INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

IMPROVING MANAGERIAL DIAGNOSIS OF PEOPLE PROBLEMS

DISSERTATION

**Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University**

By

Robert C. Litchfield, M.A., M.B.A.

*** * * * ***

**The Ohio State University
2002**

Dissertation Committee:

Professor Robert Billings, Adviser

Professor John Wanous

Professor Hal Arkes

Approved by

**Adviser
Psychology Department**

UMI Number: 3049075

UMI[®]

UMI Microform 3049075

**Copyright 2002 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.**

**ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346**

*Copyright by
Robert C. Litchfield
2002*

ABSTRACT

Managers are often called upon to deal with "people problems," in which some unproductive behavior has either taken place or is on-going. Before deciding what to do, a manager must make a diagnosis, judging the probable cause or causes of the problem. Unfortunately, observers' judgments are known to suffer from dispositional biases. Reducing dispositionalism in explanations of unproductive behavior in organizations may be desirable for at least two reasons: a) the consequences to employees who are dispositionally linked to unproductive behavior are often substantial, and b) a mis-diagnosis could create these consequences without solving the organization's problem. Of course, it is possible for dispositional explanations to be correct. Thus, any intervention to reduce dispositionalism runs the risk of leading managers away from a "true" person-based cause. A new technique for structuring diagnosis of complex people problems is introduced and tested in three laboratory experiments. A total of 332 upper-level undergraduates in industrial/organizational psychology and organizational behavior courses participated. Most participants (74%) were currently employed. The results indicate that this technique can reduce dispositionalism in initial explanations of a complex people problem, while preserving the possibility of an ultimate attribution to person-based causes. Effects of primed inferential goals on initial explanations were also found. The structured diagnosis technique increased both the desire to collect additional information before finalizing a

judgment and the breadth of the subsequent information search. Anticipated effects of structured diagnosis on confidence were not found. Further, a separate manipulation introduced to reduce confidence in one study unexpectedly failed. This research concludes that the benefits of structured diagnosis can be achieved in a relatively parsimonious format that could be easily learned and adopted by managers.

ACKNOWLEDGMENTS

I thank Aaron Bartlett, Steve Jeong, Jessica Powers, Vandana Vaidyanathan, and Rachael Woldoff for serving as judges in these studies.

I thank my committee members for their feedback throughout this process, and for providing access to research participants.

I also acknowledge the comments of three anonymous reviewers on a manuscript version of Study 1.

Finally, I thank the 332 undergraduate students of The Ohio State University who provided data for this project.

VITA

June 8, 1967.....Born - Columbus, Ohio

1989.....B.A., Economics and Business, Otterbein College

1989-1991.....Assistant Manager, Lippert, Inc.
Columbus, OH

1993..... M.B.A., The Ohio State University

1994-1996..... Manager, Greater Cincinnati Chamber of
Commerce
Cincinnati, OH

1996-present..... Graduate Teaching Associate, The Ohio State
University

FIELDS OF STUDY

Major Field: Psychology

TABLE OF CONTENTS

Abstract.....	<u>Page</u> ii
Acknowledgments.....	iv
Vita.....	v
List of Tables.....	vii
List of Figures.....	ix
Chapters:	
1. Introduction.....	1
2. Study 1 method.....	26
3. Study 1 results.....	32
4. Study 2 method and results.....	44
5. Study 3 method and results.....	59
6. Discussion.....	70
Appendices:	
Appendix A: Trouble at Big Bite.....	78
Appendix B: Solicitation scripts.....	79
Appendix C: RSVP materials.....	82
Appendix D: Judge training items.....	83
Appendix E: Debriefing.....	85
Appendix F: Responses of non-explainers.....	86
References.....	88

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1 Situational explanations by condition.....	33
2 Dispositional explanations by condition.....	33
3 Means, SD's, and correlations between willingness to act and its predictors	34
4 Results of ANOVA of RSVP and confidence on willingness to act....	34
5 Information search on views by condition.....	35
6 Information search on replications by condition.....	36
7 Situational explanations by inferential goal.....	37
8 Dispositional explanations by inferential goal.....	38
9 Situational explanations by condition.....	48
10 Dispositional explanations by condition.....	48
11 Situational explanations by condition (No explanation coded as missing)	49
12 Dispositional explanations by condition (No explanation coded as missing)	49
13 No explanations by condition.....	50
14 ANOVA of manipulations on confidence.....	52
15 Means, SD's, and correlations between willingness to act and its predictors	53
16 ANOVA of manipulations and confidence on three-item willingness to act	54
17 Alternative development and RSVP as predictors of confidence	57
18 Situational explanations by condition.....	63

<u>Table</u>	<u>Page</u>
19 Dispositional explanations by condition.....	63
20 Effects of structured diagnosis on deferring explanation.....	64
21 Means, SD's, and correlations between willingness to act and its predictors	65
22 Effects of structured diagnosis and confidence on willingness to act	66
23 Dispositional explanations by dispositional goal	67
24 Situational explanations by situational goal	67
25 Questions about views by study.....	72
26 Comparison of effects and power for the RSVP-willingness to act relationship	73

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1 Desired effects of structured diagnosis.....	4
2 Expected interaction of situational prime and structured diagnosis	24
3 Expected interaction of dispositional prime and structured diagnosis	24
4 Prediction of Hypothesis 8a.....	25
5 Expected interaction between confidence manipulation and structured diagnosis	43

CHAPTER 1

INTRODUCTION

Managers are often called upon to deal with "people problems," in which some unproductive behavior has either taken place or is on-going. Before deciding what to do, a manager must make a diagnosis of the situation, judging the probable cause or causes of the problem (Einhorn & Hogarth, 1986). Unfortunately, the simplification processes used to make judgments often systematically fail to consider relevant information, leading to biased assessments (Tversky & Kahneman, 1974). When a problem has been defined as such by the presence of unproductive behavior, a particular risk, at least in Western cultures, may be the well known tendency to prefer person-based rather than situation-based explanations for the behavior of others. This tendency, known in various forms as the fundamental attribution error, observer bias, the correspondence bias, and the overattribution bias (for a review, see Gilbert & Malone, 1995), may lead to assessments of causality which incorrectly focus on dispositional factors when situational factors may be primarily responsible. In addition, although this appears not to have been conclusively proven per se, many researchers have proposed theories and interpreted results as suggesting that individuals may tend to stop searching for causes after generating a single alternative that appears reasonable (e.g., Heath, Larrick, & Klayman, 1998; Klein, 1993; Pennington & Hastie, 1993; Shaklee & Fischhoff, 1982). As a result, initial assessments that explain a problem in terms of dispositional factors may not be corrected to take situational information into account (Gilbert, Pelham, & Krull, 1988).

Reducing dispositionalism in explanations of unproductive behavior in organizations may be desirable for at least two reasons: a) the consequences to employees who are dispositionally linked to unproductive behavior are often substantial (Mitchell & Wood, 1980), and b) a mis-diagnosis could create these consequences without solving the organization's problem. However, some difficulty exists because it is, of course, possible for dispositional explanations to be correct (Gilbert & Malone, 1995). Thus, any intervention to reduce dispositionalism runs the risk of leading managers away from a "true" person-based cause.

One potential solution to this dilemma would be to increase situational awareness enough to persuade managers to defer judgment and engage in more information search. Such a strategy would preserve the possibility of an ultimate attribution to person-based causes, while providing time for the correction processes of social inference to be enacted (c.f., Gilbert et al., 1988; Krull, 1993). Since information search that is exclusively in the area of the initial diagnosis would be unlikely to uncover additional causes, though, it is desirable to structure the subsequent search for information to cover both situation and person factors.

This dissertation reports on the development of an intervention which may have value for structuring the diagnosis of people problems, and presents results from three experimental studies designed to examine its potential. The first study establishes the basic utility of the intervention. The second and third studies evaluate boundary conditions and possible modifications. The overall goal of this research is to gain more insight into the process and effects of structuring diagnosis of complex people problems. At a practical level, I also hope to arrive at a tool which is complex enough to have utility, but simple enough that it could realistically be adopted by managers.

The remainder of this chapter is organized into three main sections. In the first, I detail the theoretical development of the primary intervention. In the second, I discuss an

alternative that may be expected to achieve some of the same effects, but not others. The third section is devoted to the presentation of hypotheses. Additional hypotheses are presented in subsequent chapters, as warranted by results of earlier studies.

RSVP: Developing a Tool for Structured Diagnosis

Figure 1 illustrates the relationships described in the introduction. A desirable tool for structured diagnosis should, according to this argument, increase situational explanations and decrease dispositional explanations. In addition, such an intervention should decrease willingness to act without further information, and it should also increase the breadth of the subsequent information search. Below, I detail the development of an intervention which may be expected to have these properties. This tool is designed to take into account the number of replications of behavior, the number of settings in which it has occurred, the number of views of the problem available, and the people who are or would be expected to be involved. These dimensions (replications, settings, views, people) are captured in the acronym RSVP. The "replications," "settings," and "people" components can be theoretically derived from Kelley's (1967) dimensions of attributions. The roots of the "views" component are found in both advice to researchers (e.g., Jick, 1979) and research on scenario generation (e.g., Gregory et al., 1982). In order to understand the development of the RSVP intervention, it is necessary to examine research on dispositionalism in more detail. After a brief review of two important causes of dispositionalism, I explain how these inform the logic behind the intervention components.

Figure 1: Desired Effects Of Structured Diagnosis

Dispositionalism

The tendency to identify behaviors with characteristics of individuals, called "dispositionalism," has been a major research area in psychology for decades (e.g., Heider, 1958; Jones, 1979; Krull, 2001). Although this tendency can be functional, it has been widely accepted that the failure to appreciate the power of situations has led to many tragic errors in judgment and decision making (Gilbert & Malone, 1995). Some details of the social inference process of behavior explanation continue to be disputed (see Trope & Gaunt, 1999; 2000), but two factors are acknowledged across descriptions as important contributors to dispositionalism: a) lack of awareness of situational constraints, and b) cognitive load.

A major cause of dispositionalism is thought to be a lack of awareness of situational constraints, which can take one of two forms (Gilbert & Malone, 1995). One problem of

awareness is the lack of visibility of situations relative to people and behaviors (Heider, 1958). Because people and behaviors are easily perceived, while situations may be tough to define, individuals are more likely to develop dispositional explanations for behavior. Thus, the fact that employees are stealing food from a restaurant may be easily captured by security cameras, but situational factors that may explain the problem, such as long shifts and low wages, are often not as visible. Heider (1958) called this property of visibility "salience." However, a complicated web of interpretations about the meaning of the term "salience" has developed in the attribution literature (for a discussion, see Gilbert & Malone, 1995). I have used the term "visibility" to distinguish what appears to be the simplest, and, according to Gilbert and Malone (1995), most appropriate use of "salience:" The degree to which something (e.g., a behavior, person, or situation) is easy to see.

A second form of lack of awareness goes beyond visibility to explore how actors and observers may draw different conclusions, even when they both appear able to see the same situation. When people act in the world, they may experience explicit behavior constraints due to task and situational factors. For example, when designing a set of questions for a test, a test writer generally seeks to examine knowledge within some pre-defined domain. In most such cases, the domain of the test prevents the asking of certain types of questions (e.g., questions about poetry in a test on biology). Thus, the task of designing a "fair" test presents explicit behavior constraints on the test writer (Ross, Amabile, & Steinmetz, 1977). However, people may also experience psychological constraints, which "do not change an actor's behavioral options so much as they change her or his understanding of those options" (Gilbert & Malone, 1995, p. 26). As a result, the situation may exert an influence on behavior that is difficult to observe from an external perspective, even if the basic environment of the decision is recognized. This suggests that, though visibility and perspective are certainly related, it may be important to distinguish between them.

One well-established example of the effects of perspective can be found in investigations of the hindsight bias. Studies of hindsight bias typically find that individuals who have information about the outcome of an event (i.e., hindsight) overestimate their ability to have predicted its occurrence (Fischhoff, 1975). Thus, we see perspective (hindsight vs. foresight) affecting judgment. This type of perspective effect has been understood to be a problem in domains as diverse as medical diagnosis (Arkes, Guilmette, Faust, and Hart, 1988), management decision making (Schwenk, 1988), and causal analysis of accidents (Reason, 1990).

The effects of perspective do not appear to be limited to temporal gaps. Attribution researchers have identified so-called "actor-observer" effects (Jones & Nisbett, 1971), which suggest that individuals engaged in behavior make different attributions than those who observe it. A variety of investigations (see, e.g., the meta-analysis and review by Burger, 1981) have also found that observers who share the same perspective as actors (such as other individuals employed in the same job) make different attributions for poor performance (more "situational") than those who observe the same situation from a dissimilar perspective (e.g., supervisors). For example, Hofmann and Stetzer (1998), in a study of supervisor and employee attributions for accidents in a large utility company, found that supervisors tended to use more dispositional explanations than employees. Hofmann and Stetzer (1998) reasoned, consistent with prior theory (e.g., Green & Mitchell, 1979), that employees shared a more similar view of the problem with the hypothetical employee involved in the accident scenario used in the research than did supervisors, even though the supervisors were also familiar with the type of situation represented in the problem. Overall, this research suggests that the effects of perspective may be a significant cause of dispositionalism.

As researchers have learned more about the underlying inference processes used to interpret social information, another major cause of dispositionalism has emerged. This

cause is related to the cognitive load individuals experience while attempting to make an explanation. Gilbert and his colleagues (e.g., Gilbert et al., 1988; Gilbert, Jones, & Pelham, 1987) have argued for a distinction between "active" and "passive" perceivers. While passive perceivers are merely listening or observing with no requirement or intent to respond (e.g., watching television), active perceivers are expecting to respond in some way (e.g., people in conversations). Because of this requirement for a response, active perceivers are thought to experience higher demands on limited cognitive resources, as they attempt to think through the information they receive, structure it to make sense, and formulate possible responses (Gilbert et al., 1988).

In the model of social inference proposed by Gilbert et al. (1988), individuals first make a dispositional inference, then correct it for situational information. When a complex problem or multiple tasks exhaust cognitive resources, individuals may stop before completing the correction, resulting in an incomplete consideration of information (Gilbert et al., 1988; Krull, 1993). The result is that, in the absence of the information processing capacity to complete the correction of the explanation, the default tendency toward dispositionalism dominates. Interestingly, Gilbert et al. (1988) found that this effect did not mean that individuals did not gather situational information. In their first experiment, subjects in the cognitive load condition were required to memorize a list of situational variables relevant to the judgment task. These individuals displayed better memory for the situational variables than did subjects who were not asked to memorize them, but they were unable to integrate this knowledge into their judgments (Gilbert et al., 1988).

In summarizing the above, it is useful to note that, though there are two major factors that are universally related to dispositionalism, the first, awareness of situational constraints, has two variants (visibility and perspective). This suggests that, in order to reduce dispositionalism, it is necessary to pay attention to both the way in which individuals perceive behavior in context and the perspective from which they do it. In addition, the

findings on cognitive load suggest that the creation of balanced inferences will require a structure to spread out, channel, or otherwise ease the thinking activities individuals must undertake. Together, these insights suggest that by improving visibility, adding perspectives, and creating a structure to channel cognition, it may be possible to improve diagnosis.

Increasing Visibility: Replications, Settings, and People

Kelley (1967) proposed that individuals use a covariation-based system for explaining whether behavior they observe should be attributed to the environment or to the person engaged in the behavior. He argued that people make this judgment by looking at variation in the behavior across persons (How do others behave?), entities (Does the behavior occur with respect to other objects?), time (Does the behavior occur repeatedly?), and modality (Does the behavior seem the same through different types of observations?). Kelley (1967) represented the attributions that people make about cause and effect covariation in the form of a cube, with sides depicting persons, entities, and time/modality (these two were compressed to a single dimension).

Kelley's (1967) definitions of these dimensions of judgment are somewhat vague, and the examples he gives are often very simplistic. While the definition of persons is relatively straightforward, it can be tough to assess what he means by the labels "entity," "time," and "modality." I'll return to this difficulty shortly, and I take up the "modality" dimension in more detail in the next section. As an example of how the person, time and entity characteristics might be used, consider a situation where a supervisor has observed poor performance by an employee. Perhaps this employee has failed to convey important information in telephone calls to customers. The supervisor might examine how other employees have done at conveying information. This corresponds to Kelley's person factor. She also might consider how often such problems have occurred, which would be an examination of time factors in Kelley's system. Finally, the supervisor could look at

whether the individual fails to convey information in phone calls to others, such as managers or friends. This would give information about whether the behavior is occurring across different entities.

Kelley (1967) continued this argument by noting that individuals frequently wish to assess the validity of the impression they have drawn. He suggested four principles of external validity in impression assessment. These were defined as follows:

1. Distinctiveness: the impression is attributed to the thing if it uniquely occurs when the thing is present and does not occur when it is absent.
2. Consistency over time: each time the thing is present, the individual's reaction must be the same or nearly so.
3. Consistency over modality: his reaction must be consistent even though his mode of interaction with the thing varies. (For example, he sees it to have an irregular outline and he feels it to be rough; or first he estimates the answer to the problem and then he calculates it).
4. Consensus: attributes of external origin are experienced the same way by all observers. (Kelley, 1967, p. 197).

Kelley (1967) also represented these concepts as three dimensions of a cube, with the consistency issues (#2 and #3) compressed to a single dimension. Thus, with the order of presentation rearranged to be consistent with the earlier person-entity-time/modality discussion, consensus corresponds to the "persons" dimension, distinctiveness to the "entities" dimension, and consistency to the "time/modality" dimension. To continue with the previous example, the supervisor would assess consensus by looking at the degree to which others also fail to convey information in phone calls to customers, with high consensus indicating that many others also have this problem. To assess consistency over time, the supervisor would look to see how often the behavior had occurred. High consistency would indicate that this behavior always or usually happened in phone calls to

customers. To assess distinctiveness, the supervisor could assess whether this behavior occurred on other tasks, such as phone calls to managers or friends. If the behavior had occurred across these diverse tasks, then it would be said to be low in distinctiveness. Kelley (1967) suggested this system as the basis for understanding attributions about a wide variety of human attitudes and behaviors.

Despite its well-deserved fame, “Kelley’s cube” is not immune to critique. Three issues are relevant to the present discussion. First, Kelley (1967) made no attempt to label the dimensions in a manner that would facilitate memory. This is understandable, since he was trying to explain attribution processes, and was not attempting to provide a debiasing tool. I will return to this issue below.

Second, Kelley’s (1967) discussion frequently makes it difficult to know how to represent a behavior across multiple situations. Part of this confusion undoubtedly stems from Kelley’s (1967) desire to squeeze a tremendous amount of information into three dimensions. However, Kelley also wavers in his discussion of what each of the dimensions represents. The careful reader will notice that the wording of Kelley’s (1967) definition of consensus, reprinted above, suggests something a bit different than the application I explained. This confusion may be minor, since Kelley (1967) applies the definition in his writings in the way that I have used it, and this is its commonly accepted meaning.

Dealing with the consistency (time) and distinctiveness (entity) dimensions is trickier. Kelley’s (1967) comments become particularly confusing when one attempts to understand how to label subtle changes in the situational factors surrounding behavior. Although changes in the situation have commonly been accepted as part of the entity/distinctiveness dimension (e.g., Robbins, 2001; Steers, 1991), it is not at all clear from Kelley’s (1967) writing that this is what he intended. In his initial definition, Kelley (1967) clearly labels one dimension “consistency over time” (p. 197). Later, though, he suggests that the “time categories” may “refer to different situations” (Kelley, 1967, pp. 202-203).

He also repeatedly labels one side of the cube “time (situation)(target)/modality” (e.g., Kelley, 1967, Figures 2 and 3, pp. 203, 210). Such statements may require too much information to be represented in the consistency dimension. Specifically, this representation does not make it clear how to separate information about the temporal characteristics of a behavior from other information about the situation (e.g., what other contextual features were present).

One possibility is that he intended to put less dramatic differences in situation into the consistency/time dimension, and more dramatic differences into the entity/distinctiveness dimension. In the context of the example I have been using, we might think of phone calls to friends and managers to be different entities (or, as they are often called, tasks), but consider information about more subtle situational characteristics, such as whether the calls were placed from the office versus from the car, to be part of the time dimension. Such a division would quickly create problems, though, since it is impossible to know where to draw distinctions between subtle and large situational differences. Kelley (1967) does not really provide enough detail to understand what he might have had in mind. Kelley (1973) eventually did clean up the exposition of the cube, dropping modality and simply labeling one axis “time,” but he still seems prone to consider situational features to be elements of time. For example, in a discussion of the reaction of a person, Paul, to a given painting, he writes:

“Was it something about Paul (person), something about the painting (entity), something about the particular circumstances (time), or some combination of these factors?” (Kelley, 1973, p. 110).

A third problem with Kelley’s (1967) framework is that it does not take levels of analysis into account. For example, other people (the consensus dimension) may be nested within groups, such that different clusters of people behave differently. Similarly, situations may be regarded as nested within points in time (or visa versa, depending on the research question). In a later paper, Kelley (1973) acknowledged this difficulty. His solution was to

create nested models, using the ANOVA principle of adding factors. This approach could be used to alleviate the problem of excess information in the consistency dimension by representing different situations as levels of a “situation” factor, with time nested within each (or visa versa). While this method is theoretically interesting, it can quickly lead to models that are quite complex.

In my opinion, the key problem with implementing Kelley’s framework lies in these sorts of levels issues. His model will work, but it’s complicated. In order to bypass such complexity, I have reorganized these dimensions in an attempt to provide a simple system that may be useful for the class of problems under investigation here. Kelley (1967) finessed some of the levels problems in the initial cube by representing information about time and situation on a single axis. The result of this is a blurring of these categories of information. My own compromise is to effectively push the blurriness from the time dimension into the entity dimension. I don’t think that this is a theoretically superior solution to Kelley’s (1967), but I think it’s likely to work better for the kind of problems I want to investigate. In addition, my perception is that this is the way the dimensions are most commonly represented (e.g., Robbins, 2001; Steers, 1991).

My system groups all information about time into a dimension I label “replications.” This dimension is equivalent to what Kelley (1967) labeled “consistency over time,” with the situational references removed. Replications are high to the extent that a behavior has occurred often or many times. I group information about the entity and situation into a single dimension, which I call “settings.” Settings are high to the degree that a behavior occurs across situations, which might include tasks, relationships, locations, or other contextual cues. This is equivalent to the distinctiveness dimension in many of the popular interpretations of Kelley’s (1967) framework (e.g., Robbins, 2001; Steers, 1991), but, as discussed above, Kelley’s own writings were considerably less clear. I have also rescaled this dimension to be consistent with the others. When settings are “high,” this

means that the behavior is occurring across multiple situations. In popular interpretations of Kelley's (1967) framework, this corresponds to "low" distinctiveness. This change is particularly important for the dual use of the framework to evaluate information search parameters, as discussed below. I retain Kelley's (1967) consensus concept in its entirety, but, since consensus has other meanings in organizational parlance (e.g., consensus as agreement with a decision), I have relabeled it "people." The people dimension is rated as high to the extent that a behavior is or could be expected to be common to many.

In examining what is going on in people problems in organizations, I propose that managers may often need to draw conclusions from states of affairs that may include "all employees in one task, location, group, or relationship engaged in a behavior" or "a few employees in each of many tasks, locations, groups, or relationships engaged in a behavior" or "lots of employees in many tasks, locations, groups, or relationships engaged in a behavior." By using the labels replications, settings, and people, we can conveniently sort the information on amount, situation, and individuals involved to increase visibility of the available information. As I argued earlier, this ease of use is a key feature of a potential debiasing system.

Perspective-Taking and Triangulation: The "Views" Dimension

Improved situation awareness, as captured by visibility data, may be insufficient to reduce dispositionalism (Gilbert & Malone, 1995). As discussed above, the psychological constraints associated with a given perspective are often invisible to observers (Gilbert & Malone, 1995), resulting in a lack of awareness of situational constraints that is tied to the vantage point of the observer. Attention to the replications, settings, and people involved does nothing to remedy this lack of awareness of perspective, and so may not cause much change in managers' diagnostic patterns. However, as Hofmann and Stetzer's (1998) study illustrates, attention to other views of the problem may be helpful. Attention to the available views of the problem has been represented in judgment and decision making research

through the practices of scenario construction and "perspective-taking" (Galinsky & Moskowitz, 2000). In a series of studies, Gregory et al. (1982) found evidence that imagining oneself in a scenario influenced both probability judgments and subsequent behavior. Koehler (1991) suggested that such effects, known as explanation biases, occur because the initial explanation acts as a focal hypothesis, which frames subsequent information search. This suggests that a successful debiasing strategy would involve cueing individuals to entertain additional hypotheses. Hirt and Markman (1995) found that this type of cueing reduced confidence in the initial hypothesis and stimulated the generation of additional possibilities. Galinsky and Moskowitz (2000) successfully used a similar strategy of asking individuals to take on other perspectives to debias social stereotypes.

Further support for this mechanism can be found in the research principle of triangulation. As summarized by Jick (1979), triangulation is achieved by obtaining multiple views of a given phenomenon by means of multiple measures, samples, etc. In essence, the idea is that the path to gaining more complete knowledge involves consideration from multiple vantage points. By prompting individuals to consider how many "views" they have of the current problem (e.g., managers', employees', co-workers', other stakeholders'), the implication is that, as in research, more views provide a richer diagnosis. Interestingly, this same idea appears to have been recognized to some degree by Kelley (1967). Kelley's (1967) original formulation of attributional dimensions contained four parts, as noted above. Two of these, consistency over time and consistency over modality, were combined on one side of the now-famous cube. However, modality seems to include more than simple information about the setting in some of his descriptions. For example, consider this statement, made in discussion of an example of consistency over modality:

"Does B report the same attribution when he thinks about the matter in different ways or when he approaches it with different observational methods?" (Kelley, 1967, p. 202).

This statement clearly describes the kind of perspective-taking associated with triangulation, but, because of Kelley's (1967) geometric compromise, this insight has become somewhat obscured. The present discussion suggests that this has been an unfortunate occurrence.

Providing Structure: RSVP and the Logic of Debiasing

Fischhoff (1982), in summarizing research on debiasing, emphasized the importance of structures to channel psychological approaches to problems where biases are likely. An obvious question, though, concerns the optimal complexity of these structures. In discussing the application of debiasing techniques to organizational contexts, Heath, Larrick, and Klayman (1998) have argued that simplicity is an important criterion in the development of "cognitive repairs" for biases in judgment. Heath et al. (1998) provide numerous examples of simple, easy-to-remember acronyms or sayings that facilitate use of debiasing in organizations, and they suggest that simple interventions require fewer cognitive and organizational resources to implement. Research on debiasing supports the contention that relatively simple techniques are often successful (for reviews, see Arkes, 1991; Fischhoff, 1982), though there are exceptions (see, e.g., the discussion of specific training as a debiasing technique in Arkes, 1991). Following this logic, I have organized the dimensions of the intervention (replications, settings, views, people) to suggest the acronym RSVP. Thus, when attempting to understand a complex people problem, managers need only remember to consult their "RSVP list." This provides a way to assess the situation, and, as a secondary benefit, it may be more likely to be used than other, more elaborate methods (c.f., the perspective-taking essays written by subjects of Galinsky & Moskowitz, 2000).

An additional benefit of the RSVP framework is that it provides a structure for subsequent information search. Thus, the same framework through which an initial explanation is formulated provides a structure to prompt specific types of information seeking. The structure for search is important, because, without it, individuals may fall prey to confirmation bias, searching only for information that validates their initial hypothesis

(Devine, Hirt, & Gehrke, 1990; Skov & Sherman, 1986). The degree to which this is a potential problem is unclear. Several studies have shown that individuals' primary concern may be to search for information that has high diagnostic value (e.g., Bassok & Trope, 1984; Devine et al., 1990; Skov & Sherman, 1986), but each of these investigations has also found a secondary tendency to search for information that confirms the initial hypothesis. In explaining this, Devine et al. (1990) suggest that the question is not whether individuals prefer to ask diagnostic questions (those that will inform about the truth of hypotheses), but whether they will ask the most diagnostic questions available. Their research suggests that individuals will ask diagnostic questions, but, when faced with a choice between equally diagnostic questions that focus on either the initial hypothesis or the alternative, questions about the initial hypothesis are preferred. Thus, individuals effectively prefer both diagnostic and confirming information. This suggests a role for a structure to channel information search to make sure that a premature focusing does not occur. Given this, tendencies toward confirmation bias, while not eliminated, may be of lesser concern.

Inferential Goals: An Alternative Route to Reduce Dispositionalism?

Given that I have accepted simplicity as an important criterion in the development of a cognitive repair for dispositionalism, an interesting question involves the limits to which this dimension can be pushed. One construct that may provide a test of this boundary is the inferential goal (Krull, 1993). This construct is particularly relevant to the present research because it builds upon the ideas of Gilbert and his colleagues (e.g., Gilbert et al., 1988), which, as described previously, form part of the basis for the development of the RSVP intervention.

A number of researchers have hypothesized that social inference processes follow a sequential logic similar to that found in the anchoring and adjustment heuristic (Tversky & Kahneman, 1974). From this perspective, individuals perceive others' behavior by making an initial explanation, which functions as an anchor. Additional information is then added in

to complete the picture and correct the initial explanation. Quattrone (1982) was, perhaps, the first to develop this insight. In two experiments, Quattrone (1982) showed that a standard paradigm for research on dispositionalism could be modified to cause overattribution to situational causes. By priming subjects to interpret features of the target person's situation, rather than the traditional prompt to infer a person's true attitude (when evaluating an essay that was written either freely or under duress), Quattrone (1982) effectively showed that overattribution to situational causes could be generated.

While others picked up on the adjustment aspects of the anchoring and adjustment paradigm for social inference (e.g., Gilbert et al., 1988), Krull (1993) focused on the situational/dispositional anchors. In his mixed model of social inference, Krull (1993) proposed, in conflict with nearly all of the attribution literature, save Quattrone (1982), that individuals may hold either dispositional or situational inferential goals when approaching interpretation of a behavior. Krull (1993) suggested that a perceiver's inferential goal determined the correction process that would be used. Thus, if the goal was to perceive information about an actor's dispositions, the initial identification would result in a dispositional characterization, which would then be corrected for situational information. On the other hand, if the goal was to perceive information about a situation, then the initial identification of behavior would result in a situational characterization, which would then be corrected for dispositional information. Previously, Gilbert and his colleagues (Gilbert et al. 1988; Gilbert, Krull, & Pelham, 1988) found that the initial characterization process was effortless, but the correction was effortful, and could be disrupted when the perceiver was under cognitive load. Building on this finding, Krull (1993) manipulated perceivers' inferential goal to be either dispositional or situational, and manipulated cognitive load to be either busy or not busy. Consistent with his predictions, Krull (1993) found that inferential goals could be manipulated, and that, in the presence of cognitive load, the adjustment processes could be interrupted, resulting in either dispositionalism or situationalism.

Subsequent research appears to have confirmed that, in Western cultures, default inferential goals are dispositional (Krull, Loy, Lin, Wang, Chen, & Zhao, 1999; Lee, Hallahan, & Herzog, 1996). However, people are thought to be capable of having either goal.

Recently, the anchoring and adjustment paradigm for social inference has come under attack. Trope and Gaunt (1999; 2000) have argued that situational and dispositional information are integrated at a single point in time, rather than being combined through a sequential correction process. Trope and Gaunt (2000) conducted three studies on the question of integration versus correction. In the first, they used a standard research task wherein subjects examine an essay that they were told was written either freely or under duress (the situational constraint manipulation), then attempt to rate the writer's true attitude toward the topic (marijuana legalization, in this case). Additional manipulations were conducted for cognitive load and "salience" of situational information. In the high salience condition, subjects listened to a tape of the experimenter's instructions to the writer. In the low salience condition, these instructions were printed in the experimental materials. Trope and Gaunt (2000) found that individuals given the high salience manipulation tended to rate the writer's attitude as less in favor of legalization in the constrained condition, regardless of cognitive load. In fact, they found that, in the cognitive load condition, individuals rated the writer's attitude as even more situational than in the no-load condition. In this study, Trope and Gaunt (2000) had effectively primed a dispositional inferential goal by asking subjects to rate the attitude of the writer. However, a powerful presentation of situational information appears to have overwhelmed the effects of the initial goal.

In their second experiment, Trope and Gaunt (2000) replicated the first study, but this time operationalized salience as a situational prime. However, in contrast to the direct primes used by Quattrone (1982) and Krull (1993), Trope and Gaunt (2000) used an indirect approach, preceding the presentation of the experimental materials with an assignment to explain a set of four proverbs in their own words and provide an example of

each. In the high salience condition, these proverbs described situational influences on behavior, while in the low salience condition, they were neutral. Consistent with inferential goal research (e.g., Krull, 1993; Quattrone, 1982), they found that situational priming produced situational inferences in the constrained essay condition. However, subjects were still instructed to make a dispositional inference. Thus, Trope and Gaunt (2000) effectively primed both inferential goals in succession. The situational prime was much more elaborate, though, suggesting one reason why it may have "carried through" subsequent instructions to make a dispositional rating.

In their third study, Trope and Gaunt (2000) found that specific situational information also had the effect of overwhelming a dispositional inferential goal. Here, their experimental task involved rating the dispositional strictness of a teaching assistant who was under either specific, general, or no pressure to be strict in grading.

Trope and Gaunt (2000) contend that, in each of their three studies, they made only minor changes to established research protocols (e.g., hearing a stimulus rather than reading), but found substantially different results. Based on the asserted subtlety of their changes, they then conclude that the most reasonable explanation for their findings is that their research provides evidence against an anchoring and adjustment paradigm. While they could be correct, I disagree that their changes were, necessarily, minor. Given the subtlety of nearly all manipulations in this domain of research, to simply assert that your changes were minor seems disingenuous. For example, the difference between the experience of looking over some written instructions buried in a packet of other instructions and questionnaires and that of taking time to listen specifically to a tape recording seems to me to be potentially substantial, as their results demonstrate. I also question whether this truly provides evidence against an anchoring and adjustment paradigm for social inference. An alternative explanation is that, when the visibility of the situation is very high, the adjustment component may be large enough to simply negate a dispositional anchor. This is consistent

with the principles behind the RSVP system, which aim to negate the effects of dispositionalism under cognitive load, in part by raising the visibility of situational information. Further, in their third experiment, Trope and Gaunt (2000) provide a situation, interpreting course grading in a statistics class, that is likely to resonate strongly with the perspective of introductory psychology students (their population). In a sense, then, Trope and Gaunt (2000) have provided data to suggest that a debiasing intervention that involves perspective-taking may be able to overcome the effects of a dispositional inferential goal when subjects are in a situation that requires cognitive effort.

Despite my disagreements with some of their conclusions, Trope and Gaunt's (2000) study is valuable for its perspective on the possible effects and limits of inferential goals. Their conclusions that individuals have a difficult time integrating multiple types of explanatory information to form a complete judgment, and that this problem may be particularly evident under conditions of cognitive load, are congruent with the present investigation.

Overall, research using inferential goals presents a dilemma. On one hand, the research conducted by Quattrone (1982), and later extended by Krull and his colleagues, suggests that it is possible to manipulate the anchors individuals use in judgment tasks, and that this manipulation impacts the mix of situational and dispositional explanations they make. On the other hand, though, Trope and Gaunt's (2000) research makes it clear that effects of the structure and presentation of the information available for the adjustment process may make it difficult to predict the effect that a particular inferential goal will have on the ultimate judgment in all but the most contrived tasks. Further, Trope and Gaunt's (2000) second study provides some evidence that simply priming both goals in succession does not produce a more balanced judgment. Thus, while inferential goals may be useful for directing judgments, some difficulty exists in learning how to employ them effectively. The present research seeks to improve understanding of the role of inferential goals by

comparing their effects against those of the more complex RSVP technique. This investigation also tests the generalizability of inferential goals to a somewhat different class of judgment tasks (c.f., Krull, 1993). As a result, while the main focus is on the RSVP technique, there is some potential to learn something about the functioning of inferential goals that is independent of the RSVP examination.

Hypotheses

The hypotheses in the present investigation are divided into two groups. In the first, I present hypotheses intended to establish the basic viability of the RSVP technique for structured diagnosis of people problems. In the second, I propose additional hypotheses relevant to understanding the potential role of inferential goals in structured diagnosis. Further hypotheses will be introduced following the presentation of Study 1.

Hypotheses Establishing the RSVP Technique

When formulating the hypotheses for this research, my first goal was to establish whether the RSVP method of structuring diagnosis could produce the effects shown in Figure 1. Examination of the possible consequences of structured diagnosis suggested five hypotheses. The first of these follows directly from the discussion of dispositionalism. If the RSVP intervention is successful in raising situational awareness and easing cognitive load by providing structure, a reduction in dispositionalism should result:

H1: Use of the RSVP technique will increase the number of initial explanations that reference situational factors.

Most studies of dispositionalism require explanations to be scored as either dispositional or situational (e.g., see the Causal Dimension Scale, Russell, 1982). However, for the present purposes, there is no need to presume that individuals cannot see both situational and dispositional factors and refrain from choosing between them. Still, it is reasonable to expect that, if the intervention succeeds in increasing situational explanations, these may come at the expense of dispositional explanations:

H2: Use of the RSVP technique will decrease the number of initial explanations that reference dispositional factors.

The major point of increasing situational explanations by attending to a diagnostic structure is to prompt individuals to recognize the difficulty of the problem and the need for more complete information. This may cause them to delay a final diagnosis in order to conduct a broader information search. A recent study of counselors suggests that, in dealing with the well-recognized difficulty of making an accurate diagnosis of clients' mental health problems, clinicians delayed making a final diagnosis when allowed to do so (Hill & Ridley, 2001). As discussed above, an important question is whether the additional information search will be biased to confirm the initial hypothesis. This problem, called confirmation bias, is known to occur in many information search tasks (Devine, Hirt, & Gehrke, 1990; Skov & Sherman, 1986). However, given the logic of structured cognition reviewed above, tendencies toward confirmation bias, while not eliminated, may be manageable. This leads to two additional hypotheses:

H3: Use of the RSVP technique will decrease willingness to act without further information search.

H4: Use of the RSVP technique will increase the breadth of information search.

Finally, although this relationship is not depicted in Figure 1, the realization that the situation is complex may also reduce confidence in the initial diagnosis. As research has found that debiasing overconfidence without explicit instructions to consider weaknesses is difficult (Fischhoff, 1982), this prediction is somewhat tentative. This hypothesis tests the limits of the "frame-breaking" effects of moving beyond the initial explanation (Hirt & Markman, 1995; Koehler, 1991) by examining whether a structure for information search is sufficient to reduce confidence.

H5: Use of structured diagnosis will decrease confidence.

Hypotheses Related to Inferential Goal

A second group of hypotheses can be formulated based on the manipulation of individuals' inferential goal. Previous research has found that it is possible to prime individuals to hold either situational or dispositional inferential goals (Krull, 1993). In a cognitively effortful task, this should result in explanations which are consistent with the prime.

H6: Use of a prime to hold a particular inferential goal (situational or dispositional) will increase the number of initial explanations that reference factors consistent with the prime.

H7: Use of a prime to hold a particular inferential goal (situational or dispositional) will decrease the number of initial explanations that reference factors inconsistent with the prime.

Structured diagnosis and inferential goal priming are, in the present framing, competing methods for accomplishing a reduction in dispositionalism. Structured diagnosis is expected to have a similar effect to that of the situational inferential goal prime. One very strong demonstration of the impact of the structured diagnosis manipulation on initial explanations would be an ability to increase situational and reduce dispositional explanations, regardless of the primed goal. Such a finding would indicate that, much like the situational manipulations used by Trope and Gaunt (2000), structured diagnosis may be powerful enough to overwhelm the effects of inferential goals. This leads to one additional hypothesis.

H8: Inferential goals will impact initial explanations in groups that do not receive the structured diagnosis manipulation, but structured diagnosis will negate the effects of the primes. Figures 2 and 3 demonstrate graphically the predictions of Hypothesis 8.

Figure 2: Expected Interaction of Situational Prime and Structured Diagnosis

Figure 3: Expected Interaction of Dispositional Prime and Structured Diagnosis

Alternatively, structured diagnosis may strictly augment situational explanations. In this case, the group that receives both the situational inference prime and the RSVP manipulation should stand out. This leads to an alternative hypothesis.

H8a: Structured diagnosis and situational inferential goals will each have main effects, but there will be no interaction. Structured diagnosis will not affect dispositional explanations. Figure 4 graphically depicts the prediction of Hypothesis 8a.

Figure 4: Prediction of Hypothesis 8a

CHAPTER 2

STUDY 1 METHOD

Study 1 utilized a 2 x 3 factorial experimental design. The independent variables were structured diagnosis (two levels) and inferential goal (three levels). Manipulations were self-contained within the written packet of materials completed by participants. All participants were randomly assigned to conditions.

Task Environment: The Big Bite Case

The task is a case that I have developed specifically for this investigation. The case, Trouble at Big Bite, presents a hypothetical people problem dealing with what appears to be large scale employee theft of food in a chain of restaurants (for the complete text of the Big Bite case, see Appendix A). The Big Bite case contains information on each of the dimensions of RSVP, but it is intentionally vague. For example, information on replications is not explicitly given, but it is hinted at in the comment about settings, in which the focal person in the case, Mary, says "the situation is basically the same across all 100 stores." No information about the number of views is explicitly given, but the entire case is presented through Mary's eyes. Information about the number of people engaged in theft must be inferred from comments such as the one about the number of stores above, along with Mary's comment that her employees are "all lazy, and they steal." The given information suggests that the number of replications, settings, and people involved may be high, but only

a small amount of information is presented, and it is all from one view. The case has been designed to represent the initial presentation of a complex problem, which will require a cognitively effortful process to interpret. The combination of ambiguity and complexity present suggests a normative conclusion that additional information search should be conducted.

Prior to the present investigation, the Big Bite case was piloted in three undergraduate course sections (two sections of personnel psychology and one section of organizational psychology). Students in these classes indicated that they found the case engaging and realistic, and class discussions suggested that a variety of interpretations of the case data were made.

Participants

Research participants were 52 students enrolled in an upper-level undergraduate course in organizational psychology at The Ohio State University. All were currently employed, with most in part-time positions. Participants were randomly assigned to receive materials for one of the six experimental conditions. Participation was voluntary, and materials were completed anonymously and returned to the instructor. No incentives for participation were offered in this study. Debriefing and discussion were conducted in the following class.

Procedures

To avoid contaminating the results with knowledge obtained from the class, the study was conducted at the beginning of the term, during a unit on research methods. After a brief introduction in which the instructor explained that the task was to provide an opportunity for the class to participate in and then, later, to critique, an actual research project (see Appendix B), a packet of materials containing the case (Appendix A), experimental materials (for those receiving the manipulation), and the questionnaire was

distributed. Participants were asked to imagine that Mary had asked them to help her solve this problem, and then respond to the questions in the packet.

Manipulation of Independent Variables

As mentioned above, the experimental manipulations were completely contained in the packets of materials given to participants. The structured diagnosis manipulation consisted of a page of information describing the RSVP framework and how to use it (see Appendix C). In conditions receiving the structured diagnosis manipulation, this page was integrated into the packet as the first page following the cover sheet. The understanding of this manipulation was checked via four items at the beginning of the questionnaire.

Subjects who were exposed to the structured diagnosis manipulation were asked: a) How many replications of the behavior have been observed, b) How many settings has the behavior been observed in, c) How many views of the problem are represented in the case, and d) How many people are engaged in the behaviors? Each response was made on a seven-point scale with anchors "one or a few" and "many." Students not receiving the manipulation were not asked these questions.

Consistent with previous studies (e.g., Erickson & Krull, 1999; Krull, 1993; Krull & Dill, 1996; Krull & Erickson, 1995; Quattrone, 1982), a relatively straightforward manipulation of inferential goals was used. The manipulation of inferential goal was based, in particular, on the strategy used in Krull's (1993) initial study. In this research, inferential goal was manipulated by asking subjects to watch a videotape of a person being interviewed, then estimate either "how anxious the woman on the tape is in her day-to-day life" (dispositional inferential goal, p. 343) or "how anxiety provoking the interview topic was" (situational inferential goal, p. 343). Although videotape tasks similar to that used by Krull (1993) have been used in several studies since (e.g., Erickson & Krull, 1999; Krull & Dill, 1996; Krull & Erickson, 1995), no study as yet appears to have attempted to examine how inferential goals might function in the consideration of a more complex, abstract problem

involving more than one target person. Thus, I attempted to find out if the kind of simple prompt used by Krull (1993) can be effective in a somewhat different type of judgment task. The manipulation was integrated with the dependent variable measure for the initial explanation. I asked participants in the "no goal" inferential goal condition to respond to the prompt "Based on the information given, what do you think is happening?" Participants in the "situational" inferential goal condition responded to the same question, but with the words "in this situation" inserted at the end of the sentence. Their question was worded "Based on the information given, what do you think is happening in this situation?" Participants in the "dispositional" inferential goal condition responded with the words "with these people" inserted. This question was "Based on the information given, what do you think is happening with these people?"

Measures of Dependent Variables

An important consideration in the measurement of dependent variables was the training of judges. This research uses judges' ratings of dispositionalism, situationalism, and information search. Judges were two graduate students in the social sciences (one in psychology, one in sociology). Neither judge was aware of either the hypotheses or conditions used in the study. Judges were trained to respond using fabricated items that were meant to simulate the different categories for judging. For example, in response to the question prompting participants for an explanation (see the next section below), a sample phrase (dispositional explanation) was "The job probably isn't worth much to them." In the section on information search, sample items included statements like "How do they recruit employees?" (which was scored as a question about the setting). Appendix D presents a full list of the training items.

Dispositionalism and Situationalism

Due to concerns that scale anchors for closed-ended items would act as a manipulation, prompting control group members to make more balanced assessments, this

study used free responses to measure explanations. Although some have criticized the use of free response measures scored by judges on the grounds that coding processes may suffer from errors similar to those they attempt to measure (e.g., Russell, 1982), the psychometric properties of available scales have also received unfavorable scrutiny (McAuley, Duncan, & Russell, 1992; Watkins, Sachs, & Regmi, 1997). McGill (1989) provides prior evidence that the method used here can be effective. An additional advantage in the present study is that judges were not required to choose between characterizing a respondent's answers as "all situational" or "all dispositional," or to try to weight the responses. Coding merely indicated the presence or absence of each individual type of explanation.

Participants were asked one variant of the "what is happening" question described above. Judges' ratings of individuals' free responses were used to assess the presence or absence of dispositional and situational explanations. Judges independently scored all responses for the presence or absence of both dispositional and situational explanations. Both judges were blind to hypotheses and conditions. After computing initial reliability estimates, the judges discussed and resolved outstanding discrepancies.

Confidence in Judgment

Confidence was assessed with a single item, "How confident are you that your assessment is correct?" Respondents answered on a seven-point, Likert-type scale, anchored by "not at all confident" (low) and "extremely confident" (high).

Information Search

In this section, all participants were prompted as follows: "Imagine that you can now ask some additional questions about the situation at Big Bite. What kinds of questions would you want to ask?" The same two judges independently scored all responses, categorizing the questions into the dimensions of the RSVP framework. Each question was scored as to whether it asked about replications, settings, views, or people. Scoring

indicated the presence or absence of questions addressing each category (i.e., multiple questions in a single category did not increase the score for that category). After computing initial reliability estimates, the judges discussed and resolved outstanding discrepancies.

Willingness to Act

A final item assessed the degree to which the available information was sufficient to take action. Participants were asked to respond on a verbally-anchored, seven-point scale (anchors strongly disagree, disagree, somewhat disagree, neither agree nor disagree, somewhat agree, agree, strongly agree) to the item, "Even without asking additional questions, the case provides enough information for me, as a consultant, to recommend a solution to Mary."

CHAPTER 3

STUDY 1 RESULTS

Study 1 contained manipulations of two independent variables, inferential goal and structured diagnosis. For the inferential goal manipulation, no manipulation checks were conducted. The hypothesis tests themselves were the only test of effectiveness. For the structured diagnosis manipulation, however, a manipulation check was used. To assess the degree to which individuals exposed to the structured diagnosis manipulation understood its dimensions (RSVP), four items (one for each dimension) were used. Participants receiving the manipulation were asked to indicate, on a seven-point scale, with anchors "one or a few" (low) to "many" (high), how many replications were observed, how many settings, how many views of the problem were represented, and how many people were engaged in the behaviors. Means for these items indicated that they were perceived as expected, given the structure of the case, with high means observed for replications (6.3), settings (5.7), and people (5.9), and a low mean (1.5) on the "views" dimension.

Tests of Hypotheses

Tests of Hypotheses Related to Structured Diagnosis

Hypotheses 1 and 2 predicted that the use of situational explanations would be increased by the structured diagnosis manipulation and that the use of dispositional explanations would be decreased by it. Initial inter-rater reliability between judges on the

presence or absence of explanations was acceptable, if unexciting ($\phi = .71$ for dispositional explanations and $\phi = .66$ for situational explanations), and all differences between judges were resolved through discussion without the participation of anyone with knowledge of hypotheses or conditions. Tables 1 and 2 show the frequency of situational and dispositional explanations, respectively. The Fisher's exact test of Table 1 was significant (two-sided $p < .05$, $\phi = .36$). Structured diagnosis increased situational explanations. The chi-square test of Table 2 was also significant ($\chi^2 = 10.03$, $p < .01$, $\phi = -.44$). Structured diagnosis decreased dispositional explanations. Hypotheses 1 and 2 were supported.

	Control Group	Structured Diagnosis	
No Situational Exp.	8	1	9
Situational Exp.	18	25	43
	26	26	52

Table 1: Situational Explanations By Condition

	Control Group	Structured Diagnosis	
No Dispositional Exp.	4	15	19
Dispositional Exp.	22	11	33
	26	26	52

Table 2: Dispositional Explanations By Condition

Hypothesis 3 predicted that structured diagnosis would reduce willingness to act without collecting further information. This was tested using ANOVA, with manipulation status (coded 0=control, 1=structured diagnosis) as the factor and willingness to act as the dependent variable. As confidence could be expected to exert an influence on willingness to

act, it was also included as a covariate. Descriptive information on these variables is shown in Table 3. Results of the analysis are listed in Table 4. Both structured diagnosis and confidence were significant predictors, and, as expected, use of structured diagnosis was negatively associated with willingness to act. Cell means for the willingness to act measure were 3.7 in the control condition and 2.6 in the structured diagnosis condition. When confidence was entered first into a regression equation, followed by structured diagnosis, structured diagnosis explained an additional 7.3% of the variance ($p < .05$). Overall, confidence and structured diagnosis explained 17% of the variance in willingness to act (Adjusted R^2). Hypothesis 3 was supported.

Variable	Mean	SD	RSVP	Confidence	Willingness to Act
RSVP	N/A	N/A			
Confidence	4.58	1.09	.21		
Willingness to Act	2.96	1.40	-.19	.36*	

* $p = .01$

Table 3: Means, SD's, and Correlations between Willingness to Act and its Predictors

Source	SS	MS	F	P
RSVP	7.62	7.62	4.40	.04
Confidence	16.95	16.95	9.79	.00
Error	83.15	1.73		

Table 4: Results Of ANOVA Of RSVP And Confidence On Willingness To Act

Hypothesis 4 predicted that use of structured diagnosis would result in more breadth of information search. Breadth was operationalized as the number of categories in the RSVP framework in which questions were asked. Few respondents ($n=3$, all in the experimental group) asked questions about the "replications" dimension, and most ($n=48$) asked questions related to the setting. Analysis, therefore, focused primarily on the "views" and "people" dimensions. Inter-rater reliability for both was acceptable ($\phi = .90$ for views, $.73$ for people), and, once again, all disagreements between raters were successfully resolved. Analysis of contingency tables showed no relationship at all between use of structured diagnosis and questions about people, with the same number of questions about people asked in each condition, but a relationship between structured diagnosis and questions about alternative views of the problem was found ($\chi^2 = 3.5$, $\phi = .26$, $p = .06$). Table 5 shows the presence of questions about views by condition. Analysis of questions about replications revealed that the effects of structured diagnosis also approached significance on this dimension of information search ($\chi^2 = 3.2$, $\phi = .25$, $p = .07$). Table 6 shows the presence of questions about replications by condition. No effects were found for settings ($\chi^2 = 1.2$, $\phi = .15$, $p > .20$). Hypothesis 4 received some support, but only with regard to the "views" and, possibly, "replications" dimensions.

	Control Group	Structured Diagnosis	Totals
Asked Q about view	4	10	14
No Q about view	22	16	38
Totals	26	26	52

Table 5: Information Search On Views By Condition

	Control Group	Structured Diagnosis	Totals
Asked Q about replications	0	3	3
No Q about replications	26	23	49
Totals	26	26	52

Table 6: Information Search On Replications By Condition

Hypothesis 5 predicted that use of structured diagnosis would decrease confidence in the initial explanation. The observed correlation between the two variables was not significant, and was in the wrong direction ($r=.21$, $p=.13$). Hypothesis 5 was not supported.

Tests of Hypotheses Related to Inferential Goals

Hypotheses 6 and 7 predicted that inferential goal priming would result in an increase in explanations consistent with the prime, and a decrease in explanations inconsistent with the prime. Because dispositional and situational explanations were measured as separate entities, rather than as poles on a scale, I specified these predictions as separate hypotheses. Table 7 shows the number of situational explanations for each inferential goal condition. Although the relationships are numerically as predicted, they are not significant ($\chi^2 = 3.3$, $p=.19$). Table 8 shows the number of dispositional explanations for each inferential goal. Again, no effect was found in the analysis ($\chi^2 = 1.3$, $p=.53$). Hypotheses 6 and 7 were not supported.

Condition	No Situational Explanation	Situational Explanation	Total
Situational Goal	1	16	17
Dispositional Goal	5	12	17
No Goal	3	15	18
Total	9	43	52

Table 7: Situational Explanations By Inferential Goal

Condition	No Dispositional Explanation	Dispositional Explanation	Total
Situational Goal	8	9	17
Dispositional Goal	5	12	17
No Goal	6	12	18
Total	19	33	52

Table 8: Dispositional Explanations By Inferential Goal

Hypothesis 8 predicted that there would be an interaction between the inferential goal and structured diagnosis. For dispositional explanations, this was not found. While there was a clear main effect of the structured diagnosis manipulation (see Table 2), results within each condition (structured diagnosis, no structured diagnosis) were nearly identical, regardless of inferential goal. Turning to situational explanations, subjects receiving the structured diagnosis manipulation were, once again, nearly identical in their explanations, regardless of inferential goal. Individuals who did not receive the structured diagnosis manipulation showed some differences. Specifically, although individuals receiving either

the situational goal prime or no goal prime were nearly identical (7/8 and 7/9, respectively, made situational explanations), individuals receiving the dispositional goal prime appeared to make less situational explanations (4/9 individuals citing situational factors).

Unfortunately, these cell sizes are too small for adequate statistical testing. Despite some slight evidence for Hypothesis 8, it must be considered unsupported.

Hypothesis 8a predicted that both structured diagnosis and situational inferential goals would have main effects on situational explanations, but that there would be no interaction. Although all individuals who received this combination did make situational explanations, so did all individuals who received the structured diagnosis manipulation combined with a dispositional prime. Thus, there was no main effect of the situational prime. Hypothesis 8a was not supported.

Summary of Study 1

In this initial study, I found evidence to support four of the five hypotheses related to structured diagnosis. In general, the RSVP technique appears to structure diagnosis in a way that is effective for debiasing dispositionalism in complex people problems. The sole exception to this was found in the lack of influence on confidence. At the same time, no hypotheses related to inferential goals were supported. This could be attributed to at least three possibilities. The first is that, due to the modest sample size, power to detect inferential goal effects was limited. It is possible that significant effects would emerge in the data captured by tables seven and eight if more responses were available. However, it is also possible that the inferential goal prime was not sufficiently strong. Participants may have simply ignored the inferential goal manipulation. Given the large number of findings supporting dispositionalism, it seems unreasonable to conclude that inferential goals have such limited utility on the basis of a prime as subtle as the one used in Study 1.

The findings suggest two additional foci for investigation. First, the fifth hypothesis in the initial study, which related confidence to structured diagnosis, was not supported.

However, confidence had a strong, positive association with the willingness to act. This suggests that the inclusion of a separate manipulation to reduce confidence may further reduce the willingness to act without collecting additional information. Below, I discuss the rationale for this manipulation in more detail, and introduce three new hypotheses. Second, the above analysis of the inferential goal results in Study 1 implies a follow-up study. In Study 3, I will attempt to rectify the problems of Study 1 by using both a larger sample and a stronger manipulation of the inferential goal.

Research on Overconfidence

Although Trope and Gaunt (1999; 2000) focus primarily on the integration process in judgment, they note that, despite the large volume of evidence suggesting that individuals have a difficult time making judgments which adequately take into account both situational and dispositional factors, people remain very confident in the judgments they do make. Indeed, research has found that, across a large variety of situations where specific, numerical estimates of verifiable hypotheses are made, people are overconfident in their judgments (Arkes, 2001; Fischhoff, 1982). In laboratory studies, this is generally demonstrated by asking subjects to answer factual, multiple-choice questions, then indicate their confidence in their answers (e.g., Koriat, Lichtenstein, & Fischhoff, 1980). Across a series of questions and answers, subjects are said to be "well-calibrated" to the extent that their confidence estimates match the percentage of correct answers given. A well-calibrated subject would, for example, answer correctly 90% of questions to which a 90% confidence level is assigned. The usual finding, however, is that people are overconfident in their abilities. That is, they correctly answer less than 90% of questions to which they have assigned a 90% confidence level.

Given its pervasiveness (Arkes, 2001), it is no surprise that much attention has been devoted to reducing overconfidence. In an early review of debiasing techniques, Fischhoff (1982) concentrated exclusively on efforts to ameliorate overconfidence and hindsight bias

(a special case of overconfidence about the likelihood of having been able to predict the occurrence of a past event). At the time, his conclusion was that, though overconfidence itself seemed to be present in many domains, debiasing it was proving to be difficult (Fischhoff, 1982). However, two techniques for debiasing overconfidence have since become accepted as effective. One strategy that has been shown to be effective is encouraging people to "consider the opposite" (Lord, Lepper, & Preston, 1984). Lord et al. (1984) presented subjects with fabricated research results prior to making attitudinal judgments about capital punishment. Subjects in the experimental group were prompted as follows:

"Ask yourself at each step whether you would have made the same high or low evaluations had exactly the same study produced results on the *other* side of the issue" (Lord et al., 1984, p. 1233, italics in original).

Initial attitude (for or against capital punishment) did not impact the evaluations of those who received this "consider the opposite" manipulation, but it did have an effect on evaluations of both control subjects and those who received instructions to be objective and unbiased.

A second, related technique was actually developed earlier by Koriat et al. (1980), but, because of initial problems in replication (Fischhoff & Macgregor, 1982), its validity was cloudy until confirmation was provided by Arkes et al. (1988). In the research by Koriat et al. (1980), subjects responded to general knowledge questions, as in the laboratory paradigm described above. Subjects in the experimental group were asked, before stating their confidence level, to generate as many reasons as possible why each of the options could be correct, and as many reasons as possible why it could be incorrect. Those receiving this treatment exhibited significantly less overconfidence than a control group. In their second experiment, Koriat et al. (1980) relaxed the extreme reason generation requirements, instead asking for only one reason that either supported or contradicted their

decision. Subjects in the contradicting condition, who were required to write the best reason they could conceive that either weighed against the alternative chosen or for the alternative not chosen, exhibited reduced overconfidence.

As mentioned above, Arkes et al. (1988) replicated these results in the domain of the hindsight bias. Interestingly, though, Arkes et al. (1988) obtained their effect without requiring the generation of reasons why the judgment might be wrong. Their study simply required the generation of a single "pro" reason for each possible alternative. Koehler (1991) integrated research on confidence with studies of explanation and imagination, and, in the process, provided a rationale for the strategy taken by Arkes et al. (1988). Koehler (1991) suggested that prompts to explain, imagine, or otherwise consider a particular possibility require a person to temporarily create a frame of reference which assumes that it is true. This causes all available evidence to be reorganized consistent with the frame of reference, which then facilitates the search for evidence consistent with the hypothesis, resulting in both confirmation bias and overconfidence. As a result, one might expect prompts to consider alternatives, such as those present in the study by Arkes et al. (1988), to be effective in reducing overconfidence, even without an explicit consideration of how individual choices could be wrong. As will be discussed below, this is particularly true because the alternatives used by Arkes et al. (1988) were all plausible alternatives. In the previous work of Koriat et al. (1980), no effort was made to ensure that subjects would find the alternative answers plausible. Indeed, given that their questions were factual knowledge questions with one correct and one incorrect answer, it is almost certain that these subjects did not always find the alternatives plausible.

Though less exhaustively developed than Koehler's (1991) theory, Arkes (1991) made a similar argument in his discussion of cueing to combat association-based errors. Arkes (1991) proposed that counterexplanation tasks (e.g., consider the opposite) essentially act by cueing additional stimuli that are not accessed when the initial explanation

is made. This process develops new causal paths, thereby reducing bias (Arkes, 1991). Hirt and Markman (1995) built on both Koheler's (1991) and Arkes' (1991) insights in their research on multiple explanation. In a series of three studies, they demonstrated that the consideration of any subjectively plausible alternative (not necessarily the opposite of the original hypothesis) produced evidence of debiasing. Hirt and Markman (1995) concluded that "consider an alternative" strategies break the frame of the initial hypothesis, clearing the way for more comprehensive judgment, as suggested by both Koehler (1991) and Arkes (1991).

In a recent review, Arkes (2001) summarized the available wisdom on debiasing overconfidence. In a single-judgment task, individuals are given two tools to reduce overconfidence. First, and foremost, consider alternatives. As Hirt and Markman (1995) have demonstrated, these do not have to be the opposite of the initial explanation; any subjectively plausible alternative will do. Second, to reinforce this, generate reasons for the alternatives (Arkes et al., 1988; Hirt & Markman, 1995; Koriath et al., 1980). By combining these two strategies, individuals will have at least some assurance of being appropriately confident of their judgments.

Hypotheses Related to Confidence Reduction

Three new hypotheses are based on the introduction of a manipulation to lower confidence. Following from the discussion above, I first propose that the results of other studies will replicate in the present context.

H9: Participants receiving the confidence reduction manipulation will display reduced confidence.

Of more theoretical interest is the effect of the combination of structured diagnosis and confidence reduction on willingness to act. In Study 1, structured diagnosis and confidence had opposite effects on willingness to act. Structured diagnosis reduced willingness to act, but confidence was associated with increases in it. It may be possible,

then, to improve the strength of the structured diagnosis intervention by adding an intervention for confidence reduction. This leads to the following hypotheses.

H10: Subjects receiving a confidence reduction manipulation will be less willing to act than those who do not receive the manipulation.

H11: Subjects using both confidence reduction and structured diagnosis techniques will be least willing to act without collecting more information. Figure 5 shows the expected relationship under Hypothesis 11.

Figure 5: Expected Interaction between Confidence Manipulation and Structured Diagnosis

CHAPTER 4

STUDY 2 METHOD AND RESULTS

This chapter reports the method and results for Study 2. Study 2 was conceived as an extension of Study 1. Specifically, the goal was to determine whether the structured diagnosis manipulation could be improved by adding a manipulation to reduce confidence. Study 2 maintained the methodology of Study 1, but, this time, questions thought to be "manipulation checks" in Study 1 were explicitly acknowledged as part of the manipulation. This change was made in response to comments from anonymous reviewers on a manuscript version of Study 1.

Method

Study 2 examined two independent variables using a 2 x 2, between-subjects factorial design. The first research factor, structured diagnosis, was divided into two conditions (presence or absence of the manipulation). The second research factor, the confidence reduction manipulation, was similarly divided. The experimental task was the same as in Study 1.

Participants and Procedure

Research participants were 63 students enrolled in upper-level undergraduate courses in organizational and personnel psychology at The Ohio State University. Fifty-four were currently employed (86%), working an average of 17.3 hours per week. Students were given

the experimental materials near the middle of the course. The full solicitation script is reprinted in Appendix B. A small inducement, in the form of extra credit points, was given to encourage participation in the personnel psychology class. Students in the organizational psychology class participated voluntarily, as in Study 1. Students in the personnel psychology class who did not wish to participate had the option of completing an alternative assignment to receive the extra points. Participants were randomly assigned to conditions. Participants completed the materials confidentially (anonymously in the organizational psychology class) and returned them to either the instructor (organizational psychology class) or the researcher (personnel psychology class). Debriefing was conducted via a written form in the personnel psychology class (see Appendix E), and via an in-class discussion in the organizational psychology class (as in Study 1). Otherwise, procedures were identical to those used in Study 1.

Manipulation of Independent Variables

As in Study 1, the manipulations used in Study 2 were written and fully contained within the materials packets. The presentation of information for the manipulation of structured diagnosis was identical to that used in Study 1. However, in Study 2, the questions formerly considered to be manipulation checks were now considered part of the manipulation.

Some differences were introduced between participants receiving and not receiving the confidence manipulation. Subjects not receiving the confidence manipulation first provided an initial explanation for what is happening in the case which corresponded to the control condition presentation in Study 1 (i.e., "Based on the information given, what do you think is happening?"). These participants then responded to the questions about the other

dependent variable measures (confidence, additional information search, willingness to act – see below for descriptions). Subjects receiving the confidence manipulation began with the same "what is happening" question, but then completed additional steps before responding to the other dependent measures. Specifically, after providing the initial explanation for what is happening, but before completing the other dependent variable items, these participants were asked: "Please write down one reason why you believe your diagnosis may be correct." After they provided this answer, they were given the following prompt:

Now, imagine for a moment that you have learned that your diagnosis of what is happening is incorrect. If you found that this was the case, what is one reasonable alternative view of what is happening?

After responding to this prompt, subjects in the confidence reduction condition were asked: "Please write down one reason why this diagnosis might be correct." These participants then completed the same dependent measures as control participants.

Measures of Dependent Variables

All measures used in this study were contained in the materials packet. The two dependent variables of primary interest in Study 2 were "confidence" and "willingness to act." Confidence was measured by a single item, using a Likert-type scale. This item asks "How confident are you that your (first) diagnosis is correct?" The word "first" appears in the question for subjects receiving the confidence reduction manipulation, as they were required to list a second diagnosis. Although Study 1 measured confidence on a seven-point scale, in Study 2, I opted to follow Hirt and Markman (1995), who used a 13-point scale, with anchors "not at all confident" to "very confident." I also changed the measurement of willingness to act. In Study 1, this variable was measured with a single item, which asked: "Even without asking additional questions, the case provides enough information for me, as a consultant, to recommend a solution to Mary." Subjects answered on a seven-point, Likert-type scale with verbal anchors ranging from "strongly disagree" to

"strongly agree." In Study 2, three additional items were added using the same scale: a) Without getting more information, I would be hesitant to take action in this case, b) If I were in Mary's shoes, I'd want to ask additional questions before making a decision, and c) If this were my decision, I'd be willing to act without searching for additional information. This four-item scale constituted the willingness to act measure. All other variables were measured as in Study 1. Judges in Study 2 were two graduate students (one in psychology, one in sociology). The psychology graduate student was different from the judge in Study 1.

Results

No additional manipulation checks on structured diagnosis were included in Study 2, though the questions formerly considered to be manipulation checks were expected to have some diagnostic value. Responses to these questions indicated that those receiving the structured diagnosis manipulation scored the Big Bite case similarly to participants in Study 1 (means 5.70, 5.90, 1.42, and 5.61 on the replications, settings, views and people dimensions, respectively). Hypothesis 9 provides both a replication of past research and a manipulation check for the confidence manipulation.

Tests of Hypotheses Establishing the RSVP Technique

Hypotheses 1 and 2 proposed that subjects who received the RSVP manipulation would display increased situational and decreased dispositional tendencies in their initial explanations. As in Study 1, ratings from two judges blind to hypotheses and conditions were used to measure the presence or absence of situational and dispositional explanations. Inter-rater reliability was good for both situational ($\phi=.90$) and dispositional judgments ($\phi=.80$), and all disagreements were resolved by the raters without interference from anyone with knowledge of hypotheses and conditions. A substantial number of individuals declined to make an initial explanation, all of whom were in conditions receiving the structured diagnosis manipulation. I analyzed the results in two ways. First, I coded these

non-explanations as zeros on both the situation and person dimensions. Tables 9 and 10 show the frequency of situational and dispositional explanations using this coding. When the analysis was completed in this fashion, structured diagnosis was found to significantly reduce the number of initial explanations referencing person factors, as predicted by Hypothesis 2 ($\chi^2(1)=13.64, p<.01, \phi=.47$). However, contrary to Hypothesis 1, using this coding scheme, structured diagnosis had no effect on the number of initial explanations referencing situation factors ($\chi^2(1)=.55, n.s., \phi=.09$). Indeed, the sign was negative in this analysis, suggesting that structured diagnosis reduced all initial explanations.

	Control Group	Structured Diagnosis	Totals
No Situational Exp.	10	14	24
Situational Exp.	20	19	39
Totals	30	33	63

Table 9: Situational Explanations By Condition (No Explanation Coded as Zero)

	Control Group	Structured Diagnosis	Totals
No Dispositional Exp.	12	28	40
Dispositional Exp.	18	5	23
Totals	30	33	63

Table 10: Dispositional Explanations By Condition (No Explanation Coded as Zero)

In order to determine how structured diagnosis affected those individuals who did make an explanation, I also computed these analyses with "no explanation" coded as missing data. This change had little effect on the person explanations ($\chi^2(1)=7.15, p<.01, \phi=.37$), but it did correct the discrepancy with respect to the direction of the effect on

situational explanations. However, with 11 records eliminated from this condition, the results fell short of significance ($\chi^2(1)=2.63$, $p=.11$, $\phi=.23$). Tables 11 and 12 show the frequencies using this coding.

	Control Group	Structured Diagnosis	Totals
No Situational Exp.	10	3	13
Situational Exp.	20	19	39
Totals	30	22	52

Table 11: Situational Explanations By Condition (No Explanation Coded as Missing)

	Control Group	Structured Diagnosis	Totals
No Dispositional Exp.	12	17	29
Dispositional Exp.	18	5	23
Totals	30	22	52

Table 12: Dispositional Explanations By Condition (No Explanation Coded as Missing)

Although the results seemed obvious, I also tested the hypothesis that the structured diagnosis manipulation would increase the degree to which individuals would refrain from making an initial explanation. This hypothesis was supported ($\chi^2(1)=12.12$, $p<.01$, $\phi=.44$). Table 13 shows this frequency information.

	Control Group	Structured Diagnosis	Totals
No Explanation Given	0	11	11
Explanation Given	30	22	52
Totals	30	33	63

Table 13: No Explanations By Condition

Hypothesis 3 predicted that structured diagnosis would reduce willingness to act. This hypothesis did not replicate here. The correlation between RSVP and willingness to act, though in the predicted direction, was small and not significant. More detail on this analysis is provided below in the discussion of confidence effects.

Hypothesis 4 predicted that the use of structured diagnosis would result in more breadth of information search. As in Study 1, breadth was operationalized as the number of categories in the RSVP framework in which questions were asked, as scored from free responses. Once again, probably owing to the structure of the Big Bite case, few respondents ($n=4$, all in the experimental group) asked questions about the replications dimension, and most ($n=53$) asked questions about the setting. Inter-rater reliability was low for the replications dimension ($\phi=.45$), but was acceptable for the others ($\phi=.83$, .76, and .70 for settings, views, and people, respectively). As before, all disagreements were successfully resolved by the judges.

Analyses of the contingency table data suggested that the use of structured diagnosis increased the likelihood of asking questions about replications ($\chi^2(1)=3.88$, $p<.05$, $\phi=.25$) and views ($\chi^2(1)=9.17$, $p<.01$, $\phi=.38$), had no effect on questions about people ($\chi^2(1)=1.27$, $p=.26$, $\phi=.14$), and decreased use of questions about settings ($\chi^2(1)=3.64$, $p=.06$, $\phi=-.24$). Once again, partial support for Hypothesis 4 was found.

Tests of Hypotheses Related to Confidence Reduction

Hypothesis 5, 9, 10, and 11, which dealt with confidence and its reduction, were included to extend beyond the prediction of Hypothesis 5, which was unsupported in Study 1. Hypothesis 9 predicted that participants receiving the confidence reduction manipulation would display reduced confidence. This hypothesis, as noted previously, simply attempts to apply the accepted techniques for debiasing confidence. Unexpectedly, this manipulation was not effective. In fact, participants who received only the confidence manipulation reported the highest confidence (mean = 9.47, vs. 8.40, 7.25, and 8.79 for the RSVP-only, combined manipulations, and control conditions, respectively). One problem in conducting the analyses related to confidence was that a significant number of individuals who received the structured diagnosis manipulation did not offer an initial explanation. Yet, they still provided confidence ratings. Since the target for these confidence ratings was supposed to be the initial explanation, it is not clear what ratings offered by those who made no explanation might mean. This is a somewhat frustrating situation, since these individuals were, arguably, the most affected by the structured diagnosis manipulation. Rather than speculate about the meaning of the confidence ratings, though, the most appropriate step seemed to be to remove them from the analyses. The means above were computed with the confidence ratings for those not offering an explanation removed, and, as a result, the cell sizes are not large (in the smallest cell, the RSVP-only condition, $n=10$ respondents). ANOVA was conducted to examine the effects of the manipulations on confidence. The presence or absence of each manipulation was dummy coded. Table 14 reports the results of the analysis. In support of Hypothesis 5, the structured diagnosis manipulation significantly reduced confidence, and the overall ANOVA explained about 8% of the variance (Adjusted $R^2=.08$). However, as mentioned, no effects were found for the confidence manipulation. Hypothesis 9 was not supported.

Source	SS	MS	F	P
RSVP	21.07	21.07	4.57	.04
Confidence Manipulation	.68	.68	.15	.70
RSVP x CM	10.43	10.43	2.26	.14
ERROR	216.74	4.61		

+n=51

Table 14: ANOVA of Manipulations on Confidence+

Hypotheses 3, 10 and 11 proposed effects of structured diagnosis and confidence reduction on willingness to act. Cronbach's alpha for the full, four-item willingness to act scale was .68. Dropping the willingness to act item used in Study 1 resulted in an improvement to .70. It is unclear why the four-item scale was unstable in this study, but it could be due to the difference in referent in the items. The willingness to act item used in Study 1 asks respondents to recommend a solution to Mary. In contrast, the other items all ask the respondent about what action they would take if they were making the decision. To check out any discrepancies, I did the following analyses using both versions of the scale.

- To test Hypotheses 3, 10, and 11, I examined the effects of the two manipulations and their cross-product on willingness to act. I used dummy coding for the confidence and RSVP manipulations, and I included confidence in the initial explanation as a covariate.

Table 15 shows the correlations between the variables.

Variable	Mean	SD	RSVP	CM	Con	4WTA	3WTA
RSVP	N/A	N/A	1.00				
Confidence Manipulation	N/A	N/A	-.05	1.00			
Confidence	8.55	2.24	-.31*	-.03	1.00		
4-Item Willingness to Act	9.34	3.50	-.10	-.13	.34*	1.00	
3-Item Willingness to Act	6.16	2.63	-.14	-.24	.27	.91*	1.00

* $p \leq .05$

+n ranges from 51 to 63, depending upon the number of complete responses available for the comparison. Confidence comparisons exclude data from individuals offering no explanation.

Table 15: Means, SD's, and Correlations between Willingness to Act and Its Predictors+

Hypothesis 10 stated that individuals receiving the confidence manipulation would be less willing to act than those who did not receive this manipulation. Using the three-item willingness to act scale as the dependent variable, the confidence manipulation was the only predictor approaching significance ($F = 3.34$, $p = .07$). Table 16 shows the results of the full model, which explained 7% of the variance in willingness to act. The mean willingness to act for those receiving the confidence manipulation was 6.9, versus 5.7 for those not receiving the manipulation. Hypothesis 10 received some support using the three-item measure of willingness to act.

Source	SS	MS	F	P
Confidence	12.78	12.78	2.32	.14
RSVP	.43	.43	.08	.78
Conf. Manip.	18.44	18.44	3.34	.07
RSVPxConf M	4.30	4.30	.78	.38
Error	253.81	5.52		

+n=51

Table 16: ANOVA of Manipulations and Confidence on Three-Item Willingness to Act+

As a check, I also conducted the analysis using the full, four-item version of the willingness to act scale. In this analysis, the confidence manipulation ceased to be a significant predictor ($F=1.16$, $p=.29$), and confidence emerged as the sole significant effect on willingness to act ($F=4.34$, $p<.05$). This left me suspicious that the original willingness to act item (as used in Study 1) was performing much differently than the rest of the scale. As noted earlier, this is plausible, since this item puts the respondent in the role of a consultant, whereas the other three items put the respondent in the role of the actor. A third analysis confirmed this. Here, confidence was a significant predictor of the score on the one-item "willingness to act as a consultant" scale ($F=4.37$, $p<.05$). No other predictors were anywhere near significance (e.g., the confidence manipulation $F=.37$, $p=.55$).

Hypothesis 11 stated that the combination of the structured diagnosis manipulation and the confidence manipulation would produce the strongest effects on reducing willingness to act. An examination of means suggests some directional evidence for this prediction. The mean willingness to act for those receiving both manipulations was the smallest of the four groups (5.3, vs. 6.1, 7.1, and 6.7 for the confidence manipulation-only, RSVP-only, and control conditions, respectively), but this analysis may have been hampered

by the low power for the structured diagnosis manipulation. As shown in Table 16, the interaction between the manipulations did not approach significance. Hypothesis 11 was not supported.

Finally, as mentioned above, Hypothesis 3, which predicted that structured diagnosis would reduce willingness to act, was not supported in Study 2. As noted in Table 15, the correlation between RSVP and willingness to act was negative but not significant. Two mitigating factors suggest that this lack of support in the present study is not surprising. First, because a third of those receiving the structured diagnosis manipulation had to be excluded from the analysis (due to the lack of interpretability of their confidence ratings, as discussed above), the power for this analysis was very low (observed power for the RSVP manipulation on the willingness to act variable was .06). Second, those individuals who were excluded were those believed to be most affected by the structured diagnosis manipulation. Together, these factors suggest that these hypothesis tests were unreliable.

Additional Analyses

The results of this study were surprising. Use of a standard confidence-reduction manipulation failed to reduce confidence. Upon further consideration, I wondered whether the subjective quality of the alternatives generated by the respondents might influence their confidence in the initial assessment in an unusual fashion. Specifically, the prompt to generate alternatives contains a potential problem. If I had presented a specific set of alternatives, they could have been pretested and adjusted until most individuals found them to be equally good in quality (or some other relationship). In contrast, when individuals generate alternatives, they will not all be of subjectively equal quality. Some people can be expected to think of alternatives that they think are excellent, while others will not. Although it was too late to ask participants what they thought of their alternatives, I examined these data to determine if it would be possible to judge the strength of alternatives that they put

forth. I developed a simple, three-point judging scale. For each participant in who was asked to generate an alternative, two raters (myself and one blind judge) compared the alternative to the original response. Some individuals provided alternatives that were clearly more developed than the initial response. For these respondents, the process of asking for another reason appeared to stimulate them to think more deeply, as proposed by Koehler (1991). These responses were scored as a "2." Most individuals simply complied with the instructions, generating an alternative that was not either substantially more or less developed than the first explanation they supplied. These responses were assigned a "1." Finally, a few participants either did not generate an alternative or clearly did not take the alternative seriously. Examples of the former case include both students who listed the same alternative twice, as well as a respondent who simply admitted that they couldn't think of an alternative. In the latter case, a few participants wrote a lengthy first response, followed by an alternative that, in essence, said something like "Maybe I'm just wrong." These responses were assigned a "0." The inter-rater reliability between my assessments and those of the blind judge was .65 (using the Spearman-Brown formula).

The correlation between my ratings and the subjects' confidence assessments was strong ($r = .67$, $n = 30$). The relationship between the blind judge's ratings and the confidence assessments was less strong, but in the same direction ($r = .23$, $n = 30$). Because my own observations were not blind (and, thus, likely to be biased), I used the blind judge's ratings for the subsequent analysis.

Results of a regression using the blind judge's rating of alternative development, along with the RSVP manipulation (dummy coded), to predict confidence are depicted in Table 17. Although the results for alternative development fell short of significance ($p < .10$, with the 95% confidence interval slightly overlapping zero), the presence of a more developed alternative appeared to be associated with reduced confidence in the initial explanation. Stated differently, less developed alternatives may be associated with greater

confidence in the original explanation. In this sub-sample, with alternative development included as a covariate, the RSVP technique remained a significant predictor. Overall, the regression predicted 19% of the variance in confidence in this sub-sample (adjusted R^2).

Predictor	Unstandardized β	Std β	Test Statistic	P-Value
Constant	12.325		$t=6.96$	.00
RSVP	-2.84	-.52	$t=-2.76$	.01
Alternative Development	-2.68	-.33	$t=-1.73$	.09

+n=27

Table 17: Alternative Development and RSVP as Predictors of Confidence+

Since there was no effect of the confidence manipulation on confidence, the explanation that quality of alternative influenced only the amount of debiasing, and not whether it occurred, seems to be ruled out. Rather, in light of these findings, it seems that the prompt to "consider alternatives" may have caused some subjects to formulate a weak alternative, which then increased their confidence in the original answer. For these respondents, the debiasing manipulation was actually increasing the bias.

Summary of Study 2

The results of this study appear to confirm some findings about the effects of the RSVP technique for structuring diagnosis. First, the RSVP manipulation seems to have some effects on initial explanations. However, the judges in this study clearly saw more subtleties in the responses than the judges in Study 1, and the strongest effect may be to cause individuals to defer making an initial explanation. Even with only a small remaining number of individuals after the non-explainers were removed, Hypothesis 2 was still supported, and the effect was in the direction predicted by Hypothesis 1.

Other results in Study 2 were even more complex. Hypothesis 5, unsupported in Study 1, was supported here. Unexpectedly, though, the confidence manipulation failed to reduce confidence. At the same time, the confidence debiasing manipulation had the predicted effect of reducing willingness to act. Given the small sample size, the villain could be a lack of power. Additional analysis suggested that the quality of alternatives is an important issue in debiasing confidence. Because previous studies have asked individuals to react to alternatives rather than generate them, this issue may not have received the attention it deserves. However, the effects of variance in alternative plausibility on confidence found here can be seen as one possible explanation for the much-discussed difficulty in debiasing confidence (c.f., Arkes et al., 1988; Fischhoff, 1982; Fischhoff & Macgregor, 1982; Koriati et al., 1980).

In order to resolve the basic discrepancies surrounding Hypotheses 3 and 5, which were supported in either Study 1 or Study 2, but not both, and to confirm the effects of structured diagnosis on deferring initial explanation, a study with more power is required. This is a major goal of Study 3. Drawing on the present findings, Study 3 will also test the following hypothesis:

H12: The structured diagnosis manipulation will increase the number of individuals who refrain from making an initial explanation.

CHAPTER 5

STUDY 3 METHOD AND RESULTS

This chapter reports the method and results for Study 3. Study 3 was conceived as a replication of Study 1, but with three important differences. First, the sample size was increased in order to alleviate the concerns about power seen in Study 1. Second, the inferential goal manipulation was strengthened. Third, in response to reviewer questions about whether the manipulation check questions were really manipulation checks, or, alternatively, part of the manipulation itself, a second questionnaire was devised. This questionnaire, given to both groups, contained only the appropriate experimental materials (the Big Bite case, and, in the structured diagnosis conditions, information on the RSVP framework) and the manipulation check items. So, in Study 3, although participants in the structured diagnosis conditions continued to receive information on the RSVP framework, no participant received any question specifically requiring them to link knowledge of the framework to the Big Bite case until after they had completed the main questionnaire. This study, then, provides a stronger test of whether the RSVP information itself, in the absence of any prompt to link it to the case data, can produce effects.

Method

Study 3 examined two independent variables using the same 2 x 3, between-subjects factorial design found in Study 1. The first, structured diagnosis, was divided into two conditions (presence or absence of the manipulation). The second, inferential goal, was divided into three conditions (dispositional goal prime, situational goal prime, no prime). The experimental task was the same as in Study 1.

Participants and Procedure

Research participants were 217 students enrolled in an upper-level undergraduate course in organizational behavior at The Ohio State University. One hundred thirty nine (64%) were currently employed, working an average of 11.5 hours per week. Students were given the experimental materials on the first day the course. The full solicitation script is reprinted in Appendix B. A small inducement, in the form of extra credit points, was given to encourage participation. Students who did not wish to participate had the option of completing an alternative assignment to receive the extra points. Participants were randomly assigned to conditions. Participants completed the materials confidentially and returned them to the researcher in subsequent class periods. At the time that the main questionnaire was turned in, participants were given the second questionnaire appropriate to their experimental condition (determined by numbers on the front of the first questionnaire). The second questionnaire, which contained only the manipulation check items, was returned during a subsequent class. Debriefing was conducted via a written form (see Appendix E). Otherwise, procedures were identical to those used in Study 1.

Manipulation of Independent Variables

As mentioned earlier, inferential goal was included as an independent variable in Study 1. Results of Study 1 suggested that the manipulation in use may have been too subtle, and, consequently, may have been disregarded by the subjects. Accordingly, a more aggressive operationalization of the manipulation seemed to be in order. In Study 3, I modified the

same basic question more substantially in the experimental groups. Participants in the "situational" inferential goal condition responded to the question "Based on the information given, what do you think is happening in this situation (the work environment at Big Bite, including both store and company factors)?" Participants in the "dispositional" inferential goal condition responded to the question "Based on the information given, what do you think is happening with these people (the employees)?" Control condition participants, as in previous studies, received the question "Based on the information given, what do you think is happening?" The manipulation for structured diagnosis was conducted as in Study 1.

Measures of Dependent Variables

Measures of dispositionalism, situationalism, and information search were identical to those used in Study 1. The measures of confidence and willingness to act were identical to those used in Study 2. Judges for the initial explanation task in Study 3 were two advanced undergraduates in psychology. Judges for the information search task were the same two graduate students used in Study 2 (one in sociology, one in psychology).

Results

Testing the effectiveness of the structured diagnosis manipulation with manipulation check items was expected to be difficult, since these items were expected, in and of themselves, to have some properties of a manipulation. That is, when respondents are prompted to attend to the replications, settings, views, and people in the problem, this may cause them to see the problem differently, and they may be able to get some of the benefits of the manipulation without the benefit of additional explanation. This concern was validated to some extent in the data. All respondents showed the same general pattern in scoring the scenario, regardless of whether or not they received the RSVP manipulation (as in Study 1, this pattern involves relatively high means for replications, settings, and people, and a low mean on the views dimension). Participants who received the RSVP manipulation showed a more exaggerated pattern, with slightly higher means for replications, settings, and

people, and a slightly lower mean for the views dimension. Only the difference on the settings dimension approached significance ($p=.06$). Three possible explanations exist for this finding. First, as expected, the manipulation check items themselves may have some effects on responses. In effect, the inclusion of these items in Studies 1 and 2 may have constituted an enhancement to the manipulation. Second, the manipulation could have no effects on how individuals see the case. In other words, it could be true that RSVP does not affect judgments, and the questions about its dimensions functioned as demand characteristics, leading subjects to the appropriate responses in the first two studies. I considered this unlikely, since the questions in these experiments that assessed the dependent variables did not reference the dimensions of RSVP in any way. Still, since this possibility was raised by two reviewers of Study 1, I felt compelled to investigate it. Third, respondents may attend to this information when prompted, but not use it in their judgments. Since manipulation check information in the present study was not presented to the respondents until after their responses to the questions testing the hypotheses were collected, the hypothesis tests themselves should allow us to determine which alternative is correct.

Tests of Hypotheses Establishing the RSVP Technique

Hypotheses 1 and 2 proposed that subjects who received the RSVP manipulation would display increased situational and decreased dispositional tendencies in their initial explanations. As in Study 1, ratings from two judges blind to hypotheses and conditions were used to measure the presence or absence of situational and dispositional explanations. Inter-rater reliability was acceptable for both situational ($\phi=.79$) and dispositional judgments ($\phi=.78$), and all disagreements were resolved by the raters without interference from anyone with knowledge of hypotheses and conditions. As in Study 2, a substantial number of individuals ($n=22$) declined to make an initial explanation to either person or situation factors. As in Study 2, I conducted the analysis two ways. First, I coded these responses as

zeros for both dimensions. Tables 18 and 19 show the frequency of situational and dispositional explanations using this coding. The results suggested, contrary to predictions, that structured diagnosis reduced situational explanations ($\chi^2(1) = 4.38$, $\phi = .14$, $p < .05$) and had no effect on dispositional explanations ($\chi^2(1) = 1.67$, $\phi = .09$, $p = .20$).

	Control Group	Structured Diagnosis	Totals
No Situational Exp.	20	33	53
Situational Exp.	89	75	164
Totals	109	108	217

+No explanation coded as zero

Table 18: Situational Explanations By Condition+

	Control Group	Structured Diagnosis	Totals
No Dispositional Exp.	44	53	97
Dispositional Exp.	65	55	120
Totals	109	108	217

+No explanation coded as zero

Table 19: Dispositional Explanations By Condition+

Since it was possible that the large number of non-explainers was exerting an influence on the results, I also performed the analysis with these records coded as missing data. In this analysis, structured diagnosis had no effects at all on the presence of either situational or dispositional explanations (χ^2 approximately zero for both). Hypotheses 1 and 2 were not supported in Study 3.

Hypothesis 12 predicted that structured diagnosis would increase the tendency to refrain from making an initial explanation. Table 20 shows the contingency table used in

this analysis. Analysis of the effects of structured diagnosis on non-explanations revealed a significant effect ($\chi^2(1)=13.11$, $\phi=.25$, $p<.01$).

In the present study, three individuals not receiving the structured diagnosis manipulation were scored as declining to make an initial explanation in terms of either person or situation factors. In order to understand these data more fully, I examined each of the responses coded as "no explanation." Appendix F shows the full responses of the "non-explainers" in the control group. Since several of the deferrals offered in the experimental condition were quite long, excerpts are provided for these individuals' responses. While some of the same sentiments expressed by the control group non-explainers are present in the responses of the experimental participants, the experimental participants also offered a variety of other data in defense of their non-answers. The most common themes in these responses were the lack of information on which to make an explanation and the specific weaknesses of the available data. Hypothesis 12 was supported.

	Control Group	Structured Diagnosis	Totals
No Explanation Given	3	19	22
Explanation Given	106	89	195
Totals	109	108	217

Table 20: Effects Of Structured Diagnosis On Deferring Explanation

Hypothesis 3 proposed that structured diagnosis would result in reduced willingness to act. Cronbach's alpha for the four-item willingness to act scale was .69, and

no item deletions would have increased it. As in Study 1, I used ANOVA to analyze the effects of structured diagnosis and confidence (included as a covariate) on willingness to act. Table 21 shows the means, standard deviations, and correlations for the variables in this analysis. As in Study 2, a problem existed in using responses with no explanation in this analysis, since it is unclear what the confidence estimates these individuals provided would represent. On one hand, it might be possible to regard these confidence estimates as representing confidence in the decision not to make an explanation. If this were the case, then confidence in an explanation would be properly at the lowest scale point. However, there is no way to know whether this is actually an appropriate scoring or not. Therefore, the most prudent decision seemed to be to leave out these items. Results of the analysis are listed in Table 22. As in Study 1, both structured diagnosis and confidence were significant predictors of willingness to act. Structured diagnosis decreased willingness to act, and confidence was associated with increases in it. The mean willingness to act for subjects receiving the structured diagnosis manipulation was 9.6, versus 10.8 for those not receiving the manipulation. The overall regression equation explained about 13% of the variance in willingness to act (Adjusted $R^2=.13$). As a check, I also ran the analysis with the data of the non-explainers included. The results remained significant. Hypothesis 3 was supported.

Variable:	Mean	SD	RSVP	Con	WTA
RSVP	N/A	N/A	1.00		
Confidence	9.01 [*]	1.98	-.11	1.00	
Willingness to Act	10.17	3.71	-.23 [*]	.33 [*]	1.00

^{*} $p<.01$

Table 21: Means, SD's, and Correlations between Willingness to Act and its Predictors

Source	SS	MS	F	P
RSVP	73.54	73.54	5.95	.02
Confidence	265.11	265.11	21.44	.00
Error	2361.33	12.36		

+n=194

Table 22: Effects of Structured Diagnosis and Confidence on Willingness to Act+

Hypothesis 4 stated that the structured diagnosis manipulation would increase the breadth of information search. As in Study 1, breadth was operationalized as the number of categories in the RSVP framework in which questions were asked, as scored by two judges blind to hypotheses and conditions. Inter-rater reliability was acceptable for all dimensions ($\phi = .62, .87, .78$, and $.77$ for the replications, settings, views, and people dimensions, respectively), and all disagreements between judges were successfully resolved. Three individuals left this section blank and were removed from the analysis, leaving 214 participants. Analysis of the contingency tables found a relationship between presence of the RSVP manipulation and questions about views of the problem ($\chi^2(1) = 6.35$, $\phi = .17$, $p = .01$), but no other dimensions approached significance. Thus, as in both previous studies, partial support was found for Hypothesis 4.

Hypothesis 5 proposed that the use of structured diagnosis would decrease confidence. As mentioned above, a substantial number of subjects declined to provide an initial explanation, and these individuals were excluded from the analysis. Structured diagnosis was weakly, negatively, and not significantly correlated with confidence. This was true whether the "non-explainers" were included in the analysis ($r = -.12$, $p = .09$, $n = 215$) or not ($r = -.11$, $p = .14$, $n = 194$). Hypothesis 5 was not supported.

Tests of Hypotheses Related to Inferential Goals

Hypotheses 6 and 7 predicted that use of a prime to hold either a situational or dispositional inferential goal would result in an increase in initial explanations that were consistent with the prime and a decrease in explanations not consistent with the prime. Twenty-two cases where no explanation was offered were excluded from the analysis ($n=195$). Though these hypotheses were not supported in Study 1, use of the strengthened situational prime in Study 3 resulted in an increase in initial explanations containing situational factors ($\chi^2(1)=4.17, \phi=.15, p<.05$). Use of the strengthened dispositional prime in Study 2 resulted in an increase in initial explanations containing dispositional factors ($\chi^2(1)=17.37, \phi=.30, p<.01$). Tables 23 and 24 show the contingency tables for these analyses.

	No Dispositional Goal	Dispositional Goal	Totals
No Dispositional Exp.	62	13	75
Dispositional Exp.	64	56	120
Totals	126	69	195

+No explanation coded as missing data

Table 23: Dispositional Explanations By Dispositional Goal+

	No Situational Goal	Situational Goal	Totals
No Situational Exp.	26	5	31
Situational Exp.	107	57	164
Totals	133	62	195

+No explanation coded as missing data

Table 24: Situational Explanations By Situational Goal+

Similarly, priming subjects to hold a situational inferential goal reduced dispositional explanations ($\chi^2(1)=6.64$, $\phi=-.19$, $p=.01$). However, the dispositional prime was less successful at reducing situational explanations ($\chi^2(1)=2.73$, $\phi=-.12$, $p=.10$). Overall, Hypotheses 6 and 7 were supported in this study.

Hypotheses 8 and 8a predicted two different types of interaction between the structured diagnosis and inferential goal manipulations, but no such relationships were found. As discussed above, the only effect of structured diagnosis on initial explanations in Study 3 was to cause individuals to seek to defer judgment until more information could be collected. Hypotheses 8 and 8a were not supported.

Summary of Study 3

The present study used a weaker manipulation than Studies 1 and 2, but benefited from a larger sample. As a result, some findings from the previous studies were upheld, while others were cast into doubt. With regard to the effects of structured diagnosis on initial explanations, the major effect of RSVP seems to be to cause individuals to defer judgment in order to collect more data. This finding was echoed in the effects on willingness to act. In this study, then, the prominent effect of RSVP was to cause individuals to delay judgment to collect more information. Once again, partial support was found for the hypothesis that RSVP would improve the breadth of information search. The dimension for which effects continued to be supported was "views."

This study also showed effects of inferential goals on initial explanations. With the weaker version of the RSVP manipulation used here, the effects of inferential goals and structured diagnosis appear to be conceptually complementary. That is, given that they have effects on different dependent variables (type of initial explanation vs. deferral of initial explanation), a case can be made that a superior strategy would be to use both manipulations in combination to minimize dispositionalism. Finally, no significant effect of structured

diagnosis on confidence was found, suggesting that these two variables may be basically unrelated.

CHAPTER 6

DISCUSSION

The three studies reported in the previous chapters provide substantial evidence that structured diagnosis can productively alter the early stages of judgment in complex people problems. My major conclusion is that the RSVP technique is a potentially valuable decision aid. A second conclusion is that inferential goal priming can affect initial explanations in a more deliberative, naturalistic, problem-oriented judgment task than the one used by Krull (1993). In this chapter, I examine each of these effects. I then turn to the lack of findings in the area of confidence, discuss limitations, and suggest directions for future research.

Effects of RSVP

All three studies reported here found evidence for the effectiveness of the RSVP technique as a means of structuring diagnosis of complex people problems. Specifically, the RSVP technique appears to have desirable effects on initial explanations, subsequent information search, and willingness to act without information search.

RSVP Affects Initial Explanations

Across all three studies, the RSVP technique was found to either decrease dispositionalism in initial explanations of a complex people problem (Studies 1 and 2) or delay an explanation altogether, pending the collection of more information (Studies 2 and

3). In Study 3, the study with the most power, the only effect was delay. I can think of two possible reasons for this difference in findings. First, Study 3 used a weaker manipulation than the one used in Studies 1 and 2. The difference in these manipulations was that, in the first two studies, participants were prompted specifically to link the framework to the case data. Without this prompt, some participants in Study 3 may have simply not used the decision aid. Failure to use decision aids has long been recognized as a nagging problem across multiple domains (e.g., Arkes, Dawes, & Christensen, 1986), and this may simply be another example. A weakness in this argument, though, is that RSVP had other effects, suggesting that, overall, it was utilized by many participants who were exposed to it. Second, Study 3 used a different administration technique. While the first two studies were conducted in-class, Study 3 was given as a take-home assignment. This may be a non-trivial difference, since it removes the time pressure that is one way of operationalizing cognitive load. To the extent that individuals were able to utilize this time to ease their cognitive load, Gilbert and his colleagues (e.g., Gilbert et al., 1988) would predict that the participants would correct their inferences to effectively take into account both situational and dispositional information.

RSVP Affects Information Search

In all three studies, use of the RSVP technique caused participants to search for information from different perspectives. Table 25 shows the total proportions of respondents asking questions about views in each study. Overall, 29% of participants exposed to RSVP asked questions seeking others' perspectives, while only 11% of control participants sought such information.

Study	No-RSVP Conditions	RSVP Conditions	Totals
1	4/26 (15%)	10/26 (38%)	14/52 (27%)
2	1/30 (3%)	11/33 (33%)	12/63 (19%)
3	13/108 (12%)	27/106 (25%)	40/214 (19%)
Totals	18/164 (11%)	48/165 (29%)	66/329 (20%)

Table 25: Questions About Views by Study

In the present studies, the RSVP technique did not have consistent effects on the other dimensions of information search (e.g., replications, settings, people). This may be an artifact of the structure of the Big Bite case. Recall that, across all three studies, participants scored the case as relatively high in information on replications, settings, and people, and low on views. Thus, it makes sense that main difference in information search would be concentrated on the views dimension. A possible critique of the findings in Studies 1 and 2 is that the pre-questionnaire prompt to those receiving the RSVP manipulation, which instructed them to score the Big Bite case on the degree to which it contained information on each dimension, may have constituted a demand characteristic. Study 3 answered this critique by moving the scoring of the case on the RSVP dimensions to a second questionnaire, completed after the dependent variable measures were collected. As noted in the chapter reporting Study 3 results, participants in both RSVP and non-RSVP conditions, when prompted, tended to score the case similarly. Yet, the differences shown above for Study 3, which were measured prior to these prompts, still occurred. This rules out the possibility that demand characteristics were responsible for these findings.

RSVP Reduces Willingness to Act Without Additional Information Search

A slightly more contentious conclusion is that RSVP reduces willingness to act. Supported in Studies 1 and 3, this finding was not significant in Study 2. However, the

effects were in the same direction in all three studies, and power was a problem in Study 2. Table 26 shows the sample, effect size, and power information for the relationship between RSVP and willingness to act.

Study	Overall n	Effect Size r	Usable n for Analysis	Observed Power
1	52	-.19	51	.54
2	63	-.14	51	.06
3	217	-.23	194	.68

Table 26: Comparison of Effects and Power for the RSVP-Willingness to Act Relationship

In addition to the obvious problem of a small effect size, the "power outage" experienced in Study 2 was also affected by the fact that a third of the participants who received the structured diagnosis manipulation declined to make an initial explanation, and, therefore, had to be excluded from the analysis. As a result, only 22 individuals were present in the RSVP cell of the experiment. I did re-run the analysis with the full sample, using the confidence estimates that the non-explainers supplied, and I got a stronger, though still not significant, effect for structured diagnosis. This analysis is flawed, though, in my opinion, since there is just no way to know for sure what the confidence estimates provided by the individuals who did not offer an explanation represent.

It must also be noted that the measure of willingness to act is not identical across these three studies. Study 1 used a single-item measure which put the respondent in the role of a consultant. In Studies 2 and 3, three additional items were added to create a four-item measure of willingness to act. However, in Study 2, the measure with the best scale reliability was a three-item subset that did not include the item used in Study 1. Including

the fourth item did not make a difference in whether or not RSVP predicted willingness to act (it did not in either case). In Study 3, the four-item scale had the best reliability.

Overall, the two studies which had the power to detect effects did. Further, in Study 3, where power was the highest, the manipulation was the weakest, and the measure of willingness to act with the broadest construct coverage (four items) was found to be the most reliable, structured diagnosis had a significant effect on willingness to act.

Effects of Inferential Goal Priming

Although Study 1 found no effects of inferential goals, in Study 3, using a stronger manipulation, inferential goal did affect initial explanations as predicted. This required what, in my opinion, was a fairly heavy-handed prime, which, in the situational-prime condition, prompted respondents to consider specifically "the work environment at Big Bite, including both store and company factors." Even then, the effect of the situational prime was not large. However, there is reason to suspect that the effect size found in Study 3 may be a conservative estimate of the true effects of a situational prime. The Big Bite case appears, on average, to be seen, at least by the present population of college students, as having situational explanations. Of the 195 participants who offered an initial explanation in Study 3, 164 (84%) cited situational factors in their explanations (the figures were 82% and 75% in Studies 1 and 2, respectively). In fact, 80% of participants not primed with a situational goal in Study 3 (107/133) cited situational factors in their explanations. With such a high base rate, the room for the manipulation to operate to increase situational explanations was not large. Indeed, almost all individuals (57/62, or 92%) primed with a situational inferential goal made situational explanations.

In contrast, the dispositional goal prime was much more effective. While the Big Bite case also appears, on average, to be seen as having dispositional explanations, this tendency is smaller and less consistent. For example, in Study 3, 120 out of 195 participants (62%) who offered an initial explanation cited dispositional factors. This

figure was consistent with Study 1 (63%), but, in Study 2, only 37% of participants cited dispositional factors in their explanations (the lower figure in Study 2 may be due to the lack of a dispositional goal manipulation). This lower base rate left more room for the dispositional goal prime to work, and this may have contributed to the larger effect ($\phi=.30$).

Confidence

Research participants were, on average, pretty confident of their initial explanations. Means on the confidence measures were above the midpoint of the scale in all three studies. The findings related to confidence were mixed. On one hand, confidence was associated with willingness to act in Studies 1 and 3, and, depending on the measure used, it may have had an association in Study 2 as well. On the other hand, support for hypotheses about confidence was spotty. Despite some evidence of effects in Study 2, in Study 3, when plenty of power was available, the correlation between structured diagnosis and confidence was weak and not significant. Coupled with the finding in the wrong direction in Study 1, I think the best conclusion is that any effects of structured diagnosis on confidence are trivial.

Most surprising was the minimal effectiveness of the confidence manipulation in Study 2. With no effect at all on confidence, and only a marginally significant effect on willingness to act, one wonders what is going on! I believe that the key lies in the supplementary analysis of subjective alternative quality, but the present research was not designed to convincingly demonstrate anything about alternative quality. This one will have to wait for another day, but, given the ambiguity of the Big Bite case, the fact that so many research participants were so confident of their explanations highlights the value of pursuing this line of research in the future.

Limitations and Future Directions

The present research is limited in several ways. As in all laboratory research, compromises in realism were made for the sake of control. Despite efforts to ensure that the vignette was realistic, it is possible that these effects will not generalize to real problems

in real managerial contexts. A further limitation stems from the use of students. Since the majority of these students work, they may represent the population of managers more closely than many student samples. Also, neither employment status nor hours worked was significantly correlated with any variable under study in either Study 2 or 3 (all students reported being employed in the Study 1 population, and hours were not collected). Still, it is certainly possible that a population of seasoned, professional managers might respond differently. Perhaps the strongest limitation of the present research is the use of a single experimental case representing a single type of people problem. Other types of people problems may provoke different responses.

Future research should, of course, attempt to generalize this research to other settings and populations, using other problem types. An important step would be to secure access to a population of experienced managers. Given the wide variety of findings on dispositionalism cited here, however, including some using field samples (e.g., Hofmann & Stetzer, 1998), it would seem that there is reason to be optimistic about the utility of these results.

Aside from the issue of generalization, there are several other potentially productive research projects suggested by the studies reported here. As mentioned above, much work remains on the issue of debiasing confidence. To some extent, the RSVP technique appears to bypass the issue of confidence, mitigating its effects on willingness to act without really affecting the thing itself. Given its significant association with willingness to act, however, additional work on debiasing confidence directly would appear to be in order. Results from Study 2 suggest that looking at subjective alternative quality is one place to start.

Another issue worth looking at in more detail is the improvement of information search. While the RSVP technique had significant effects on search for information from multiple perspectives, I note with dismay that, even though almost all participants, regardless of condition, were able to recognize the lack of multiple perspectives in the Big Bite case

when prompted, only about a third of participants who received the structured diagnosis manipulation sought out more information on views. That was more than double the rate seen in the control groups, but, given the potential importance of multiple perspectives in understanding people problems, it still seems disappointingly low. Future research might examine how structured diagnosis could be strengthened to further encourage this type of search behavior.

A third avenue for study would be the utilization of information in making a final judgment. The present research did not carry through to the final judgment stage. Obviously, though, the ultimate test of a decision aid is whether it results in better judgments, with "better" meaning judgments that utilize more relevant (diagnostic) information. While research cited above seems to indicate that individuals will often use diagnostic information, future attention could be given to whether or not the framework of the RSVP technique aids in the utilization of information.

Conclusion

The studies reported here contribute to knowledge by providing development and initial validation of a new tool to improve managerial diagnosis of people problems. Overall, the results suggest that the RSVP framework, particularly if used in combination with a situational goal prime, should reduce dispositionalism, encourage delays in initial judgment, increase the breadth of information search, and decrease willingness to act without searching out needed information in complex people problems. In addition, the RSVP technique is easy to learn and use. Given the available knowledge about shortcomings in human judgment, use of the RSVP framework may, therefore, provide valuable benefits to managers.

APPENDIX A

TROUBLE AT BIG BITE

Mary is a district manager at Big Bite, a fast food burger chain. Mary tells you that she can't get any decent employees. "They're all lazy, and they steal! They know that we have security cameras in the restaurants, but we catch them stealing food and beverages anyway. They know that they're supposed to pay for the food, but they just take it. Then we have to fire them, which makes our labor shortage worse. If only I could have some honest employees!"

Mary is in charge of about 100 Big Bite stores in three states, so she doesn't get to any particular one that often. But, she tells you, "Each store has its own store manager, and the situation is basically the same across all 100 stores, so it can't be the managers' fault."

APPENDIX B

SOLICITATION SCRIPTS

Study 1 and 2: Organizational Psychology Class

As part of our unit on research methods in organizations, we are going to engage in a little in-class research. In a minute, I am going to pass out a short case study to each of you that describes a fictional "people problem" in a work setting. Attached to this scenario is a questionnaire, which asks you for your impressions about the problem presented. My goal is to use your responses to this questionnaire for two purposes: 1) to improve our understanding, as a class, of the role of research methods in responding to organizational problems, and 2) to collect some data that may be used for my own research.

Because of this second goal, I want to be clear in telling you that you do NOT have to participate in filling out this questionnaire. This is strictly voluntary, and there are no penalties of any kind for choosing not to participate. If you do choose to participate, please do not write your name or otherwise identify yourself in any way on the questionnaires.

Now, here's my plan for how we'll use this: First, I'll pass out the questionnaires, and those who are willing to participate will complete them and turn them in. By turning in a questionnaire, you are consenting to participate in the research. Again, no names please – just fill them out anonymously. Even if you are not willing to participate in the questionnaire, you'll probably want to look over the case study in the packet, since we're going to go over it in class.

When you read the case, you'll see that you are being asked to take on a consulting role. Please try to imagine yourself in the role. This is useful for the research, but most important for us to make the best use of the information for your learning, because research suggests that the opportunity to participate in an experience and receive feedback on it enhances the learning of skills in many situations. There are several versions of the questionnaire that differ slightly, so please just focus on the material you have in your packet.

After those who are willing turn in the questionnaires, I'll review the study design, the hypotheses, and we'll go over the case itself. Then, next time, I'll bring back the aggregated data (no individual responses will be shown) and we'll get some insight into the kind of feedback that comes out of applied research. We don't have a lot of people here, so it's possible we'll have inconclusive data, but that's a part of the reality of applied research that we have to address as well. We'll talk about what the options are in that case, too.

Any questions or concerns?

Study 2: Personnel Psychology Class

My name is Rob Litchfield, and I am a graduate student in industrial/organizational psychology. I am interested in how individuals solve problems they may encounter at work. In a minute, I am going to pass out a short case study to each of you that describes a fictional "people problem" in a work setting. Attached to this scenario is a questionnaire, which asks you for your impressions about the problem presented.

The main purpose of this exercise is to collect data for research purposes. Because of this purpose, I want to be clear in telling you that you do NOT have to participate in filling out this questionnaire. This is strictly voluntary, and there is no penalty for choosing not to participate.

Why might you want to do this? Well, aside from satisfying your altruistic impulse to help me out, your instructor has generously offered to provide extra credit points for participation in the study. You may still earn these points in another way, however, and I have copies of an alternative assignment here for anyone who prefers not to participate in research. The alternative assignment will also get you the same extra credit. For those of you who are about to ask, I'm sorry, but you cannot earn double extra credit by doing both the research and the alternate assignment.

For those of you who are willing to participate, let me highlight a couple of things:

First, when you read the case, you'll see that you are being asked to take on a consulting role. Please try to imagine yourself in the role. There are several versions of the packet that differ slightly, so please just focus on the material you have, and do not worry about whether anything is missing or consult others who may have different information to respond to. Please do this on your own.

Second, and most important, please write your name (CLEARLY) on the cover page ONLY. I need your name in order to provide you with credit, but I will not enter your name into any research database. This cover sheet will be discarded after you have been awarded credit. Please do not identify yourself on any page other than the cover sheet.

When you turn in your packet, you'll receive materials which discuss the research and relate it to management topics.

Any questions?

Study 3

My name is Rob Litchfield, and I am a graduate student in industrial/organizational psychology. I am interested in how individuals solve problems they may encounter at work. In a minute, I am going to pass out a short case study to each of you that describes a fictional "people problem" in a work setting. Attached to this scenario is a questionnaire, which asks you for your impressions about the problem presented.

The main purpose of this exercise is to collect data for research purposes. Because of this purpose, I want to be clear in telling you that you do NOT have to participate in filling out this questionnaire. This is strictly voluntary, and there is no penalty for choosing not to participate.

Why might you want to do this? Well, aside from satisfying your altruistic impulse to help me out, your instructor has generously offered to provide extra credit points for participation in the study. You may still earn these points in another way, however, and I have copies of an alternative assignment here for anyone who prefers not to participate in research. The alternative assignment will also get you the same extra credit. For those of you who are about to ask, I'm sorry, but you cannot earn double extra credit by doing both the research and the alternate assignment.

For those of you who are willing to participate, let me highlight a few things:

First, when you read the case, you'll see that you are being asked to take on a consulting role. Please try to imagine yourself in the role. There are several versions of the packet that differ slightly, so please just focus on the material you have, and do not worry about whether anything is missing or consult others who may have different information to respond to. Please do this on your own.

Second, please remember to bring this with you to turn in during the next class period. If you forget or cannot attend, you may still turn it in on either April 8 or 10. If you do not complete the assignment by April 10, you can still earn extra credit through the alternative assignment, but your chance to receive credit for completing the research materials will be lost.

When you turn in the main questionnaire, there is a second questionnaire that you will pick up. This one is VERY short, and should take you only 3-4 minutes to complete. I will have collection boxes available for these questionnaires as well, beginning next Monday.

Finally, and most important, please write your name (CLEARLY) on the cover page ONLY. I need your name in order to provide you with credit, but I will not enter your name into any research database. This cover sheet will be discarded after you have been awarded credit. Please do not identify yourself on any page other than the cover sheet.

When you turn in your packet, you'll receive materials which discuss the research and relate it to management topics.
Any questions?

APPENDIX C

RSVP MATERIALS

RSVP is a mechanism for structuring understanding of people problems. It has four components:

- 1) **Replications** – the number of times a given behavior has been observed.
- 2) **Settings** – the number of situations, relationships, or locations in which the behavior is observed.
- 3) **Views** – the number of perspectives on the problem (e.g., managerial, employee, customer). Note, this is not the number of people who have seen the problem. It is the number of different perspectives represented in the diagnosis. For example, in an argument between two groups of employees (like sales people versus accounting staff), each group has a perspective. A better diagnosis takes both sides into account.
- 4) **People** – the number of people engaged in (or, who would be expected to engage in) the behaviors.

RSVP is intended to be used as a tool to aid in understanding complex problems in the workplace. By critiquing your representation of an actual problem using the framework, you may be able to get a better idea of what's happening. If a proposed description of a problem is weak in one or more of the dimensions of RSVP, then this may be a cue to get more information before making a decision.

To use RSVP to understand what is going on in a problem involving people in the workplace, follow these steps:

- 1) Go through the information presented to you in the problem
- 2) Evaluate the information you have in terms of each dimension of RSVP. Do you have information on all of the dimensions? How much?
- 3) If you have only limited information on one or more aspects of RSVP, this is an indication that you may want to ask more questions, get more details, etc.
- 4) When you have information from multiple views on the number of replications, the number of settings, and the number of people, then you are in a better position to accurately assess what is going on. The number of views is particularly important, since what looks ridiculous or irrational from one perspective may be very logical from another.

APPENDIX D

JUDGE TRAINING ITEMS

Initial Explanations

The trained scorings are in parentheses after the items, with S=situation, P=person, B=both, and N=neither. Items were presented to judges in sets, then discussed (not all items below were rated at once). All items marked with "N" were added for training for Studies 2 and 3 only.

Training items used for Study 1

The employees feel like they should get the food because they are not paid enough. (P)
The employees are not being treated fairly. (S)
The employees probably don't get paid enough to live. (S)
The employees probably don't think of it as stealing, because the food is there and they need to eat. (B)
The job probably isn't worth much to them. (P)
Not using a good hiring process. (S)
They need to weed out bad employees. (P)

Training items added for Studies 2 and 3

Mary needs to spend more time visiting the stores. (N)
It's a training problem: Managers and employees are not properly trained on what they can and can't do. (S)
Employees don't realize that taking food could get them fired. (S)
Managers need to have better relationships with employees. (S)
They need to do reference checks before hiring. The workers have low education, and are probably not good quality. (B)
Mary is over-generalizing this problem. (N)
Mary is just not aware of everything that is going on. She only sees the bad things. (N)
The rules are not tight enough. I wouldn't be surprised if managers are stealing too. (S)
The employees probably have to take the food because of their low wages. (S)
Mary says the employees are dishonest. She thinks that this is causing the labor shortage.(N)
Big Bite needs some employees who want to work. (P)
They may be hiring the wrong kind of employee, but they also need to fix their management system. (B)
They need to think about how they are hiring. (S)

People are stealing food and getting caught. I think Mary needs to look for more information. (N)
 Employees think they can get away with stealing, so they do it. (P)
 Big Bite probably does not pay enough money for the employees to eat, so they steal. (S)
 The employees don't respect their job, the pay is low, and the managers are probably stealing too. (B)
 Mary isn't there enough to know what is going on. She's biased. (N)
 I think that the employees probably don't know that they aren't allowed to have the food.
 Most fast food restaurants let you eat for free. (S)
 Food and beverages are being stolen. Big Bite is losing money. It could be based on other things. (N)
 Employees feel that they should get the food since they make it. (P)
 The workers need more training. (S)
 The stores may be located in poor neighborhoods, which attract poor workers. (P)
 They are hiring bad employees. (P)

Information Search

Replications: questions dealing with frequency of behavior (stealing).

1 replication= 1 theft by 1 person 1 time

How often? How many times?

Settings: questions about the situation at Big Bite. Wages/hrs. managers, training, policies, hiring systems, security cameras, pay, etc.

Views: questions soliciting perspectives other than Mary's. These must be explicit in seeking out others' perspectives.

Can I interview the employees? What do the managers think? What is the employees' perspective?

People: questions about the number and characteristics of individuals who are or could be expected to be engaged in stealing. Questions here try to get at whether there are or could be expected to be a lot of people stealing - but not because of the setting.

How many people are stealing? What percentage of employees steal? What percentage are caught? What is their education level?

Other sample items

Are they paid enough? (S)
 Have you tried discounting the food? (S)
 What do managers think? (V)
 How do you know it isn't the managers' fault? (S)
 How do they recruit employees? (S)
 How long do employees stay on average? (P)
 Do employees get breaks? (S)
 How do you catch them? (S)
 Who fires them? (S)
 Has this happened for a long time? (R)

APPENDIX E

DEBRIEFING

Thank you for your participation in this study. The purpose of this research is to develop tools to assist managers in solving complex "people problems" in organizations. In responding to the Big Bite case, you may have been given one or more tools to use in your diagnosis. This study is designed to assess whether such interventions have an impact on the way in which people respond to the case. If you did not receive any information about a diagnostic tool, this is because you were part of the "control group," which is used for comparison purposes.

Here's an example to clarify things: The primary tool tested in this study is a diagnostic structure called RSVP. Briefly, RSVP proposes that attention to the number of replications of a behavior, the number of settings in which it occurs, the number of views of the problem present, and the number of people involved are all important for determining what is going on in a given problem. You were randomly assigned to be either prompted or not prompted to pay attention to the components of RSVP when making your diagnosis. By comparing the diagnoses made by those of you who were given the RSVP information to the diagnoses made by those of you who were not, we hope to be able to assess whether this technique for structuring diagnosis alters the explanations people make.

If you have any questions about this study, please contact Rob Litchfield, Litchfield.4@osu.edu, or Robert S. Billings, Billings.2@osu.edu. Thanks again for your help!

APPENDIX F

RESPONSES OF NON-EXPLAINERS

Control Group Participants (Full Responses Quoted):

#11: Big Bite, fast food restaurant, has a really big problem with management of employees behaviors. (stealing and laziness).

#52: I believe that the employees are indeed stealing food like she stated. Having worked in a restaurant for two summers while in high school, I myself witnessed the same situation.

#104: I will organize a kind of championship between different stores. The most efficient store will earn an amount of money. One of the goal of this store is to limit steals. So BB managers will explain to their employees that they've got interest not to (unintelligible word).

Experimental Group Participants (Excerpts Quoted):

#112: ... "The RSVP analysis tells me that this issue is a serious problem that goes beyond the individual stores and is something that Mary needs to solve on the executive level."

#128: ... "There seem to be two perspectives. It seems to be a management versus employee problem."

#129: "Because Mary has so many stores, she needs to have a meeting with the store managers to discuss the situation."

#130: ... "Mary has made an assumption about all employees based on limited information."

#133: ... "Since Mary is in charge of such a large number of stores, her generalization about everyone can't be correct."

#138: Blank (remainder of questionnaire, including confidence and questions, filled out).

#140: ... "Mary may not have an accurate view of what is happening."

#141: ... "The information is ambiguous, one-sided, and insufficient to support the broad claim that Mary makes"...

#159: "They are stealing food."

#161: "It is not enough info to conclude what is happening to the employees. They were not asked, weren't interviewed, and it is hard to tell here."

#167: "They are all being grouped into one category."

#184: "All the managers (including Mary) didn't try to find out the reason why employees steal"...

#191: ... "Mary is generalizing that this happens at "all of them."

#197: "Replications: The information does not specify the number of times the behavior has been examined"... (similar for each dimension)

#200: "It is hard to tell the severity of the problem without asking more questions."...

#202: "I think Mary has a "one way" attitude with all of her employees. Just because she had a few people steal from her, doesn't mean everybody will steal from you. Also she says that basically all of the stores are the same, but she has not seen how the other ones operate. So, she has no right to say that they are the same."

#203: "There are not views, since it is only from the perspective of Mary"... (discussion of each dimension)

#207: "Only some of the employees will steal food"...

#216: "Employees for some reason are not following Big Bite policies which result in the loss of their jobs"...

LIST OF REFERENCES

- Arkes, H. R., (1991). Costs of judgment errors: Implications for debiasing. Psychological Bulletin, 110, 486-498.
- Arkes, H. R. (2001). Overconfidence in judgmental forecasting. In J. S. Armstrong (Ed.), Principles of forecasting: A handbook for researchers and practitioners. Boston, MA: Kluwer Academic.
- Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. Organizational Behavior and Human Decision Processes, 37, 93-110.
- Arkes, H. R., Guilmette, T. J., Faust, D., & Hart, K. (1988). Eliminating the hindsight bias. Journal of Applied Psychology, 73, 305-307.
- Bassok, M., & Trope, Y. (1984). People's strategies for testing hypotheses about another's personality: Confirmatory or diagnostic? Social Cognition, 2, 199-216.
- Burger, J. M. (1981). Motivational biases in the attribution of responsibility for an accident: A meta-analysis of the defensive-attribution hypothesis. Psychological Bulletin, 90, 496-512.
- Devine, P. G., Hirt, E. R., & Gehrke, E. M. (1990). Diagnostic and confirmation strategies in trait hypothesis testing. Journal of Personality and Social Psychology, 58, 952-963.
- Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.
- Erickson, D. J., & Krull, D. S. (1999). Distinguishing judgments about what from judgments about why: Effects of behavior extremity on correspondent inferences and causal attributions. Basic and Applied Social Psychology, 21, 1-11.
- Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288-299.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. Cambridge, UK: Cambridge University Press.

Fischhoff, B., & Macgregor, D. (1982). Subjective confidence in forecasts. Journal of Forecasting, 1, 155-172.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. Journal of Personality and Social Psychology, 78, 708-724.

Gilbert, D. T., Jones, E. E., & Pelham, B. W. (1987). Influence and inference: What the active perceiver overlooks. Journal of Personality and Social Psychology, 52, 861-870.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. Psychological Bulletin, 117, 21-38.

Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. Journal of Personality and Social Psychology, 55, 685-694.

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. Journal of Personality and Social Psychology, 54, 733-740.

Green, S. G., & Mitchell, T. R. (1979). Attributional processes of leaders in leader-member interactions. Organizational Behavior and Human Performance, 23, 429-458.

Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? Journal of Personality and Social Psychology, 43, 89-99.

Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. Research in Organizational Behavior, 20, 1-37.

Heider, F. (1958). The psychology of interpersonal relations. New York: Wiley.

Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. Journal of Personality and Social Psychology, 69, 1069-1086.

Hoch, S. J. (1984). Availability and inference in predictive judgment. Journal of Experimental Psychology, 10, 649-662.

Hofmann, D. A., & Stetzer, A. (1998). The role of safety climate and communication in accident interpretation: Implications for learning from negative events. Academy of Management Journal, 41, 644-657.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. Administrative Science Quarterly, 24, 602-610.

Jones, E. E., (1979). The rocky road from acts to dispositions. American Psychologist, 34, 107-117.

Jones, E. E., & Nisbett, R. E. (1971). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. Morristown, NJ: General Learning Press.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), Nebraska symposium on motivation. Lincoln, NB: University of Nebraska Press.

Kelley, H. H. (1973). The process of causal attribution. American Psychologist, 28, 107-128.

Klein, G. A. (1993). A recognition-primed decision model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsombok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corp.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. Psychological Bulletin, 110, 499-519.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.

Krull, D. S. (1993). Does the grist change the mill? The effect of the perceiver's inferential goal on the process of social inference. Personality and Social Psychology Bulletin, 19, 340-348.

Krull, D. S. (2001). On partitioning the fundamental attribution error: Dispositionalism and the correspondence bias. In G. B. Moskowitz (Ed.), Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition. Mahwah, NJ: Lawrence Erlbaum.

Krull, D. S., & Dill, J. C. (1996). On thinking first and responding fast: Flexibility in social inference processes. Personality and Social Psychology Bulletin, 22, 949-959.

Krull, D. S., & Erickson, D. J. (1995). Judging situations: On the effortful process of taking dispositional information into account. Social Cognition, 13, 417-438.

Krull, D. S., Loy, M. H., Lin, J., Wang, C., Chen, S., & Zhao, X. (1999). The fundamental attribution error: Correspondence bias in individualist and collectivist cultures. Personality and Social Psychology Bulletin, 25, 1208-1219.

Lee, F., Hallahan, M., & Herzog, T. (1996). Explaining real-life events: How culture and domain shape attributions. Personality and Social Psychology Bulletin, 22, 732-741.

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. Journal of Personality and Social Psychology, 47, 1231-1243.

McAuley, E., Duncan, T. E., & Russell, D. W. (1992). Measuring causal attributions: The revised causal dimension scale (CDSII). Personality and Social Psychology Bulletin, 18, 566-573.

McGill, A. L. (1989). Context effects in judgments of causation. Journal of Personality and Social Psychology, 57, 189-200.

Mitchell, T. R., & Wood, R. E. (1980). Supervisor's responses to subordinate poor performance: A test of an attributional model. Organizational Behavior and Human Performance, 25, 123-138.

Pennington, N., & Hastie, R. (1993). A theory of explanation-based decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. Zsombok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing

Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. Journal of Personality and Social Psychology, 42, 593-607.

Reason, J. (1990). Human error. Cambridge, UK: Cambridge University Press.

Robbins, S. P. (2001). Organizational behavior (9th ed.). Upper Saddle River, NJ: Prentice-Hall.

Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social perception processes. Journal of Personality and Social Psychology, 35, 485-494.

Schwenk, C. R. (1988). The cognitive perspective in strategic decision making. Journal of Management Studies, 25, 41-55.

Shaklee, H., & Fischhoff, B. (1982). Strategies of information search in causal analysis. Memory & Cognition, 10, 520-530.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmation strategies, and perceived hypothesis confirmation. Journal of Experimental Social Psychology, 22, 93-121.

Steers, R. M. (1991). Introduction to organizational behavior (4th ed.). New York, NY: HarperCollins.

Trope, Y., & Gaunt, R. (1999). A dual-process model of overconfident attributional inferences. In S. Chaiken & Y. Trope (Eds.), Dual-process theories in social psychology. New York, NY: Guilford Press.

Trope, Y., & Gaunt, R. (2000). Processing alternative explanations of behavior: Correction or integration? Journal of Personality and Social Psychology, 79, 344-354.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Watkins, D., Sachs, J., & Regmi, M. (1997). Confirmatory factor analysis of the revised causal dimension scale: A Nepalese investigation. Psychological Reports, 81, 963-967.