psychology: Human Learning and Memory, 7(5), 344-354.

Hoel, P. G., Port, S. C., & Stone, C. J. (1987). Introduction to stochastic processes. Los Angeles: Waveland Press, Inc.

Kayser, A. (1985). Heads. New York: Abbeville Press.

Kindermann, R. P., & Snell, J. L. (1980). On the relation between Markov random fields and social networks. Journal of Mathematical Sociology, 7, 1-13.

Korte, A. O. (1990). A first order Markov model for use in the human services. Computers in Human Services, 6, 299-312.

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating Social Stereotypes: A Functional Analysis. Journal of Experimental Social Psychology, 30, 370-389.

Morganstein, S. (1974). Stochastic theory of memory: Three state Markov processes: Primacy and recency effects. Catalog of Selected Documents in Psychology, 4, 38.

Nisbett, R. R. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84(3), 231-259.

Nosofsky, R. M., Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. Psychological Review, 104(2), 266-300.

Pettiway, L. E., Dolinsky, S., & Grigoryan, A. (1994). The drug and criminal activity patterns of urban offenders: A Markov chain analysis. Journal of Quantitative Criminology, 10, 79-107.

Robertson, C. (1982). Modelling recall order in free recall experiments. British Journal of Mathematical and Statistical Psychology, 35(2), 171-182.

Rojek, D. G., & Erickson, M. L. (1982). Delinquent careers: A test of the career escalation model. Criminology: An Interdisciplinary Journal, 20, 5-28.

Rothbart, M. & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. Journal of Social Issues, 41(3), 81-104

Sandusky, A. (1971). Signal recognition models compared for random and Markov presentation sequences. Perception & Psychophysics, 10, 339-347.

Schweickert, & Townsend, J. T. (1989). A trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. Journal of Mathematical Psychology., 33(3): 328-347

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction time experiments. American Scientist, 57(4), 421-457.

Townsend, J. T., & Ashby, F. G. (1982). Experimental test of contemporary mathematical models of visual letter recognition. Journal of Experimental Psychology: Human Perception and Performance, 8, 834-864.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. Journal of Mathematical Psychology, 33(3), 309-327.

Tversky, A. (1977). Features of similarity. Psychological Review, 84(4), 327-352.

Ward, L. M. (1973). Use of Markov-encoded sequential information in numerical signal detection. Perception & Psychophysics, 14, 337-342.

Footnotes

[1] An evolutionary example of a Markov model stemming from mathematics is a "branching chain" which was created to investigate the probability of extinction. The model is based on the fact that the probability of extinction of an nth generation may be dependent only on the nth-1 generation (Hoel, Port & Stone, 1987). In economics, Markov models have been used to estimate factor demands and also to study equilibria of security markets and game theory (Duffie, Geanakoplos, Mas-Colell & McLennan, 1994; Gordon, 1996). In sociology, scientists have analyzed the Markov property in regard to patterns of criminal activity (Pettiway, Dolinsky & Grigoryan, 1994; Rojek & Erickson, 1982) and social networks (Kindermann & Snell, 1980). Markov models have also been applied to population mobility (Korte, 1990).

[2] For instance, our colleague, Dr. John Kruschke, pointed out that from certain a priori considerations alone, both responses could be considered to be "decisions" in some sense and both could be considered in some sense to be "categorizations." The condition and test put forth here provide a probing of this idea.

[3] It has been suggested by one reviewer that subject reports might provide additional information on the order of categorization and decision making. Nisbett and Wilson (1977) make the case, however, that subject reports on higher-order cognitive processes are highly suspect; their reports seem to be based on a priori causal theories.

[4] As one reviewer indicated, related research by Rothbart and John (1985) concerns the conditions under which such stereotypes might be "unlearned."

Figure Captions

Figure 1. Round faces with thin lips had a 60% chance of being labeled as an “Adok” and a 40% chance of being called a “Lork.”

Figure 2. Narrow faces with thick lips had a 60% chance of being a “Lork” and a 40% chance of being an “Adok.”

Figure 3. "Categorization->Decision" Markov model of interactions: categorization is followed by decision, decision is based solely on categorization. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face

Figure 4. Assuming conditional independence, as the Markov property does, the probabilities can be multiplied to form the expected proportions as above. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face

Figure 5. "Decision->Categorization" Markov model of interactions: decision is followed by categorization, categorization is based solely on decision. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face

Figure 6. Assuming conditional independence, as the Markov property does, the probabilities can be multiplied to form the expected proportions as above. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face

Figure 1. Round faces with thin lips had a 60% chance of being labeled as an “Adok” and a 40% chance of being called a “Lork.”

Figure 2. Narrow faces with thick lips had a 60% chance of being a “Lork” and a 40% chance of being an “Adok.”

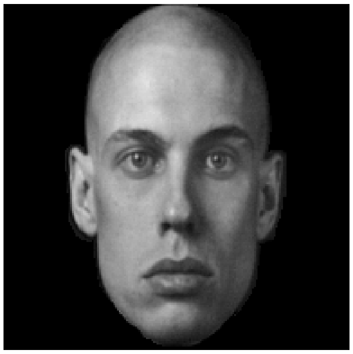
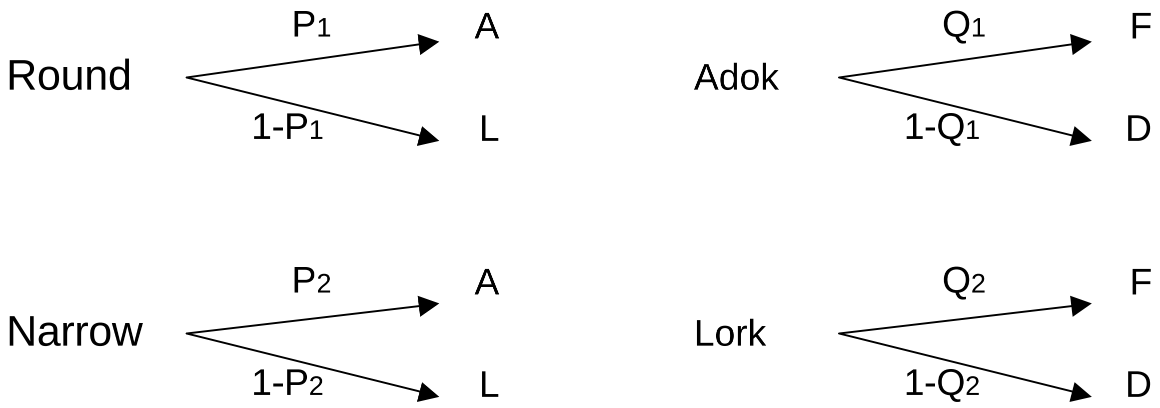


Figure 3. "Categorization->Decision" Markov model of interactions: categorization is followed by decision, decision is based solely on categorization. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face



P_1 = P (categorization = Adok | round face)

$1-P_1$ = P (categorization = Lork | round face)

P_2 = P (categorization = Adok | narrow face)

$1-P_2$ = P (categorization = Lork | narrow face)

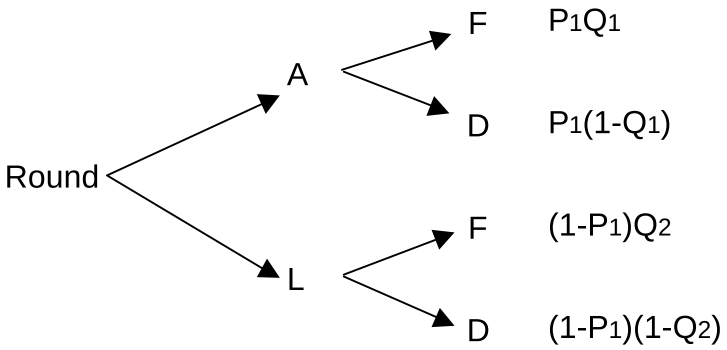
Q_1 = P(decision = friendly | categorization = Adok)

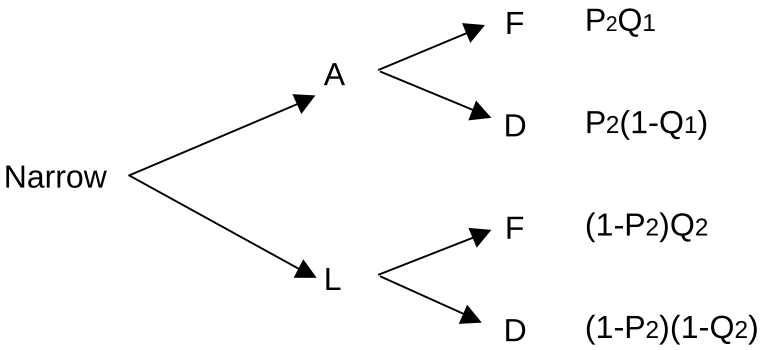
$1-Q_1$ = P(decision = defensive | categorization = Adok)

Q_2 = P(decision = friendly | categorization = Lork)

$1-Q_2$ = P(decision = defensive | categorization = Lork)

Figure 4. Assuming conditional independence, as the Markov property does, the probabilities can be multiplied to form the expected proportions as above. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face





Participant k

Round

		Interaction Decision		
		D	F	
Participant Categorization	A	mAD	mAF	mA
	L	mLD	mLF	mL
		mD	mF	M

Narrow

		Interaction Decision		
		D	F	
Participant Categorization	A	nAD	nAF	nA
	L	nLD	nLF	nL
		nD	nF	N

$P1 = (mAD + mAF) / M \quad mA / M$

$1\text{-}P1 = (mLD + mLF) / M : \quad mL / M$

$P2 = (nAD + nAF) / N = \quad nA / N$

$1\text{-}P2 = (nLD + nLF) / N = \quad nL / N$

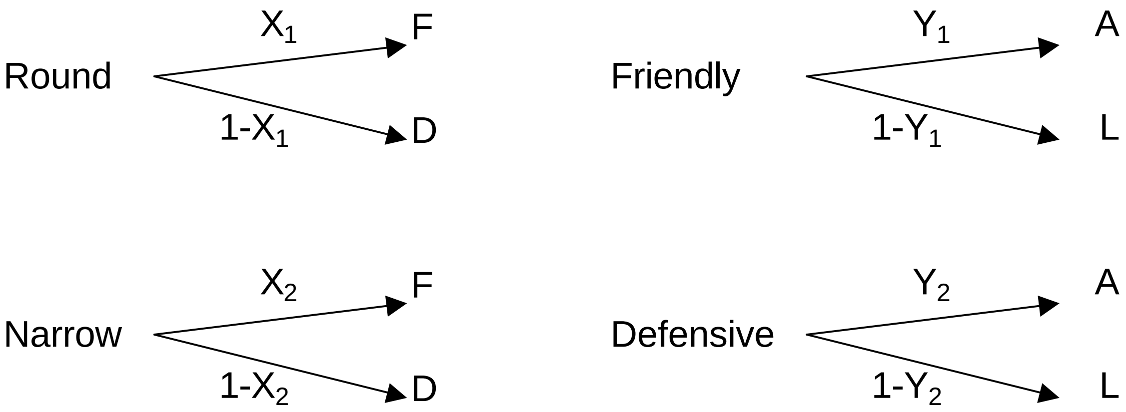
$Q1 = (mAF + nAF) / (mA + nA)$

$1\text{-}Q1 = (mAD + nAD) / (mA + nA)$

$Q2 = (mLF + nLF) / (mL + nL)$

$1\text{-}Q2 = (mLD + nLD) / (mL + nL)$

Figure 5. "Decision->Categorization" Markov model of interactions: decision is followed by categorization, categorization is based solely on decision. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face



X_1 = P (decision = friendly | round face)

$1-X_1$ = P (decision = defensive | round face)

X_2 = P (decision = friendly | narrow face)

$1-X_2$ = P (decision = defensive | narrow face)

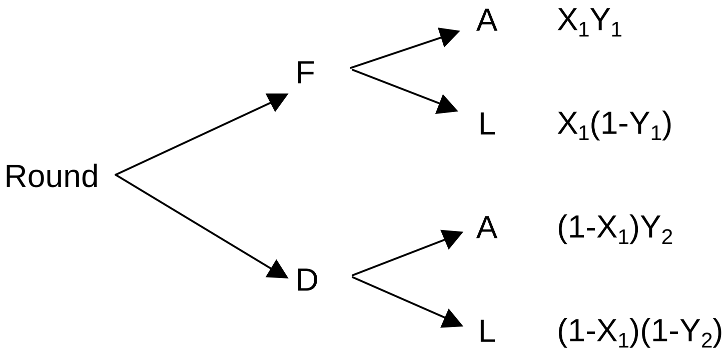
Y_1 = P (categorization = Adok | decision = friendly)

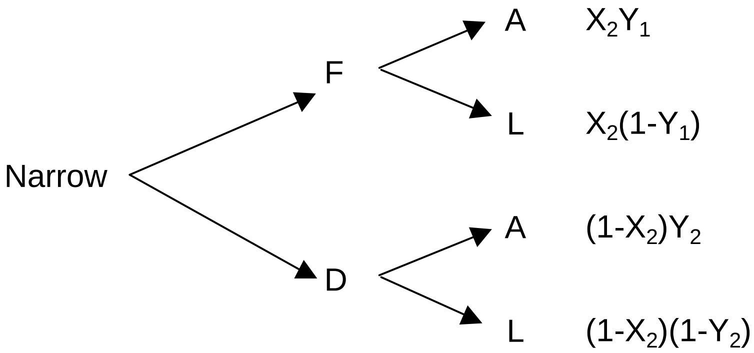
$1-Y_1$ = P (categorization = Lork | decision = friendly)

Y_2 = P (categorization = Adok | decision = defensive)

$1-Y_2$ = P (categorization = Lork | decision = defensive)

Figure 6. Assuming conditional independence, as the Markov property does, the probabilities can be multiplied to form the expected proportions as above. F = adopt a friendly demeanor, D = adopt a defensive demeanor, A = categorize face as an Adok, L = categorize face as a Lork, R= round, N = narrow face





Participant k

Round *Interaction Decision*

Participant Categorization

	D	F	
A	m_{AD}	m_{AF}	m_A
L	m_{LD}	m_{LF}	m_L
	m_D	m_F	M

Narrow *Interaction Decision*

Participant Categorization

	D	F	
A	n_{AD}	n_{AF}	n_A
L	n_{LD}	n_{LF}	n_L
	n_D	n_F	N

$X_1 = (m_{AF} + m_{LF}) / M$

$1-X_1 = (m_{AD} + m_{LD}) / M$

$X_2 = (n_{AF} + n_{LF}) / N$

$1-X_2 = (n_{AD} + n_{LD}) / N$

$Y_1 = (m_{AF} + n_{AF}) / (m_F + n_F)$

$1-Y_1 = (m_{LF} + n_{LF}) / (m_F + n_F)$

$Y_2 = (m_{AD} + n_{AD}) / (m_D + n_D)$

$1-Y_2 = (m_{LD} + n_{LD}) / (m_D + n_D)$