ABSTRACT
        Three influences on the outcomes of literature
searches undertaken as part of integrative research reviews were
examined: (1) the degree of expertise of the searcher; (2) the amount
of information available (i.e., keywords, bibliographics, abstracts);
and (3) the cognitive characteristics of the searcher. Participants
were presented with descriptions of documents varying in topic area
(related or not related to their expertise), amount of information,
and whether the document was relevant or irrelevant to the search.
Results revealed no difference between experts and nonexperts in
their accuracy of judging relevant documents but experts were better
able to discern that irrelevant documents were, in fact, irrelevant.
Judgments based on abstracts were more accurate than judgments based
on keywords or bibliographics, which did not differ, and this effect
was more pronounced for relevant than irrelevant articles.
Participants who had previously conducted more literature searches or
published more research reviews made more accurate judgments as did
searchers high in cognitive complexity or tolerance of ambiguity.
Data are also presented on how the independent variables related to
confidence in relevance judgments, familiarity with the literature,
and general beliefs about topic areas. Results are discussed in
regard to prevailing beliefs about the literature search process and
how the process might be improved. (Author/THC)

Influences on the Outcome of Literature Searches for

Integrative Research Reviews

Harris Cooper and Ronald Ribble

University of Missouri-Columbia

Running Head: INFLUENCES ON LITERATURE SEARCHES

Abstract

Three influences on the outcomes of literature searches undertaken as part of integrative research reviews were examined: (a) the degree of expertise of the searcher; (b) the amount of information available (i.e. keywords, bibliographics, abstracts); and (c) the cognitive chacteristics of the searcher. Participants were presented with descriptions of documents varying in (a) topic area (related or not related to their expertise), (b) amount of information and (c) whether the document was relevant or irrelevant to the search. Results revealed no difference between experts and nonexperts in their accuracy of juuging relevant documents but experts were better able to discern that irrelevant documents were, in fact, irrelevant. Judgments based on abstracts were more accurate than judgments based on keywords or bibliographics, which did not differ, and this effect was more pronounced for relevant than irrelevant articles. Participants who had previously conducted more literature searches or published more research reviews made more accurate judgments. Searchers high in cognitive complexity or tolerance of ambiguity made more accurate judgments. Data was also presented on how the independent variables related to: (a) confidence in relevance judgments; (b) familiarity with the literature, and; (c) general beliefs about topic areas. Results are discussed in regard to prevailing beliefs about the literature search process and how the process might be improved.

Influences on the Outcomes of Literature Searches
for Integrative Research Reviews

Since the publication of Smith and Glass' (1977) meta-analysis
on the effects of psychotherapy, psychologists have taken a keen
interest in quantitative procedures for summarizing and integrating
empirical research findings. Proponents and opponents have argued
over whether employing statistical procedures in research reviews
makes their outcomes more or less valid.

Regardless of how one views meta-analysis, a positive
consequence of the surrounding debate is that it has increased the
attention paid to other aspects on integrative research reviewing
(Cooper, 1984). For instance, research reviewers are now expected
to take greater care in conducting their literature searches and
in reporting their searc. procedures. Psychologists are realizing
that the validity of a review's conclusions is as much a function
of the legitimacy of the searching procedures as it is of the
integrating process.

This paper reports a study examining some of the factors that
influence literature searches conducted as part of integrative
research reviews. While information scientists have closely
scrutinized certain aspects of the literature searching process,
psychologists are largely unaware of these efforts and their
implications for the types of reviews they conduct. In addition,
crucial aspects of searches performed for integrative research
reviews have largely been ignored by both disciplines.

The present study focused on four aspects of the literature
searching process -- the amount of information a reviewer has when
making searching decisions, and the topic expertise, search'ng
experience, and cognitive characteristics of the searcher.

Amount of Information

The information about documents a searcher initially
encounters comes in three forms: keywords, bibliographic
information, and abstracts. Keywords are critical terms meant to
describe the contents of the documents. These may be supplied by
the author, attached by indexers employed by the literature
database service (using a thesaurus of accepted index terms), or
located in titles and abstracts by computer scanning.

Intuitively, the accuracy of a relevance decision should be
greatly influenced by the amount of available information. In two
previous studies, the effect of information on relevance decisions
was examined. These studies defined as "relevant documents" those
judged applicable to a search based on a reading of the entire
document. Rath, Resnick, and Savage (1961) found little difference
in relevance judgments based on sentences chosen systematically
from documents (e.g. the first and last five percent of sentences)
and on readings of the full document. They did find, however, that
agreement between full document readings and titles only was low.

On the other hand, Dym (1967) found that the probability of a
judgment based on bibliographic information corresponding exactly
with one based on the entire text was .69. The addition of an
abstract to the bibliographic information only raised the
probability to .73. No previous research has compared the utility
of keywords to that of titles or abstracts.

In the present study, all three levels of information
encountered in the course of performing a search were examined.
Whether a search is manual or computerized, the task begins with
the specification of keywords and an examination of the longer
lists of keywords attached to articles. Next, during a computer
search, searchers can examine only the bibliographic information
(authors, title, publication source) of documents, or they may
retrieve the abstract along with the bibliographic information.
In manual searches, the abstract and bibliographic information are
typically retrieved together, after keywords have led the searcher
to believe the document may be relevant. In searches employing
techniques other than literature databases (e.g. scanning
journals), bibliographic information is likely to be the first type
of information encountered.

The keywords, bibliographic information, and abstracts used in
this study were all actual contents of literature databases, were
associated with empirical and nonempirical journal articles, and
had been located as part of previously published integrative
research reviews.

Topic Expertise

The study of how expertise affects decision making has
stimulated considerable psychological research. For instance,
research in the medical domain often focuses on how experts and
novices formulate diagnostic questions, how the nature of these
questions may differ, and how the resulting information is
utilized or combined (Elstein, Shulman, and Sprafka, 1978).

Despite their surface dissimilarities, medical diagnosis is in
many ways analogous to decision-making during a literature search.
Searchers formulate inquiries (e.g. decide on a topic's breadth)
and choose the channels of scientific communication through which
they will gather information (e.g. research literature databases
and index terms to search, journals to scan, colleagues to
contact). It is more difficult in the context of literature
searching than medical diagnosis, however, to assess the "utility"
or "accuracy" of decisions during these initial phases. For
example, Cooper (1985) found that experts in a topic area differed
from nonexperts not only in how they conducted a literature search
but also in how they construed the tasks -- experts were less
interested in writing research reviews that exhaustively covered a
literature. In medical diagnosis the task is more invariant.

The outcome of the initial phase of a literature search is a
series of incomplete descriptions (i.e., keywords or index terms,

bibliographic information, abstracts) of potentially relevant pieces of information (i.e., journal articles or other documents). From these descriptions the searcher makes numerous discrete decisions about whether to pursue the individual documents further. Medical experts might be confronted with a similar series of decisions. However, medical decision-makers may follow different paths, of indeterminant accuracy, to arrive at a similar overall diagnosis. The literature searcher's decision about each piece of information, in contrast, can be assessed for its individual accuracy, that is, whether it leads to a document containing information relevant to the research review. It is this second phase of the literature search, involving decisions about the relevance of documents based on incomplete information, that is the focus of this study.

Research in the information sciences indicates that judges with greater knowledge of a topic area may show greater agreement among themselves concerning whether particular articles are relevant to a search. For example, Cuadra and Katter (1967) divided senior psychology majors and first year graduate students into four levels of expertise, based on their number of academic credits in psychology and their knowledge of psychology authors. Students then judged the relevance of nine journal abstracts for answering eight questions. The average correlation of judgments within-groups indicated that individuals in the more knowledgeable groups agreed more often about the relevance of the abstracts to the search. However, for all four groups the average correlations varied only between .41 and .49.

In another study, Rees and Schultz (1967) employed a similar research design but judges varied in medical, rather than psychological, expertise. They found that not only did medical scientists, the group with the greatest expertise, show the greatest within-group agreement but this group also judged the fewest number of documents as relevant to the search.

Research examining the effects of topic expertise on literature searching has largely been abandoned for the past two decades. Two reasons account for this development. One involves problems associated with defining the term "relevance", and by implication "judgmental accuracy", in a manner that is meaningful across all the forms of inquiry encountered by information scientists (Saracevic, 1970, 1975). Thus, in both studies mentioned above, interjudge agreement was scrutinized in lieu of judgment accuracy.

In the present study a satisfactory operational definition of accuracy could be realized because of the unique nature of the integrative research reviewing task. The accuracy of judgments was made measurable by using as stimuli documents known, a priori, to be either relevant or irrelevant to the searchers' task. Specifically, searchers were asked to estimate the likelihood that documents contained empirical data relevant to the topic of an integrative research review. Whether the documents did in fact

contain such data was manipulated as an independent variable. Therefore, "greater accuracy" in this study describes a finding when documents that actually contained empirical data were assigned a greater likelihood of containing such data and/or when documents without empirical data were assigned lesser likelihoods. "Lesser accuracy" describes the opposite situations.

The second reason for the abandonment of topic expertise studies was that the advent of computerized literature databases created a whole new area requiring critical investigation. One question given added importance by the computerization of literatures was whether experience with this searching mechanism leads to more satisfactory search outcomes.

Search Experience

How past experience affects the searching process has recently been a much studied question in information science (Fenichel, 1980; Howard, 1982; Oldroyd, 1984). Most studies demonstrate that experienced searchers perform more precise and cost-effective searches than inexperienced searchers. According to Bates (1977) search experience leads to a familiarity with subject headings and how they are applied. However, these studies are confined to the initial phases of searching, that is, to the definition of queries, the choice of databases to search, and the selection of keywords. No studies have examined the effects of searching experience on the searchers ability to distinguish relevant from irrelevant documents based on incomplete information. In addition, past research has used as searchers librarians or library science students, typically with different levels of computer searching experience or varying degrees of exposure to particular literature databases (e.g. ERIC, PsychInfo).

The present study examined the effect of searching experience on the accuracy of relevance decisions based on keywords, bibliographic information, and abstracts. The searchers were "end-users", that is, the individuals who intend to use the outcome of the search to produce new scholarly documents. In addition to search experience, this study also examined whether past experience at writing and publishing integrative research reviews also affected relevance decisions.

Cognitive Characteristics

While it could be expected that the cognitive characteristics of searchers would influence decisions concerning the relevance of documents, surprisingly few studies have examined this issue. Davidson (1977) noted that while "individuals selecting information from alternatives do so in a somewhat idiosyncratic fashion (existing library systems deal) with users as though they were all the same" (p. 273).

The present study examined two cognitive characteristics of searchers: tolerance of ambiguity and cognitive complexity.

Budner (1962) defined intolerance of ambiguity as "the tendency to perceive (i.e. interpret) ambiguous situations as undesirable" (p. 29). Allport (1954) maintained that intolerance of ambiguity is manifested through a need for concreteness and definiteness. Persons who are intolerant of ambiguity would be more uncomfortable with unclear or unfamiliar situations. Frenkel-Brunswik (1949) held that people intolerant of ambiguity value a clear perception more than a correct one.

With regard to cognitive complexity, persons who are low on this dimension attempt to reduce everything to a few simple components (Honess, 1976). In contrast, cognitively complex individuals possess more cognitive categories in more complex interrelation. They are able to perceive differences in a more sophisticated manner, and to better integrate information, both positive and negative, into a coherent whole (Campbell, 1960; Bieri, 1968).

Hypotheses

In ronclusion, the primary hypotheses of this study, based on the above research and, when research was not available, on experience and some intuition, were:

(a) Judgments of the relevance of an article, when searching for empirical studies to include in an integrative research review, will be more accurate if they are based on bibliographic information and an abstract than on bibliographic information alone. Likewise, bibliographic information will lead to more accurate judgments than the keywords used to index the articles. Confidence in judgments of relevance will follow a similar pattern.

(b) Experts will make more accurate judgments than nonexperts of the relevance of articles in their topic area. Experts will also be more confident about their judgments.

(c) Persons with more experience at performing literature searches and/or at publishing research reviews will be more accurate judges of the relevance of articles to their search.

(d) Persons high in cognitive complexity and/or tolerance of ambiguity will generally judge articles as more relevant to their search than will persons low in these characteristics.

Methods

Overview of Design

The basic study design involved the manipulation of three independent variables associated with performing a literature search: (a) the topic area of studies, either the relation of locus of control to academic achievement or the effects of drug treatments on hyperactivity; (b) the amount of information about

an article given to participants, either keywords only, bibliographic citation, or bibliographic citation and abstract, and; (c) the article's content, that this, whether or not the article was relevant to the participants' searching goal. All three independent variables were manipulated within-subjects, constituting a 2x3x2 repeated measures design.

The major dependent variables were the participants': (a) estimate of the likelihood that a given article was relevant to their search, and; (b) confidence in their likelihood estimates. In addition, measures of the participants' familiarity with the literature and general beliefs about the two topic area were collected.

Several individual difference measures were also collected to determine if they moderated the effects of the manipulations. These included: (a) the expertise of the participants in the two topic areas; (b) the number of integrative research reviews the participants had published; (c) the number of literature searches the participants had conducted; (d) the participants' cognitive complexity, and; (e) the participants' tolerance of ambiguity. Each of these individual differences was added to the basic research design as predictors of responses.

Literature Search

Both the selection of participants and the stimulus materials used in this study were facilitated by two literature searches conducted in connection with previously published integrative research reviews. The first review, entitled "Locus of control and academic achievement: A literature review," was published by Findley and Cooper (1983). The initial bibliography for this review was constructed through an online computer search of Psychological Abstracts, Resources in Education, and Dissertation Abstracts International. The keywords used in the search were crossings of the terms achievement or performance with locus of control or internal-external. The three databases contained 802 studies that had both of these terms in their title or abstract. A reading of the abstracts led the initial reviewers to examine 208 articles in their entirety. Of the 208, 98 articles were deemed relevant to the goal of the review. To be considered relevant, a study had to include (a) a measure of locus of control, (b) a measure of academic achievement, and (c) a test of the relation between the two measures.

The second review, entitled "Drug treatment of hyperactivity in children," was conducted by Ottenbacher and Cooper (1983). Studies potentially relevant to this topic were obtained through computer searches of Index Medicus, Psychological Abstracts, Current Index to Journals in Education, Exceptional Children Educational Resources, and Dissertation Abstracts International. A total of 831 study abstracts were retrieved. Of these studies, 61 met the relevance criteria of (a) investigating the effect of a drug treatment on children diagnosed as hyperactive; (b) using a

two group comparison including some form of control group; (c) employing random assignment and a dcuble blind procedure as part of the design, and; (d) reporting the results in a fashion that allowed quantitative analysis.

In the discussion that follows, the term "relevant" is used to describe articles that were chosen for inclusion in either of the two literature reviews because they contained empirical data relating to the direction and size of the relation being reviewed, that is between locus of control and academic achievement or the effect of drug treatment on hyperactivity. The term "irrelevant" is used to describe an article not included in either of the two reviews and which contained no analyzable empirical data. These articles, of course, could be pertinent to the topics in ways not related to the present study or to the judgment the participants were asked to make. The irrelevant articles were chosen from articles uncovered by the literature searches but not used in the research reviews.

Participants

Participants for the study were recruited through mail solicitation. Of the 98 first authors of locus of control and academic achievement studies, 67 were invited to take part in the study. Of the 61 hyperactivity first authors, 50 were invited to take part. Invitations were based on the availability of the author's address. Authors were paid a $25 honorarium for time spent on the project.

For the locus of control sample, 24 of the 61 invitees did not respond, 19 declined the invitation, 7 accepted the invitation but never returned the materials, and 17 (25%) returned completed questionnaires. For the hyperactivity sample, 18 of the 50 invitees did not respond, 15 declined the invitation, 3 accepted the invitation but did not return the materials, and 14 (28%) returned completed questionnaires.

In sum, 31 authors of empirical articles participated in the study, with 17 having topical expertise in locus of control and academic achievement and 14 in drug treatment of hyperactivity. Topical expertise, then, was initially defined as a participant having published one article relevant to one of the topics. The expertise labelling of participants was later cross-validated by responses to a section of the questionnaire which asked authors to tally their publications and to rank their "standing among scholars" in both fields. A six point scale, ranging from 1, "a person new to the area", to 6, "a leader in the topic area" was used. Only two authors claimed to have made published contributions to the topic area to which they had not been assigned (both were hyperactivity experts claiming contributions to the locus of control and achievement literature) and no participant claimed a standing among scholars in their nonexpert area within two scale points of their expert area.

In general, hyperactivity experts claimed somewhat higher standing in their field than did locus of control experts. Hyperactivity experts averaged a rating on the "standing among scholars" scale of 5, equivalent to "an active, continuing contributor". Locus of control experts average 3.8, with 10 authors calling themselves "a former contributor, no longer active" and 6 calling themselves "a person who made a single contribution".

Stimulus Materials

The stimulus materials were presented to participants in two separate sections, one focusing on locus of control and one on hyperactivity. Each section began with a background sheet which briefly introduced the substantive topic under scrutiny and the nature of the task.

The locus of control and academic achievement section began with the following description of the topic:

Locus of control refers to people's feelings about whether they control the events that occur in their life. Some people, labeled internals, feel personally responsible for the things that happen to them. Other people, labeled externals, feel that their outcomes in life are determined by forces beyond their control, for example, fate ur other people. Locus of control beliefs are not so discrete, however--people form a continuum ranging from highly internal belief systems to highly external ones, with most belief systems falling somewhere in between.

One topic that has intrigued both psychologists and educators is the relationship between locus of control beliefs and achievement in school. The prediction that a stronger belief in internal locus of control will be associated with higher academic achievement is both logical and consistent with relevant theories. People who feel more able to control the outcome of events should exert more effort to do so. Also, internals should experience more pride when they succeed and more shame when they fail. Thus, the rewards and punishments experienced in school by internals should be stronger and, therefore, more motivating than those experienced by externals.

The drug treatment of hyperactivity section began with the following description of the topic:

Hyperactivity, a disorder involving a deficit in attention, has been estimated to afflict between 4% and 20% of school-aged children. Due to its frequent occurrence, both educators and physicians have vigorously pursued effective treatments for hyperactivity. The proposed treatments range from dietary restrictions and supplements to behavior modification training. By far the most prevalent treatment of hyperactivity, however, is pharmacological management and almost all children identified as hyperactive receive some form of drug therapy in the course of their treatment.

In spite of the fact that drug treatment of hyperactivity is so frequently prescribed and evidence for its effectiveness is widespread, there is still substantial controversy over the advisability of its use (e.g., its ethicality and its potential susceptibility to placebo effects). In addition, there has been some question about   relative effectiveness of different types of drugs (stimulants and nonstimulants) and about any drug's ability to affect behaviors beyond simple motor or perceptual performance (e.g., social adjustment and educational performance).

After each topic description, the nature of the participants' task was described as follows (the changes in wording associated with the two topics appears in brackets):

Assume you have decided to perform a literature review on the topic of [locus of control and academic achievement/drug treatment of hyperactivity]. The review will have the following goals, in descending order of importance: (a) to describe, summarize and interpret the findings of empirical research meant to uncover the direction and magnitude of the [locus of control and academic achievement link/drug treatment effect on hyperactivity]; (b) to identify important issues associated with the topic (e.g., methodological and conceptual issues, interacting variables); and (c) to suggest topics for future research.

Your goal is to uncover as much literature on the relation as you possibly can. This does not mean, of course, that every relevant work will be cited in your paper. However, before you begin the summarizing and interpreting process, you want your personal file of relevant works to be as exhaustive as possible. To help you construct this file, you ran computerized literature searches of the [PsychInfo and ERIC/Index Medicus, Psychological Abstracts, ERIC and Exceptional Children Education Resources] abstracting services, using the keyword(s) ["achievement" or "performance" crossed with "locus of control" or "internal-external"/"hyperactivity"]. The searches yielded a total of (802/831) citations, so it is impractical for you to examine all of the full reports. Assume also that the abstract service information that follows represents a random sample from the full list (and only the beginning of a much longer task).

Finally, assume you intend to submit the resulting review paper for publication in a scholarly journal like Psychological Bulletin or [the Review of Educational Research/Developmental Medicine and Child Neurology]. Thus, your primary audience is other researchers interested in the topic.

On the pages that follow, you will find abstracting service information on 30 articles uncovered by your computer search. For 10 articles only the keywords or index terms associated with the manuscript are given, for 10 articles the bibliographic reference is given, and for 10 the reference and abstract are given.

The order in which the three levels of information are presented corresponds with how information is typically encountered in actual searches. That is, first you are asked to make judgments based only on keywords, then judgments based on bibliographic references and then on abstracts.

Below each entry you are asked several questions concerning whether or not you anticipate the article contains relevant information. Please complete these questions in as realistic a manner as possible. Complete them in the order in which they are presented.

After reading the locus of control and academic achievement topic and task description, participants were presented with thirty pages, each of which contained information about a single locus of control article. The thirty article descriptions varied in (a) the amount of information about the article that was presented and (b) whether or not the article described was relevant or irrelevant to the task (called article content).

As stated in the instructions, the first ten pages contained only the keywords (i.e. index terms used by the computer) associated with each of the ten articles. The second ten pages contained the bibliographic citation (i.e. author(s), year of publication, title, journal, volume, and page numbers) associated with each article. The final ten pages contained the bibliographic citation and abstract associated with each article.

Within each amount of information condition, five of the ten articles were preselected so as to be relevant to the task of the participant (i.e. to contain empirical information about the locus of control and academic achievement relationship). The other five articles contained only irrelevant information (to the empirical relation, though other types of important insights may have been contained in these articles). Relevant and irrelevant articles were alternated within each condition. Of course, participants were not informed that article content was being manipulated.

All subjects also completed a section containing a topic and task background sheet and thirty articles concerning drug treatment of hyperactivity. The section on hyperactivity articles were constructed in a manner identical to the locus of control section.

The thirty articles in each topic section were chosen from a pool of sixty articles (thirty relevant and thirty irrelevant) and each article appeared an equal number of times in the study. No article appeared twice, that is in more than one amount of information condition, in the materials for a single participant, though articles appeared in all conditions across participants. Half of the participants completed the locus of control section first and half completed the hyperactivity section first, without regard to the participants' topical expertise. Precautions were also taken to minimize the chance that participants would be asked

to make judgments about studies they had personally authored, although this did occur twice in the 1860 article ratings.

Individual Difference Measures

Publication and Literature Searching Background. The number of research reviews participants had published was measured by asking them if they had every published an integrative research review and, if so, how many. On average, participants claimed 3.83 published reviews, with a range from 0 to 31 and a median of 2. For use as an individual difference variable, labelled review publications, participants were divided, using an approximate median split, into those with less than two published research reviews and those with two or more published reviews.

Participants were also asked how many computerized literature searches they had conducted (or had conducted for them) and how many manual searches of an abstracting service they had performed. Participants claimed an average of 6.80 (sd=10.06) computerized searches and 8.23 (sd=14.06) manual searches. The number of computerized and manual searches were summed for each participant in order to form a single search experience variable. Participants were then divided, using a median split, into those who had conducted less than sixteen searches and those who had conducted sixteen or more searches.

Cognitive Complexity. A revised version of the Paragraph Completion Test (PCT; Schroder and Streufort, 1962) was employed to measure cognitive complexity. Participants were presented five sets of words or phrases that began an incomplete sentence and were asked to complete the sentence and write at least one additional sentence following from the first. The five incomplete sentences were, "When scholars disagree with one another . . . ", "When I am in doubt . . . ", "Rules . . . ", "When others criticize me it usually means . . . ", and "Confusion . . . ".

Scoring was accomplished in accordance with Schroder's (1971) seven-point scaling method. For a score of 1, representing low complexity, the response must be generated by a single fixed rule, with a high degree of certainty. A score of 3 (medium complexity) is given if the response clearly contains reference to the availability of alternate means of perceiving the event. A score of 5 (medium-high complexity) indicates that in addition to the consideration of alternatives the respondent proposes some kind of joint outcome. Finally, high complexity (a score of 7) means that the response reveals the functional relationships in the person's "world view".

Each response was scored by two independent judges. Interjudge reliability for the sum of the five responses was r=.43, similar to reliabilities found in other studies (see Tom, Cooper, and McGraw, 1984). The average of the two judges ratings was used as the measure of a participants' cognitive complexity. The mean total score for the sample was 17.30 (sd=3.96), or 3.46

per item. The cognitive complexity variable was created by performing a median split dividing the sample into high complexity and low complexity groups based on total scale scores.

Tolerance of Ambiguity. Budner's Scale of Tolerance-Intolerance of Ambiguity (STIA; Budner, 1962) was used. The STIA contains sixteen Likert-type items accompanied by five-point scales ranging from "strongly agree" to "strongly disagree". This instrument has shown acceptable intercorrelations with other measures of tolerance of ambiguity (r=.46; see Robinson and Shaver, 1973).

Only eight of the items on the STIA were used in this study. An example of an item used reads, "It is more fun to tackle a complicated problem than it is to solve a simple one". Items were counterbalanced so that agreement could reflect both tolerance or intolerance of ambiguity.

Individual item responses were summed to create a single tolerance of ambiguity score, with higher scores denoting greater tolerance. The average score in the sample was 22.87 (sd=3.84). As with the other individual difference variables, a median split was used to create high and low tolerance of ambiguity groups.

Vannoy's (1965) study of 18 cognitive complexity measures demonstrated an intercorrelation between the PCT and STIA of r=-.06. Tom, et.al. (1984) found r=.08. In the present study, the intercorrelation was r=.21.

Dependent Variables

The two primary dependent variables were the participants' (a) assessment of the likelihood that an article contained empirical data relevant to their search, and (b) confidence in the likelihood estimate.

For locus of control and academic achievement studies, the question assessing the likelihood of containing empirical data was worded, "How likely is it that this paper contains empirical data (e.g., group comparisons, correlations) relevant to the direction and size of the relation between locus of control and academic achievement?" For drug treatment of hyperactivity studies, this question was worded identically, except that "the effect of drug treatment of hyperactivity" was substituted for the reference to locus of control. Participants responded on an eleven-point scale, with 0 labelled "Not likely at all" and 10 labelled "very likely".

For both locus of control and hyperactivity studies, the question assessing confidence in the likelihood estimate was worded, "How confident are you of the above estimate?" Participants again responded on an eleven-point scale, with 0 labelled "not confident at all" and 10 labelled "very confident".

In the (a) bibliographic and (b) bibliographic and abstract information conditions, participants were also asked "Are you familiar with the content of this article". A "yes" or "no" response was requested. This question was omitted from the keyword information condition since it was extremely unlikely that a particular article could be recognized by participants based solely on its associated index terms.

Ancillary measures. In addition to the primary dependent variables, participants were asked several questions concerning their general beliefs about the two topic areas.

Participants were asked to estimate (a) the average size of the relation between locus of control and academic achievement and (b) the average effect of drug treatment on hyperactivity. Participants expressed the locus of control and academic achievement relation in terms of the average correlation coefficient to be expected in a study. The drug treatment effect was expressed in terms of the U_2-index, where U_2 equals the percent of children in the drug-treated group who are less hyperactive than the median child in the untreated group (see Cohen, 1977). Participants were also asked to rate their confidence in their estimate, using the same confidence scale described above.

Participants responded to three questions meant to measure more contextual beliefs about each topic area. These questions, presented as attitude statements, were worded as follows (the changes in wording associated with the two topics appear in brackets): (a) The link between [locus of control and academic achievement/drug treatment and hyperactivity] is highly complex, involving many variables that mediate the relationship; (b) In general, problems in measurement mean that the outcomes of studies relating [locus of control and academic achievement/drug treatments and hyperactivity] are not reflective of the link as it exists in nature, and; (c) The [relation between locus of control and academic achievement/impact of drug treatment on hyperactivity] is an important topic for [psychologists and educators/parents and the medical community]. All questions were responded to on an eleven-point scale, with -5 labelled "strongly disagree" and 5 labelled "strongly agree".

Analytic Design

Since participants responded to a total of sixty articles representing twelve distinct crossings of the three within-subjects variables (two topics x three amounts of information x two types of article content) the basic units of analysis were the average responses to the five articles that appeared in each condition. If any of the three within-subjects factors was omitted from a particular analysis, responses were also aggregated over the omitted factor. Newman-Keuls post hoc means tests were employed where appropriate (see Myers, 1972). Degrees of freedom sometimes varied from one analysis to another because of missing data.

All analyses conducted on the data used the Analysis of Variance program from the Statistical Analysis System computer package (SAS, 1985).

Results

Topic Area, Topic Expertise, Amount of Information, and Article Content

A series of ANOVAs containing one between-subject factor (topic expertise) and three within-subject factors (topic area, amount of information, and article content) were performed on the three primary dependent variables (estimates of the likelihood that articles contained empirical data, confidence in the estimates, and familiarity with articles).

Likelihood estimates. The ANOVA performed on estimates of the likelihood that an article contained data relevant to the search revealed numerous significant effects. The three-way interaction, involving topic area, topic expertise, and article content, underlied several of these effects. The associated means are displayed in Table 1.

A significant main effect for topic area revealed that hyperactivity articles (M=6.81) were more likely to be judged as containing empirical data than were locus of control articles (M=6.22; F(1,29)=10.17, p .0034). A main effect for article content revealed that empirical articles (M=7.50) were judged more likely to contain empirical data than nonempirical articles (M=5.55; F(1,29)=107.88, p .0001). A topic area by topic expertise interaction revealed experts were less likely to believe articles in their area contained empirical data than article in their nonexpert area (F(1,29)=4.34, p .0461).

A three-way interaction (see Table 1) revealed that expertise had little effect on estimates of the likelihood that empirical data was contained in articles that actually did contain empirical data. However, expertise did effect estimates for nonempirical articles. Specifically, compared to nonexperts, experts judged nonempirical articles in their area as being less likely to contain empirical data (F(1,29)=23.68, p .0001).

Place Table 1 About Here

Two significant effects involved the amount of information variable. A main effect revealed that article abstracts led to higher estimates that an article contained empirical data (M=6.94) than did bibliographic information (M=6.32) or keywords (M=6.28; F(2,58=)5.26, p .008). An interaction between article content and amount of information indicated that abstracts increased participants' estimates that empirical data was contained in articles that were in fact empirical but did not affect estimates

for nonempirical articles (F(2,58)=9.96 p .0002). Table 2 presents the means underlying this interaction.

Place Table 2 About Here

Confidence in Judgments. The ANOVA conducted on the participants' ratings of confidence in their likelihood estimates also revealed several significant effects. Two main effects revealed that participants were more confident about their judgments concerning hyperactivity articles (M=7.81) than they were about locus of control articles (M=7.28); F(1,29)=13.18, p .0011) and more confident about judgments concerning empirical (M=7.74) than nonempirical articles (M=7.35; F(1,29)=10.26, p .0033).

An interaction between topic area and topic expertise indicated that experts were more confident of judgments made about articles in their own area than their nonexpert area (for locus of control articles, locus of control experts M=7.47, hyperactivity experts M=7.07; for hyperactivity articles, hyperactivity experts M=7.95, locus of control experts M=7.67; F(1,29)=5.43, p .027).

The amount of information factor entered into three significant effects. A main effect showed that participants were more confident about judgments based on more information (keywords M=6.92, bibliographic M=7.34, abstract M=8.36; F(2,58)=35.72, p .0001), with the addition of an abstract precipitating a large increase in confidence. An interaction between amount of information and article content indicated that the addition of information had a greater impact on the confidence in judgments made about empirical articles than nonempirical ones (F(2,58)=4.59, p .0142).

Place Table 3 About Here

Finally, an interaction involving amount of information and topic area revealed that the effect of information was greater on the confidence ratings for hyperactivity articles (keywords M=6.64, bibliographic M=7.34, abstract M=8.55) than for locus of control articles (keywords M=7.18, bibliographic M=7.34, abstract M=8.18; F(2,58)=3.38, p .0407).

Familiarity. The ANOVA performed on the measure of familiarity with the articles showed that participants were more familiar with hyperactivity articles (M=1.51) than locus of control articles (M=0.63;F(1,29)=13.81, p .0009) and with empirical articles (M=1.19) than nonempirical ones (M=0.95; F(1,29)=12.06, p .0016).

The means underlying two significant interactions appear in Table 4. An interaction between topic area and topic expertise revealed that participants were more familiar with articles in their own topic area (F(1,29)=46.59, p .0001). An interaction involving topic area, topic expertise, and article content showed that experts were more likely to be familiar with empirical than nonempirical articles in their own area. In their nonexpert area, however, no different existed between their familiarity with empirical and nonempirical articles (F(1,29)=13.25, p .0011).

Individual Difference Measures.

ANOVAs with two between-subject factors (review publications and search experience) and two within-subject factors (article content and amount of information) were conducted on participants' estimates that an article contained empirical data and their confidence in these judgments.

The analysis of likelihood estimates revealed an interaction between review publications, search experience, and article content (F(1,24)=4.97, p .0354). The interaction indicated that while all groups judged empirical articles more likely than nonempirical articles to contain empirical data, participants with less than two published reviews and fewer than sixteen searches gave more similar estimates to empirical and nonempirical articles than any other group. Table 5 presents the means underlying this effect.

Place Table 5 About Here

With regard to confidence in judgments, one two-way interaction involving review publications and article content proved significant. Participants with more than two published reviews showed a greater difference in their confidence about judgments of empirical and nonempirical articles (empirical M=15.6, nonempirical M=14.4) than did participants with less than two published reviews (empirical M=15.3, nonempirical M=15.1; F(1,24)=4.45, p .0456), though not significantly so by the Newman-Keuls test.

ANOVAs similar to those just described were conducted substituting the participants' cognitive complexity and tolerance of ambiguity for review publications and search experience.

An interaction between cognitive complexity and article content revealed that participants high in cognitive complexity showed a greater difference between their estimates that empirical and nonempirical articles contained empirical data than did estimates by participants low in cognitive complexity (F(1,26)=6.28, p .0188), though not significantly so by the Newman-Keuls test. Table 6 presents the associated means. An interaction between tolerance of ambiguity and article content

indicated that participants high in tolerance of ambiguity showed a greater difference in their estimates that empirical and nonempirical articles contained empirical data than did estimates by participants low in tolerance of ambiguity (F(1,26)=4.60, p .0415), though not significantly so by the Newman-Keuls test. Table 7 presents the underlying means. The direction of effects indicated that while all subjects estimated empirical articles were more likely to contain empirical data than nonempirical articles, participants who were higher in cognitive complexity or tolerance of ambiguity tended to show this effect to a greater degree.

Place Tables 6 and 7 About Here

The analysis of the confidence in judgments revealed no significant effects involving cognitive complexity or tolerance of ambiguity.

Ancillary Measures.   The ancillary measures were examined employing three sets of ANOVAs containing groupings of the between-subjects factors similar to those described above and the topic area as the sole within-subjects variable.

A topic area main effect indicated that participants believed drug treatment of hyperactivity was a more important area to parents and the medical community (M=9.19) than locus of control and academic achievement was to psychologists and educators (M=8.06;F(1,29)7.46, p .0106).   A topic expertise main effect revealed that locus of control experts believed that highly complex links involving many variables were needed to explain the two relations (M=8.74) more than did hyperactivity experts (M=7.47;F(1,29)=4.52, p .0421).

Three of the measures revealed significant topic area by topic expertise interactions. These indicated that (a) experts were more confident than nonexperts about their estimates of the size of the relation in their topic area (F(1,29)=31.68, p .0001); (b) experts were less likely than nonexperts to believe that the link between their topic area variables was complex and involved many mediators (F(1,29)=8.59, p .0065), and (c) experts were less likely than nonexperts to agree that measurement problems rendered studies in their area invalid (F(1,29)=15.80, p .0004).   Table 8 presents the means underlying these three interactions.

Place Table 8 About Here

Finally, an interaction between tolerance of ambiguity and topic area indicated that for locus of control articles participants high in tolerance of ambiguity had more confidence in their estimates of relation strength (M=6.73) than those low in

tolerance of ambiguity (M=5.33), while for hyperactivity articles participants low in tolerance of ambiguity were more confident (M=6.47) than those high in tolerance for ambiguity (M=4.93;F(1,26)=4.55, p .0425).

No effects involving the review publications or search experience individual differences proved significant.

Discussion

In the discussion that follows, constant reference will not be made to the fact that participants in this study were performing a specific type of literature review, namely an integrative research review. This omission should not be construed to mean the results are assumed to be generalizable to other forms of review (see Cooper, 1985). The question of generalizability is open for future research.

Consistent with expectations, experts in a topic area were found to be more accurate judges than nonexperts of whether articles contained empirical data relevant to their topic. Unpredicted, however, was the finding that nonexperts were as accurate as experts in judging empirical articles as relevant to their search. The greater overall accuracy of experts was due solely to their ability to judge nonempirical articles as irrelevant.

One possible explanation for experts' greater ability to dismiss irrelevant articles is that they are more familiar with the irrelevant (in this case, nonempirical but potentially related) literature. The measure of familiarity with the literature provided some qualified support for this notion--experts were more familiar than nonexperts with the nonempirical literature in their area. Curiously, however, experts were also more familiar with the empirical literature, but this did not enhance their ability to identify relevant articles.

An explanation for this asymmetry in the effect of familiarity might be that the information contained in the research literature databases gives clear clues that certain articles contain empirical data. When these clues are not present, the articles are, in fact, usually nonempirical. Experts, due to familiarity, rule out certain of these irrelevant articles. Nonexperts, on the other hand, are less likely to rule these articles out since such a judgment would be based solely on the fact that the available information does not affirm the article as relevant. However, if this explanation were perfectly adequate, a statistical interaction, involving the topic area, expertise of the searcher, and the amount of available information, would have been expected. The "advantage" of expertise would be greatest when more information was available about nonempirical articles, that is, when articles were most recognizable and, therefore, the added familiarity of experts would have the most pronounced affect on judgments.

One important implication of the findings regarding expertise deserves special mention. As noted above, topic expertise had no impact on the searchers' ability to accurately detect articles that contained empirical data. Based on this, it might be concluded that experts are unlikely to become aware of more empirical references than nonexperts conducting similar searches. Such a conclusion is not wholly justified because experts often obtain information about their topic through channels other than literature searches, including invisible colleges and subscriptions to topic-related journals. It might be concluded, however, that if research reviews written by authors with previous expertise in an area contain more empirical references than reviews written by nonexperts this may be due to their access to other mechanisms for literature retrieval, rather than any advantage they have in using the research literature data bases. Instead, the advantage of expertise in searching may manifest itself solely in how long it takes experts and nonexperts to conduct their task - experts may spend less time examining articles irrelevant to their task.

It would also seem important to ask why both locus of control and hyperactivity experts were more accurate at judging empirical articles on hyperactivity than on locus of control. This finding probably involves the length and specificity of the information contained in the databases. In general, the number of keywords and the length of abstracts associated with hyperactivity articles were greater than those associated with locus of control articles. In addition, articles dealing with drug treatments of hyperactivity tended to deal only with this topic while locus of control and academic achievement articles often included tests of other hypotheses, most frequently other predictors of achievement.

With regard to the amount of information available to searchers, it was somewhat surprising to find that keywords and bibliographic information did not differ in their ability to elicit accurate judgments. While the title of an article contains an author-supplied list of keywords, the number of important identifiers in a title is usually less than the number supplied by the database indexers. Apparently, these additional keywords were enough to offset any advantage bibliographic citations might have because they provide author and journal names, and the year of publication.

Having an abstract in addition to the bibliographic information was clearly an asset for accurately judging relevant articles but it was of no help for judging irrelevant articles. Again, a possible key to understanding this finding is that reference to empirical data in abstracts makes judgments of most empirical articles relatively straightforward. The absence of reference to empirical data does not rule out its presence in the article, thus leaving the searcher with some guesswork. This interpretation is supported by the finding that participants expressed the greatest confidence in their judgments made about empirical articles based on abstract information. It is also

true, however, that participants expressed greater confidence in their judgments of nonempirical articles based on abstracts than on keywords or bibliographic information. Yet, their accuracy for nonempirical articles based on abstracts was no better than that based on less information, meaning this confidence was misplaced.

This last finding withstanding, it is still apparent that searchers display greater overall accuracy when basing judgments on abstracts rather than less information. A second implication of this study, then, is that persons conducting manual searches of the research literature should not rule out studies based on keywords or bibliographic information unless the article is clearly irrelevant. Likewise, persons conducting computer-assisted searches should always request that abstracts accompany the computer's output.

With regard to other individual differences in search outcomes, greater experience with either the searching process or with publishing reviews was found to have a beneficial affect on judgment accuracy. As noted in the introduction, both effects are probably attributable to greater experience with the searching process, which apparently facilitates a person's ability to predict which specific terms to look for when searching the literature.

There was also some tentative evidence that persons high in cognitive complexity or tolerance of ambiguity, may be more accurate searchers. Though this result was not predicted its cause readily discernible, it is not inconsistent with what might be anticipated. Those high in tolerance for ambiguity are by definition (Frenkel-Brunswik, 1954) better able to deal with indefinite and irregular information, while those high in cognitive complexity (Scott, 1962) have greater flexibility and tend to comfortably incorporate and use discrepant (incomplete, vague) information in making judgments. Regardless of the precise process underlying this particular result, it is clear that one cannot overlook the "human" dimension of the interface between a searcher and the information system, a process that may sometimes be viewed as purely mechanical. Apparently, persons confronted with the same task and information can obtain different search outcomes based on deeply-rooted differences in how they relate to their environments.

Finally, the results concerning general beliefs indicated that experts saw their own topic areas as requiring fewer mediators and producing more valid results than nonexperts. Thus, familiarity with a topic area did not produce a sense of complexity and difficulty. Or perhaps, only those who are not discouraged by a topic's difficulties go on to publish research relevant to it.

Based on the results of this study, then, future research might fruitfully be aimed at addressing the following questions:

(a) Is there something in addition to familiarity that accounts for experts greater ability to accurately reject material that is irrelevant to their search?

(b) Why do keywords and bibliographic information not differ in the accuracy of the judgments they elicit?

(c) How do personality variables affect searching and what can be done to counteract detrimental personality effects on search outcomes?

(d) Are the results of this study generalizable to literature reviews involving tasks other than research synthesis?

References

Allport, G. (1954). The nature of prejudice. Reading, MA: Addison-Wesley.

Bates, B. (1977). Factors affecting subject catalog search success. Journal of the American Society for Information Science, 28, 161-169.

Bieri, J. (1968). Cognitive complexity and judgment of inconsistent information. In Abelson et al. (Eds.). Theories of cognitive consistancy: A sourcebook. New York: Rand McNally.

Budner, S. (1962). Tolerance of ambiguity as a personality variable. Journal of Personality, 30, 29-50.

Campbell, D. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. American Psychologist, 15, 546-553.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (revised). New York: Academic Press.

Cooper, H. (1985). Literature searching strategies of integrative research reviewers (Report No. 366), Columbia, MO: Center for Research in Social Behavior.

Cooper, H. (1984). The integrative research review: A systematic approach. Beverly Hills: Sage.

Cuadra, C., & Katter, R. (1967). Experimental studies of relevance judgments. Santa Monica: Systems Development Corp.

Davidson, D. (1977). The effects of cognitive style on judgments of document relevance. Journal of the American Society for Information Science, 28, 273-284.

Dym, E. (1967). Relevance predictability: I. Investigation background and procedures. In A. Kent et al. (Eds.). Electronic handbook of information: Testing and evaluation. Washington, D. C.: Thompson.

Elstein, A., Shulman, L, & Sprafka, S. (1978). Medical problem solving. Cambridge, MA: Harvard University Press.

Fenichel, C. (1980). An examination of the relation between searching behavior and searcher background. Online Review, 4, 341-347.

Findley, M., & Cooper, H. (1983). Locus of control and academic achievement: A literature review. Journal of Personality and Social Psychology, 44, 419-427.

24

25

Frenkel-Brunswick, E. (1954). Further explorations by a
    contributor to "The authoritarian personality."

Frenkel-Brunswik, E. (1949).   Intolerance of ambiguity as an
    emotional and perceptual personality variable. In J. Bruner
    and D. Krech (Eds.), Perception and personality: A symposium.

Durham, NC: Duke University Press.

Honess, T. (1976). Cognitive complexity and social prediction.
    British Journal of Social and Clinical Psychology, 15, 23-31.

Howard, H. (1982). Measures that discriminate among online
    searchers with different training and experience. Online
    Review, 6, 315-327.

Myers, J. (1972). Fundamentals of experimental design (2nd ed.).
    Boston: Allyn and Bacon.

Oldroyd, B. (1984). Study of strategies used in online searching
    5: Differences between the experienced and inexperienced
    searcher. Online Review, 8, 233-244.

Ottenbacher, K., & Cooper, H. (1983). Drug treatments of
    hyperactivity in children. Developmental Medicine and Child
    Neurology, 25, 353-357.

Rath, G. Resnick, A., & Savage, T. (1961). Comparisons of four
    types of lexical indicators. American Documentation, 12,
    126-130.

Rees, A., & Schultz, D. (1967). A field experimental approach to
    the study of relevance assessments in relation to document
    searching. Cleveland: Center for Documentation and
    Communication Research, Case Western Reserve University.

Robinson, J., & Shaver, P. (1973). Measures of social
    psychological attitudes. Ann Arbor, MI: University of
    Michigan.

Saracevic, T. (1975). Relevance: A review of and a framework
    for the thinking on the notion in information science. Journal
    of the American Society for Information Science, 26, 321-343.

Saracevic, T. (1970). The concept of 'relevance' in information
    science: A historical review. In T. Saracevic (Ed.).
    Introduction to information science. New York: Bowker.

SAS Institute. (1982). SAS Users guide: Statistics. Cory, NC:
    SAS Institute Inc.

Schroder, H., & Streufort, P. (1962). The measurement of four
    systems varying in levels of abstractness. Technical Report
    No. 11, Office of Naval Research.

Scott, W. A. (1962). Cognitive complexity and cognitive
    flexibility. Sociometry, 25, 405-414.

Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy
    outcomes. American Psychologist, 32 752-760.

Tom, D., Cooper, H., & McGraw, H. (1984). Influences of student
    background and teacher authoritarianism on teacher
    expectations. Journal of Educational Psychology, 76, 259-265.

Vannoy, J. (1965). Generality of cognitive complexity   --
    simplicity as a personal construct. Journal of Personality
    and Social Psychology, 2, 385-396.

Author's Notes

Table 1

Estimates That an Article Contains Empirical Data as a Function of Topic, Topic Expertise, and
Article Content

		Expertise			
		Locus of Control		Hyperactivity	
	Topic	Locus of Control	Hyperactivity	Locus of Control	Hyperactivity
Content	Empirical	7.16a	7.80a	7.08a	7.99a
	Nonempirical	4.83c	6.12b	5.86b	5.38c

Note. Means not sharing common subscript differ significantly by the Newman-Keuls test
(p .05).

29

Table 2

Estimates That an Article Contains Empirical Data as a Function of

Information and Article Content

		Content	
		Empirical	Nonempirical
	Keywords	7.05b	5.54c
Information	Bibliographics	7.17b	5.47c
	Abstract	8.26a	5.63c

Note. Means not sharing common subscript differ significantly by the

Newman-Keuls test (p .05).

Tab.e 3

Confidence in Judgment as a Function of Article Content and Information

		Content	
		Empirical	Nonempirical
	Keywords	7.02c	6.82c
Information	Bibliographics	7.43bc	7.25c
	Abstract	8.75a	7.98b

Note. Means not sharing common subscript differ significantly by the

Newman-Keuls test (p .05).

Table 4

Familiarity With an Article as a Function of Topic, Topic Expertise and Article Content

| | | Expertise | | | |
| | | Locus of Control | | Hyperactivity | |
	Topic	Locus of Control	Hyperactivity	Locus of Control	Hyperactivity
Content	Empirical	1.28c	0.40e	0.14e	3.00a
	Nonempirical	0.87d	0.44e	0.17e	2.41b

Note. Means not sharing common subscript differ significantly by the Newman-Keuls test (p .05).

32

Table 5

Estimates That an Article Contains Empirical Data as a function of

Published Reviews, Searching Experience, and Article Content

		Published Reviews			
		<2		>2	
	Searches	<16	>16	<16	>16
Content	Empirical	7.06ab	7.63a	7.75a	7.61a
	Nonempirical	6.26bc	5.35c	5.51c	5.38c

Note. Means not sharing common subscript differ significantly by the
Newman-Keuls test (p .05).

33

Table 6

Estimates That an Article Contains Empirical Data as a Function of

Cognitive Complexity and Article Content

		Cognitive Complexity	
		Low	High
Content	Empirical	7.46	7.62
	Nonempirical	5.83	5.15

Table 7

Estimates That an Article Contains Empirical Data as a Function of

Tolerance of Ambiguity and Article Content

		Tolerance of Ambiguity	
		Low	High
Content	Empirical	7.41	7.59
	Nonempirical	5.83	5.24

Table 8

General Beliefs About Topic Areas as a Function of Topic Expertise and Topic Area

	Expertise			
	Locus of Control		Hyperactivity	
Topic	Locus of Control	Hyperactivity	Locus of Control	Hyperactivity
Confidence in				
Estimate[a]	7.00	4.93	4.00	7.73
Complex Linkages				
Numerous Mediators[b]	2.38	2.60	3.12	1.33
Measurement Problems,				
Outcomes Invalid[b]	-0.31	0.87	1.69	1.20

Notes. a. Higher scores mean greater confidence.

b. Positive scores mean agreement.

36