The Deconstruction of a Study: Toward More Effective Evaluation of Research Studies in Cognitive Social Psychology

by

José A. Montelongo

Robert E. Kennedy Library

California Polytechnic State University

San Luis Obispo, CA 93407

jmontelo@calpoly.edu

(805) 756-7492

Submitted to

Cal Poly Digital Commons

February 26, 2008

The Deconstruction of a Study: Toward More Effective Evaluation of Research Studies in Cognitive Social Psychology

Abstract

Information literacy modules produced by academic libraries to facilitate the research process typically use the criteria of relevance, timeliness, reliability, coverage and accuracy to assess the various information resources undergraduate students use to write research reports. These same criteria are applied to the wide spectrum of research sources that may range from popular magazines to research journal articles.

In the field of Cognitive Social Psychology, many research questions necessitate the use of psycholinguistic stimuli (word lists, paired-associates, sentences, stories, etc.) as their treatments. This paper investigates the ability of information literacy modules based on the standards set forth by the Association of College and Research Libraries (ACRL) to assist students in evaluating empirical studies investigating social cognitive behavior.

A study of social balance schemas was deconstructed and analyzed. Using the evaluation module based on ACRL standards, this study was evaluated as relevant, reliable, authoritative, and accurate. Similarly positive assessments of the study have been reached by experts in the field of social cognitive psychology. However, the evaluation of the study using questions grounded in experimental methodology and a basic understanding of psychological theory and statistical methods proved to be contradictory. A new set of analytical questions for evaluating research studies using psycholinguistic materials was generated from the errors in the experimental study.

The Deconstruction of a Study: Toward More Effective Evaluation of Research Studies in Cognitive Social Psychology

The Association of College and Research Libraries (ACRL) has defined information literacy as those abilities required "to recognize when information is needed and the ability to locate, evaluate, and use effectively the needed information." Toward this end, the ACRL (2003) has set forth five standards, each with its own learning outcomes and performance indicators, for developing the information literacy of college students. The five standards for an information literate student are:

1. defines and articulates the need for information
2. selects the most appropriate investigative methods of information retrieval systems for accessing the needed information
3. evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system
4. uses information effectively to accomplish a specific purpose
5. understands many of the economic, legal and social issues surrounding the use of information and accesses and uses information ethically and legally

The focus of the present paper is on the third of these five standards—the *evaluation* of journal articles, specifically those that include experimental investigations in the area of cognitive social psychology for which the stimuli used are words, sentences, paragraphs, or stories.

The ACRL (2003) has enumerated seven performance indicators for the evaluation standard:

1. summarizes the main ideas after having read the text

2. articulates and applies initial criteria for evaluating both the information and the sources
3. synthesizes main ideas to construct new concepts
4. compares new knowledge with prior knowledge to determine the value added, contradictions, or other unique characteristics of the information
5. determines whether new knowledge has an impact on the individual's value system and takes steps to reconcile differences
6. validates understanding and interpretation of the information through discourse with other individuals
7. determines whether the initial query should be revised

Academic libraries often create web-based instructional modules to help their students evaluate information. Some of these sites concern themselves only with the evaluation of websites, while more ambitious ones extend theirs to the evaluation of articles in periodicals and journals. Such modules take the learner through the stages of research from defining a topic to evaluating a scholarly article to incorporating the article into one's knowledge base.

Most library-created information literacy modules, such as the CSU Information Competence Project (Swanson, 1999) shown in Figure 1 use the criteria of *reliability*, *validity*, *accuracy*, *authority*, *timeliness*, and *point of view* for evaluating information. *Reliability* refers to the credibility of the source. For evaluating the reliability of information, a reader is taught to ask, "Who is the author of the information?" To determine the *validity* of the information in an article, one must ascertain if the research findings have been discredited or become obsolete. To assess a*ccuracy*, it is necessary

determine whether the information is factual or an opinion. The type of publication in which the work is published provides a measure of its *authoritativeness*. The reader must determine if the article is published in a popular magazine, trade publication, or scholarly peer-reviewed journal. *Timeliness* is gauged by examining the source's date of publication. *Point of view* is a consideration of any bias in the information. *Relevance* is synonymous with appropriateness. Among the questions related to relevance are: "Is the information a primary or secondary source?" and "Is the information directed toward a general or a specialized audience?" Finally, *coverage* concerns the depth and breadth of the information. Readers will note that the questions used in the evaluation of information from printed sources in library modules are those used in evaluating Internet websites.

Insert Figure 1 about here

The Special Case of Social Cognition Experiments Using Linguistic Stimuli

Researchers into the field of social cognitive psychology often encounter research investigations dealing with language phenomena. Verbal materials (nonsense syllables, paired-associates, sentences, and stories) have been used to study information processing, learning, and forgetting.

Evaluating research articles concerned with verbal behaviors is often problematic for students in social cognitive psychology. This is because they are not exposed to studies with verbal materials in their research methods classes. Instead, they are taught experimental design and analysis for evaluating research studies dealing with subject variables such as gender, age, and intelligence. In these classes, students learn to judge

the experimental methods for selecting and assigning their participants to the treatments. They also learn to statistically analyze the results of experiments and the importance of the generalizability of results to the *subject* population.

Students develop those abilities required to find evidence of experimental bias and they become adept at pinpointing procedural missteps related to subject variables such as those dealing with the ways subjects are sampled and assigned to the experimental conditions. However, students in psychology remain underprepared for evaluating empirical studies that utilize verbal materials as their treatments.

In studies investigating subject variables, researchers are concerned with generalizing their results to the subject population. Since not every subject can participate in their experiment, they select their participants with the intent of maximizing the representativeness of the sample for purposes of generalizing the results to the subject population. Such care is warranted because a study whose results hold only for those particular subjects is of limited value to the field. For studies that employ verbal materials as their stimuli, there exists a parallel rationale. Investigators take steps to employ a *representative stimulus sample* to maximize the possibility that the obtained results will generalize to the language population. Ideally, every word from the entire word population would be included for study. Realistically, however, it is often too impractical for a researcher to exhaust the entire category. The best a researcher can do is to use as much of *both* the word and subject populations as possible and apply inferential statistics tests to generalize findings to *both* the subject and language populations (Coleman, 1964).

In the seminal article on generalizability to the language population, E. B. Coleman (1964) suggested that many researchers investigating verbal behavior were using *nonrepresentative* language samples in their studies. He observed that researchers often restricted their ability to generalize to the language population by imposing limits on their stimuli with respect to such things as the number of syllables, the length of the stimulus word, and the initial letters of words. Coleman pointed out that with each control, the researcher restricts the number of potential stimuli to only a subset of the population and thus limits the potential for generalizing the findings to the language population. On this matter of limited generalizability, Coleman wrote:

> "…(studies) of verbal behavior are scientifically pointless if their conclusions have to be restricted to the specific materials used in the experiment." (p.219)

Another criticism of research in verbal learning and cognitive psychology results from the fact that many researchers did not provide a statistical analysis of results with the stimuli as the unit of analysis (Coleman, 1964; Clark, 1973). Ordinarily, researchers provide statistical analyses of their results for generalizing to the subject population, but neglect to provide an analysis of results for the stimulus sample. As pointed out by Coleman and Clark, researchers are obligated to provide the results of inferential statistical tests conducted on the language sample. As Coleman observed:

> "This unhappy state of affairs seems important enough to bear paraphrasing and italicizing: *There is little or no statistical evidence that the many experiments investigating language characteristics could be successfully replicated if different samples of language materials were used*." (p. 220).

Coleman's observations have implications not only for the execution of language experiments, but also for the evaluation of empirical studies in social cognition where the verbal materials are the treatments. As is the case in the evaluation of experiments investigating subject variables, it is essential for students to learn to question the experimental methodology in cases where it is apparent that the experimenter imposed so many controls that the results of the experiment are limited in generalizability to the language materials used in the experiment. Students must also learn to critically analyze the designs of the experiments, the procedures for presenting the verbal materials, and the statistical analyses applied to the results.

In this paper, the strategy will be to first deconstruct an actual empirical study in the area of social cognition and then use the observed flaws in the study to develop an evaluation module. The resulting module will then be compared with that of a library-produced evaluation module. The implications for information literacy modules will be discussed.

The Deconstruction of a Study: Toward More Effective Evaluations of Research

An article, "Cognitive Organization and Coding of Social Structures," by Picek, Sherman, and Shiffrin (1975) was selected for study. This study is an excellent test case for several reasons. First, it affords consideration of a small, but *exhaustive language population* of social story structures. This permits the reader to consider an entire language population with which to judge the representativeness of the sample of memory structures actually used in an experiment. This study also allows one to consider those variables that may affect the results of experiments utilizing psycholinguistic materials such as orders of presentation and retention intervals. The study was chosen because of

its experimental design, the procedures used to administer the treatments, and the statistical evaluations of results.

The specific research question investigated by the Picek, Sherman, and Shiffrin (1975) was whether participants use schemas to encode information contained in stories about social situations. A schema is a cognitive structure in memory that can be used to process and store information. The type of schema investigated by Picek, Sherman, and Shiffrin (1975) is derived from social balance theory and is used to explain the dynamics in social groups. More specifically, balance theory is used to predict conflict and tension in social situations. It is an attempt to *mathematically* distinguish between two types of social situations, balanced and imbalanced. *Balanced* situations are those harmonious states characterized by triads in which the number of mutually positive sentiments among its members is an odd number (one or three). *Imbalanced* situations are those states of tension and conflict where the number of *mutually positive* sentiments in a triad is an even number (zero or two). To illustrate, imagine triad of male employees, Joe, George, and Charlie, meeting to discuss a project. There are four potential situations. These are presented in Figure 2.

Insert Figure 2 about here

The situations in the figure range from the balanced Situation 1, in which all of the characters like each other to the imbalanced situation 4, in which in which the characters all dislike each other. Situation 2 is also balanced because it consists of one (an odd number) positive pairwise sentiment. However, Situation 3 is imbalanced since

there are two (an even number) positive sentiments. As can be seen, the degree of balance does not correlate with the number of mutually positive sentiments, but with whether or not the number of positive sentiments in a particular triad is even or odd. Because positive pairwise sentiments may occur in a balanced or imbalanced situation, it is the triad and not the mutual individual sentiments that is the unit of analysis.

To test the psychological reality of balance schemata, Picek, Sherman, and Shiffrin (1975) used stories about the social interactions of four persons instead of the usual three-person situations. Situations involving groups of four persons are slightly more complicated than those in Figure 2. While the stimuli used in the experiment were stories typed on paper and presented in booklets, graphic structures represent the social information contained within the stories. These graphic representations afford the reader quick-and-easy summaries of the social information included in the stories. One should always keep in mind, however, that these graphic organizers represent the stories presented in textual formats for the participants in the study. A balanced and an imbalanced structure are each graphically depicted both in their entirety and into their component parts (a, b, c, d) in Figure 3.

Insert Figure 3 about here

The set of complete four-person structures and their component triads represent the pairwise sentiments among the four characters in a story. Each letter (A, B, C, or D) represents a character in the story. The solid lines signify that the characters mutually

like each other. The dashed lines indicate that the characters mutually *dislike* one another. As can be seen, the first complete figure is balanced because all of the triads that comprise it contain an odd number of positive sentiments (one or three). The second complete figure is imbalanced because *at least one* of the triads is not balanced.

For their study, these researchers conducted two memory experiments using stories about persons in social situations. Picek, et al. hypothesized that those stories about social situations which conformed to balanced theory rules would be easier to learn and remember than stories that were imbalanced. They presented their participants with two "somewhat interesting and humorous" stories about the social interactions among four same-sexed characters. One of the stories was a *complete* story, which meant that all of the sentiments among the four characters were provided in the story. The complete story was either *balanced* or *imbalanced*. These researchers also wanted to explore the possibility of a "drift toward balance." They reasoned that if balance schemas are indeed "psychologically real" and guide the processing of information, then subjects might also exhibit the tendency to commit errors in a manner consistent with the hypothesized pre-existing balance schema. Therefore, the experimenters included a second *incomplete* story for the participants to read. For the *incomplete* stories, not all of the mutual sentiments among the four characters were provided. These stories could be either *balanceable* (none of the triads was already imbalanced) or *non-balanceable* (one of the triads comprising it was already imbalanced). In the second experiment, Picek, et al. (1975) utilized only the incomplete structures. In this experiment, each participant received both a balanceable story and a nonbalanceable story. The four structures used by Picek, et al. are depicted graphically in Figure 4.

Insert Figure 4 about here

Readers will recall having seen the complete balanced figure before. This is because it was the example of a *balanced* structure in Figure 3. It is balanced because *all of the triads* that comprise it contain an odd number of positive sentiments. On the other hand, the *imbalanced* figure consists of two balanced triads, ABC and ACD, and two imbalanced triads ABD and BCD. Since the entire structure is the unit of analysis and not the individual triads that comprise the structure, it only takes one imbalanced triad for the entire figure to be imbalanced.

The incomplete structures are also included in Figure 3. The *balanceable* figure consists of one complete triad CBD and it is already balanced. The *nonbalanceable* figure is not balanceable because it already has one imbalanced triad, CBD. By leaving the sentiments that would complete the other triads (CAB, BAD, and CAD) unstated, Picek, et al. (1975) wanted to test the possibility that subjects would complete the incomplete triads in a manner consistent with that predicted by balance theory.

Evaluation

Picek, Sherman, and Shiffrin (1975) used 216 participants in their two experiments. There were 144 subjects in Experiment 1 and 72 in Experiment 2. All of the

participants were tested in large groups of 40 and were afforded eight minutes to read each story. An interpolated math task (to eliminate the possibility of retrieval from short-term memory) separated the presentation of the stories. After the subjects had read the two stories and solved arithmetic problems for 2 minutes, they were provided with the names of the pairs of characters and asked to recall the sentiments between each of the pairs.

The balanced and imbalanced versions of one of the two stories used by Picek, et al. (1975) are included in Appendix A. In the stories, character "A" is Captain William Wave, "B" is Dr. Elmer Cleft, "C" is Professor John Stone, and "D" is Colonel Sam Far. The stories were each eight paragraphs long. The first paragraph provided an introduction of the characters, the setting and the predicament the characters faced. Subsequent paragraphs provided additional personal information about each of the four persons. The last sentence of every paragraph revealed the *critical* information: whether or not the characters "liked" or "disliked" each other. No reasons for the way characters felt toward each other were furnished. The final paragraph provided the resolution of the story.

The results of the first experiment were consistent with Picek, et al's main hypothesis. For the complete structures, the sentiments included within the balanced structure were easier to recall (Mean=83.6%) than those contained in an imbalanced structure (Mean= 73.6%). The results were statistically significant. For the incomplete stories, however, recall for the sentiments comprising the balanceable stories

(Mean=.72.9%) were not statistically different from those in nonbalanceable ones (Mean=72.7%). From this first experiment, Picek, et al. (1975) concluded that:

> "The concept of social balance and imbalance is real and significant to subject, He (sic) has verbal labels and images to discriminate instances of balance from those of imbalance. (p. 767)"

Furthermore, the authors suggested:

> "The codes for balanced situations are easier, shorter, and more available than those for imbalanced situations. (p. 767)"

Careful examination of the graphic organizer representing the complete balanced structure reveals that the code for this particular structure is indeed easier and shorter. The information in the balanced structure *may be reduced* to the rule, "Everyone likes each other except for person C, who dislikes and is disliked by everyone." For the balanced example in Appendix A, the rule is: "Everyone likes each other except for Professor John Stone, who dislikes and is disliked by everyone." Readers may experience this for themselves by reading the balanced story in the Appendix and focusing on Professor John Stone, the character who dislikes and who is disliked by everyone. However, no such rule exists for the imbalanced structure. For this structure, the participants must use rote learning (or some other idiosyncratic method) to remember the "likes" and "dislikes" for each of the four characters in the story.

While the code for this particular balanced structure is shorter and easier for subjects to learn than the imbalanced structure it was tested against, the larger question

for evaluating the representativeness of the language sample is whether or not the presence of a reduction rule is a distinctive feature of balanced structures in general (as Picek et. al. suggest) or if it is idiosyncratic to this particular balanced-imbalanced comparison. In order to judge whether or not these researchers confounded structural balance with ease-of-learning, readers must consider the *entirety* of the language population from which these structures were drawn.

The Evaluation of a Linguistic Sample

Picek, Sherman, and Shiffrin (1975) utilized these two particular complete structures in order to *control* for the number of positive and negative mutual sentiments. They used structures that included equal numbers of positive and negative sentiments in order to control for subjects' predisposition to use positive and negative sentiments equally. This meant having to employ balanced and imbalanced structures that contained three positive and three negative sentiments. The desire to control for guessing, however, led to the *exclusion* of other structures in the population that might have also been tested. Those structures with *six positive* sentiments and *no negative* ones and their converse, structures with *five positive* sentiments and *one negative* sentiment and their converse, structures with *four positive* sentiments and *two negative* sentiments and their converse, and any other *three positive* and *three negative* structures in the population were excluded from study. Moreover, there was no *theoretical* reason to restrict the sample to stimuli that contained equal numbers of positives and negatives as there is no stipulation in balance theory that states that structures must have an equal number of positive and negative sentiments.

Considering the Entire Stimulus Population

A consideration of the *entire* population of four-person balanced and imbalanced structures is necessary for evaluating the study. The total population is provided in figure 5. The eleven complete structures range from one in which everyone likes each other (figure 5a) to its matched imbalanced mate where everyone dislikes each other (figure 5b).

Insert Figure 5 about here

Inspection of Figure 5 reveals that there are three possible balanced structures (5a, 5e, and 5i). The other eight structures are imbalanced. The information in some balanced and some imbalanced structures may be reduced to a simple rule, while rote learning must be used to learn and remember the other structures. For instance, figure 5c may be reduced to the rule that every likes each other except for persons A and B, who dislike each other. For figure 5d, the converse rule holds. For the matched balanced and imbalanced pair of figures, 5e and 5f, the learner must remember the two specific pairs that like each other (figure 5e) or dislike each other (figure 5f) while every other pair feels the opposite. The imbalanced figures 5g and 5h are similar to figures 5e and 5f, except for the fact that in the former, the subject must remember that character “A” likes two people (Figure 5g) or dislikes two characters (Figure 5h).

An important fact uncovered by the consideration of the entire linguistic population is that there are imbalanced structures (figures 5b, 5f, and 5j) that are the *exact* matched mates of the balanced figures (5a, 5e,5i, respectively). Matched mates are equal in terms of the rule-learning they entail, although opposite in meaning. For instance, Figure 5j, deconstructed earlier in figure 3, is the matched mate of the balanced structure used by Picek, et al. (1975). Like its matched balanced mate, the information about the social sentiments among the four persons may also be reduced to a simple rule: "Everyone dislikes each other except for C, who likes and is liked by everyone."

The discovery of a matched imbalanced structure reducible to a code seemingly equal in ease-of-learning to the balanced structure used by Picek, Sherman, and Shiffrin (1975) is an important one. Had these researchers included this matched imbalanced structure along with the two structures they actually used, alternative conclusions may have been reached. For instance, if the balanced structure had shown itself easier to recall than the two imbalanced ones, this would have provided some evidence for the psychological reality of balance schemas. However, if the balanced and its matched imbalanced mate had led to equal levels of learning and both were superior to the difficult imbalanced structure, this would suggest that recall for stories about social situations is dependent upon the availability of a rule for making learning and remembering easier.

This step of considering the entire population of complete structures reduces the validity of the arguments presented by Picek, Sherman, and Shiffrin (1975) with respect

to complete structures. The next step is to extend the analysis to the *incomplete* balanceable and nonbalanceable structures.

The investigators included the incomplete stories in their study was to see if there existed a “drift toward balance.” For if subjects do have a mental representation of social situations that follows from balance theory, then, subject errors might also reflect the existence of social balance schema.

There are several types of errors that subjects can make with respect to the incomplete structures. First, participants may mistakenly recall a sentiment in which the characters “disliked each other” as one in which the characters “liked each other.” Or, they may make a similar error with those characters “who liked each other.” The third type of mistake is the intrusion. Intrusions are the result of a subject recalling a sentiment when, in fact, none had been overtly stated. Since intrusions are believed to be the result of a pre-experimental source of interference, the schema, their occurrence can be used to provide evidence for the existence of a balance schema if they balance the resulting structure.

Although there were no statistically significant differences in recall for the incomplete balanceable and non-balanceable structures in Experiment 1, balanceable stories were recalled better (Mean=.785) than the nonbalanceable ones (Mean=.692) in Experiment 2. This superiority in recall of the balanceable story relative to its counterpart seemingly provides the first indisputable measure of support for the existence of a balance schema. In order to determine the validity of this claim, however, the reader

should carefully examine the population of incomplete structures for the possibility of an alternative explanation.

The incomplete structures, the balanceable and nonbalanceable structures used in both Experiments 1 and 2 by Picek, Sherman, and Shiffrin (1975), were depicted graphically in Figure 2. At first glance, these incomplete structures appear equivalent. Both structures are comprised of two negative, two positive and two unstated sentiments. However, there exists a subtle difference in the stated sentiments for person "C." In the balanceable figure, both of the sentiments for character C toward characters B and D are negative. C and B dislike each other and C and D dislike each other. For the nonbalanceable structure, however, the participant had to recall a positive sentiment for D and a negative sentiment for B. That is, C and B dislike each other, but C and D like each other. This subtle difference may have accounted for the statistically significant superiority of the balanceable structure. For both the balanceable and nonbalanceable structures, the reader had to recall that no sentiment was stated for the relationship between characters C and A.

The incomplete balanceable and non-balanceable structures in the Picek, et al. study were similar to the complete structures except for the fact that the mutual sentiments between two of the six pairs of the characters were not overtly stated. Before considering the entire population of such structures, it should be noted that there again existed no *theoretical* reason for omitting the sentiments between two pairs of characters. In addition to other structures in which four mutual sentiments were stated and two mutual sentiments were excluded, these researchers might have employed stories where

five sentiments were declared and one omitted, three declared and three omitted, two declared and four omitted, etc. The researchers chose to omit two sentiments because they wanted an equal number of undeclared positive and negative sentiments. However, by choosing to limit their investigation to just one of these possibilities, these researchers greatly reduced their ability to generalize their results to the stimulus population from which these structures were drawn.

There are fifteen incomplete four-person structures in the balanceable/non-balanceable population for which the sentiments of two of the pairs are omitted. The entire population is presented in Figure 6. As was the case with the complete structures, there exists a non-balanceable structure (Figure 6n) that is the matched counterpart of the balanceable structure (Figure 6m) that was used by these researchers. Both of these structures differ from the non-balanceable structure (Figure 6o) actually used by the experimenters. Both figures (6m and 6n) appear to be easier than the structure (Figure 6o) because subjects only had to recall one (as opposed to two) type of sentiments for person C in addition to recalling the one sentiment that was omitted.

Insert Figure 6 about here

Picek, Sherman, and Shiffrin (1975) again utilized two structures that were needlessly unequal in terms of the amount of information to be learned. That is, they confounded structural balance with ease-of-learning. This error in experimental practice

might have been avoided had these researchers considered the population of incomplete structures.

The evaluation of the Picek, Sherman, and Shiffrin (1975) study has been limited to the story structure sample and the population from which it was drawn. However, for each one of these complete and incomplete structures, there are an infinite number of *stories* about an infinite number of *characters* in an infinite number of *social situations*, all of which may have been used. Yet, for generalizing to this vast story and social situation population, the authors used only these two "somewhat humorous" stories. Furthermore, the stories were about groups of people of the same sex and the stories were humorous and not serious. However, there was no *theoretical* reason to limit the stories by sex or degree of humor. By choosing to limit their stimulus sample to these two stories, these researchers greatly restricted their potential to generalize to the language population.

The analysis of the linguistic stimuli used by Picek, Sherman, and Shiffrin (1975) has served to point out the importance of considering the entire language population from which the sample is drawn. It is important not only for the researcher, but also for the reader evaluating the article. This exercise has demonstrated that questions about the representativeness of the language sample need to be included as part of evaluation modules, web-based or otherwise, for investigations in which the stimuli are verbal materials.

Evaluation of the Experimental Design

Information literacy also includes the ability to criticize the *experimental design* of an experiment. Picek, Sherman, and Shiffrin (1975) utilized two conditions in their initial experiment. One group of participants read a balanced story and a non-balanceable story. A second group of subjects was presented with the converse: an imbalanced story and a balanceable story. Procedurally, Picek et al correctly had one-half of each group read the incomplete story prior to the complete story, while the other half read the complete story prior to the incomplete one. However, two other possible experimental groups of participants were missing. Omitted were the *balanced-balanceable* and the *imbalanced-nonbalanceable* conditions. These two conditions are important because many experiments in the area of verbal learning and cognitive psychology have repeatedly demonstrated the greatest transfer effects (proactive facilitation) when two similar stimulus conditions follow each other. Contradistinctively, the *balanced-nonbalanceable* and *imbalanced-balanceable* conditions are antagonistic to each other. Reading a balanced story would not seem to facilitate the learning of an imbalanced story as much as the prior reading of another balanced story. Therefore, Picek et al. may be criticized for not including these other two experimental conditions that might have maximized the potential for observing transfer effects. The failure to observe a "drift toward balance" and the limited occurrence of intrusions, therefore, may have been due to the fact that the researchers did not use an experimental design conducive to the observance of these phenomena and not to the absence of a "drift."

Picek, Sherman, and Shiffrin (1975) might have includes a *balanced-neutral* and *imbalanced-neutral* conditions and their converse in their design. Instead of using incomplete structures that were balanceable or nonbalanceable, they could have employed an incomplete neutral structure. Figure 6i, for example, is comprised of two positive, two negative, and two "nothing stated" sentiments. The use of a neutral structure such as this one might have provided a better assessment of transfer than the balanceable and nonbalanceable structures.

In most verbal learning studies requiring the learning of successive lists or stories, an immediate recall test is administered after the reading of the first story and prior to the reading of the second story. Such a design permits the investigators to assess and quantify the interference effects occasioned by the learning of the second list or story, by permitting the tracking of individual items from pre-test to the post-test. It also permits the recording of those items that were initially incorrect on the immediate test, but correct on the delayed test. Since Picek, et al. only administered a test after the participants had seen both stories, it is not possible to precisely gauge the effect the learning of one story on the story that either preceded it or followed it.

The evaluation of the experimental design raises further doubt about the validity of the Picek, Sherman, and Shiffrin (1975) study as an investigation of cognitive schemata. Besides using nonrepresentative story structures (complete and incomplete), these researchers failed to employ an experimental paradigm that permits the observation of transfer effects. Information literacy modules need to include questions about the experimental design of a study to aid in the evaluation of empirical studies.

Evaluation of Procedures

While the evaluation of the linguistic sample and the experimental design are critical factors to be considered in the judgment of an experiment where linguistic materials are used as the stimulus treatments, there are still other factors to consider. Among these are the boundary conditions employed by the experimenters. These would include such factors as the orders of presentation and the length of the retention intervals.

In addition to the fact that Picek, Sherman, and Shiffrin (1975) used experimental conditions that were not conducive to the observance of intrusions, their failure to find evidence of intrusions or a "drift toward balance" may have also been the result of the particular boundary conditions employed by the investigators. The subjects were given eight minutes to read each story and then were tested almost immediately after having read both stories. Most verbal learning studies find greater evidence for intrusions from prior learning at retention intervals ranging from one day to several weeks. By using only an eight-minute study period for reading each story and by limiting their study to but one retention interval, Picek, et al. limited their ability to observe a "drift toward balance."

A remarkable finding claimed by Picek, Sherman, and Shiffrin (1975) concerned the evidence for the serial processing and recall of the social information included in the stories. In their general discussion of results, Picek, et al. (1975) remarked:

> Subject constructs a social code serially in the order of presentation of the relationship. In a story, he (sic) attempts to add new social information to the structure that has been constructed to that point (p. 767).

As evidence, the authors included the figures depicting the recall of relationships by order of presentation. The figure graphically depicting recall for the complete structures is shown in Figure 7.

Insert Figure 7 about here

As can be seen from figure 7, recall for the balanced structure was significantly superior to that for the imbalanced structure from the third position onward. It is because of this third position differentiation that Picek, et al. (1975) concluded that subjects were coding the information using a balanced schema. However, the fact that these investigators utilized only one order of presentation (C-D, B-D, B-C, A-B, A-C, and A-D) and only one order for testing the recall of the sentiments (C-D, B-D, B-C, A-D, A-C, and A-B) permits an alternative explanation. The reader will recall that the sentiments comprising the balanced structure could be reduced to the rule that, ”Everyone dislikes “C”, but everyone likes everybody else.” Examination of the performance graph for the balanced structure reveals that the three points of highest recall are found at the first, third, and fifth positions. From the presentation order presented by the authors, these positions correspond precisely to those relationships in the balanced structure having to do with person “C,” for whom there is a reduction rule. This fact raises the possibility that the superiority of recall for the balanced structure relative to the imbalanced one may have had nothing to do with serial encoding at all. Rather, the effect may have been due to the fact that subjects remembered the information about person “C.” In other words,

subjects may have been responding to the recall questions on the basis of the rule about person “C” and not on the basis of a serial ordering.

A careful reading of the experimental methods section also uncovers the fact that the order of presentation was different for the balanced and balanceable structures from the one used for the imbalanced and non-balanceable ones. The ordering for the latter was (C-D, B-C, B-D, A-C, A-B, and A-D) and reflects a difference from the presentation of the former at the second, third, fourth, and fifth positions. The differences in ordering were necessitated because the authors wanted to “alternate between like and dislike relations for all structures” to guard against biases toward the positive. As with other variables, employing only one presentation order severely limits the potential to uncover potential interactions between order and types of structures, as well as the generality of results, especially when the information in one of the structures can be reduced to a rule.

The claim for having discovered a serial position effect also presupposes the ability to pinpoint precisely when the information was available to the participant. As noted earlier, most psychological studies employ an immediate recall test following a reading of the first story and a prompted recall test at the end. In the Picek, et al. (1975) study, the procedure was such that subjects read two stories (and completed an interpolated math task) before attempting to recall the information about the social relationships from the stories. To test their memory, the subjects could answer in any order they wanted and there was no record of the order in which the information was retrieved. The lack of an immediate recall test following the presentation of the first story, especially in the case of the complete stories, raises the possibility that subjects

might have processed the information in many ways before ever writing down their responses than the authors would have us believe. For instance, the subjects who were presented the balanceable structure with the reduction rule may not have arrived at the rule until they had already finished reading the second story or after they had completed their interpolated mathematics task. They may not have realized the rule until they had written down all of their answers. Under this set of experimental procedures it is difficult to ascertain with precision what the subjects knew and when.

The reason for using only one delayed recall test resulted from the authors' desire to *control* for yet another variable. The investigators did not want participants to know that they would be tested for recall of the social relationships. In order to see whether this sacrifice in precision was warranted, these researchers might have conducted another experiment or pilot experiment to observe the priming effects of a pre-test, if indeed they exist.

The findings for the incomplete structures, balanceable and nonbalanceable, in terms of presentation order are provided in Figures 8a and 8b. Once again, there is a drop in performance for the third position of the nonbalanceable structure in both experiments. However, the authors do not explain why there is a rebound in performance for the fourth relationship in Experiment 1 and not in Experiment 2. Nevertheless, Picek, et al. seized upon this third position drop-off as evidence for a social schema encoding explanation of balance. An alternative explanation for this result lies in the observation that the subjects had to learn more about the character "D" in the nonbalanceable structure than about the person "C" in the balanceable one. The third pairwise sentiment presented in the

nonbalanceable structure was the one between characters “B” and “D.” The third sentiment in the balanceable structure was that between “B” and “C.” For both the balanceable and nonbalanceable structures, the subjects had to learn that “B” disliked both “C” and “D.” In the nonbalanceable story, person “D” disliked “A,” but like “B” and liked “C.” In the balanceable story, person “C” disliked “B” and “D” and nothing was said about “C's relationship with A. Because the sentiments for “C” in the balanceable story are both negative and because the subjects had to remember the two people who liked “D” and also the person who disliked him, there was a higher level of performance for relationships for person “C” than for person “D.” However, with the differences in the amount of information to be learned, the differences in presentation order, and the effects of having to read a different complete story, it is difficult to disentangle what actually happened. These alternative explanations for the claimed “serial position” effects serve to remind us of the importance of knowing how to read graphs and charts is for true information competence.

Insert Figures 8a and 8b about here

Evaluation of Statistical Analyses

Readers must also be able to judge the appropriateness and completeness of the statistical analyses presented in those articles as part of their information literacy. In the

analysis of their first experiment, Picek, et al. found two statistical comparisons that reached statistical significance. These were:

a) balanced structures were recalled better than imbalanced ones; and,

b) complete structures were recalled better than incomplete structures.

However, the subjects did not read the stories in isolation. They were presented with a complete story and an incomplete story, or the converse. Missing from the analyses in this study was a statistical evaluation of the interactions required by virtue of the fact that the stories, complete and incomplete, could either be read first or last. Not only is it important to know if the reading of one story has an effect on the reading of a subsequent story, but the discovery of differential effects might serve to inform theoretical discussions of the results.

In their results section of their paper, Picek et al. (1975) avoid a discussion of the possible interaction effects caused by the fact that they presented subjects with one story (either complete or incomplete) and then another (incomplete or complete). The authors dismissingly wrote:

> "Since all subjects received one type of complete and one type of incomplete structures, interaction which involved the type of structure would be difficult to interpret. Consequently, the data from the complete and incomplete structures were analyzed separately." (p. 762)

While it is not possible to submit their data to statistical analyses without the raw data, readers might organize the data to determine the plausibility of the authors' conclusions. The mean percent recall for the complete/incomplete story combinations are re-presented in Figure 9 to reflect the interactions of the complete and incomplete story combinations. There were two types of complete/incomplete story combinations, *balanced-nonbalanceable* and *imbalanced-balanceable* and two possible orders of presentation (first and second). The means for the four groups are depicted in Figure 9. The means were abstracted from the data provided by Picek, et al. in their separate analyses of complete and incomplete structures.

Insert Figure 9 about here

One of the striking features of the results in Figure 9 is the high level of performance exhibited by the subjects. It will be recalled that subjects had only to indicate whether or not the character pairs liked each other or disliked each other for the complete stories. For the incomplete stories, subjects had a third alternative, "nothing stated," for those pairs of characters whose sentiments toward one another were not mentioned.

One of the expectations from a design in which one story is read before another is that the most recent story will be recalled more accurately than the earlier story on tests of immediate recall—the so-called "recency effect." As expected, in three of the four

cases in the Picek study, the second story was recalled better than the first story. The lone exception occurred for the group that read the balanced story first and then the nonbalanceable story. In fact, for both groups presented the balanced-nonbalanceable combination, the balanced story is recalled better. For the imbalanced-balanceable groups, superiority of recall is always demonstrated for the latest story.

One way to account for this finding is to imagine the differential types of learning available to subject in both cases. The reader will recall that the balanced structure could be reduced to the rule that "Everyone likes each other, except for character "C", who dislikes and is disliked by everyone." No such rule is available for the non-balanceable structure. For this structure, the subjects must learn all of the sentiments through rote learning. Since the subject has only to remember the rule about person "C," this story is not only easier to learn than the nonbalanceable one, but it is more resistant to forgetting regardless of whether it is read first or second. Such is not the case for the second type of complete/incomplete combination. Both the imbalanced and balanceable structures require rote learning of the sentiments for all of the characters. It would follow, then, that learning and remembering all of the sentiments is more difficult than that case for which a rule is available.

There is suggestive evidence that the participants inferred the rule available for person "C." The subjects presented with a balanced structure recalled "dislike each other" sentiments (Mean=.873) better than they did "like each other" sentiments (Mean=.796). Furthermore, although the participants presented with an imbalanced story recalled "dislike each other" and "like each other" sentiments equally well (Mean=.736),

the disparity between the two complete structures for correct “disliked each other” is pronounced (Mean=.873 to Mean=.736). This suggests that some subjects presented with a balanced story used that portion of the rule dealing with the unpopularity of character “C.”

Although the imbalanced-balanceable combination demonstrated the predicted “recency effect” for the superiority of recall for the second story, an anomaly appears in the data that defies explanation on theoretical grounds. As can be seen from Figure 8, the means for three of the groups range from a low of .704 to a high of .847. One of the four subject groups, however, was markedly different from the others. The group that was presented with a balanceable structure first and then read an imbalanced story had means that ranged from only .630 to .680. The level of performance is so different from the other groups (especially the very similar imbalanced-balanceable condition) that the reader is left to wonder what caused this. These low results would be expected on the basis that both stories required the rote learning of the sentiments between the characters. However, the group that received the Imbalanced and balanceable stories in converse order exhibited relatively high levels of performance. An examination of the relation data provided by Picek, et al. reveals that for the low-performing balanceable group, many subjects had a problem with the “nothing stated” category. While this provides a partial answer for the anomaly, it does not explain why there was a higher level of recall for the Imbalanced story when it was presented first and not when it was the most recent story. Why is it that a story which required the learning and recall of individual items of information for which there was no apparent reduction rule was recalled better when it preceded another difficult story than when it followed the difficult story?

These results may be due to other factors besides the two stories or the interaction of the two stories. One possibility is that just by chance one of the groups was comprised of subjects who were either better readers and/or guessers. Or, maybe one group included the worst readers and/or guessers! The possibility that this is what occurred is made more realistic by the realization that the authors did not report on whether or not their subjects were randomly assigned to the various conditions in their methodology section.

The analysis of this study's statistical analysis of results and the conclusions derived from this re-analysis point to the need to include questions about the statistical treatment of the data by the investigators in information literacy modules. In this case, the review of the statistical analyses revealed the possibility that the processing of rule-governed structures was more resistant to forgetting than those having to be learned in rote fashion. Furthermore, it raised the possibility that the assignment of subjects to the four conditions may not have been random.

Most of the analyses performed by Picek, Sherman, and Shiffrin (1975) dealt with the recall of the individual items. These researchers analyzed their data for the "like each other" and "dislike each other" responses for the various conditions in an effort to inform the inferential process. However, the unit of theoretical importance in this study is the entire structure and not the composite triads or the individual sentiments embedded within those triads. These investigators should have devoted a considerable portion of their data analysis to the *actual structures* subjects generated in their recall of the information. An analysis of each subject's responses for type of structure would have

provided evidence for or against the existence of a balance schema. Were the structures generated by subjects balanced? Imbalanced? Balanceable? Non-balanceable?

To their credit, Picek, et al. (1975) examined some of the responses for the incomplete structures for evidence of a drift toward "balance." However, they only examined the responses for those items for which no sentiment was overtly stated.

Referring to the balanceable structure in figure 4, an error was deemed "in the direction of balance" if:

1. A-B was misrecalled as "liked each other" and A-C was misrecalled as "disliked each other."

2. A-B was misrecalled as "liked each other" and A-C was correctly recalled as "nothing stated."

3. A-B was correctly recalled as "nothing stated" and A-C was misrecalled as" disliked each other."

The analysis revealed that 29 of the 36 errors were in the direction of balance in the first experiment. The results were statistically significant using a chi-square test. A similar analysis of the results of the second experiment fell just short of statistical significance. To guard against response bias, the authors checked to see that the same responses were not used for the nonbalanceable structure. The subjects did not make the above errors with the nonbalanceable structure as they had with the balanceable one.

Thus, the investigators concluded that there was a “drift toward balance” for the balanceable structure.

Picek, Sherman, and Shiffrin (1975) then analyzed the results for the nonbalanceable structure to see if subjects demonstrated a “drift toward imbalance.” The authors examined only those triads containing unstated sentiments. These results revealed a “drift toward imbalance!” In their summary of these two analyses of the incomplete triads, the authors concluded that subjects’ responses for information not explicitly stated is dependent upon the structure: if the structure is balanceable, the subjects tend to make responses to the unstated links in the direct of balance. If the structure is nonbalanceable, subjects make their responses in the direction of imbalance.

The reader must consider the appropriateness of these two analyses. In this particular case, an analysis of only those pairwise items for which nothing was overtly stated, may be misleading. Just because a subject provided a “like” or “dislike” response for one of these items does not necessarily mean that it was an intrusion. As was discussed earlier, some errors might not be “intrusions” reflective of the existence of a schema, but rather, instances caused by the subject’s confusion of the various characters. Given this possibility, the authors needed to have analyzed the entire structure to decide if the mistake was really an “intrusion.”

In their final analysis, Picek, Sherman, and Shiffrin (1975) examined the structure of triads resulting from the subjects’ responses without regard for the correctness of the response for only the two triads that were missing only one pairwise sentiment. Unlike the previous analyses, which employed the individual sentiment as the unit of analysis,

this assessment of the data examined *two* of the four triads as the unit of analysis. These were the ABD and ACD triads (see Figure 4). The triad BCD was already complete. The remaining triad, ABC, was missing two pairwise sentiments.

The results using the triads as the unit of analysis revealed the trend noted earlier: if the structure is balanceable, the subjects tend to make responses to the unstated links in the direct of balance. If the structure is nonbalanceable, subjects make their responses in the direction of imbalance.

There are several problems with this analysis. First, the proper unit of analysis is the entire structure, not the structure of the component triads that comprise the structure or the individual pairwise sentiments that make up the triads. Secondly, there is no theoretical reason for analyzing only the resulting structures of the two triads that lacked only one sentiment. This selective analysis is inappropriate because half of the information of the entire structure is omitted. Just because there was a tendency for subjects to balance the balanceable triads or imbalance the nonbalanceable ones for those missing one sentiment, does not mean that this was the case for the other two triads for which there is no analysis. For instance, suppose that a subject balances the ABD triad. The possibility exists that the subject may have also unbalanced the ABC, ACD, or BCD triads. If the subject unbalanced any of these, then, the resultant structure is unbalanced. Without an analysis of the entire structure itself, the analysis put forth by these authors borders on the pointless.

The discussion of the Picek, Sherman, and Shiffrin (1975) analyses of results further demonstrates the need for information literacy guides to include exercises

designed to evaluate the analysis of results provided by an article. Information literate readers must develop the ability to point out incorrect or incomplete assessments by researchers.

Evaluation of the Picek, Sherman, & Shiffrin (1975) using ACRL Standards

The Picek, Sherman, and Shiffrin (1975) study was evaluated using the criteria suggested by the ACRL (2003). The analysis of the article is provided in Table 1.

Insert Table 1 about here

As can be seen, the Picek, Sherman, and Shiffrin (1975) study meets or surpasses the standards established by the ACRL for evaluating information. The study's reliability is almost unquestionable given the scholarly reputation of two of the authors, Sherman and Shiffrin. Although Yahoo and Google searches of James Picek proved fruitless, Steven Sherman and Richard Shiffrin are prominent members in their respective fields of social psychology and cognitive psychology. Professor Sherman is one of the nation's most cited researchers with over three hundred articles to his credit (ISIHighlyCited.com, 2001). Professor Shiffrin is a renowned psychologist and a member of the prestigious Society of Experimental Psychologists, the National Academy of Sciences and the American Academy of Arts and Sciences. The *Journal of Personality and Social Psychology*, the journal in which the article is published, is a periodical published by the American Psychological Association that is widely circulated.

With regard to coverage and accuracy, this article was not an abridged version of an original study. It provides a review of the extant literature including a discussion of the various experimental paradigms used to investigate the psychological reality of schemata dealing with social situations. Finally, the conclusions reached by the authors were grounded in the statistical analysis of their data. Although written thirty years ago, the article's timeliness is not at issue since recent social cognition studies continue to cite it (for example, Janicik & Larrick, 2005; von Hecker, 2004). Using the evaluation criteria published on the ACRL and associated websites, the Picek, Sherman, and Shiffrin (1975) study represents a solid investigation of social schemas. The article possesses relevance, timeliness, reliability, coverage, and accuracy.

In addition to meeting the evaluation criteria suggested by the ACRL (2003), the Picek, Sherman, and Shiffrin (1975) study has received positive reviews in social cognition textbooks, handbooks, annual reviews, and other journal articles. Some of these reviews are included in Table 2. The sources for the reviews were drawn from social cognition textbooks and handbooks.

Insert Table 2 about here

The reader can see that all of the comments in Table 2 are positive. Hastie, Park and Weber (1984), for instance, list the Picek et al. (1975) study as an example of a "more sophisticated (and probably more sensitive) research method" for testing

predictions emanating from balance theory. In Higgins & Bargh's (1987) article in the *Annual Review of Psychology*, the authors suggest that the Picek et al. study is a "nice example" of memory schemata influencing processing information without evidence of intrusions. In summary, the assessments of the Picek, Sherman, and Shiffrin (1975) article using the criteria published by academic librarians and their associations for the evaluation of information and those provided by investigators in the field of social cognition concur in their positive estimations of the study.

The present analysis suggests that the field of social cognition was too uncritical in its assessment of this study. Decades after its publication, the study is still inexplicably lauded for its experimental sensitivity and its findings. Undoubtedly, much of this is due to an underconcern for generalizing to the language population on the part of editors and reviewers.

Toward More Powerful Evaluation Modules

The deconstruction of the Picek, Sherman, and Shiffrin (1975) study has exposed the weaknesses of the ACRL evaluation criteria for evaluating experiments utilizing psycholinguistic stimulus materials. The results of the present analysis make clear that such modules need to be amended, at least for social cognition experiments where the treatments are verbal materials. For such studies, readers must first and foremost consider the representativeness of the language population. In order to accomplish this, a reader should generate as much of the total language population as possible and then compare a study's language sample to it. Conclusions reached by authors of studies that utilize only a tiny subset of a language population should be critically analyzed.

Readers should also consider the completeness of the experimental design as part of their information literacy. They must guard against designs and procedures that are not consistent with extant theory and practice. Attention should also be paid to the procedures used for presenting the stimuli, as well as to the boundary conditions under which performance will be observed. Finally, readers should examine the appropriateness of the statistical treatments of results to ensure that the experimenters are conducting valid assessments of their data. As a first step, a list of questions for evaluating experimental studies that includes psycholinguistic material is presented in Figure 10.

Insert Figure 10 about here

Many of the questions in the module reflect those asked in the deconstruction of the Picek, Sherman, and Shiffrin (1975) study. They are clustered in categories reflecting concern for generalizability to the linguistic population, the appropriateness of procedures, the evaluation of the statistics applied to the results, and an overarching evaluation of the study itself.

The first set of related questions (1a-6a) concerns the representativeness of the language sample. As is the case with the subject sample, students must judge the adequacy of the stimulus sample for generalizability to the language population. Students should try to imagine instances of the stimulus population which would seem to follow

uphold the findings of the study at hand or imagine instances of the language population for which the stated results might not hold. This process led the present author to a consideration of the entire four-person stimulus population and to question this study.

The second set of questions (1b-4b) asks about any potential biases in the language sample other than the independent variable being tested that may have influenced the results. This follows from the Picek, et al. study, where a reduction rule was available for the balanced structure and not for the imbalanced structure even though the investigators "controlled for" the number of positive and negative sentiments. While a few language characteristics (frequency, imagery, etc.) are mentioned, the list is by no means exhaustive.

The next set of questions (1c-4c) relate to the treatment and selection of the subjects. As to whether the groups are treated individually or in groups, the preferred and more rigorous manner is obviously that whereby the subject is administered the treatment individually. When participants are treated in groups, many subjects may not put forth their best efforts because of the anonymity they enjoy. In the Picek, et al. study, for example, anomalous performance levels not attributable solely to stimulus variables were observed. Variables such as age, level of education, sex, and handedness (indicative of brain dominance) have all been shown to affect performance on linguistic tasks and thus should be considered in evaluating psycholinguistic studies.

As was observed in the Picek, Sherman, and Shiffrin study, readers should also take into account the procedures used by the experimenters. Many list-learning experiments have long demonstrated robust "primacy" and "recency" memory effects:

items that occur early in a list (primacy) and later occurring items (recency) are recalled better than those in the middle. Students should know that the most powerful studies are those that strive to exhaust the population of presentation orders.

The use of various retention intervals is similarly important because of interference effects. Prior knowledge can proactively facilitate or interfere with new learning. Newly acquired information can retroactively inhibit old knowledge. There is also the phenomenon of "spontaneous recovery," whereby temporarily forgotten prior knowledge is remembered. Thus, items exhibiting superior recall on an initial recall test may become forgotten at later retention intervals as prior knowledge items regain in strength. Students must be alert to the possibility that certain behaviors need to run the course of time in order to be observed. Just because some phenomenon was not in evidence during the time it was observed, does not necessarily mean that it will not be observed at longer retention intervals.

The fourth set of questions (1d-4d) concern themselves with the statistical analyses of results. For these questions the student must guard against meaningless analyses that do not relate to the research question(s) being considered. They must also be wary of selective analyses that fail to consider the totality of the data. In the Picek, Sherman, and Shiffrin article, the authors failed to include analysis that seemed relevant and crucial to their hypotheses. Furthermore, information savvy readers must possess the ability to conduct their own analyses when the author of a study fails to provide it. We saw that by ignoring the potential interaction effects for the balanced-nonbalanceable group, Picek, et al. (1975) failed to pick up the observation that the balanced story was

recalled at a higher level than the nonbalanceable story regardless of whether it was presented first or last. This may have suggested to the researchers that different types of learning (rule vs. rote) had occurred .

Picek, Sherman, and Shiffrin (1975) also failed to make a statistical appraisal of the structures generated by the subjects. That is, these investigators failed to assess their results using the appropriate unit of analysis. Furthermore, these investigators selectively chose to assess the structure created for only half of the triads when it made just as much theoretical sense to evaluate all of them. The information literate reader should possess the capacity to perceive faulty analyses of behavior and to decide if those that are missing are crucial to behavior.

The questions concerning series of experiments are crucial in the overall assessment of research. The best examples of research are those that build upon prior experimentation ever seeking to generalize their findings to other stimuli, other boundary conditions, other subjects, and even other dependent variables (performance measures). Information literate readers should not only be able to follow the reasoning behind the progression of a series of experiments, but more importantly should possess the ability to suggest their own experiments.

The list of questions is not exhaustive. Further modifications are necessary as others contribute to this enterprise. Notably lacking in the present treatment is a consideration of the dependent variable. Just as there are retention intervals, presentation orders, and presentation rates that make the observation of certain phenomena difficult, there are also performance measures that are likewise insensitive. In the case of the Picek,

Sherman, and Shiffrin (1975) study, for example, it is possible that a reaction time measure would have uncovered more differences than the recall task that was actually used.

In the end, the research enterprise reflects the investigator's desire to understand the phenomena under study. The opportunity to make discoveries is greatest when a researcher seeks to exhaust the language population. The chances of finding out something new are further enhanced when that researcher employs different boundary conditions, orders of presentation, alternative dependent (performance) measures. For the information literate reader, this means having the ability to understand the researcher's purpose and also the critical thinking skills to not only critique a study but also to conceive of further experimentation.

Summary

The deconstruction of an experiment that included the use of verbal materials as stimuli has uncovered many of the processes involved in reading and evaluating such studies. The exercise has also demonstrated the inadequacy of the initial criteria suggested by the ACRL (2003) and other information literacy training modules such as those found in the CSU Information Competence Project (Swanson, 1999) or OASIS (2004). The evaluation of a study requires much more than a consideration of relevance, timeliness, reliability, completeness, and statistical grounding.

The exercise of deconstructing a study can inform the process for creating a more potent evaluation model. While no one article can exhaust the entire population of

methodological, statistical, or theoretical questions that may be asked of experimental investigations, evaluations of studies such as that of Picek, Sherman, and Shiffrin (1975) may provide a foundation for creating more potent information literacy teaching modules that will result in the development of readers capable of assessing the experimental design and procedures, the representativeness of the linguistic stimuli, and the adequacy of the experimental design and statistical analyses.

One of the purposes of the present paper was to demonstrate the importance of evaluating the linguistic sample of a study. The Picek, et al (1975) study proved a more than apt example for this purpose because of the investigators' almost total absence of concern for generalizing to the stimulus population. This was apparent from the observation that so few complete structures were used, especially when there was no theoretical reason for not including the other nine structures in the experiment. Additionally, the investigators confounded structural balance with ease-of-learning. There existed a reduction rule that reduced facilitated the learning and retention of the balanced structure. To learn the social relationships included in the imbalanced structure required the rote learning of the information.

Similarly, only two out of a possible fifteen *incomplete* structures were utilized. While these two stimuli were matched for number of positives and number of negatives, the more precise match was not used. Again, these researchers did not equate their stimuli for ease-of-learning. Additionally, these investigators were guilty of using only two out of an infinite number of stories about an infinite number of characters in an infinite number of social situations for generalizing their findings to the language population.

Information literacy also extends to a consideration of the experimental design. The particular design employed by Picek, et al. (1975) suffered theoretically from the omission of two experimental conditions typically found in studies examining the effects of prior learning on new learning. As was noted earlier, many experiments in the fields of verbal learning and cognitive psychology have included positive transfer or "priming" conditions as part of their design in an effort to observe any transfer effects. For this study, this would have required the inclusion of a *balanced-balanceabl*e condition as well as an *imbalanced-nonbalanceable* condition. The inclusion of a balanced –balanceable and imbalanced-nonbalanceable may have permitted the observation of transfer effects. Furthermore, the investigators failed to include an immediate recall test that would have permitted a more detailed understanding of proactive and retroactive interference effects in the recall of social sentiments.

True information literacy also involves close examination of procedural details. In their study, Picek, Sherman, and Shiffrin (1975) used only one order of presentation out of many. They also used only one order for testing the recall of the sentiments between the four characters. With respect to testing, only one rather short interval intervened between the time the participants finished reading the article and the time they were tested.

Information literacy extends to the analysis of the statistical treatments reported in the journal articles. The statistical evaluations presented in the Picek, et al. study were found to be lacking in several respects. Since the participants read two stories consecutively and not one story in isolation, the authors should have included an analysis

of the various interactions. The analysis of the interactions in this paper suggested that the possibility that subjects encoded the social information in the balanced structure as a rule and not through rote learning as was probably the case for the nonbalanceable story. A second problem with the study was the finding that one of the four experimental groups exhibited a level of performance that cannot be explained by stimulus variables. There is the possibility that this anomaly was due to the fact that subjects were not randomly assigned to groups

Finally, Picek, Sherman, and Shiffrin (1975) reported the statistical evaluations of their data using the individual sentiments, and an analysis of only some of the triads. The theoretically appropriate unit of analysis is the entire four-person structure itself and not the triads or the individual sentiments since balance is defined at the *structural level*. Information literate students must learn to recognize those instances where the reported statistical analyses are not congruent with the research question.

In the end, the task of the information literate reader is to determine whether or not a particular study adds to his/her knowledge. Considered in its totality, the Picek, Sherman, and Shiffrin (1975) study is clearly an example of what Coleman (1964) had in mind when he suggested that many studies using verbal materials were “pointless” because the authors provided no evidence that the findings could be generalized beyond the stimulus sample and particular boundary conditions actually used.

References

Association of College and Research Libraries (2003). Information literacy. http://www.ala.org/ala/acrl/acrlissues/acrlinfolit/informationliteracy.htm, visited December 9, 2006.

Clark, H. H. (1973). The language-as-a-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335-359.

Coleman, E.B. (1964). Generalizing to a language population. *Psychological Reports*, *14*, 219-226.

Freedman, J, Sears, D., & Carlsmith, J. (1978). *Social psychology*. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

Hastie, R., Park, B., & Weber, R. (1984). Social memory. In R. Weyer, and Skrull, T. (Eds.), *Handbook of Social Cognition*, v.2 (pp. 151-212). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Higgins, E., & Bargh, J. (1987). Social cognition and social perception. *Annual Review of Psychology*, *38*, 369-425.

ISIHighlycited.com (2001). http://www.isihighlycited.com/?AddInterest=1284, visited Feb 28, 2006.

Janicik, G. & Larrick, R. (2005). Social network schemas and the learning of incomplete networks. *Journal of Personality and Social Psychology*, *88*, 348-364.

Markus, H., and Zajonc, R. (1985). The cognitive perspective in social psychology. In G. Lindzey, and Aronson, E. (Eds.), *The Handbook of Social Psychology* vol. 1 (pp. 137-230). New York: Random House.

OASIS (2004). Online advancement of student information skills. J Paul Leonard Library, San Francisco State University. http://oasis.sfsu.edu/index.html, visited December 9, 2006.

Ostrum, T., Skowronski, J., and Nowak, A. (1994). The cognitive foundations of attitudes: It's a wonderful construct. In P. Devine, Hamilton, D., and Ostrom, T. (Eds.), *Social Cognition: Impact on Social Psychology* (pp. 195-258). San Diego, Ca: Academic Press.

Picek, J.S., Sherman, S.J., & Shiffrin, R.M. (1975) Cognitive organization and coding of social structures. *Journal of Personality and Social Psychology*, 31, 758-768.

Schmidt, T. & Hitchon, J. (1999). When advertising and public relations converge: An application of schema theory to the persuasive impact of alignment ads. *Journal of Mass Communication Quarterly*, *76*, 433-455.

Swanson, J. (1999). CSU information competence. http://www.lib.calpoly.edu/infocomp/modules/ , visited December 9, 2006.

von Hecker, U. (2004). Disambiguating a mental model: Influence of social context. *Psychological Record*, *54*, 27-43.

von Hecker, U. (1993). On memory effect of Heiderian balance – a code hypothesis and an inconsistency hypothesis. *Journal of Experimental Social Psychology*, *29*, 358-386.

von Hecker, U., Crockett, W, Hummert, M. & Kemper, S. (1996). Social cliques as mental models. *European Journal of Social Psychology*, *26*, 741-749.

Weyer, R., and Gordon, S. (1984). The cognitive representation of social information. In R. Weyer, and Skrull, T. (Eds.), *Handbook of Social Cognition*, v.2 (pp. 73-150). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Table 1. The evaluation of the Picek, Sherman, and Shiffrin (1975) study using the CSU training module.

RELEVANCE	**Picek, Sherman, and Shiffrin (1975)**
Is the format or medium of the information useful for your assignment?	Yes. Articles take from journals are acceptable for a research paper on balance schemata
Is the information a primary or secondary source?	Primary. These researchers collected the data for themselves. That is, they are not taking the data from another source.
Is the information comprehensive enough for your needs?	The study certainly provides more than a brief summary of schemas for social situations.
Does the information express a particular point of view?	This study is purported to be objective. The parameters and purpose of the study are clearly defined. Furthermore, statistics are applied to objectively arrive at the conclusions.
Is the information directed toward a general or a specialized audience?	The information is directed to that specialized audience interested in social cognition; specifically those investigating schemas for social balance.
TIMELINESS	
When was information created? compiled? published?	1975. Most instructors would probably wish for more current information, but a citation search reveals that less than twenty articles on social balance have been published in the last thirty years.
Is the information regularly updated and how often?	Not applicable to this particular type of information source: experimental investigation.
Is the information still valid for your topic?	Yes. To date, no one has published information questioning this study's validity.
How stable is the information?	Not applicable. This information will remain the same. Alternative conclusions about the results may continuously change.
RELIABILITY	
Is there an author/producer given for the information you retrieved?	Yes. There are three authors. The first author was a doctoral student in Psychology at Indiana University. The second and third authors were faculty members of the IU faculty.
What is the level of expertise if the author/producer on the topic?	<u>James S. Picek</u> • 1976 I.U. dissertation: Effects of reward uncertainty and ability information on attributions of intrinsic motivation

	Steven J. Sherman: • Ph.D. Michigan (1967) • editor of the Journal of Personality and Social Psychology from 1984-1990 • over 300 refereed publications • received a Senior Scientist Award from NIH in 1999 • recognized as one of the most highly cited researchers in the world by the Institute for Scientific Information (2004) Richard Shiffrin: • Ph.D. Stanford (1968) • Society of Experimental Psychologists in 1981 • editor of Journal of Experimental Psychology: Learning, Memory, and Cognition from 1980-1984. • elected to National Academy of Sciences (1995) • elected to American Academy of Arts and Sciences in 1996 • over 100 refereed publications In 2001, he received the prestigious $100,000 award, the David E. Rumelhart Prize, presented by the Glushko-Samuelson Foundation and Cognitive Science Society for outstanding contributions to human cognition.
Where is the information published?	The Journal of Personality and Social Psychology publishes original papers in all areas of personality and social psychology. It emphasizes empirical reports but may include specialized theoretical, methodological, and review papers.
COVERAGE	
Information abridged?	The information is not abridged.
Breadth of coverage?	The authors present different hypotheses and various experimental situations that have led to studies on social schemas.
ACCURACY	
Is the information documented fact, opinion, or propaganda?	The authors' interpretations of the results are grounded in the statistical analyses of their data.
References to verify accuracy of the information?	The authors use other studies on balance schemata to ground their arguments. Furthermore, the article was refereed before it was published.
What kind of language is used?	The language is systematic and not emotionally charged.

Table 2. Comments on Picek, Sherman, & Shiffrin (1975) in textbooks and articles.

Author(s)	Citations of Picek, Sherman, & Shiffrin (1975)
Freedman, Sears, & Carlsmith (1978)	"Such research has shown these balancing forces affecting a number of cognitive operations. People *learn* balanced situations more readily than they learn imbalanced ones; they *recall* balanced triads more readily than they recall imbalanced ones (Picek, Sherman, and Shiffrin, 1975);… (p. 176)
Hastie, Park, & Weber (1984)	"Recently more sophisticated (and probably more sensitive) research methods have found more evidence for Heider's predictions. For example, Picek, Sherman, and Shiffrin (1975) gave subjects short stories describing social relationships among four people. When subjects attempted to recall the relationships presented in each of the stories, balanced stories were better recalled than imbalanced ones. The authors were even able to localize the difficulty in recall at the third relationship which was the first unbalanced relationship in the unbalanced structure." (p. 184)
Weyer & Gordon (1984)	Nevertheless, two studies provide strong evidence that when balance principles are applicable, they do have schematic properties of the sort implied by implicational molecule theory. In one (Picek, Sherman, and Shiffrin, 1975), subjects read stories describing four different persons' liking for one another. In some instances, not all relations were actually stated in the story….These data not only suggest that a balance principle was used to organize the information in memory, but also indicate that the organization resulting from its application had schematic properties. (pp. 88-89)
Markus and Zajonc (1985)	"This recall of schema-consistent information was superior to their recall of irrelevant or inconsistent information. The inference here is that schema associated with the occupational label influenced what common features of the various segments of the film received from the subject. There are numerous similar studies:…Picek, Sherman, and Shiffrin, 1975…

Higgins & Bargh, 1987	"A nice example of how memory schemata influence the encoding of social information without necessarily producing intrusions in episodic memory is provided by **Picek et al.** (1975)" p. 372.
von Hecker, 1993	"Moreover, **Picek et al.** (1975) as well as Hummert et al. (1990) propose that the storing of situations should proceed by building up code-like representations of their balance state, that is, information about simply whether a given structure is balanced or imbalanced. As these authors demonstrate experimentally, such a code is sequentially constructed in the course of incoming information about a social situation. The code contributes to effective storing, because information can be reduced to simple social models." (p. 360-361)
Devine, Hamilton, & Ostrom (1994)	"One of the first topics of research in social cognition was an examination of whether balanced structures were stored as a schematic structure in long-term memory. Schematic structures have two properties relevant to our present concerns: (1) they are stored as a single unit and (2) they can be used to fill in missing information. Research has generally supported the idea that balanced structures are more schematic than imbalanced structures. Balanced structures are processed more rapidly (Sentis & Burnstein, 1979) and are learned more readily (Picek, Sherman, and Shiffrin, 1975) than unbalanced structures." (p. 235)
von Hecker, Crockett, Hummert, and Kemper (1996)	Balanced patterns, i.e. triads of persons containing exactly zero or two sentiment relations of negative value, were often found to be better remembered than unbalanced patterns, i.e. triads of persons containing one or three negative sentiment relations (… Picek, Sherman, & Shiffrin, 1975; p. 742
Schmidt & Hitchon, (1999)	In general, numerous researchers have found that information congruent with prior expectations or a schema is better remembered. ..James S. Picek, Steven J. Sherman, and Richard M. Shiffrin,

Figure 1. The learning objectives from the CSU Information Competence Project (Swanson, 1999).

The 9 online tutorials listed below provide guidance and practical exercises on Information Competence.

- **Define the Research Topic**
- **Determine the Information Requirements for the Research Question**
- **Locate and Retrieve Relevant Information**
- **Use the Technological Tools For Accessing Information**
- **Evaluate Information**
- **Organize and Synthesize Information**
- **Communicate Using a Variety of Information Technologies**
- **Understand the Ethical, Legal, and Socio-Political Issues Surrounding Information and Information Technology**
- **Use, Evaluate, and Treat Critically Information Received From the Mass Media**

Figure 2. Balanced and Imbalanced Three-Person Situations.

Balanced

Situation 1

Charlie and George like each other. (+)

George and Joe like each other. (+)

Charlie and Joe like each other. (+)

Situation 2

Charlie and George like each other. (+)

George and Joe dislike each other. (-)

Charlie and Joe dislike each other. (-)

Imbalanced

Situation 3

Charlie and George like each other. (+)

George and Joe like each other. (+)

Charlie and Joe dislike each other. (-)

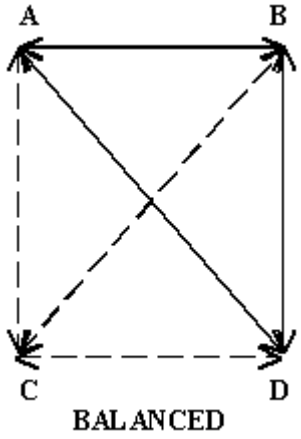
Situation 4

Charlie and George dislike each other. (-)

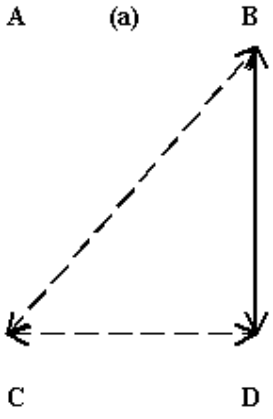
George and Joe dislike each other. (-)

Charlie and Joe dislike each other. (-)

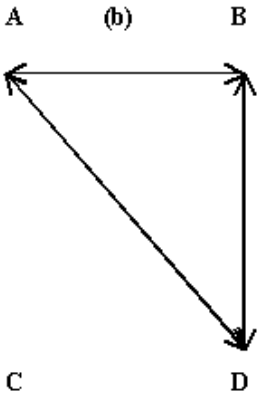
Figure 3. Complete Structures deconstructed into triad components.



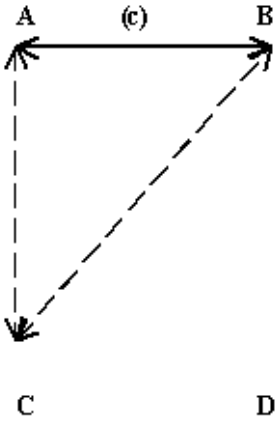
(Complete Structure)



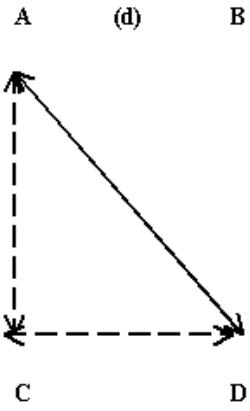
C and D dislike each other.
D and B like each other.
C and B dislike each other.



A and D like each other.
D and B like each other.
A and B like each other.

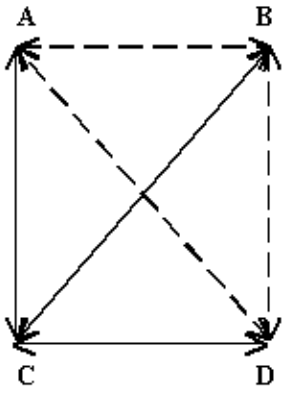


C and B dislike each other.
C and A dislike each other.
A and B like each other.

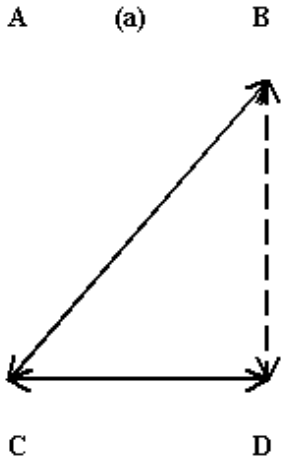


C and D dislike each other.
D and A like each other.
C and A dislike each other.

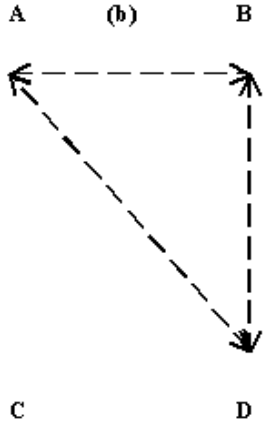
Component Structures
a,b,c,d



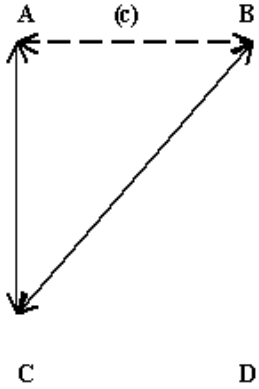
(Complete Structure)



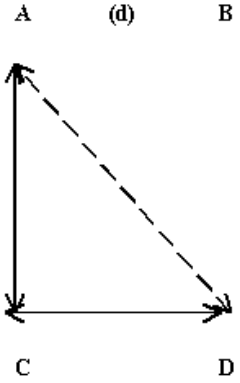
C and D like each other.
D and B dislike each other.
C and B like each other.



A and D dislike each other.
D and B dislike each other.
A and B dislike each other.



C and B like each other.
C and A like each other.
A and B dislike each other.

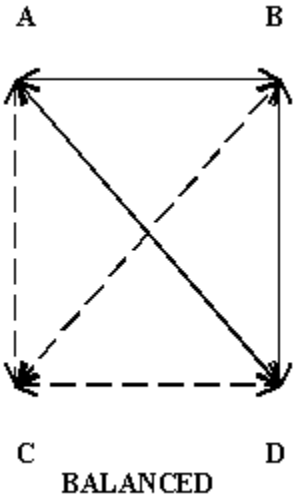


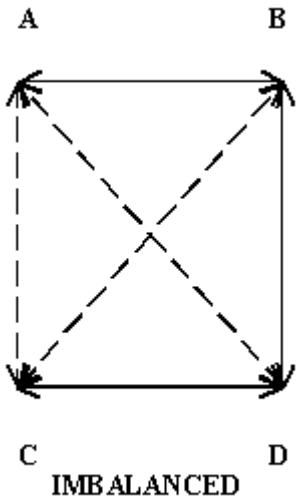
C and D like each other.
D and A dislike each other.
C and A like each other.

Component Structures
a,b,c,d

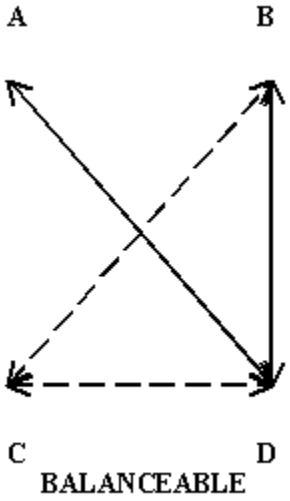
Figure 4. The figures used by Picek, Sherman, and Shiffrin (1975)

COMPLETE FIGURES:





INCOMPLETE FIGURES:



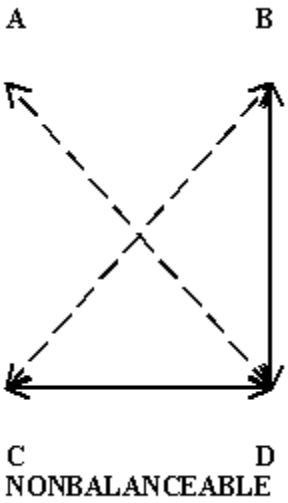
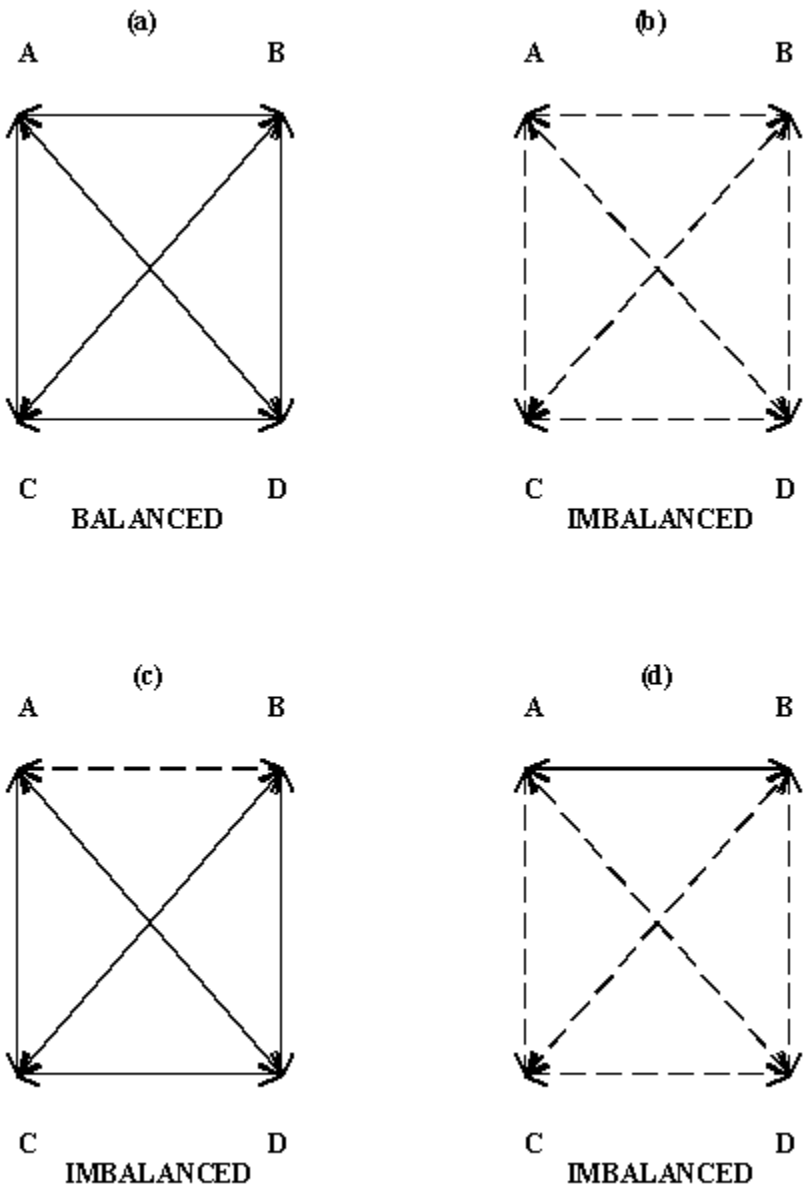
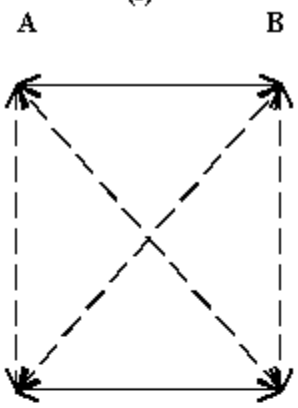


Figure 5. Entire Complete Structure Population.

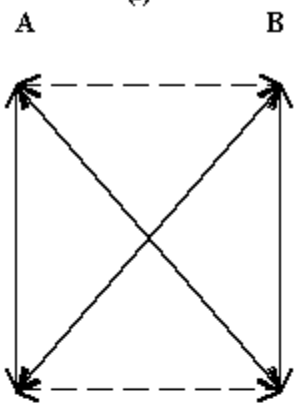


(e)



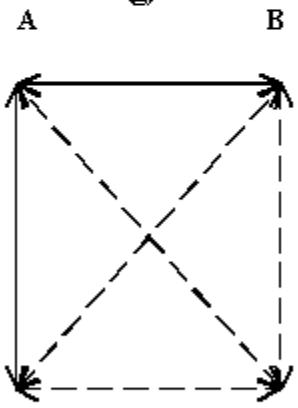
BALANCED

(f)



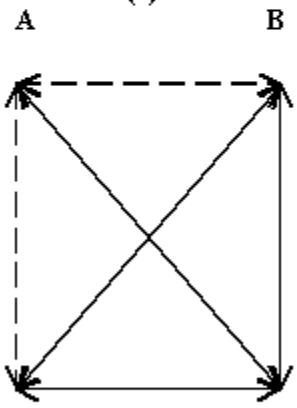
IMBALANCED

(g)



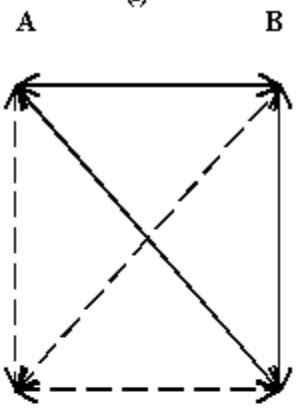
IMBALANCED

(h)



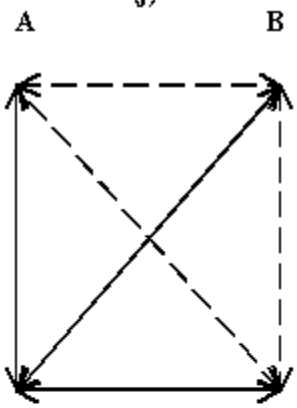
IMBALANCED

(i)



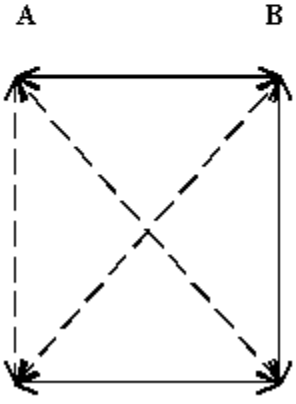
BALANCED

(j)



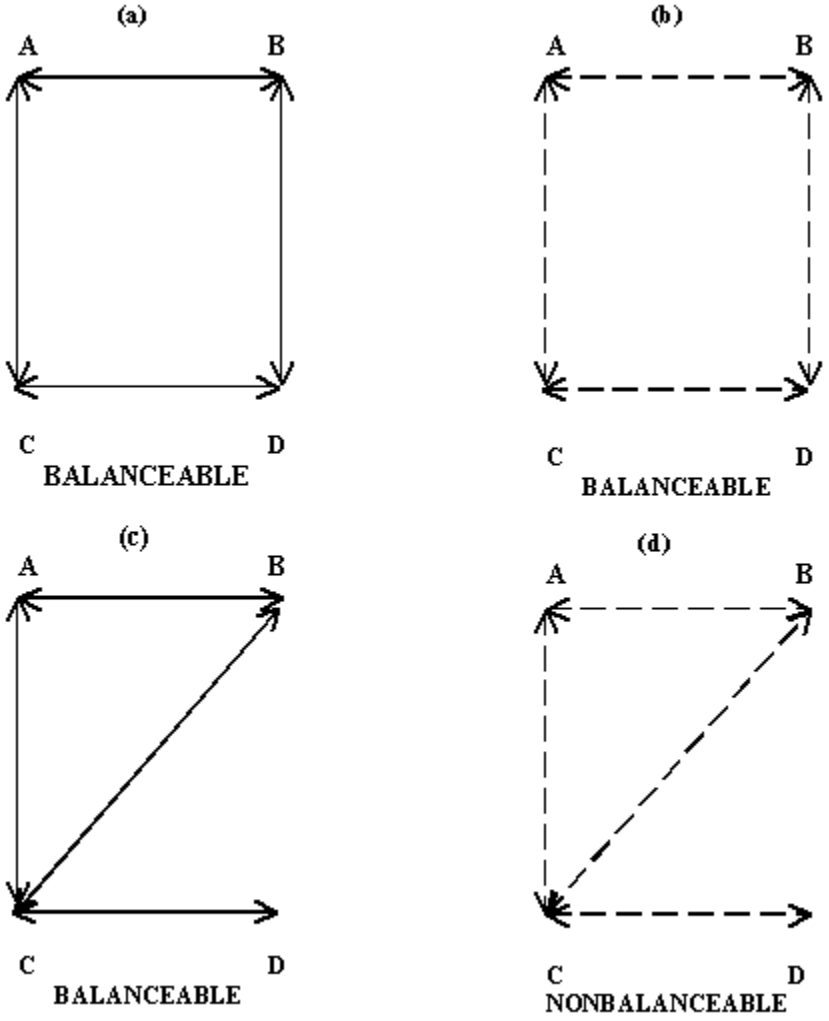
IMBALANCED

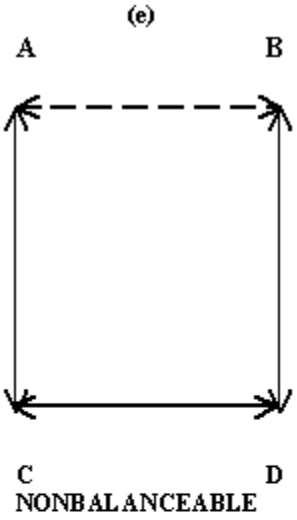
(k)

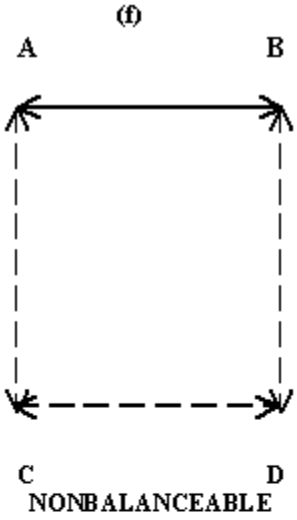


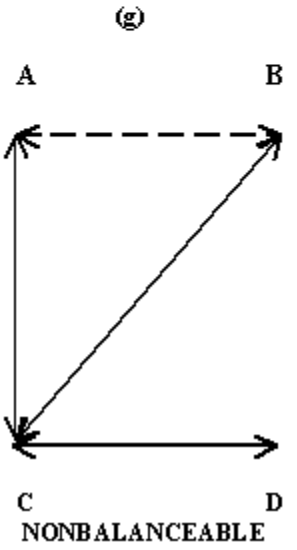
IMBALANCED

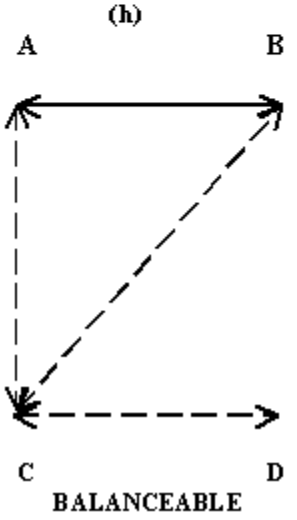
Figure 6. Entire Incomplete Structure Population

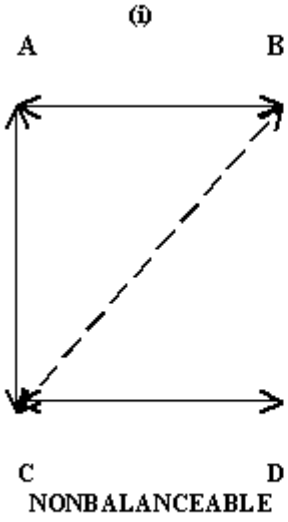


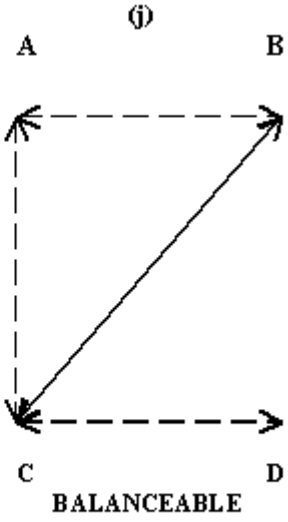

(e)
A
B
C
D
NONBALANCEABLE

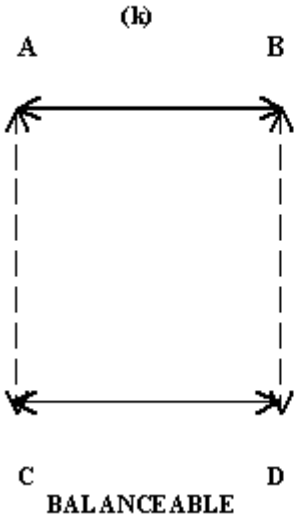

(f)
A
B
C
D
NONBALANCEABLE

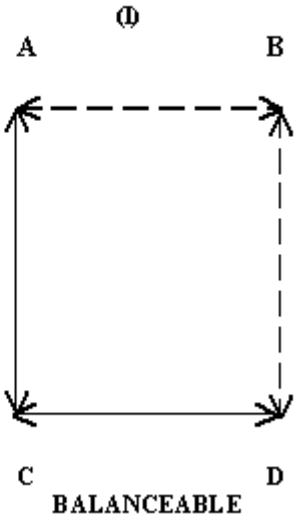

(g)
A
B
C
D
NONBALANCEABLE


(h)
A
B
C
D
BALANCEABLE

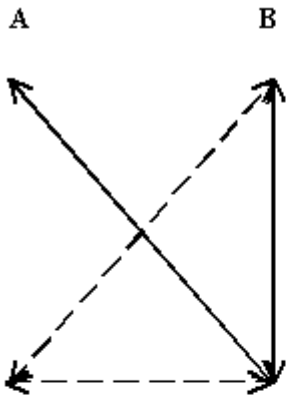

(i)
A
B
C
D
NONBALANCEABLE


(j)
A
B
C
D
BALANCEABLE


(k)
A
B
C
D
BALANCEABLE

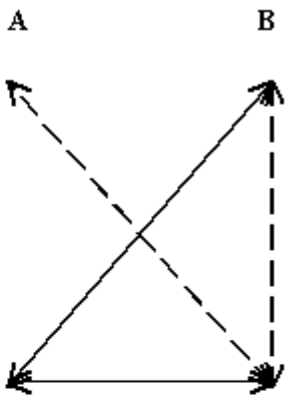

(l)
A
B
C
D
BALANCEABLE

(m)



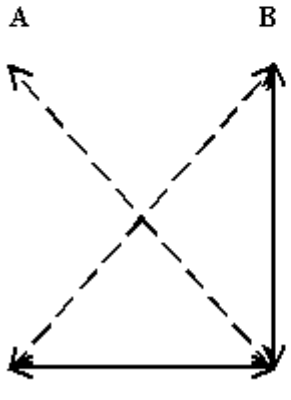
BALANCEABLE

(n)

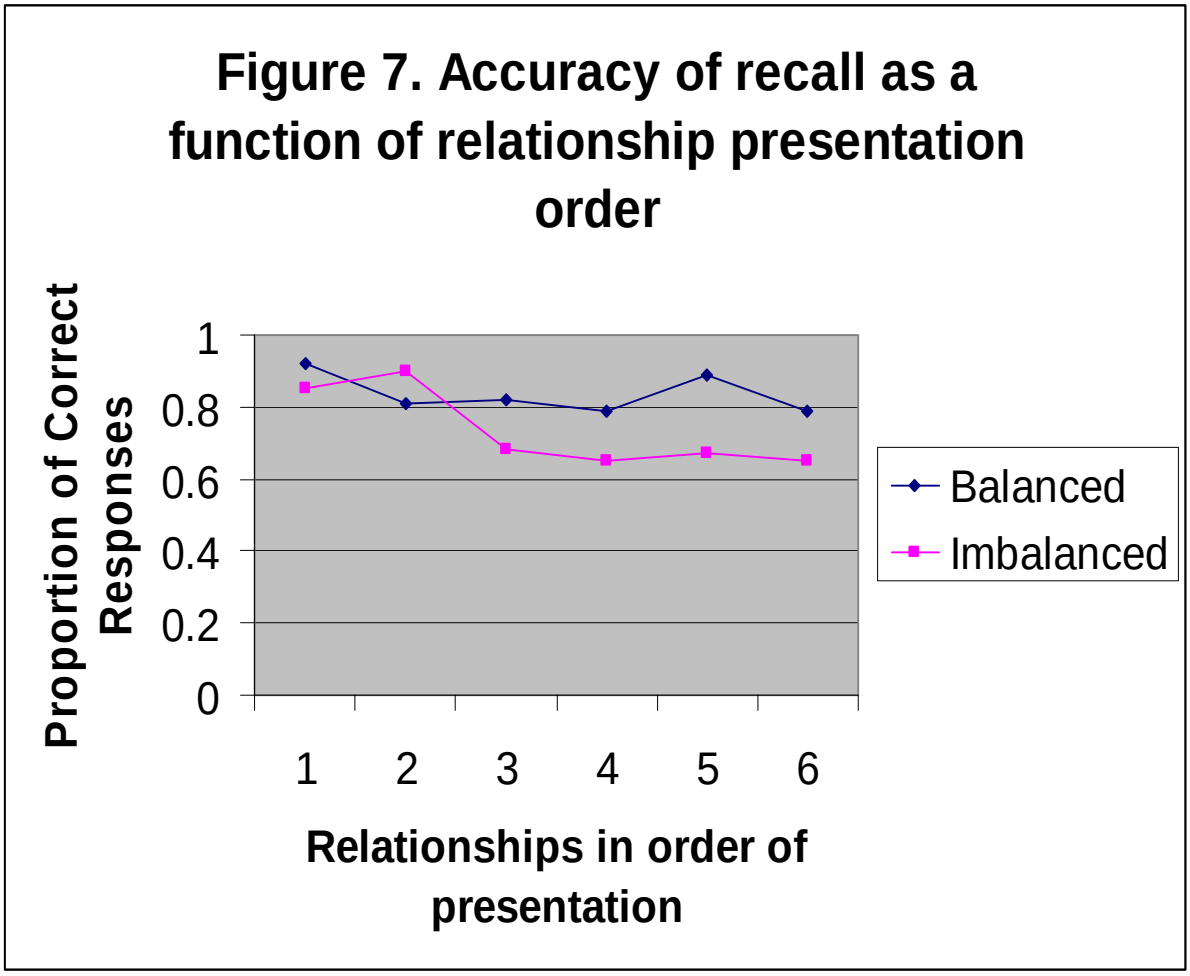


NONBALANCEABLE

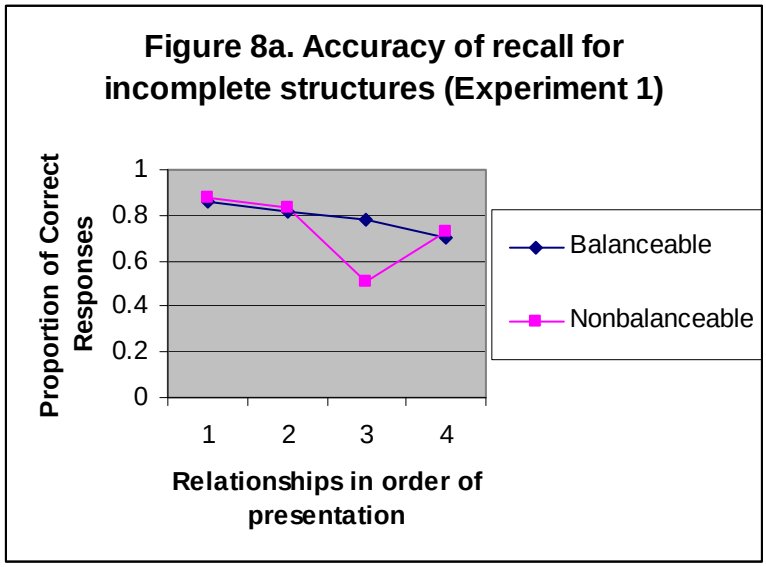
(o)

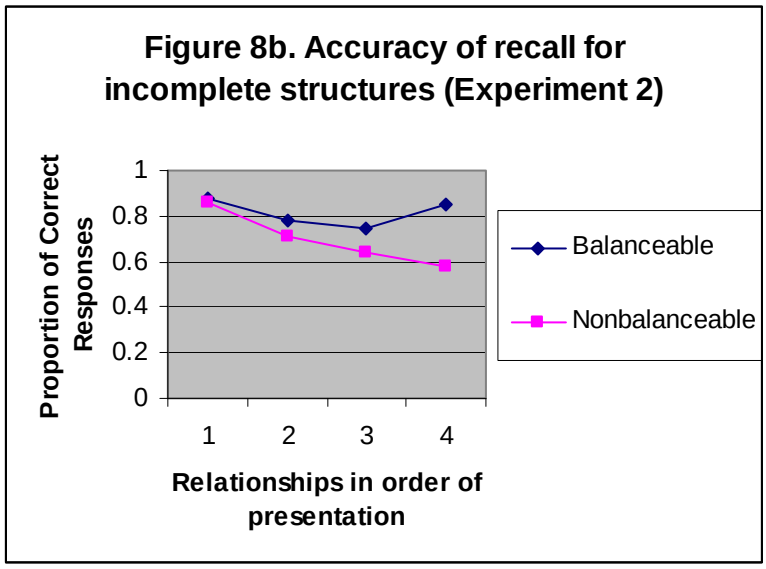


NONBALANCEABLE



After Picek, Sherman, and Shiffrin (1975)





After Picek, Sherman, and Shiffrin (1975)

Figure 9. Mean Proportion Recalled for Complete/Incomplete Structure Interactions

First Story/Second Story	**First Story/SecondStory**
Balanced/Nonbalanceable	**Imbalanced/Balanceable**
.824/.750	**.792/.828**
Non-Balanceable/Balanced	**Balanceable/Imbalanced**
.704/.847	**.630/.680**

Figure 10. Sample Questions for Information Literacy Tutorial Module for Experiments Utilizing Linguistic Stimuli

1. Does the study include a large language sample?
2. If lists are used, are there a lot of lists?
3. If sentences are used, are there many sentences?
4. If stories are used, are there enough used to convince you of their generalizability to other stories?
5. Is every subject given a different list, set of sentences, story?
6. Can you imagine linguistic stimuli that could be used to show the opposite result?

1. Are the psycholinguistic materials matched appropriately?
2. For word frequency?
3. For Imagery and Concreteness?
4. Along another important dimension?

1. Do the investigators test each subject (participant) individually or in groups?
2. Are the subjects (participants) matched for the important factors such as age, level of education, sex, handedness?
3. Do the investigators use different orders of presentation?
4. Do the investigators use various retention intervals?

1. Are the statistical analyses relevant?
2. Does the investigator provide a statistical test for generalizing to the language population in addition to one testing for generalizing to the subject population?
3. The researcher(s) conducted a series of experiments with each subsequent experiment building upon the previous one?
4. The researcher(s) sought to broaden the generalizability of the previous experiment's result to other language samples and boundary conditions, in addition to other subject samples?

1. In your estimation, does the investigator seek to generalize to all the important variables?
2. In your estimation, does the investigator seek to control all variables? .

Appendix A

The Astronaut Story (Balanced)

Four astronauts were launched from Cape Kennedy for a lunar expedition to collect rock specimens and conduct experiments. The astronauts were returning from the mission and in orbit around the earth when the retrorockets and subsidiary engines failed to fire so that they were stranded in this orbit. The situation is such that they have only enough food and oxygen to last eight days in this orbit. Since these are the last eight days for the astronauts, psychologists are interested in the relationships among them.

One of the astronauts, Captain William Wave is an ex-Navy frogman. The second astronaut, Doctor Elmer Cleft, is a brain surgeon. Professor John Stone, the third astronaut, is a geologist. The fourth astronaut, Colonel Sam Far, is an Air Force test pilot.

Colonel Far is a three-time Frisbee champion and a pioneer theorist in its aeronautics. Doctor Cleft spends his leisure time playing pinball machines and once won a weekend jaunt to Peoria, Illinois, for his runaway victory on "Triple Action." Colonel Far and Doctor Cleft like each other.

Professor Stone was a promising sculptor until struck on the head by a collapsing nude. Amnesia erased six years of his fine arts schooling, but on his return to consciousness, he immediately published his highly acclaimed dissertation on the porosity of Camera marble. Colonel Far began his aviation career barnstorming in a rebuilt Stuka. This stint ended when he crashed into a barn. Professor Stone and Colonel Far dislike each other.

Captain Wave was recently hospitalized following a hasty ascent from a dangerous assignment off the Mercedes Islands. Wave's Case was published in Lancet as "The

Mercedes Bends." Doctor Cleft served for three years at Fort Ord as chief lobotomist for the Officer Candidate School. Captain Wave and Doctor Cleft like each other.

Doctor Cleft is, originally from Seattle, Washington, where he was employed as an apprentice apple picker before entering medical school. Professor Stone divides his time between geology and a little sheep farm in Montana. Since he had finished his spring fleecing, Professor Stone agreed to go on the moon mission. Doctor Cleft and Professor Stone dislike each other.

Colonel Far -has amassed a fortune in royalties since his offer from the Ephemera Toy Company to autograph their 10¢ balsa wood planes. Captain Wave developed several new techniques of underwater photography, most notably the deployment of electric eels for flash effects. Colonel Far and Captain Wave like each other.

Professor Stone lost his left thumb and pinky while attempting to gather specimens of molten lava. Captain Wave resigned his frogman commission to be an astronaut after nearly being harpooned by a Japanese whaling vessel. Professor Stone and Captain Wave dislike each other.

On the eighth day, through ingenuity, more than a little luck, and instructions from NASA, the four astronauts finally got the, retrorockets to work but on the way down they crashed into the recovery helicopter and were all killed.

The Astronaut Story (Imbalanced)

Four astronauts were launched from Cape Kennedy for a lunar expedition to collect rock specimens and conduct experiments. The astronauts were returning from the mission and in orbit around the earth when the retrorockets and subsidiary engines failed to fire so that they were stranded in this orbit. The situation is such that they have only enough food and oxygen to last eight days in this orbit. Since these are the last eight days for the astronauts, psychologists are interested in the relationships among them.

One of the astronauts, Captain William Wave is an ex-Navy frogman. The second astronaut, Doctor Elmer Cleft, is a brain surgeon. Professor John Stone, the third astronaut, is a geologist. The fourth astronaut, Colonel Sam Far, is an Air Force test pilot.

Colonel Far is a three-time Frisbee champion and a pioneer theorist in its aeronautics. Doctor Cleft spends his leisure time playing pinball machines and once won a weekend jaunt to Peoria, Illinois, for his runaway victory on "Triple Action." Colonel Far and Doctor Cleft like each other.

Professor Stone was a promising sculptor until struck on the head by a collapsing nude. Amnesia erased six years of his fine arts schooling, but on his return to consciousness, he immediately published his highly acclaimed dissertation on the porosity of Camera marble. Colonel Far began his aviation career barnstorming in a rebuilt Stuka. This stint ended when he crashed into a barn. Professor Stone and Colonel Far like each other.

Captain Wave was recently hospitalized following a hasty ascent from a dangerous assignment off the Mercedes Islands. Wave's Case was published in Lancet as "The Mercedes Bends." Doctor Cleft served for three years at Fort Ord as chief lobotomist for the Officer Candidate School. Captain Wave and Doctor Cleft like each other.

Doctor Cleft is, originally from Seattle, Washington, where he was employed as an apprentice apple picker before entering medical school. Professor Stone divides his time between geology and a little sheep farm in Montana. Since he had finished his spring fleecing, Professor Stone agreed to go on the moon mission. Doctor Cleft and Professor Stone dislike each other.

Colonel Far has amassed a fortune in royalties since his offer from the Ephemera Toy Company to autograph their 10¢ balsa wood planes. Captain Wave developed several new techniques of underwater photography, most notably the deployment of electric eels for flash effects. Colonel Far and Captain Wave dislike each other.

Professor Stone lost his left thumb and pinky while attempting to gather specimens of molten lava. Captain Wave resigned his frogman commission to be an astronaut after nearly being harpooned by a Japanese whaling vessel. Professor Stone and Captain Wave dislike each other.

On the eighth day, through ingenuity, more than a little luck, and instructions from NASA, the four astronauts finally got the, retrorockets to work but on the way down they crashed into the recovery helicopter and were all killed.