A COMPARISON OF LECTURE AND INTERACTIVE TRAINING DESIGNED TO REDUCE THE INFLUENCE OF INTERFERING MATERIALS: AN APPLICATION TO SOIL SCIENCE

by

GARY J. GAETH

B.S., Northern Michigan University, 1974
M.S., University of Wyoming, 1977

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER'S OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1980

Approved by:

Major Professor

ACKNOWLEDGEMENTS

First, I would like to thank the thirteen soil judges who were willing to to give five hours of their valuable time to this project. I would also like to thank their coach, Dr. C. W. Bidwell, soil scientist, for his invaluable guidance from the conception to the completion of this project. In this case the old cliché is true; without his help this project would not have been possible. During several frustrating times committee memebers, Dr. Frank (Skip) Saal, and Dr. Richard J. Harris, provided encouragement and critical comments. I would also like to thank my advisor, Dr. James Shanteau, for his useful commets, suggestions, and his faith in my abilities. Fianlly, I dedicate this manuscript to my father, John H. Gaeth.

Table of Contents

List of Figures

List of Tables

INTRODUCTION

A great body of literature and research has dealt with "expert" judges in terms of such issues as how accurate their judgments are, how well their judgments can be modeled, and what causes erroneous judgments. However, little research has been directed at the question of how decision makers learn to make expert decisions; and specifically what types of training procedures are most effective. This study considered this issue through evaluation of two training programs designed to further improve the accuracy of previously experienced decision makers through the reduction of the adverse influence of interfering materials. One of the training procedures was strictly lecture, much like a classroom lecture, while the second procedure involved "interactive" practice of the verbal training suggestions.

The research presented here systematically measured the amount of learning which took place after two different training procedures were given to judges experienced in the applied agricultural area of soil judgment. Specifically, both of the training procedures were given to experienced soil judges and represented attempts to reduce the adverse influence of materials which are irrelevant to the judgment of soil.

In general, the systematic study of various applied training techniques, and especially the documentation of successful training programs has been scarce, and the results of many of the studies have been far from satisfactory (Goodman, Fischhoff, Lichtenstein & Slovic, 1978). In fact, this persistent lack of success has lead to a certain amount of pessimism toward the discovery of training techniques which can be applied in a real world setting and can lead to nontrival learning (Slovic, Fischhoff & Lichtenstein, 1977). In the few cases where training has increased learning the setting is usually the laboratory (Hammond, 1971) and the task to be learned is quite often artificial (Brehmer, 1977). This research was an attempt to rectify the shortcomings of previous research by providing a systematic study, in an applied agricultural setting, of the impact of two training programs on subsequent learning. This research was an attempt to rectify some of the shortcomings found in previous studies.

Review of the Literature

Because of the cross-disciplinary nature of this research, the introductory review which follows contains four somewhat separate sections. First, the research dealing with the impact of training on judges is considered. Special attention is given to relationships which might exist among various training procedures, their success, the setting in which the training is given, and the content of the material to be learned. Two training procedures, outcome feedback and process feedback (to be discussed below), which have been tested in applied settings and seem to have generality across topic areas, are of primary interest in this review.

Second, the general problem of nondiagnostic (irrelevant) information is defined and evidence is presented which shows that nondiagnostic information may adversely influence the accuracy of judgments. Although very little research has been aimed at the question of what role nondiagnosticity plays in decision making, a number a studies which can be reinterpreted in terms of nondiagnosticity are discussed, in addition to the few available studies which deal with the question directly.

Third, the specific task of in . soil texture-classification is outlined. Included in this section are: (a) A brief description of the soil judgment task; (b) A brief review of the available research on what might best be called the psychology of soil judgment; And (c), evidence that some materials which naturally occur in the soil, but are irrelevant to the soil-texture, interfere with accurate soil judgment, even for experienced soil scientists.

Fourth, the purpose of this study and the rationale behind the choice of the particular training procedures are given. Specific hypotheses which were tested by this study are outlined in this last section.

Outcome versus Process Feedback Training Procedures

No attempt will be made to review all of the available literature concerning the various investigations of the effectiveness of training techniques. The scope of this review has been narrowed for two reasons. First, a distinction has been drawn between training techniques, (Hammond, 1971), defined as those techniques which are designed to result in the trainee learning something new, and decision aids or techniques which are designed for use after the original decision has been made (see for example, Fischer, Edwards, & Kelly, 1978; Humphreys, 1979). Only training techniques which result in the decision maker learning how to make better judgments will be considered here. It was thought that training techniques would be more effective and versatile in reducing the influence of nondiagnostic information, and hence were the procedures chosen to be tested in this study. The training techniques selected here did not reflect any one specific approach developed in previous research, rather they were a composite of a number of techniques.

This review has been somewhat restricted because much of the training research has involved naive participants and may have limited generality to more applied judgment situations. However, whenever available, training techniques which have involved experienced decision makers as participants will be discussed rather than a comparable study involving naive participants who were taught to perform an unfamiliar task.

Roughly speaking, there appear to be two basically different procedures that have been used to improve the accuracy (validity and reliability) of the judgments made by decision makers. Both approaches depend upon the use of some form of feedback as the primary training device, but the two approaches differ in the form of this feedback. It is this difference which distinguishes the two training techniques. The first procedure has been called outcome feedback training; i.e., the training depends on feedback which tells the subjects how valid their judgments are with respect to some external standard or criterion. The second approach differs from the first in that the judges are trained through process feedback training; i.e., subjects are given some type of information about a "model" of their judgmental processes. Process feedback training may take a variety of forms, from complete model feedback (based on an optimal or normative model) to subjective model feedback (derived from the subject's own decision strategy).

Outcome Feedback Training. Outcome feedback is probably the most common form of training (Hammond, et al., 1977), and if an appropriate criterion is available: It is certainly the simplest training to explain to the judge and the easiest to employ.

Unfortunately, outcome feedback is beset with two major problems which decrease its effectiveness and versatility. First, in many (most?) applied settings the final validity based on the outcome of the decision is simply not determinable. For an agricultural example, consider the livestock judge. The judge may quite confidently rank order a collection of animals based on their judged slaughter potential. Whether the judged order is the "correct" order is not an answerable question, because the judges themselves define the correctness of the outcome by defining the criterion. For an other example, an audit may be considered accurate simply because the senior partner of the audit firm was the person who conducted the audit (Ettenson, Note 1). A second, and perhaps more disturbing problem with outcome feedback, even encountered when an appropriate criterion exists, is that this intuitively compelling procedure has often been found to be surprisingly ineffective (Einhorn & Hogarth, 1978). Despite long and repeated outcome training sessions, the amount of learning may be only moderate, and frequently fails to generalize to other very similar settings (Brehmer & Svensson, 1976).

Research on debiasing, that is, removing the adverse influence of various heuristics (Tversky & Kahneman, 1974), has shown that outcome feedback, consisting of right-wrong information, has been repeatedly unsuccessful as a training device (Kahneman & Tversky, in press). It should be noted that these studies have typically employed undergraduates and used tasks which are of questionable interest to the participants (Kahneman & Tversky, 1974). Outcome feedback training has failed even when combined with specific verbal or written instructions to avoid the detrimental effects of these biases (Fischhoff,1977). The ineffectiveness of outcome training found in these studies may arise from the fact that the decision maker is not able to localize the reasons why a decision lead to the particular outcome. For example, consider a prototypical example in which the participant is asked to judge the probability that x will occur, and it is consistently found that base rate information is ignored. Repeated feedback that the probability estimate is wrong does not allow the judge to easily discover that the source of the problem is that he/she has been ignoring base rate information. In addition, it is not difficult to think of situations in which bad outcomes may originate from good (even optimal) decisions, or good outcomes may result from a bad decision. In this situation outcome feedback may actually be detrimental.

In contrast to the generally unsuccessful training studies discussed above which typically used undergraduates and artificial tasks, outcome feedback has been employed with some success to teach both experienced and inexperienced decision makers how to be better calibrated. Calibration can be defined in various ways. The definition below is from Lichtenstein, Fischhoff and Phillips (1977), and is the most commonly accepted one. Formally, a judge is well-calibrated if, over a large number of occurrences, for all propositions the judge assigns a certain probability of occurrence, the proportion of these occurrences which are true is equal to the probability given. For example, if an investment broker repeatedly predicts that a stock has a 25% chance of returning earnings, and over a period of time these stocks do return earnings 25% of the time, then the broker is perfectly calibrated. Weather forecasters represent a group of decision makers who have been found to be generally well-calibrated (Murphy & Wrinkler, 1977). It is hard to deny the critical impact of outcome feedback as a training device to improve the calibration of these experienced experts. The outcome validity of a particular forecast is often quite visible the next day. Clearly, weather forecasters have a great deal of outcome training because of this frequent day-to-day evaluation of their predictions. It must also be remembered that weather forecasters are likely to be highly motivated, and generally quite experienced.

This relationship between experience and effectiveness of outcome feedback training seems to extend to other areas as well. A review of a great number of calibration studies seems to indicate that judges are better calibrated and more easily trained for a task with which they have had previous expertise and interest (Lichtenstein, et al., 1977). Judges are more poorly calibrated and training is often unsuccessful when the task is of less interest. An exception to this, worthy of note, is the recent series of training studies conducted by Sarah Lichtenstein and Baruch

Fischhoff. They were able to train naive participants to have improved calibration for answers to two-alternative general-knowledge questions using outcome feedback training (Lichtenstein & Fischhoff, 1978). They found that the subjects' calibration improved when given what Lichtenstein and Fischoff called "performance" feedback for only 200 items (a total of 2200 items were presented to the participants over the course of the entire experiment). Limited generality to other very similar types of probability tasks was also seen. The performance feedback used by Lichtenstein and Fischhoff (1978) involved both outcome feedback and suggestions specifically designed to show subjects how to improve their calibration.

Finally, one form of outcome feedback training called "discovery" has achieved some success in the undergraduate laboratory setting. In discovery training the judges are asked to think about, and then create for themselves, outcomes which might result from their decision. It is thought that having subjects discover an outcome may result in deeper processing and therefore result in better retention (Moscovitch & Craik, 1976). Presumably, the judges are providing themselves with a form of "pre-decision outcome feedback," that is, "what might happen if I do this," and thus causing this feedback to be processed more deeply. Compared to strict outcome feedback, Slovic and Fischhoff (1977) found that having subjects discover the effects of the hindsight heuristic (over confidence for memories of past events) was a more successful training technique. This type of self-generation of outcomes was also somewhat successful in training subjects to make the difficult distinction between directly asserted and implied information (Bruno & Harris, in press). These training procedures have the distinct disadvantage that the participants may not be able to make the critical discoveries for themselves. In most of the typical applications the subjects are eventually told what they should have discovered on their own (see, for example, Bruno, in press). This makes it difficult, if not impossible, to generalize the procedure to areas in which a criterion is not available.

Process Feedback Training. The other approach which has commonly been taken to training has been to provide the decision maker with feedback about the processes used to arrive at the judgment. In this case the feedback can be given with respect to a criterion (an optimal process) or simply as a description of the judge's own decision process. Of course, process feedback can be supplied with or without outcome feedback.

As a fairly representative example, process feedback is usually the approach taken by researchers working within the framework of the len's model (Brunswick, 1956).

In the len's model the weights obtained from a multiple-regression analysis are used to provide judges with feedback of their own judgment model. Hammond (Hammond, et al., 1977) has repeatedly argued, rather convincingly, that process feedback is considerably more effective than outcome feedback. First, Hammond notes that outcome feedback often does not arrive in time to help the decision maker (i.e., especially in cases where one or only a small number of decisions are to be made). And second, outcome feedback does not teach the decision maker what should have been considered as a part of the decision and what should not have been. In contrast a len's model analysis provides that judge with information about how each cue was used (or not used) in the judgment.

In a series, of studies Norman (see Norman, 1976, for a summary) used information integration procedures (Anderson, 1974) to compare outcome and process feedback training. He found that strict outcome feedback caused general changes in the integration model for both the weight and the model form (Norman, 1974a; Norman, 1974b), while process feedback resulted in more specific changes to particular weights in the model. This may indicate that when specific training impact is desired, such as in the training of experienced decision makers, process feedback may be more "finely tuned."

Although, it seems that both outcome and process feedback may be useful training techniques, exactly how to implement the feedback is not always clear. The studies discussed here which have used experienced decision makers seem to indicate that in general an attempt should be made to tailor the type of training specifically to the judges being trained and the material to be learned. In addition, the interest and experience of the decision maker may be a very important variable in the success of the training, when either type of training is used.

Impact of Nondiagnostic Information

The central problem faced by most decision makers is to evaluate and make a judgment based on a number of stimuli, each of which contains a wide variety of information. Unfortunately, not all of the available information should be given equal importance. Much of the information is likely to be quite diagnostic; that is, information which is relevant to making the best judgment. Other information will be less diagnostic. Finally, a part of the information may be completely irrelevant to the judgment problem at hand, that is, nondiagnostic, and should not be used by the decision makers to make the best possible judgment.

Decision makers may either knowingly, or unknowingly, rely on this nondiagnostic material to help them reach a judgment. Such reliance on irrelevant information may, at worst, result in a judgmental error and, at best, dilute the capacity of the decision maker. Obviously, in either case, decision makers should avoid any use of nondiagnostic information, and instead rely solely on relevant or diagnostic information. Formally, information is nondiagnostic if it should not be used as part of the input in a decision. Determination of nondiagnosticity can be through definition (as will be seen to be the case in soil judgment), by convention (the livestock judge does not look at the curliness of the pig's tail), or by the law (discrimination is not to be made by race).

Evidence of the adverse influence of nondiagnostic information has been found in a variety of research settings; examples are given below from research done in the areas of decision making (inference judgments), social psychology (attitude formation), research on aging (concept formation), and perception (recognition tasks). However, in general nondiagnosticity has been studied only as an issue tangential to the major point of these studies. Following consideration of a number of examples of research in which nondiagnosticity was only a tangential issue the few studies which directly investigated the influence of nondiagnosticity will be discussed. Finally, this review of the influence of nondiagnostic information will conclude with a discussion of what little research has been conducted using experts.

In a number of studies of the type in which the participant is asked to draw inferences from a set of cues, nondiagnostic information has been included, either intentionally or unintentionally, as one of the cues. For instance, Castellan (1973), using a multiple-cue probability learning task, reported that the presence of irrelevant cue dimensions retarded learning. He suggested that the decreased learning may have been due to subjects failing to ignore the irrelevant cues.

In their classic demonstrations of errors generated by dependence on the representativeness heuristic, Tversky and Kahneman (1974) showed that some of these mistakes may occur because subjects rely on information from nondiagnostic samples to make their judgments. It appears that this reliance on nondiagnostic information coupled with the ignoring of relevant information resulted in very poor overall judgment accuracy.

In a study of job-performance evaluation, Norman (1976) included both relevant and irrelevant types of information. In this situation he found that subjects almost always used both the irrelevant and relevant information. This inappropriate dependence on nondiagnostic information was further increased when the amount of other information was restricted.

Finally, Wallsten (1979) has observed nondiagnosticity effects in a pair of studies involving a variety of inference tasks. He asked participants to make similarity inferences based on cues having 1, 3, or 5 dimensions. There was a tendency for the subjects to concentrate on one aspect of the stimuli, without regard to diagnosticity, to the possible exclusion of more diagnostic cues. These studies by Wallsten do not point to a consistent directional effect of nondiagnostic material, but rather the effects seen in his studies seem to depend on the context.

In the social psychological area of attitude formation, Youngblood and Himmelfarb (1972) found that a neutral message received prior to a positive (or negative) communication resulted in less extreme attitude judgments than did exposure to positive (or negative) statements alone. Although it may not be, in fact, irrelevant it seems quite reasonable to view the neutral message as nondiagnostic with respect to the information which is to come later, and from this perspective nondiagnostic information seems to have been given inappropriate importance; that is, the participants seemed to make use of the information. In another social psychological study (Fromkin, Goldstein & Brock, 1974), subjects were exposed to both relevant and irrelevant derogation of a third party (a cab driver). The irrelevant derogation, which should have had no influence on the amount of frustration felt toward the cab driver, was found to have as much effect as relevant derogation. Although no other social psychological studies will be reviewed, it should be noted that a number of social-psychological issues, for example prejudice, could be meaningfully considered within the framework of nondiagnosticity.

Recent studies seem to point to a role that nondiagnostic information may play in explaining the difficulties faced by the elderly. For example, in a concept formation task, Hoyer, Rebok, and Sved (1979) used three groups of subjects: young, average age 20.6; middle-aged, average age 52.4; and elderly, average age 72.6. They found that for the elderly there was a disproportionate increase in both reaction time to solution and number of errors as the number of irrelevant dimensions increased. Evidence has accumulated suggesting that the ability to ignore irrelevant information may decrease with age (Layton, 1975; Rabbitt, 1964; Rabbitt, 1965).

Of the few studies which were directly concerned with nondiagnosticity, probably the clearest demonstration comes from the "bookbags-and-pokerchips" paradigm (Edwards, 1968). In one early study Shanteau (1975) showed that nondiagnostic samples (those with likelihood ratio of 1.0) had the effect of producing less extreme probability judgments; logically the judgments should not have been influenced at all. Troutman and Shanteau (1977) used a number of different bookbags-and-pokerchips tasks and found that nondiagnostic information presented in a number of ways had a definite impact on the inference judgments. In related research, Lichtenstein, Earle and Slovic (1975) found that when the second of two numerical cues was less diagnostic than the first, a substantial number of the inferences became less extreme. In this particular task the judgments should have become more extreme if the diagnosticity of the cues were correctly (optimally) interpreted.

A number of studies have also been conducted which investigate the role of irrelevant cues in perception. It has been found that irrelevant interactive cues which physically surround the target stimuli delay the search and recognition time of subjects in reaction time experiments (Krueger, 1973; Dixon & Just, 1978; Williams, 1974). These studies show that even when irrelevant information is not used as part of the final decision process, it may adversely influence the decision simply by requiring more processing.

Taken together, the results of this diverse set of laboratory studies seem to point to one consistent finding; nondiagnostic information has the potential to be an adverse influence on human judgment.

The question remains, however, as to whether similar nondiagnosticity effects are observed outside of the psychological laboratory. Unfortunately, the evidence on this question is sparse. Nevertheless, there is some research which suggests that even experienced experts may be subject to the adverse influence of nondiagnostic information. As early as 1971 Tversky and Kahneman reported that experienced psychological researchers used many of the same misleading heuristics (such as representativeness) which lead naive subjects to use nondiagnostic information (Tversky & Kahneman, 1971). They used members of the Mathematical Psychological Group and of the American Psychological Assocation to estimate the probability of significance of two hypothetical studies, based on different sample sizes. They found that most psychologists have a tendency to attach too much diagnosticity to sample size.

Similarly, Wallsten (1979) observed the some of the same tendencies to use nondiagnostic information in medical settings that were found in laboratory studies. Other research (Krischer, 1976), also in the medical setting, has shown that the surgical treatment for cleft palate may in some cases be decided based on nondiagnostic information (eg., the sex of the child).

Two recent studies have investigated the influence of nondiagnostic information in applied settings using information integration techniques (Anderson, 1974). The first, a re-analysis by Shanteau (1979) of Phelps' (1977) research on the training of livestock judges, showed that more experienced judges are better able to ignore the irrelevant factors, for example curliness of a pig's tail, than less experienced judges. A second recently completed but unpublished study by Nagy[1] has shown that when making personnel decisions, Master's level students rely on more of the irrelevant information (eg., sex of applicant), than do more experienced personnel directors.

Thus, while research which addresses the problem of the irrelevant information is limited, the evidence is certainly suggests that nondiagnostic information does produce many of the same problems for experts that it does for naive subjects.

Finally, given the established influence of nondiagnostic information, it is curious that there seems to be no attempt to formulate a theory describing how it influences the decision making process. The study of nondiagnostic information has reached a point where it is not enough simply to demonstrate its influence, rather, it is time to move on to the problems of understanding and reducing its influence.

<u>Soil Judgment</u>

This section consists of three parts. First, how soil is assigned one of twelve possible textural classes will be defined. Then the literature describing the accuracy of soil judgment will be reviewed. Finally, irrelevant materials which occur naturally in the soil will be discussed as a possible interfering source causing erroneous soil textural judgments.

Although soil scientists must make many judgments concerning various aspects of soil and soil conditions, only soil texture assessment will be considered here. It seems appropriate to comment briefly on the uses of soil-texture analyses. Soil classification is of great practical importance in a wide variety of agricultural, construction, and engineering settings. For instance, the choice of location for a building or the construction site of a dam can vitally depend on the soil-texture involved. For the farmer, irrigation practices, the type of crops, and amount of fertilizer are only a few of the decisions which depend on soil texture.

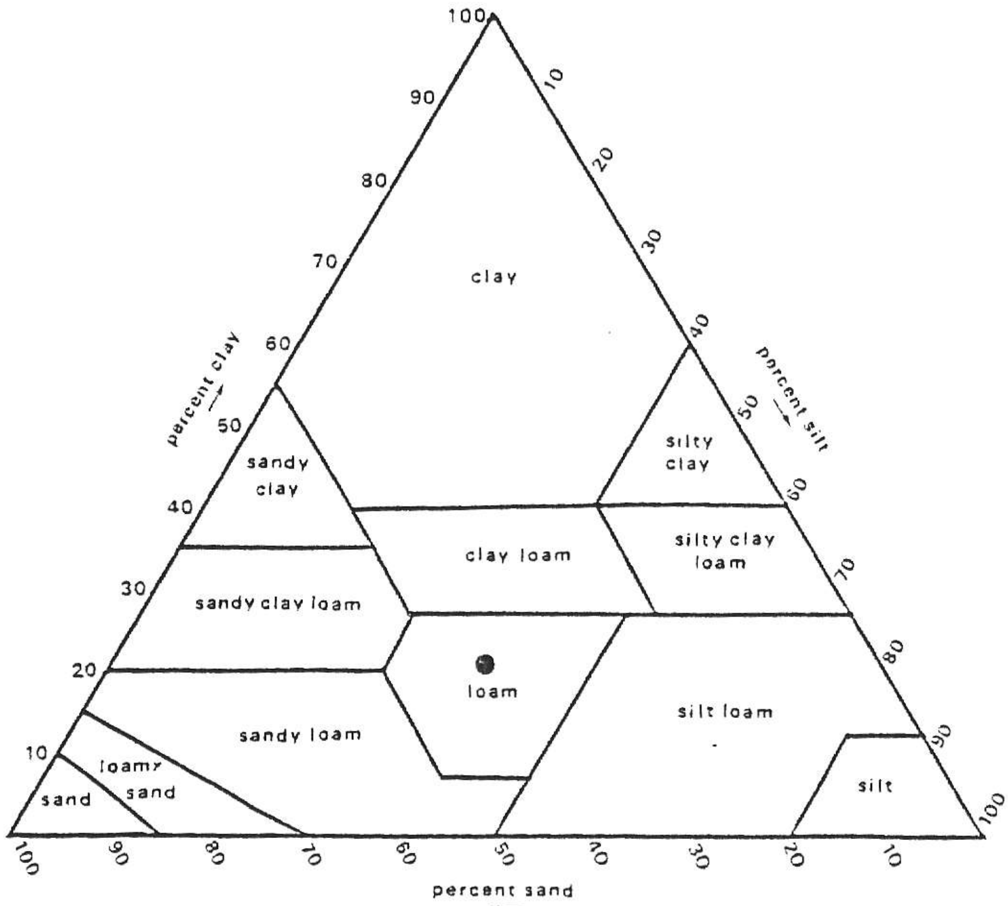
Soil Texture-Classification Assessment. It is not necessary to become an expert on soil science in order to understand the essentials cf the soil texture-classification system. For textural classification purposes, all type of soils can be assigned to one of twelve textural classes labelled; clay, sandy-clay, silty-clay, sandy-clay-loam, clay loam, silty-clay loam, loam, silt lcam, sandy-loam, loamy sand, silt, and sand. For the purpcses of assigning soil to one of ttese twelve textural classs, only three fundamental components of the soil are relevant: sand, particles sized from 2.0mm-.05mm; silt, soil particles sized from .05mm-.002mm; and clay, soil particles sized less than .002mm. The textural class of a particular soil is defined by the amount (percentage) of sand, silt, and clay which compose the soil (because sand, silt, and clay define the soil-texture, the sum of these percentages is 100%).

Soil texture was originally assessed cnly by hand, and the texture classes were solely intended to reflect field use of the soil. Based on these purely practical applications, the first textural classes were assigned and labelled with names which corresponded to the field usage of the soil. Subsequently means became available to mechanically analyze soil samples with laboratory techniques. The percentages of sand, silt, and clay could be determined independently of human judges. Naturally this development necessitated relating the older "field" texture scheme to the results of the laboratory analyses. Although a number of such systems were tried, the USDA Textural Triangle (1951) is currently accepted almost exclusively in this country. See Figure 1 for the triangular representation of the textural classification scheme. Formally, the triangle represents a three-dimensional space with the three coordinates representing the % of sand, silt, and clay. In addition, a restriction that the sum of the three dimensions must add up to 100% is imposed. This must occur because, by definition, soil-texture is dependent only upon sand, silt, and clay.

The USDA textural traingle was developed as an attempt to retain the old texture names and corresponding field characteristics. This, tc a large extent, explains the peculiar shape of many of the texture classes (see, for example, sandy loam in Figure 1). Given the values of percentage sand, percentage silt, and percentage clay, the location of this point is determined on the triangle, and its location defines the soil textural classification. For example, soil containing 40% sand, 40% silt, and 20% clay is plotted in Figure 1. As can be seen from its location, the soil would be given a textural class of loam.

Soil texture can be determined in two ways: with mechanical procedures carried out in a laboratory, and with the use of the feel (sometimes called field) method performed by hand. Therefore, when assessing soil-texture in the field, the job of the soil surveyor is to determine the textural classification of the soil as it would be

Figure 1 USDA soil textural classification triangle: Soil with 40% sand, 40% silt, and 20% clay plotted.


100
90
80
70
60
50
40
30
20
10
percent clay
10
20
30
40
50
60
70
80
90
100
percent silt
100
90
80
70
60
50
40
30
20
10
percent sand
clay
sandy
clay
silty
clay
clay loam
silty clay
loam
sandy clay loam
loam
silt loam
sandy loam
loamy
sand
sand
silt

classified if the analysis were done mechanically in a laboratory. However, because of the constraints imposed by the rather high cost of mechanical soil sample analysis (often as high as $100 per soil sample and the additional time involved, only a very small portion of the total number (approximately 2-5%) of field texture determinations can be sent to a laboratory for analysis. For the psychologist this situation is advantageous, since the field assessment of soil-texture is purely subjective, while the laboratory analyses provide an objective criterion. It should be noted that over time the standard laboratory procedures have occasionally even been changed, and have recently been seriously questioned (Hawando, 1978) on the grounds that in most laboratory analyses the clay particles have been assumed to be round, however, this is often not the case and rather the clay is flat. This mistake may result in considerable error in the clay percentage. However, the commonly used laboratory assessments are generally quite reliable (if not valid), and are currently the accepted standard.

Accuracy of Soil Texture Assessment. The accuracy of the subjective feel method of texture assessment has not gone untested; rather, most of the evaluation has been done at an informal level, especially during the creation of the USDA textural-triangle. A few rigorous studies do exist, however. In research conducted by Foss, Wright, and Coles

(1975) and Leenheer, Appelmans and Caestecker (1954), accuracy was measured by counting the frequency with which field textural classifications made by soil survey personnel agreed with a laboratory analysis of the same soil. Despite time and location differences between these two studies, both Foss et al. (1975) and Leenheer, et al. (1954) concluded that soil-texture assessments agreed with the laboratory roughly 50-60 percent of the time (1 of 12, or approximately 8% would be expected by chance).

In the two studies discussed above accuracy was defined in terms of only the textural classifications. In comparison, Hodgson, Hollis, Jones and Palmer (1976) measured accuracy based solely on individual judge's specific estimates of the percentages of clay and silt. They found that in 80% of the cases the soil scientists were within ±8% of the standard for clay and ±12% for silt. Although the basis of their analyses was somewhat arbitrary, it did indicate the there was room for improvement.

More recent research using soil scientists at Kansas State University examined the validity and reliability of soil-texture judgments. In a study involving 8 experienced soil judges (members of the soil judging team, faculty, and staff), 160 soil-texture assessments were obtained. Forty-seven percent of these judgments agreed with a mechanical analysis conducted by the National Soil Survey Laboratory (Shanteau & Gaeth, 1979). In this same study, reliability of the expert's judgments was also measured, on an intra-judge basis. Shanteau and Gaeth found that intra-judge reliability was roughly 50%.

In another validity study, field survey data were reanalyzed from entries in the Soil Survey Investigations Report No.4 (1966) of Kansas. It was found that 57% of 320 field textures agreed with the subsequent laboratory analyses (Gaeth & Shanteau, 1979). In a similar study of accuracy, texture judgments coming from the 1978 Soil Judging Contest were analyzed for validity. In this case 38.8% of the judgments agreed with the mechanical analysis. Any interpretation of this result must be tempered by the fact that these judges were students and not yet professionals.

Interfering Materials. A considerable amount of research indicates that soil judges are typically correct roughly 50% of the time in their texture judgments. This certainly leads to the question of how the accuracy of soil textural-classification can be improved. One approach to improved accuracy is to consider what causes errant judgments. Sources have suggested that difficulties could originate due to the presence of material other than sand, silt, and clay in the soil (Clarke, 1936; Leenheer, et al., 1954; Russell, 1973 and other sources mentioned below). Specifically, Foss et al. (1975) indicate that coarse fragments, free iron, organic material, and very coarse or very fine sands might cause increased difficulty assessing the texture of soil. Hodgson, et al. (1976) indicated that organic carbon and calcium carbonate seemed to be related to errors in texture judgments.

To further investigate the effects of such interfering materials, a survey was taken of ten soil scientists in Kansas each of whom had a great deal of professional experience assessing soil-texture in the field. They were asked to rate the effect that a number of naturally occurring materials had on the accuracy of their soil-texture judgments. All of the interfering materials discussed above were included in the survey. Each was rated as having some detrimental effect (see Table 1 for a summary). In addition, overly moist soils, tilled soils, coarse silt, clay type, and other possible interfering factors were suggested by the participants. Based on these results and personal communication with soil team coaches (O. W. Bidwell, Kansas State University; J.E. Foss, University of Maryland; R. Pennock, Pennsylvania State University) evidence points to interfering material as a likely cause for reduced accuracy even in expert soil judgment.

Table 1

Ratings[1] of the Influence of Materials Irrelevant to Soil Textural Classification

Material	Kansas Soil Scientist										
	1	2	3	4	5	6	7	8	9	10	Avg.
Excessive Moisture	5	7	4	6	4	6	4	4	6	6	5.2
Coarse Sand	4	1	3	6	6	4	2	1	3	7	3.7
Coarse Fragments	2	7	1	4	3	6	1	7	6	7	4.4
Fine Sand	5	7	1	1	2	5	6	2	1	7	3.7
Calcium Carbonate	4	5	2	5	2	6	4	7	4	4	4.3
Organic Material	6	5	2	NE	4	5	5	NE	5	2	4.3
No Moisture	1	1	1	1	1	1	1	7	4	6	2.4
Free Iron	3	NE	NE	NE	NE	NE	NE	NE	NE	2	2.5

Note. NE = no experience.

[1]Ratings were taken using a scale from 1 (little or no adverse interference) to 7 (considerable or likely adverse interference).

Purpose and Rationale for This Study

This study was designed to test three major hypotheses. First, interfering material in the soil adversely influences the accuracy of the soil texture-classification judgments made by experienced soil judges. Second, the two training procedures (lecture and/or interactive) will help soil judges reduce the influence of the interfering materials. And third, reduction of the adverse influence of the interfering materials results in improved accuracy of the soil texture-classification judgments. Details of each of these hypotheses are given below.

The literature concerning research with nondiagnosticity shows a clear need for more empirical research involving experienced decision makers. In addition, as discussed above, there is a need to move on to the problem of understanding and reducing the influence of the irrelevant information. Because of the well-defined nature of the soil texture-classification problem, and the apparent adverse influence of irrelevant (i.e., nondiagnostic) materials, soil texture-classification judgment was chosen as the area to study the effects of nondiagnostic materials on accuracy and to attempt to test training techniques designed to reduce the influence of the irrelevant materials. Thus, the first issue addressed in this study was the question of whether irrelevant materials can be shown to cause systematic errors in soil-texture judgments. The two materials chosen to be tested were coarse fragments (particles larger than 2 mm.), and excessive moisture present in the soil (above the level of moisture which is normally used in hand textural classification assessment).

Second, to reduce the presumed interference of the irrelevant materials, two training procedures were used; these were designed to improve accuracy of the texture judgments through reduction of the influence of the irrelevant materials. As discussed in the literature review, no training procedures were found which were specifically designed to reduce the influence of nondiagnostic information. However, based on a combination of the two general types of training discussed earlier, two different training procedures were chosen to be tested here. Neither involved the use of direct outcome feedback. One, a strictly verbal series of instructions, called "lecture training", was selected to parallel the type of training soil judges typically receive in the classroom. The second training procedure, called "interactive training", is more in line with the previous work using the discovery method to produce deeper processing. In this training procedure the judges were asked to practice the suggestions given in the verbal training and encouraged to discover the influence of the irrelevant materials for themselves. This training is called interactive to stress the active role played by the trainee. Both the lecture and the interactive training are forms of process feedback in that the judges are taught suggestions how to better texture soils, and are not forms of outcome feedback training.

It must be noted that the names chosen for the two training procedures are intended only as convenient labels. There seem to be no appropriate labels available from previous research.

Third, it is possible that the first hypothesis is true and irrelevant materials do adversely influence the judgments made by soil judges. Further the second hypothesis may be verified, and some combination of the training procedures does result in a reduction of the adverse influence of the irrelevant materials. This could occur if the irrelevant materials some how aided the judgment (this seems unlikely, however, it is possible). Nevertheless, the accuracy of the texture-classification judgments does not improve. Hence the third hypothesis specified that reduced influence of the irrelevant materials would lead to improved performance for the soil judges.

METHOD

The judge's general task was to assess the textural classification and estimate the percentages of sand, clay, and silt for 16 soil samples in each of three evaluation sessions. In addition, each judge was trained with one of two training programs. The training programs contained different content but followed parallel patterns. The content of one of the two training programs was either designed to reduce the influence of excessive moisture, or to reduce the influence of coarse fragments. Each training program consisted of two training sessions; the first type of training was lecture training and the second was interactively based. To clarify terminology for future use, training procedure will refer to either the lecture or interactive training procedures, while training program will refer to the separate content (i.e., either designed to reduce the adverse influence of excessive moisture or coarse fragments) of the training procedures. Thus, each judge was given both the lecture and the interactive training procedures, and either the coarse fragments or excessive moisture training program.

The impact of the two training procedures was measured through the use of three evaluation sessions; thus yielding a total of five sessions. In addition, both groups received the same series of five sessions in the same temporal order. The first session was an assessment of the initial baseline ability of each of the judges to ignore the influence of the interfering material. The second session contained lecture training designed to reduce the adverse influence of the interfering material. The third session was an evaluation of the lecture training. The fourth session consisted of interactive training of the soil judge using actual soils. The fifth and final session was an evaluation of the cumulative impact of the two previous training procedures. The remainder of this section contains a description of the participant soil judges, and the set of soil samples used as stimuli followed by a detailed discussion of each of the five sessions.

Participants and Stimuli

Participants. Thirteen soil judges (2 females, 11 males) were recruited from an advanced Soil Morphology class (#415), taught in the fall of 1979 at Kansas State University. The judges were paid $12.00 for their completion of this study; the total time requirement was from 3.5 to 4.0 hours. Each of the judges had taken at least one soils class in which texture-classification assessment of soil was taught. All judges had additional experience in the determination of soil texture in the field, either through various combinations of personal work experience, laboratory work associated with a soil morphology class, or previous participation on a soil judging team. The judges were randomly assigned to one of the two different training groups (that is, they either received the excessive moisture or the coarse fragments training program), based on a pseudo-random schedule arranged prior to the beginning of the experiment. The schedule was designed to keep the size of each of the two training groups as equal as possible during the course of the experiment. This was necessary because the total number of judges that would be available was not known prior to the end of the experiment, and equal group size facilitated the ease of analysis. One of the thirteen judges was dropped because of inadequate familiarity with soil texture-classification; this judge had recently come from Europe and was not experienced with the USDA system of soil classification.

Soil Stimuli. Ten Kansas soil samples were used in this experiment. These ten soils included 5 soils containing coarse fragments and 4 filler soils which did not contain coarse fragments. All of the soils were either collected by the researcher or provided by a USDA Survey Soil Scientist[2]. Each of the soils was found within a 200-mile radius of the Manhattan, Kansas area. All soils used in the experiment were analyzed by the Front Range Laboratory in Fort Collins[3]. The results of these analyses are given in Table 2. This set of ten soils provided the basis for the creation of two collections of stimuli; the evaluation collection of 16 soils (more throughly described below), and the two training soils. All of the soils were chosen to have reasonably similar amounts of sand, silt, and

Table 2[1]

Percentage of sand, silt, clay, coarse fragments, and excessive moisture for each evaluation and training soil.

Soil #	Usage	%Sand	%Silt	%Clay	%C.F.(Natural)	%C.F.(Used)[2]	%E.M.(Used)[2]
2	test	5(5)	66(70)	29(25)	30	23	20
4	test	17(20)	60(60)	23(20)	25	26	30
6	test	52(20)	29(55)	19(25)	38	33	25
1	filler	62	20	28	--	--	--
7	filler	22	49	29	--	--	--
8	filler	8	63	29	--	--	--
9	filler	7	65	28	--	--	--
3	training	23	61	16	44	--	--
5	training	26	48	26	82	--	--

Note. C.F. = Coarse Fragments, E.M. = Excessive Moisture. Values in ( ) reflect estimates made by the soil scientist.

[1]This is based on a mechanical analysis conducted by the Front Range Laboratory, Fort Collins, Colarado (see text footnote 3).

[2]Soil used in perceptual training procedures.

clay, but yet to also be distinguishable. The training soils were used in the interactive training procedure and will be discussed in more detail in the section on active training.

Following collection, all of the soils were dried, hand crushed, and passed through a 2-mm sieve. It should be remembered that 2 millimeters is the cutoff size between the largest sand particle and coarse fragments. The percentage of coarse fragments was determined by weighing the coarse fragments which did not pass through the sieve and dividing this into the weight of the soil which did pass through the sieve (see Table 2 for the percentage of coarse fragments in each soil). In this way the coarse fragments were removed from the soil samples and the "naturally occurring" percentage of coarse fragments was determined for each soil. Henceforth, each of the sieved soils (soil with coarse fragments removed) will be called base soils.

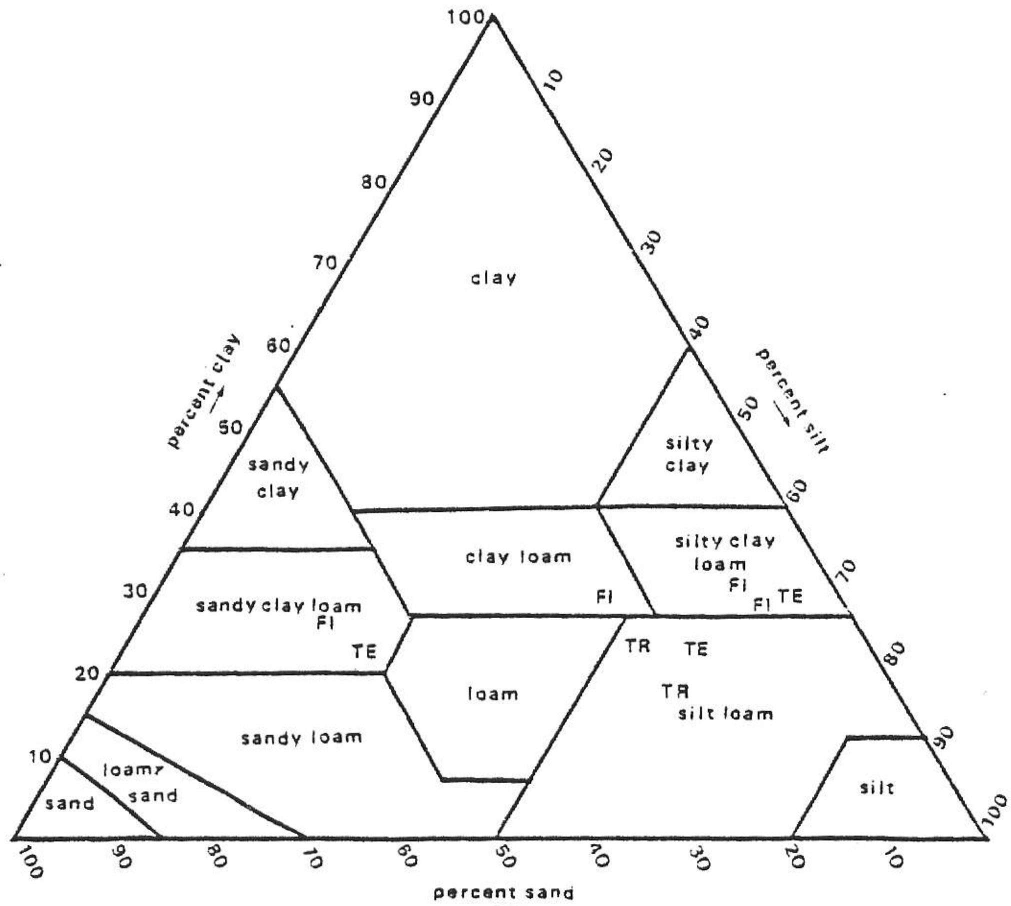
The test set, which comprised twelve of the 16 soils, was created by preparing 100 grams of each of the 3 base soils as determined by a 2-levels of coarse fragments by 2-levels of moisture complete factorial design. For the coarse-fragment factor, the low level of coarse fragments was chosen as the base soil containing no coarse fragments. The high level of coarse fragments was roughly the same amount found in the soil naturally. Because there was only a limited amount of each soil, the exact proportion of coarse fragments which was used depended on the availability of the soil, and in all cases the amounts were rounded off to simplify the mixing process. Only coarse fragments which were originally sieved from a soil were ever returned to that soil. That is, the evaluation soils may not have contained exactly the percentage of coarse fragments which occurred naturally, but all fragments added to the soil were indigenous. See Table 2 for the exact percentages of coarse fragments returned to the soils for use in the evaluation sessions.

For the moisture factor, the low level of moisture was chosen as the dried base soil with virtually no moisture in it. Actually, because extensive oven drying was not carried out (extensive drying changes the appearance of the soil substantially), a small amount of moisture remained in the base soil samples. Despite the presence of this moisture, the soils were considerably drier than the level of moisture typically used when a hand texture is taken. The high level of moisture was arrived at through the help of an experienced professional soil scientist*. This level was chosen in an attempt to duplicate the conditions which might prevail in a field situation after a heavy rain or due to the presence of high ground water. For all of the base soils in the evaluation set, the high level of moisture content was deliberately chosen to be in excess of the level of moisture normally used to assess the textural classification of soil by hand. See Table 2 for the exact amounts of moisture used in each of the evaluation soils.

Thus, each of the 3 base soils was prepared in 4 different conditions yielding a set of 12 soils to be used in each evaluation session. These 12 soils will be called the test soils. That is, each test soil was either dry, excessively moist, contained original coarse fragments and dry, or contained original coarse fragments and excessively moist. Added to these 12 test soils were 4 filler soils to complete the evaluation set of 16. An attempt was made to select four filler soils which were reasonably similar to the test soils, without duplicating them exactly. Figure 2 shows the textural classification and location on the textural-triangle of all of the 16 evaluation soils.

Soil Preparation. In order to ensure that the correct amount of moisture was in the soils, they were prepared "fresh" each day they were to be used. Only the base soils required no preparation before presentation to the soil judges. In order to simplify the preparation process, the moisture or coarse fragments mixture was always created from 100 grams of the base soil. Of course, only those coarse fragments which were originally obtained from the particular base soils were added back to the soil. After preparation the soils were stored in an 8-ounce plastic cup sealed with a water tight lid. Each container was coded with a color to indicate whether it was from a pre- (yellow), mid- (white), or final-evaluation (red) set, and with a code number to indicate which of the 16 test soils was contained inside.

Figure 2 Textural classification of the 3 test (TE) soils, 4 filler (FI) soils, and 2 training (TR) soils.


percent clay
percent silt
percent sand
clay
silty clay
sandy clay
clay loam
silty clay loam
sandy clay loam
loam
silt loam
sandy loam
loamy sand
sand
silt
FI
TE
TR
100
90
80
70
60
50
40
30
20
10

Procedure

General Procedure

There were two quite different procedures used in the five sessions, one procedure for the three evaluation sessions, and one for the two training sessions. The temporal order of session one through five was: Session 1 (Pre-Evaluation); Session 2 (Lecture Training); Session 3 (Mid-Evaluation); Session 4 (Interactive Training); Session 5 (Final-Evaluation). This description of the procedure will be divided according to the same pattern. First, the general evaluation procedure common to all three of the evaluation sessions will be explained. Next, the general rationale and sequence of events will be described for the two training sessions. Following these general descriptions of the procedures, each of the five sessions will be considered individually in the order actually experienced by the participants.

General Evaluation Procedure. All three evaluation sessions (pre, mid, and post), given in sessions 1, 3, and 5, were executed in the same fashion. The judges were run through the evaluation sessions singly, with the occasional exception that one judge may have been finishing up while a second judge was starting. When this occurred the judges were separated by at least 20 feet and were never allowed to communicate.

The order in which the 16 soil samples were presented to the judge for textural classification analysis was determined using the following 3 steps. First, one of the filler soils was randomly chosen as the initial soil. Second, one soil was randomly chosen from each of the three groups of test soils. The next 4 soils were randomly selected, one from each of the three test soils and one from the four filler soils. The next 4 soil samples selected in the same fashion as were the last 4 with the additional restriction that the final soil was not a filler. In the evaluation sessions after the first (i.e., mid- and final-), the same technique as just described was used with the additional restriction that no soil could appear in the same sequential position as it had in an earlier evaluation session.

After the order of presentation was determined, a response booklet was prepared. The booklet contained 16 ordered pages each listing which soil was to be textured. During an evaluation session the experimenter obtained the container corresponding to the number indicated in the booklet and handed it to the soil judge, who was told to estimate the percentages of sand, silt, and clay, and to assess the textural classification of the soil. In all cases water, paper towel, a pencil, paper, and a USDA textural-triangle were available for use during the textural classification process. After the judge had completed the textural classification assessment, the experimenter asked the judge to verbally report the texture-classification, and estimates of the percentages of sand, silt[5], and clay. These responses were recorded in the booklet by the experimenter. Next the time required for the texture assessment was recorded, after which the judge's responses were repeated back in order to ensure accurate reproduction of the intended responses.

Then the judge was asked to indicate how confident he/she was in the texture assessment just given. The judges were instructed to make their confidence judgment based solely on their texture-classification judgment, not on their estimates of sand, silt, and clay. The confidence response was obtained with the use of a Galton bar with a millimeter ruler on the back. The bar was labeled at each end; left end as "incorrect" texture, and the right end as "correct" texture. The judge was told that the end labeled incorrect texture indicated 100% certainty that the texture assessment given was incorrect. While correct texture indicated that the judge was 100% certain that the texture assessment given was the correct one. The judge was asked to position the slider somewhere between the two ends to describe how much confidence he/she had in the textural classification given. The resulting value was read from the reverse side of the Galton bar and recorded in the booklet with the other responses for the soil. This value was not reported back to the judge. Up until this time the judge was allowed to make changes or corrections in his/her judgments for the soil which was being assessed. A judge was never allowed to see or change a response to a soil seen earlier.

When this was completed the "used" soil was placed in a separate container, and the judge washed his/her hand. Following this the experimenter obtained the soil sample whose number was next in the booklet. This next soil was then given to the judge for texture assessment and the procedure was repeated.

Each of the 16 soil samples was presented singly as described above. A soil sample typically required about 4 minutes for the judge to assess the texture, evaluate the percentages of sand, silt, and clay, indicating the confidence in the textural classification, and for recording of these values. Thus, each evaluation session consisting of 16 soils, lasted about 70 minutes, however, 90 minutes were scheduled for each evaluation session so that the judge did not feel any time pressure.

General Training Procedure. The two training groups were given training which followed similar patterns in sessions 2 and 4. First both groups received the lecture training procedure, and then the interactive training procedure. The content of the training procedures differed, however, and represent the two different training programs. One group, called the coarse fragments training group, was given lecture and interactive training designed to reduce the adverse influence of coarse fragments. The other group, called the excessive moisture training group, received training parallel in form to that received by the coarse fragments groups, but dealing specifically with the problem caused by excessive moisture.

Because of the similarity of the two training programs, in the following sections they will be referred to as training for interfering materials, unless it is necessary to distinguish between the training program for excessive moisture and coarse fragments. See Appendix for the complete protocol for both the lecture and interactive training procedures.

The content of the training programs was developed with the help of Dr. Bidwell, an experienced soil scientist. The programs contained what were essentailly reminders of information the judges were (or at least should have been) already familiar with.

Specific Procedures

Pre-Evaluation. Each participant was given a pre-evaluation prior to any training. The pre-evaluation was conducted as described above in the general evaluation section. Based on earlier pilot research, very short instructions were found to be adequate. Basically, the judges were instructed (a) to texture the soil using the feel method as they would typically do in the field, (b) to take as much time as they felt necessary, since they were being timed only for my information, (c) that high accuracy was the most important consideration, (d) how to use the Galton bar, and (e) that some of the soils to be textured would be given to them in various conditions, and that this was done to simulate conditions which might be faced in the field. Finally, they filled out an informed consent form which described the length of the experiment, their right to quit the experiment at any time, and the amount they would be paid upon the completion of the experiment. The judges were asked if they had any questions concerning the procedure, however, because of their previous experience with soil texture assessment, only the use of the Galton bar was new to them, and this never caused any reported misunderstandings (the

At the conclusion of the pre-evaluation session judges were scheduled for complete protocol of the instructions in included in the Appendix). the next session (lecture training) and instructed not to discuss what they had done with anyone else until the end of the entire project.

Lecture Training. Both of the lecture training programs, in session 2, consisted of three sections which were recited to the judge by the experimenter (a set of concealed cue cards was used to assure that the same instructions were given each time). The lecture training required approximately 30 minutes. A different lecture training protocol was used for the two groups; however, the structure of the two lecture training programs was made as identical as possible. Each contained the same number of points and used reasonably parallel wording.

The lecture training procedure was divided into three parts. The first part contained evidence that the interfering material actually adversely influenced even the judgments of expert soil judges. The second part provided the judges with a definition of the particular interfering material. The third, and final part of the lecture training was a list of seven suggestions designed to reduce the adverse influence of the interfering materials. Details of each of the separate parts of the lecture training procedures are given below.

The lecture training began with evidence that the particular interfering material (coarse fragments or moisture) was a cause of erroneous texture-classification assessment, even for expert soil scientists. This was accomplished with 5 pieces of evidence, each designed to convince the judge that the interfering material was a problem. During the lecture presentation of this evidence a short summary of each piece of evidence was written on a cardboard sheet ard covered by another cardboard sheet. Then, as each piece of evidence was discussed, the summary sentence was uncovered fcr the judge to see (see Figure 3 for a list of the summary sentences). When the complete set of 5 points had been discussed, the judge was requested to repeat back the 5 pieces of evidence and asked if there were any questions. The next section of the lecture training was used to formally define the particular interfering material relevant for each of the training groups. Coarse fragments were defined to be particles in the soil larger

Figure 3 Evidence used in lecture training to show adverse influence of interfering materials.

THIS BOOK CONTAINS NUMEROUS PAGES WITH THE ORIGINAL PRINTING BEING SKEWED DIFFERENTLY FROM THE TOP OF THE PAGE TO THE BOTTOM.

THIS IS AS RECEIVED FROM THE CUSTOMER.

EVIDENCE THAT EXCESSIVE MOISTURE AFFECTS THE ACCURACY OF TEXTURE JUDGMENTS

1. Experience of soil team in Iowa.
2. Dr. Holzhey's comments.
3. Research in soil science journals.
4. Survey of USDA soil scientists.
5. Analysis of the data in the Kansas Soil Description.

EVIDENCE THAT COARSE FRAGMENTS AFFECT THE ACCURACY OF TEXTURE JUDGMENTS

1. Experience of soil team in New Mexico.
2. Dr. Holzhey's comments.
3. Research in soil science journals.
4. Survey of USDA soil scientists.
5. Analysis of the data in the Kansas Soil Description.

than 2 mm. Excessive moisture was defined to be the presence of moisture above the amount normally in the soil when a hand texture-classification analysis is conducted. In all cases the definition given the judge was consistent with previous training.

Finally, the last section contained seven suggestions designed to help the soil judge deal with the problem of the particular interfering material for which he/she was being trained. Each of the 7 suggestions was discussed in detail, and again, a short list of summary sentences was shown simultaneously with the oral presentation. This list of summary suggestions is given in Figure 4 for both the coarse fragments training and the excessive moisture training groups. After the recitation of the list of suggestions was complete, the judge was asked to summarize each of the points in his/her own words. Any deletions or misunderstandings were corrected, and the judge was again asked if he/she understood each of the suggestions. Any additional questions concerned directly with the training suggestions were also answered. After completion of the lecture training, the judges were scheduled for the mid-evaluation.

Figure 4 Seven suggestions used in the interactive training.

SEVEN SUGGESTIONS TO HELP YOU TEXTURE SOILS CONTAINING EXCESSIVE MOISTURE

1. REMOVE THE EXCESS MOISTURE.
2. ROUGHLY ESTIMATE THE AMOUNT OF EXCESSIVE MOISTURE.
3. DETERMINE THE CONDITION OF THE SOIL-MOISTURE MIXTURE.
4. TAKE A REPRESENTATIVE SAMPLE OF THE MIXTURE.
5. ESTIMATE THE SAND CAREFULLY.
6. ESTIMATE THE CLAY CAREFULLY.
7. BREAK DOWN THE COMPACTED CLAY.

SEVEN SUGGESTIONS TO HELP YOU TEXTURE SOILS CONTAINING COARSE FRAGMENTS

1. REMOVE THE VERY LARGE FRAGMENTS.
2. ROUGHLY ESTIMATE THE AMOUNT OF COARSE FRAGMENTS.
3. EVALUATE THE SIZE OF THE REMAINING FRAGMENTS.
4. DETERMINE THE SHAPE OF THE FRAGMENTS.
5. ESTIMATE THE SAND CAREFULLY.
6. ESTIMATE THE CLAY CAREFULLY.
7. BREAK DOWN THE COMPACTED CLAY.

Mid-Evaluation. The mid-evaluation, session 3, occurred within 3 days of the lecture training. The same procedure as described for the pre-evaluation was used except that the judges were briefly (approximately 3 minutes) reminded of the instructions and the color of the soil containers was changed. After the judge completed the texture assessment of the 16 soils, an appointment was made for the next session (interactive training).

Interactive Training. The interactive training procedure, session 4, involved "hands on" practice of the same seven suggestions presented in the lecture training session. The interactive training consisted of four sections, and required approximately 30 minutes to complete. The training began with a review of the seven suggestions given in the lecture training procedure. The second section was designed to demonstrate to the soil judge that the interfering material caused errors in his or her classification assessments. The third section of the interactive training procedure consisted of a series of practice soils containing the interfering material. The fourth section involved an attempt to have the soil judge gain confidence in his/her ability to deal with soils containing interfering material.

In the first section the judge was shown the list of 7 suggestions (see Figure 4) from the lecture training and then asked to paraphrase each of the suggestions. If any of the suggestions were forgotten or incorrectly recalled, they were repeated back to the judge. After this review the judge was told that this session would be used to practice the seven suggestions when actually making soil texture-classification judgments.

The second section of the interactive training involved the judge assessing the texture of a soil containing interfering material. The first interactive training soil was chosen as one which had been texture-classified in the two earlier evaluation sessions (sessions 1 and 3) when it was used as a filler soil. The judge's two responses to this soil were available to the experimenter at the time of the interactive training. After completion of the texture assessment the judge was told that he/she had analyzed the soil the first time when it did not contain any interfering material. Then the experimenter reported to the judge the previous responses which were the most disparate with respect to the current texture assessment values. The point which was made to the judge can be paraphrased as follows: "when you saw this same soil in an earlier session it had no coarse fragments (excessive moisture) in it. Then you said ____________, but now you say ____________. As you know these two texture assessments should be identical. Therefore it must be the coarse fragments (excessive moisture) which changed your analysis." Using this as a motivating factor, the judge was told that the remainder of the session would be devoted to practice in reducing the effects which interfering materials have on the accuracy of the texture assessments.

The third section of the interactive training involved the presentation of a set of 3 soils which had never before been textured by the judges. The training techniques used for each of the 3 soils were similar. First, the judge was asked to determine the texture of the soil without any interfering material. Then, the judge was asked to assess the texture of the soil when it contained the interfering material mixed in. The point was strongly emphasized that the soil texture itself had not changed with the addition of the interfering material. Therefore, any differences they may feel in the soil <u>should</u> <u>not</u> influence the texture assessment. The same soil was shown two more times with increasing amounts of the interfering material present. At each stage the judge was asked to estimate the amount of interfering material present in order to further emphasize its presence. Essentially the same procedure was used for the two remaining training soils.

The fourth and last section required the judge to assess the textural classification of a soil containing a relatively large amount of the interfering material. Again, this was a soil which had been used in the earlier evaluation sessions (as a filler soil), and previous responses were available to the experimenter. Upon completion of the judge's texture assessment, the responses were compared to the texture responses made earlier. In this instance, the values which were closest to the current responses were reported back to the judge. This was done to provide encouragement and hopefully result in the judge leaving the interactive training session with a feeling of accomplishment. The interactive training session required about 30 minutes and upon completion the judge was scheduled for the final-evaluation session.

Final-Evaluation Session. The final-evaluation session, session number 5, followed within 3 days of the interactive training. The same procedure as described for the mid- evaluation was used again, except that the color of the containers and the code numbers on the soil samples were changed. Following the usual evaluation procedure, the judges were asked a number of questions concerning how they rate (using a 1 to 7 scale) their own ability to determine the soil texture-classification and estimate the percentage of sand, clay, and silt, which of the sessions and/or training techniques they felt were most effective, and finally their attitude toward the experiment and the experimenter was determined. At this time the judges were read a very short debriefing statement explaining that they would receive a complete feedback package later on. They were asked if they had any specific questions concerning the experiment which they wanted answered immediately. They were then paid the $12.00 participation fee.

Feedback. A feedback package was prepared for each judge as partial payment for his/ her participation. This was given to the participants after all the judges had completed the entire experiment, in all cases within three weeks of the conclusion. This feedback included: A brief explanation of the rationale behind the experiment; a set of graphs which were plots of the judge's average responses to each soil compared to a criterion value (either a laboratory analysis or an assessment made by a professional soil scientist[6]); and, an explanation of how to interpret the graphs and possibly improve the accuracy of later texture assessments.

RESULTS

In the first major section of the results, the raw-score results (%-of-sand, %-of-clay, and the textural classification judgments) will be compared to the hypotheses which were originally proposed (see Introduction). The %-of-silt estimates will not be considered as dependent measures because the soil judges determine the percentage of silt throgh subtraction of the %-of-clay and %-of-sand estimates from 100% (see footnote 5). The raw-score results provide tests of the _effect of soils_, and the _influence of the interfering materials_. The raw-score analysis also provides tests of the _overall impact of the training procedures_, and the _specific impact of the training programs_ as interactions with the evaluation factor. The results obtained from the confidence ratings and the time measures will be briefly considered.

To more efficiently test for the impact of training, a series of difference-score analyses, based on within-evaluation session differences, will be considered for the %-of-sand, and %-of-clay estimates, and the texture category judgments. Thus, the second section of the results will include evaluation of the _effect of soils_, the _overall impact of the training procedures on the percentage estimates_, and the _overall impact of the training procedures on the texture judgments_.

The third major section will be an evaluation of difference-score results for the individual judges. Two approaches are used: A statistical evaluation of the impact of training, and a graphical consideration of the impact of training on the difference-score data.

Because the original hypotheses predicted an improvement in performance, the fourth section of the results will detail a series of tests of the original hypotheses with respect to a standard of performance defined by a professional soil scientist and a primary instructor of the soil judges.

Each of the four sections will be structured following the same general pattern. A description of the nature and source of the data values will be given first, followed by consideration of the major results, preceeding from general to specific. All results reported as reliable were significant at $p < .05$.

Tests of Original Hypotheses

Raw-Score Results

The results which follow in this section are based on the raw-scores of the %-of-sand and %-of-clay as given by the soil judges for each of the 12 test soil samples. Recall that the design was a mixed factorial with soils, coarse fragments, excessive moisture, and evaluation sessions as the within-subjects factors and training as the only between-subjects factor. A multivariate analysis of variance using the %-of-sand, and %-of-clay estimates as dependent variable was considered. However, because each dependent measure was of separate interest, the analysis was not performed. However, the statistical tests of training impact suffer from a number of weaknesses (more fully discussed in the next section) and will only be briefly considered for completeness here.

Effect of Soils. An elementary examination of the task suggests that across different soils, the %-of-sand (and %-of-clay) estimates should be different. Specifically, it was predicted that the judges would give different %-of-sand, and %-of clay estimates for each of the three soils. This was tested by examining the main effect due to soil. This was indeed found to be the case for both sand and clay estimates, with $F(2,20)=35.78$, and $F(2,20)=3.96$, respectively. A Newman-Keuls comparison of the three means showed that for the %-of-sand estimates soil 2 differed from both soil 4 and soil 6. Soils 4 and 6 did not differ from each other. For the %-of-clay estimates, only responses to soil 2 and 4 differed significantly.

Influence of Interfering Materials. It was originally hypothesized that the influence of the two interfering materials, excessive moisture and coarse fragment, collapsed over sessions, would emerge as a main effect on the judgments of the %-of-sand and %-of-clay estimates, respectively. For the %-of-clay estimates both excessive moisture and coarse fragments significantly influenced the judgments, $F(1,10)=8.50$, and $F(1,10)=21.13$, respectively. This did not occur in the case of the %-of-sand estimates.

Collapsed across evaluation sessions, neither excessive moisture nor coarse fragments had a reliable effect on the sand estimates.

Despite the lack of a main effect due to either the excessive moisture or the coarse fragments factors, a significant excessive moisture by coarse fragments interaction was found in the %-of-sand estimates, $\underline{F}(1,10)=9.22$, but not in the clay. While this did not conform to the specific a priori predictions, it does show that excessive moisture when combined with coarse fragments did influence the %-of-sand estimates. Thus, for both the %-of-sand and %-of-clay estimates, the interfering materials were found to have an influence on the percentage judgments, either in the form of a main effect (for clay) or an interaction (for sand).

In addition to the predicted influence of each of the interfering factors as a main effect, it was originally believed that excessive moisture would cause errors only in the %-of-clay estimates, and coarse fragments would cause errors in the sand estimates. This did not occur. Finally, a significant soil-by-coarse fragments interaction was found. This result presents no difficulty to the interpretation of the critical effects, and is very reasonable from a soil science prespective. However, it was not of direct concern to this study, and will not be considered further.

<u>Overall Impact of Training Procedures</u>. As discussed above, it was expected, and found, that both excessive moisture and coarse fragments would influence the judgments obtained in the pre-evaluation session. If the lecture training which occurred after the pre-evaluation session affected the influence of the interfering factors, then the main effects of the interfering materials should be reduced or eliminated. Similarly, if the interactive training had an impact on the influence of the interfering factors, the critical main effects should be further reduced in the final-evaluation session. Taken together these predictions indicate that an evaluation session-by- moisture and an evaluation session-by-coarse fragments interaction should both be present. These interactions did not emerge as significant for these interactions in either the either the %-of-sand or the %-of-clay variables. The excessive moisture by coarse fragments interaction, found in the %-of-sand estimates, did not change across the evaluation sessions.

It was further hypothesized that the training program factor should interact with the evaluation-by-interfering material interactions. Neither of the two critical triple interactions were found to be significant. Thus, for the fairly weak raw-score analysis, the training was not found to have a significant impact.

Finally, there was a main effect due to evaluation sessions for the %-of-sand estimates which is rather difficult to interpret in isolation, $\underline{F}(2,20)=5.68$. A general inspection of the data indicates that the overall magnitude of the %-of-sand estimates decreased significantly across the evaluation sessions. No such trend was seen in the %-of-clay estimates. This result will be reconsidered in the context of the performance analyses.

Confidence and Time Measures. For the confidence measures three factors emerged as significant. First, and most critical, confidence increased across the evaluation sessions, $\underline{F}(2,20)=3.57$; with evaluation means of 55.4, 59.6, 62.3 respectively. Excessive moisture emerged as a significant effect on the confidence ratings, $\underline{F}(1,10)=8.40$, with a decrease in confidence seem when the excessive moisture was present. No such effect emerged for coarse fragments. In addition, soils had a significant influence, $\underline{F}(2,20)=3.7$, with the judges feeling least confident for soil 2 and most confident for soil 1.

For the time measure no factors were significant. There was a nonsignificant trend for less time to be taken across the evaluation sessions, with means of 3.1, 3.1, 2.7 for evaluation sessions 1, 2 and 3 respectively.

Difference-Score Analyses

As discussed above, the original hypotheses concerning the impact of training were evaluated based on the dependent measures of the raw %-of-sand and %-of-clay estimates. This approach has the advantage of being straightforward and simple. However, the raw-score analyses have two serious shortcomings. First, general changes (eg., the trend to give smaller sand estimates) which might occur between evaluation sessions that are unrelated to the training increase the error variance and hence decrease the sensitivity of the analysis. Second, when raw-score data are used, errors caused by the interfering materials may cancel out. For example, if the influence of coarse fragments on the %-of-sand judgment is not strictly directional (a likely situation, especially between judges), then an error of +5% made for one judge and an error of -5% by another judge will cancel to be an error of zero in the raw score analysis.

These deficiencies can be corrected using an analysis based on difference-scores calculated from data within each evaluation session. This technique has been used in past research in soil science (Gaeth & Shanteau, 1979).

For these analyses, the difference-score data are obtained by using as a <u>standard</u> the percentage estimate given to the <u>base</u> soil (this is done separately for the sand and clay dependent measures) in each evaluation session. Then, the absolute difference between the standard estimate and each of the other three estimates given to the same soil when it contained interfering material is calculated. The set of difference-scores is determined separately for each judge and soil in each evaluation sesson. The difference-score data were then subjected to an analysis of variance which was otherwise identical to the one described for the raw-score data. In addition to correcting the two deficiencies discussed above, the difference-score transformation has the added advantage that the influence of the interfering material can be tested directly through main effects. This is because smaller means for the %-of-sand and the %-of-clay difference-score estimates reflect less influence of the interfering materials.

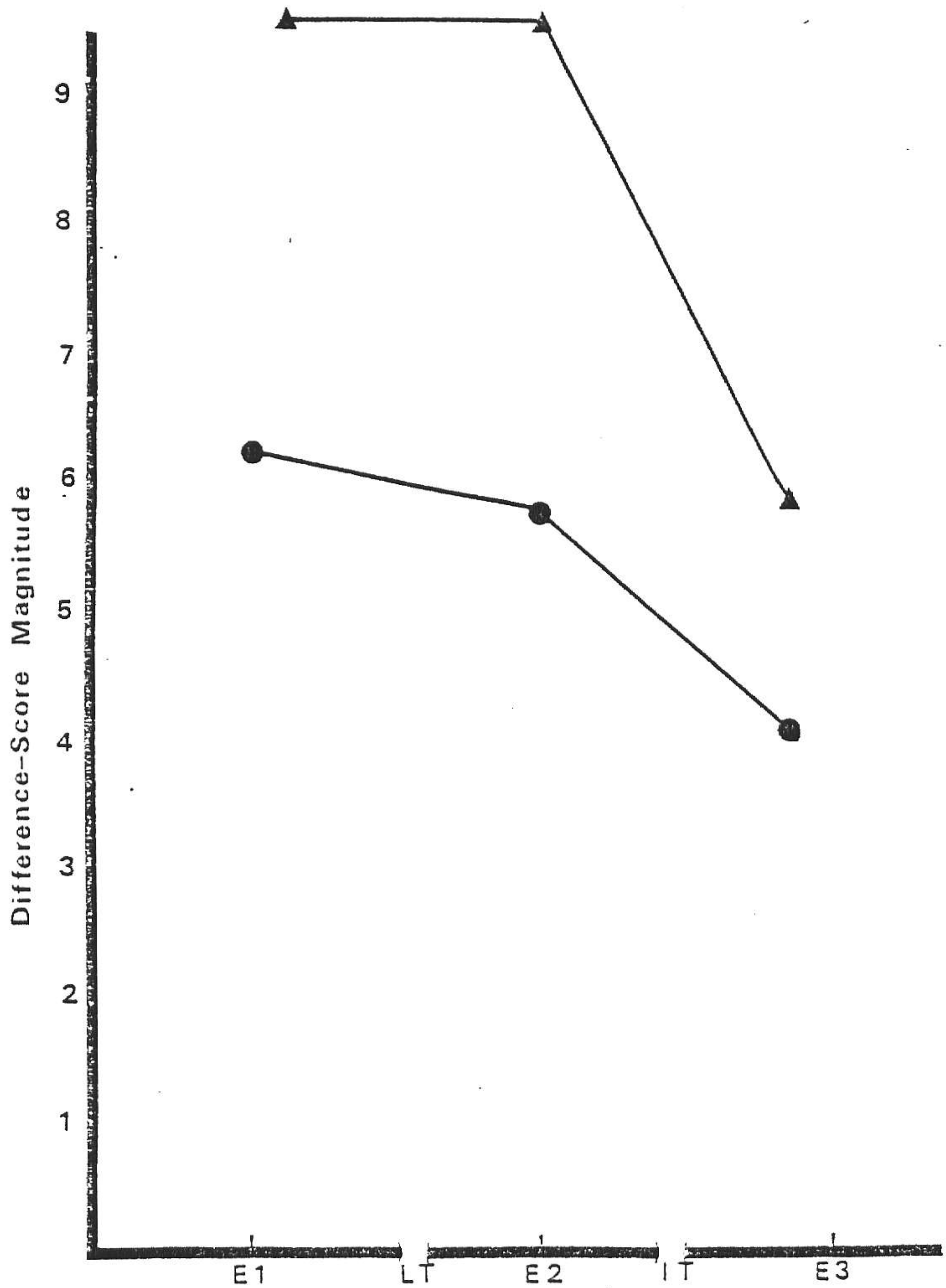
The texture categorization judgments can be analyzed in a similar way, again using difference-scores. The texture-classification judgment given to each of the base soils is used as the standard. Each of the texture judgments (in each evaluation session) for the same soil in the three other conditions is then compared with the base. If the textural classifications agree, then a difference-score is assigned a value of zero. If the textures do not agree, the difference-score is one. These zero-one difference-scores are then used in an analysis of variance in order to provide a reasonably powerful test of the critical training factors. Again, smaller means reflect less influence of the interfering materials on the texture category judgments.

The %-of-sand, %-of-clay, and the texture category difference-score data present one problem, however. Because of the use of one cell in the design as the standard, the mean of this cell is always zero. Thus a large difference between this cell and the other cells usually exists. Hence, only the effects for the interfering materials are meaningful.

<u>Overall Impact of Training on Percentage Estimates</u>. After training the difference between the responses given to a soil, both with and without the presence of the interfering material, should be reduced. The difference-score analyses specifically test for this type of impact of the training. If the training had an impact, then one would predict a significant main effect due to the evaluation session factor. Moreover, the means should become smaller across the evaluation sessions, reflecting a change in the influence of the interfering materials. This was indeed found to be the case, for both the %-of-sand and %-of-clay differences, $\underline{F}(2,20)=4.41$, and $\underline{F}(2,20)=3.92$ respectively.

The significance of the evaluaton session factor shows that the training procedures had an overall impact of significantly <u>altering</u> the influence of the interfering materials. However, the main effect of evaluation session does not guarantee that the impact of the training <u>reduced</u> the influence of the interfering materials. Figure 5 provides a graphical description of the changes in the evaluation session means. Visual inspection of the three evaluation means shows that the values decreased across the evaluation sessions, and almost all of the reduction of influence of the interfering factors came in the third evaluation session, (the means from the first, second and third evaluation sessions were,respectively, 9.42, 9.41, and 5.75 for the %-of sand estimates, and 6.10, 5.67, and 3.99 for the %-of-clay estimates. In both cases a Newman-Keuls

Figure 5 Change in difference-scores across the evaluation sessions. ▲= SAND, ● = CLAY.


Difference-Score Magnitude
9
8
7
6
5
4
3
2
1
E1
LT
E2
IT
E3

test revealed that the only significant differences in the means occurred between the first and third and second and third evaluation sessions. This indicates that either the interactive training alone or the cumulative impact of the interactive and lecture training combined to reduce the influence of the interfering materials.

<u>Overall Impact of Training on the Texture Judgments</u>. The texture category difference-scores can be evaluated in the same way as the percentage estimates. Again, if the training has a significant impact, then a main effect due to evaluation sessions should be present. This was found to be the case, $F(2,20)=7.28$. This indicates that the influence of the interfering factors was reduced by the training. Inspection of the texture data indicates that the major reduction of the influence of the interfering factors occurred after the second evaluation session (after the interactive training). This trend can be seen in the mean difference-scores of .56, .54, and .37 for evaluation sessions one, two and three, respectively. A Newman-Keuls analysis of the means showed that the only significant differences between the means were for the last and first and the last and second, again pointing to the fact that the majority of the training impact was seen in the final-evaluation session, which followed the interactive training. This seems to indicate the presence of substantial training impact after the interactive training.

Individual Judge Analyses

Previous research on training has shown that there are often great differences with respect to the impact of training on individual judges (Slovic, Fischhoff, & Lichtenstein, 1977). This study was intentionally designed so that statistical tests could be made on the raw-scores for the impact of training at the level of the individual judge. In general, the individual judge results for the raw-score data were consistent with the group results.

Since training had a significant impact, and reduced the influence of the interfering materials for the difference-scores, it is of interest to consider individual judges with respect to the difference-score data. Unfortunately, no appropriate error terms are available to test these results. Therefore, this section will involve graphical consideration of the individual judge difference-score data. These results should be thought of simply as characteristic examples.

Graphical Consideration of the Difference-Score Results

Separate consideration of the impact of the lecture and the interactive training can be done by looking at the pattern of evaluation means. For these reasons, the difference-score results will be considered for individual judges. This is best done graphically, using the following method for plotting. Percentage-of-sand and %-of-clay mean responses are plotted for each of the evaluation sessions. Lines are plotted separately for the presence and absence of the interfering materials. In general, convergence of the two lines at an the evaluation session indicates that the influence of the particular interfering material was reduced by the preceding training. The individual judge data tended to follow one of four patterns across the evaluation sessions: Either impact of the lecture training, impact of the interactive training, impact of both trainings, or no or detrimental impact of the trainings. The pattern of the means across the evaluation sessions were visually inspected for each of the twelve judges, on the two dependent measures (sand and clay), and placed in one of the training impact categories. A count for each of the four training-impact-categories is given in Table 3. In addition, representative examples for each of the patterns are graphed in Figures 6-8. It should be remembered that these graphs do not represent statistical tests, but rather are used as characteristic examples of the individual judge results.

Impact of Lecture Training. A few of the judges seemed to indicate that the lecture training reduced the influence of the interfering materials. If the lecture training had such an impact, one would expect that the two lines would converge toward each other, indicating that the presence of the interfering factor did not influence their judgment. For example, see Figure 6 which shows two of the individual judges who seem to show a considerable impact of the first training session as reflected in the second evaluation session.

Table 3

Number of Judges Displaying Different Forms of Training Impact

	Training Impact			
Estimate	Lecture	Interactive	Both	Neither
Sand	2	5	2	3
Clay	2	3	2	5

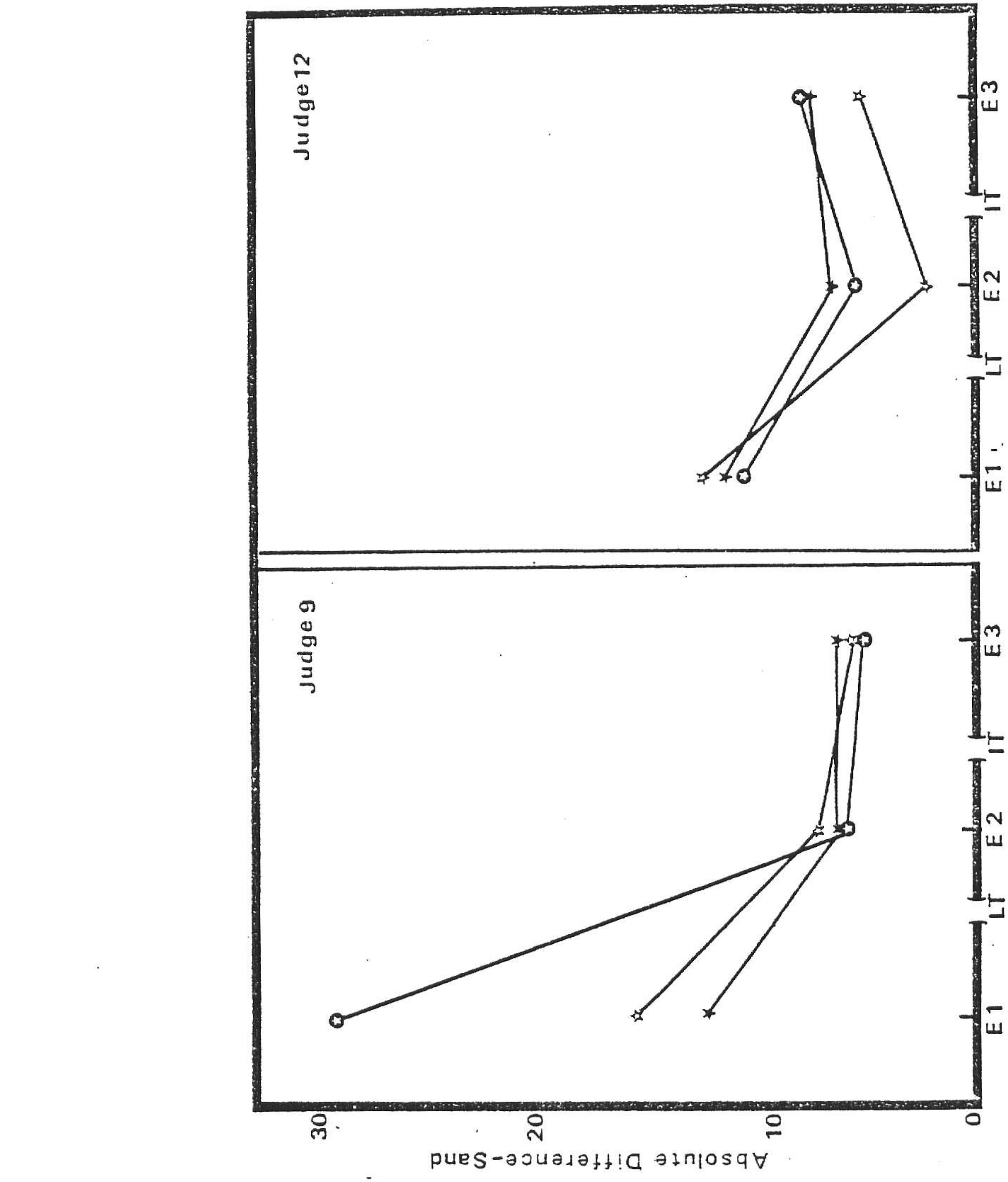
Note: Each Judge was assigned one of the four possibilities based on visual inspection of the graphical results.

Figure 6 Two soil judges showing impact of lecture training.

★ = excessive moisture not present, coarse fragments present

☆ = excessive moisture present, coarse fragments not present

✪ = excessive moisture present, coarse fragments present


Judge 9
Judge 12
E1
LT
E2
IT
E3
30
20
10
0
Absolute Difference-Sand

Impact of Interactive Training. More judges showed the indication of an impact due to the interactive training procedure than to the lecture training. Two representative examples are graphed in Figure 7. The impact of the interactive training can be seen in the convergence of the lines only in the third evaluation session (which follows the interactive training).

Impact of both Lecture and Interactive Training. Some individual judges show the impact of both training procedures, as reflected in a gradual reduction in the influence of the interfering factors across both the interactive and the lecture trainings, and hence a gradual convergence of the lines on the graphs. A representative example is given in Figure 8.

Miscellaneous Individual Judge Results. In some cases the training seemed to have no impact or perhaps even a detrimental impact on the influence of the interfering materials. A representative example can be found in Figure 8.

Tests of Convergence to a Standard

Although the analyses discussed above permit determination of the impact of training on the influence of the interfering factor, it is not possible to know whether reduced influence of the interfering factor necessarily leads to improved accuracy. Two different, but appropriate standards of accuracy are available. Because laboratory

Figure 7 Two judges showing impact of interactive training.

★ = excessive moisture not present, coarse fragments present

☆ = excessive moisture present, coarse fragments not present

✪ = excessive moisture present, coarse fragments present

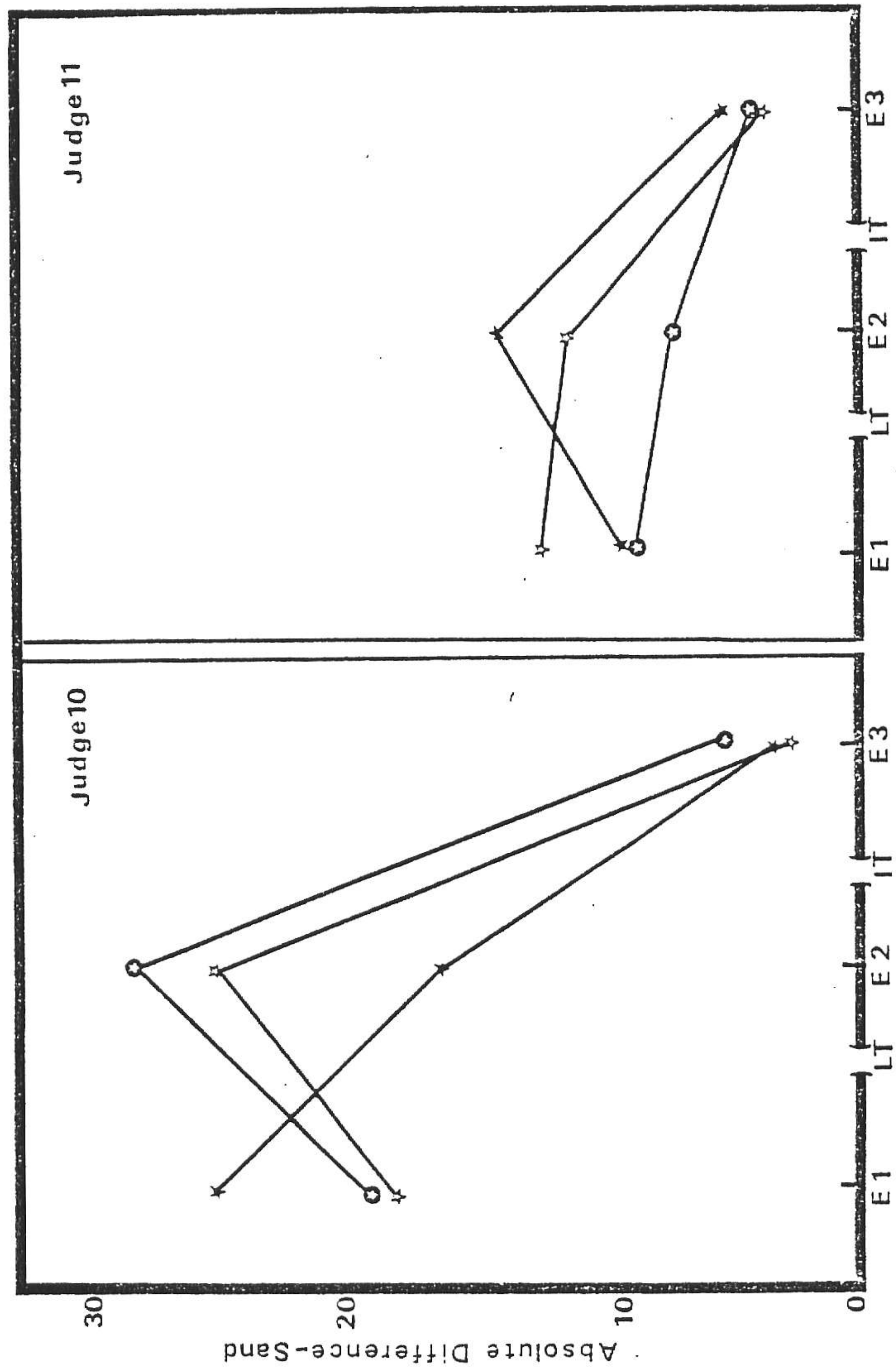
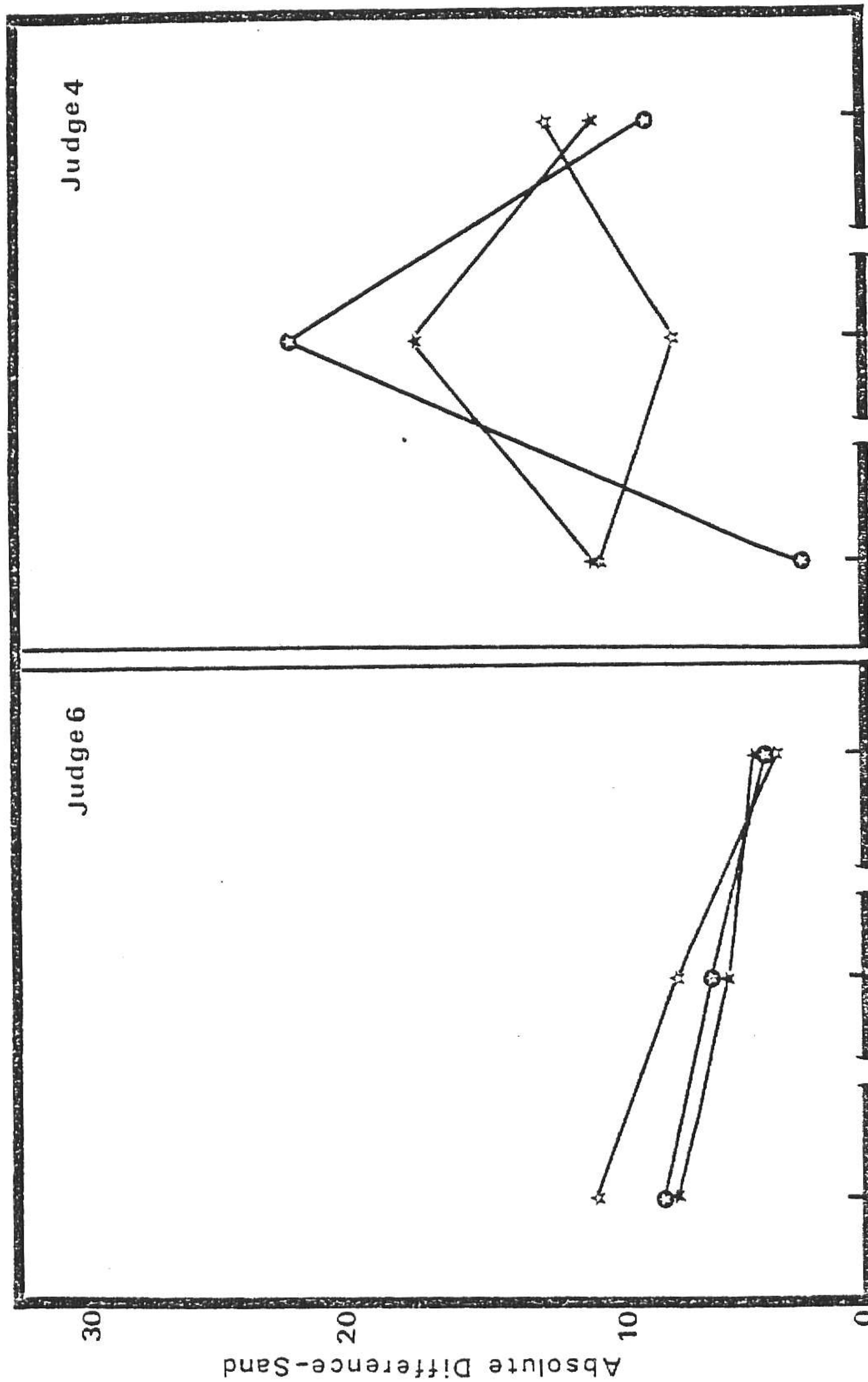

Judge10
Judge11
E1
LT
E2
IT
E3
0
10
20
30
Absolute Difference-Sand

Figure 8 Two judges, one showing impact of both trainings, and one showing impact of neither.

★ = excessive moisture not present, coarse fragments present

☆ = excessive moisture present, coarse fragments not present

✪ = excessive moisture present, coarse fragments present


Judge 4
Judge 6
E1
LT
E2
IT
E3
0
10
20
30
Absolute Difference-Sand

analyses are expensive most soil judges are trained to conform to the standard defined by a soil scientist, usually the soil judge team coach and/or instructor. This was the case for the soil judges used in this study. They had little exposure to laboratory standards. Thus, the first accuracy standard considered will be defined as _performance_, and consist of an evaluation of the soils as made by the instructor of the soil morphology class taken by the judges. The second accuracy standard, _validity_, is defined by the laboratory analyses. The original response data was compared to these two criterion, and discussed separately below.

Performance

This section represents an attempt to determine if the estimates made by the soil judges converged to the standards defined by an expert soil scientist, who was also the instructor of all of the soil judges. Throughout their training the basic goal for these soil judges has been to achieve agreement with an experienced soil instructor. The judgments of this expert were used here to provide a criteria to determine the absolute magnitude of the errors. Using this performance data, a series of analyses parallel to those discussed above, for the difference-scores, were carried out.

Effect of Soils. No main effect was expected on the magnitude of the errors made across the various soils. That is, the errors were anticipated to be roughly equivalent across the three soils. This was not the case. Soils had a significant impact on the magnitude of the errors made for the %-of-sand estimates, $F(2,20) = 3.74$, but not the %-of-clay estimates.

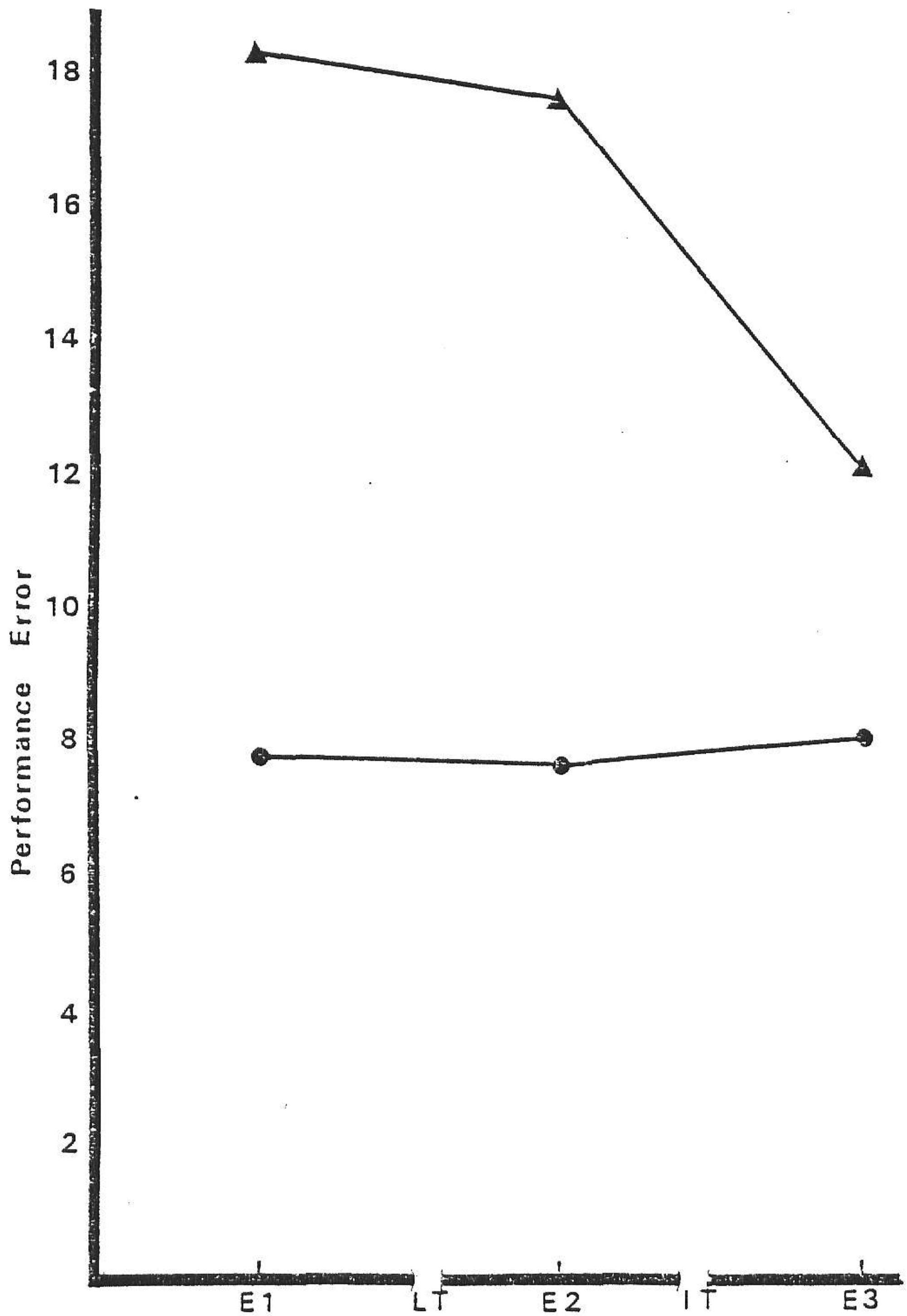
Influence of Interfering Materials. Both excessive moisture and coarse fragments were expected to have a main effect of the magnitude on the errors made by the judges. The influence of the two interfering factors differed between the sand and clay estimates. For the %-of-clay estimates excessive moisture had a reliable influence, $F(1,10) = 14.33$, while neither coarse fragments nor the interaction of the interfering factors was significant. In contrast, for sand, neither excessive moisture nor coarse fragments had a significant influence on the size of the errors, although the interaction between coarse fragments and excessive moisture was significant, $F(1,10) = 8.72$.

Again these results were not as predicted. However, just as in the case of the raw-score results, these data indicate that the interfering materials had a significant influence on the performance. Apparently, for the %-of-clay estimates, excessive moisture alone has an influence on performance: while for the %-of-sand estimates excessive moisture and coarse fragments, when present together, have an influence.

Overall Impact of Training on Percentage Estimates. Because these data reflect absolute agreement with a standard, the main effect of evaluation session may be thought of as reflecting overall changes in performance. Evaluation session was a significant factor in the validity data for sand $\underline{F}(2,20)=4.85$. Changes in the means for the %-of-sand difference sccres are plotted in Figure 9. Inspection of the means of the three evaluation sessions, 18.32, 17.73, and 12.19 respectively, clearly shows that the main effect was due to reduction in the absclute difference. A Newman-Keuls comparison of the three means confirms that only the mean absolute errcrs for evaluaticn sessions 1 and 2 differed from evaluation 3. As can be seen, no main effect was found for the %-of-clay estimates. was not found.

Overall Impact of Training on Texture Judgments. The training had no significant impact cn the judges' performance with respect to the textural classification data. However, this analysis suffered frcm two problems. First, overall the textures were in agreement with the standard only 21% of the time. This may have created a floor effect in the data. Second, when the textural classification judgments are transformed into a 0-1 score, a great deal of informaticn is lost. The mean number of error can be seen to decrease primarily across the second and third evaluation sessicns, with values, .81 and .74 respectively. The trend seen in these means for the

Figure 9 Changes in performance over the three evaluation sessions ▲ = SAND, ● = CLAY.



18
16
14
12
10
8
6
4
2
Performance Error
E1
LT
E2
IT
E3

texture-difference scores across the evaluation sessions was consistent with findings discussed earlier.

Comparison of The Two Training Programs. Again, the two training programs can be compared with respect to their impact on the two interfering materials. As discussed above, a training-by-excessive moisture-by-evaluation session interaction and a training-by-coarse fragments-by-evaluation session interaction was predicted. Neither of these two triple interactions were significant, for either the %-of-sand or %-of-clay estimates.

Validity

It was originally planned that a parallel set of analyses using results from the laboratory analysis as a standard would be presented and discussed. This turned out to be generally unnecessary because of the small difference between the set of standards given by the soil scientist and the laboratory results (See table 2 for a comparison). The major disgreement occurred for soil 6 with respect to the percentage of sand (and hence silt). According to the laboratory analyses this soil contained 30% very fine sand, which was apparently evaluated as silt by the soil scientist (O. W. Bidwell, Personal Communication). Because of the overall high degree of similarity between the performance data and the validity data, only the results which pertain to training impact will be reported here.

Overall Impact of Training of the Percentage Estimates. The overall impact of training with respect to the laboratory standard was not significant ($F[2,20]=1.14$). That is, the means across the evaluation sessions did not become significantly smaller. However, the interaction between evaluation sessions and soils was significant, $F(4,40)= 5.20$. This interaction is graphed in Figure 10, and visual inspection clearly shows that the reponses to soils 2 and 4 converge to the validity standard, while they diverge from the standard for soil 6. It should be remembered that the criteria the soil judges are working toward is defined by the soil scientist, and only indirectly by the laboratory.

DISCUSSION

The first section of the discussion contains a brief summary of the results. The second section is a brief description of four limitations which should be imposed on any interpretation of these results. The third section of the discussion briefly outlines psychological implications, while the fourth section is a consideration of some general suggestions to decisions makers. This is followed by a series of suggestions specifically aimed at the task of soil judgment. The sixth section presents general comments and outlines a number of possibilities for future research. The last section of the discussion summarizes the major conclusions which may be drawn from this study.

Summary of Results

Four major results emerged from this study. First, the raw-score analyses showed that the judges responded differently to the different soils, thus indicating that the judges were able to discriminate among the soils.

Second, based on both the raw-score analyses and the standard-based analyses, interfering materials influenced both the percentage estimates, and the texture category judgments.

Third, the training had the impact of significantly reducing the influence of the interfering materials. This effect appeared to concentrated after the interactive training.

And fourth, reduction of the influence of the interfering materials was accompanied by a convergence to a standard of performance defined by an experienced soil scientist.

Limitations

At least three limitations should be considered before any conclusions can be discussed. First, although experienced, the participating soil judges certainly do not represent the highest possible level of soil science expertise. It is not intended that the results discussed here should necessarily be generalized to those experts.

Second, the sample size of the soil judges was quite small. In addition, the sample was very homogeneous because all twelve judges had been trained by the same soil judge; they were all enrolled in the same soil morphology class, and they were all at essentially the same level of experience. Thus generality, even to other student soil judges may be limited.

Third, the soil samples were purposely not chosen to reflect a diverse collection, and in fact represented a very limited range of textural classes. This may restrict generality to other soils. In addition, the soils were presented to the soil judges indoors, under artificial light (soil textural classification is normally done out-of-doors), and without the usual contextual cues (eg., surrounding landscape, other soil horizons, etc.). Although these limits on the soils may well have influenced the overall performance results, it seems quite unlikely that the influence of the interfering materials and the impact of the training would be affected.

Psychological Implications

It seems reasonable to begin this section by repeating the three important questions addressed in this study. First, is there any evidence that the irrelevant materials, which have thus far been called *interfering*, really deserve the title? Second, what impact did the training, which was designed to reduce the influence of the interfering materials, have? Finally, did the training result in improved performance?

Were the Irrelevant Materials Really Interfering Materials? One of the most critical aspects of this investigation concerned the extension of previous laboratory findings indicating that irrelevant information may adversely influence human judgment in "real world" judgment situations. The answer to this question, based on the results from this study, is yes. However, the relationship it bears to other research concerned with irrelevant information is complicated by the explicitness of the design. In most earlier research dealing with nondiagnosticity, the irrelevant dimension was not specifically controlled and tested (see Troutman & Shanteau, 1977, for an exception). For example, in studies on the use of simplifying heuristics conducted by Kahneman and Tversky (1973), they showed only that the irrelevant information changed judgments qualitatively. They did not investigate the magnitude or direction of the influence. Because the irrelevant factor was of direct interest in this study, tests for its influence were more explicit than in earlier research. Despite the result that the interfering factors of excessive moisture and coarse fragments did not always emerge as main effects, in all cases the irrelevant factors had some definite and significant influence. Specificaly, for the %-of-sand estimates the two interfering materials led to a significant interaction. While for the %-of-clay estimates the irrelevant materials emerged as significant main effects.

Thus, to summarize, the a priori hypothesis that the irrelevant materials would influence judgments was directly confirmed for the clay estimates and indirectly for the sand estimates. Thus, it is possible to conclude that the irrelevant materials did influence the jugdment of the soil judges and may appropriately be called interfering materials.

Because of the somewhat inconsistent influence of the interfering materials (i.e., both main effects and interactions), since some concern may be expressed over the lack of agreement with previous studies of nondiagnosticity. It should be noted that if this study had considered only the qualitative influences of irrelevant materials, the main effects and interactions would have been indistinguishable (Kahneman & Tversky, in press). In this case the results would have been completely consistent with earlier research.

It has been shown that the interfering materials thus change the estimates made by the judges; but perhaps these materials actually make them more, rather than less accurate. Consideration of the performance data indicates that this is not the case. That is, compared to the judgments (both percentage estimates and the texture category) made of the base soil, the soil judgments diverge from the performance standard when coarse fragments or excessive moisture (or both) are present. This can be seen in Figure 10, which shows the relationship between performance and the presence of the interfering factors.

Inspection of Figure 10 may also provide some explanation as to why the two interfering factors produced a significant interaction for sand percentage judgments. The performance did not become worse when the two factors were present together; rather, the performance improved. Two explanations for these interesting results seem plausible. First, it may be that when the two interfering factors are present together, the soil judge is simply "extra careful." When the judges were questioned after the completion of the experiment a number indicated that when the soil contained both of the interfering materials, they were especially difficult samples to texture-classify. A second possibility may be that a ceiling effect occurred, and when both factors were present the judge thought "it feels like x% of sand, but that is just too much, so I will cut down on my estimate." These explanations are worthy of future

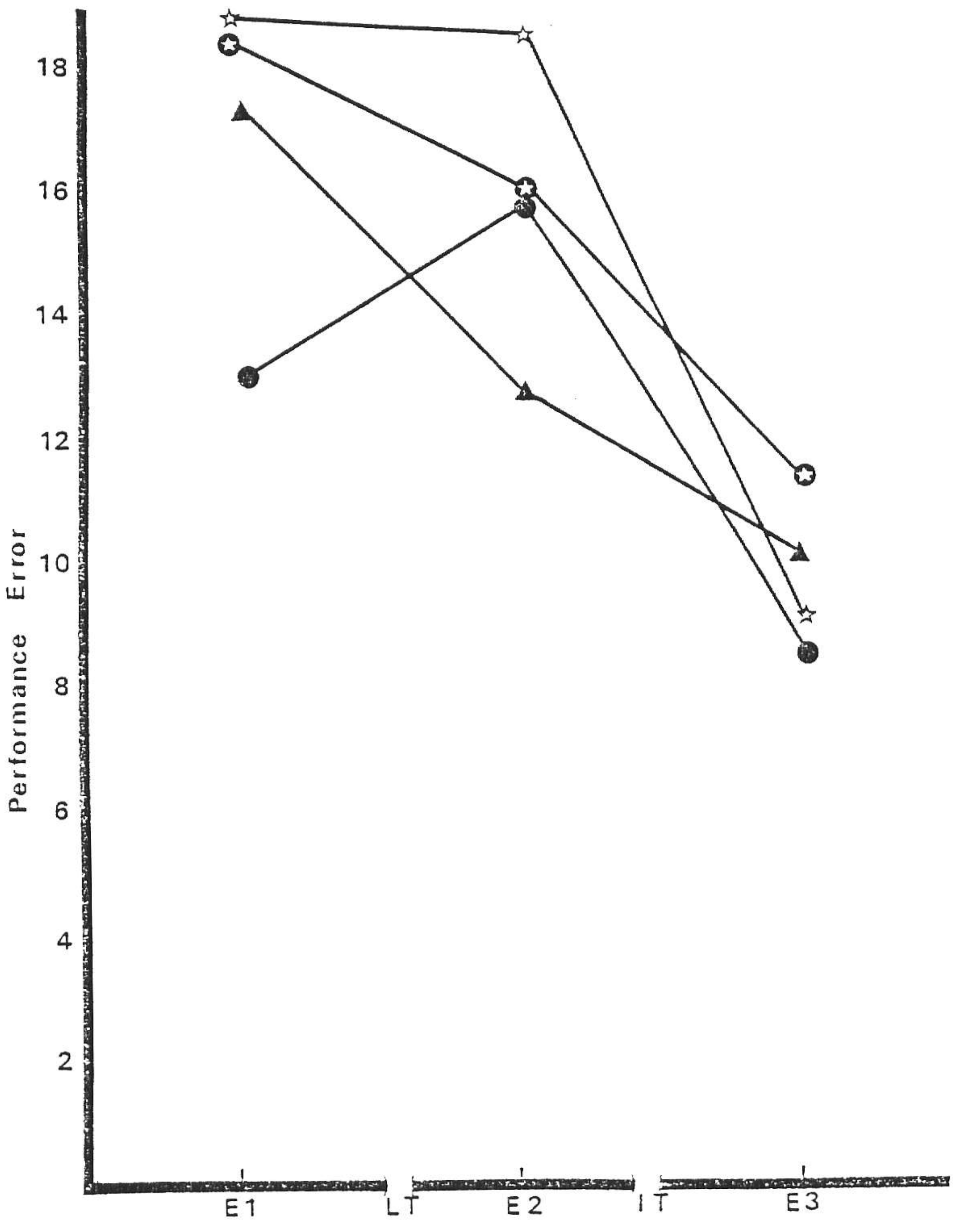
Figure 10 Influence of excessive moisture and coarse fragments on performance across evaluation sessions.

● = excessive moisture not present, coarse fragments not present

✪ = excessive moisture present, coarse fragments not present

☆ = excessive moisture not present, coarse fragments present

▲ = excessive moisture present, coarse fragments present


18
16
14
12
10
8
6
4
2
Performance Error
E1
LT
E2
IT
E3

research.

Did the Training Have any Impact? The overall impact of the training programs was evaluated in a number of ways. The raw-score analyses did not provide evidence of training impact. This may have been due to the procedure used to test for the impact of training in the case of the raw-scores. The power of the tests used to evaluate the training impact was limited for three reasons. First, the critical prediction used to measure the impact of training was a three-way interaction having only 2 degrees of freedom in the numerator. This is a very limited test. Second, when the analyses is conducted on raw-scores, the direction of an error becomes a factor and errors may cancel to zero, as discussed in the results section. Third, changes which occur across the evaluation sessions, which are unrelated to training, tend to inflate the error terms and again reduce power.

These three problems were remeded through the use of difference score analyses based on absolute differences. All three of the difference-score score analyses (the %-of-sand, %-of-clay and textural classifications) point to a significant impact of training. These analyses, based on within-session difference scores, reflected what might most appropriately be called a change in the influence of the interfering factors. In these instances, the influence of the two interfering factors is reflected in their tendency to produce different responses to the same soil when it contains interfering factors. Training resulted in a decrease in this tendency across the three evaluation sessions. Further, as shown by all of the Newman-Keuls comparison of the means discussed above, this reduction is seen primarily after the interactive training, in the last evaluation session. Parallel results are seen in the %-of-sand, %-of-clay, and the texture-classification judgments, strongly pointing to impact of training after the interactive training. Although these dependent measures are partly related, the consistentcy of these results considerably increases confidence in their reliability. Thus, based on these results, it is reasonable to conclude that the training did reduce the adverse influence of the interfering materials.

Consistent with the result that training reduced the infnuence of the interfering materials was the significant increase in confidence which occurred across the evauation sessions. This supports the notion that as the influence of the interfering materials is decreased, the judge's confidence appropriately increases. This relationship is only correlational, however.

Did Training Improve Performance? Together these results allow for the conclusion that training did reduce the impact of the interfering materials. But did the training necessarily improve the performance? Consideration of the accuracy analyses confirms this. There was a significant increase in the agreement with the standard for both the percentage estimates across the evaluation sessions. And, although nonsignificant, the same trend was seen in the texture-category judgments. In both cases, this occurred primarily after the interactive training as reflected in the third evaluation session.

In summary, the training procedures did improve performance, and the improvement occurred primarily after the interactive training. The question as to why this improvement was seen only after the interactive training is of great interest, both theoretically and practically. Unfortunately, as was known before this study was run, it is not possible to separate qualitatively four quite different explanations.

First, the procedure used in the interactive training even isolated from the lecture training may be the cause of the reduced influence of the interfering materials and the subsequent improved performance. This would be consistent with the notion that interactive training is more effective than simple lecture warnings (see, for example, Hammond, 1971; Slovic, Fischhoff, Lichtenstein, 1977).

A second possibility, which would explain the impact of training seen only in the last evaluation session, depends simply on the repetition of the lecture training which occurred as part of the interactive training. This would be consistent with the results of Bruno and Harris's (Bruno & Harris, in press) research in which repeated lecture instruction alone was found to reduce susceptibility to implications.

Third, the correct reason for the training impact may be a combination of the two explanations given above. It could be that neither lecture training nor interactive training alone would be successful, but together they interact to achieve impact.

Fourth it is possible that the change in accuracy is unrelated to the training sessions entirely, and simply due to practice which occurred during the evaluation sessions. This latter explanation does not seem likely, however. In addition, when the changes in validity (based on the laboratory analyses) of the filler soils are considered, no significant reduction in errors is seen, $F(2,22) = .43$ for sand, and $F(2,22) = 1.82$ for clay. Means for the three evaluation sessions were 11.63, 11.73, and 12.81 for sand, and 7.94, 9.96, and 7.67 for clay, respectively. This seems to indicate that simple practice effects are not responsible for the training impact.

Did the Different Training Programs have Differing Impact? A fundamental a priori prediction stated that the two training programs (ie., for coarse fragments and excessive moisture) should separately have impact on the influence of the two interfering factors. The accuracy of this prediction was resoundingly disproved. It is encouraging that the training procedures have generality across similar interfering materials. There was evidence that the different training programs each decreased the influence of coarse fragments and excessive moisture. It is not known why the training programs had such general impact. The reason for this unusual finding must await future research, however, is possible to speculate on a possible reason for the lack of specific training impact. There is no question that all of the judges were aware of both of the two interfering materials when present in the soil. It seems quite likely that when a particular judge was given training for one of the interfering factors, say coarse fragments, that he/she was immediately aware of the presence of the other factor as also being interfering. Then, because the two training programs were so similar in form, and to some extent in content, the seven lecture suggestions and the interactive training given for the coarse fragments may have carried over quite directly to the excessive moisture factor. Conversations with the soil judges after the experiment confirmed that in all cases they were aware that the factor that they were not trained on was also interfering. However, in most cases they had not explicitly attempted to use the suggestions to help them with both interfering factors. This explanation could be tested by including other interfering factors which are less obvious (such as organic material) and/or by making the training suggestions more specific to the interfering factor and thereby reducing the chances of carry-over.

Implications for Applied Judgments

The results of this experiment can be utilized to apply to "real world" decision making. First, the judge should recognize the possibility that something which he/she knows should be ignored as nondiagnostic may still be used as part of the decision process. It is very easy for the decision maker to say "I would not do that"; however, evidence suggests that it may be otherwise. As demonstrated in this study, an experiment can be contructed which systematically explores the use of nondiagnostic information. This type of demonstration may help decision makers discover that they do use the nondiagnoastic information.

Second, if in a particular judgment task, an irrelevant dimension is found to have an adverse influence, it appears that the influence may be reduced through training consisting of lecture suggestions and/or interactive practice. Because the training procedures used in this study do not rely on outcome feedback, their implementation should be fairly easy in other areas. For example, it is suggested that the judge be confronted with a series of stimuli which both do and do not contain interfering material. Then using some type of blind presentation system, the stimuli are judged. From the responses to the stimuli, the judge can then determine if the interfering material is adversely influencing his/her judgment. If this is indeed the case, then the stimuli can be used for interactive practice in a fashion similar to the procedure used in this experiment.

Implications for Soil Judgment

The results of this study can be directly applied to problems faced by soil judges. First, because both coarse fragments and excessive moisture were shown to interfer with the accuracy of the texture-classification judgments made by the soil judges special care should be taken when such soils are judged. This may include the precautions and suggestions develcped in this study, in addtition to techniques used by individual judges. As discussed previously, it would be possible for the soil judge to create a collection of soil samples similar to the one used here and perform their own individual training.

A second implication for soil judgment pertains to a more general problem. Judges from a wide variety of area seem to have a common problem of overconfidence in their judgments and judgmental abilities. An important aspect of this study is that it was demonstrated that materials which are irrelevant to judgments may interfer with the accuracy of a judgment, whether the judge is aware of it or not. In the specific instance of soil judgment, it appears to be important that the judges are aware not only that some materials are irrelevant, but also that they, themselves, may be adversely influenced by the irrelevant information.

Lastly, the difference between the performance results and the validity results, almost exclusively caused by soil 6, is striking. This demonstrates the importance of the choice of a standard. Periodic retraining, based on a laboratory standard, is imperative if the soil judge is to remain accuracte. It is especially important in preparation for soil judging contests that the judges are trained using the same standard as will be used in the competition.

General Comments and Suggestions for Future Research. Based on this and other findings it has been shown that information which is irrelevant to a particular judgment is not always ignored, but rather may interfer with the judgment. Specifically, in this study it was found that the interfering, but irrelevant, materials of coarse fragments and excessive moisture may adversely influence soil texture-judgments. Therefore, an important direction for future research would be to investigate how widespread the problem of irrelevant information is for experienced decision makers. One contribution of this study is that it provides a successful way in which to measure the influence of interfering factors through the use of a systematic design. This was accomplished by including the interfering materials as a within-subjects factor, allowing the researcher to qualitatively test for the influence of the irrelevant dimensions. This approach should be relatively easy to apply to other content areas.

Assuming the presence of adverse influence due to the interfering factors, the question becomes what can be done about it? Despite the general lack of a specific training impact found in many judgment studies, the approach used here succeeded in both reducing the influence of the interfering materials and in improving the performance of the judges, possibly through reduction of the influence of interfering factors. Two general properties of the training procedures used here are worthy of mention and future research.

First, the influence of the interfering factor was reduced by simply making the judge aware of the interfering factor. This may seem counterintuitive, because if the decision maker is to ignore irrelevant information then one might imagine that the training should stress what the judge is supposed to pay attention to. In contrast, the training employed here depended on not only making the judge aware of the irrelevant material, but also giving some specific directions as to how to eliminate any adverse influence of the materials by adjusting cognitively and perceptually. This is not unlike the situation faced by explanations of selective attention; how can an observer choose not to attend to a stimuli without first attending to it at some level? Perhaps the soil judges in this experiment learned how not to be influenced by the interfering factors by first attending to them and then learning how "not to attend." What is needed is a move toward the development of a theory of nondiagnostic interference. Because of the similarities between selective attention and the influence of nondiagnostic information use of the theories and research on selective attention may be a valuable aid in the development of a theory of nondiagnosticity.

A second property of the training used in this study is its unique lack of dependence on feedback. No validity feedback was given, and only a limited amount of "self-reliability" feedback was given in the interactive training session. Both of the training procedures were dependent upon only lecture suggestions and having the judge practice on identical soils both with and without the critical irrelevant factor. Thus, this training may be viewed as a self-help program in which a criterion is not necessarily ignored. This independence from a criterion increases the possibility of effective generality to other applications and decison making problems.

Third, the problem of separating the impact of the lecture training from the interactive training is an important one for future research. A study designed to tease apart the effects of the two training procedures could easily be devised. The most obvious approach would invlove a three-group extension of this study, in which lecture, interactive, and no training is given independently to each group. This design would allow separation of simple practice effects, the lecture training procedure, and the interactive training procedure.

Another worthwhile direction for future research would involve determination of the duration of the training impact. First, the question of how long the training impact remains could easily be answered using a series of followup studies. If some impact of the training is lost (quite likely), then a related question would be what is the most effective type of retraining? Various forms of retraining might be attempted and tested. For example, if is shown that interactive training has the greatest impact, but that the impact is lost over time, it may be possible that retraining can be acomplished using only lecture training. Or, because outcome feedback is not a necessary part of the training used here it may be possible to teach the judges how to use lecture or interactive training techniques to retrain themselves.

Conclusions

Three important conclusions may be drawn from this study. First, "interfering materials" do, in fact, interfer with soil judges' assessment of percentages and textures. Given that this is the case, soil judges should be made aware of this and also be taught techniques to avoid the adverse influence on the interfering materials. In addition, this result provides further evidence that nondiagnostic information adversely influences the judgment of even experienced decision makers.

Second, the interference of the irrelevant materials can be reduced through lecture and interactive training procedures. It has been repeatedly shown in past research that strictly verbal (i.e., lecture) training is not adequate as a training technique (for example, see, Bruno, in press; Hammond, 1972; Slovic, Lichtenstein, & Fischhoff, 1977). These findings were repeated here. However, encouragingly, the combination of lecture and interactive training used in this study provided significant training impact. The demonstration of a successful training procedure outside the laboratory is especially interesting.

Third, reduction of the interference of the irrelevant materials can be accomplished through training. And, at least in this case, reduction of the adverse influence of the interfering materials resulted in improved accuracy. There is an increasingly large body of literature demonstrating the adverse influence of irrelevant information (see introduction). Reduction of this influence, in itself, would be interesting, but would not necessarily yield a worthwhile tool for training decision makers to make better judgments. However, the evidence provided here, showing that reduction of the interference caused by the irrelevant information leads to increased accuracy, may provide a beginning for improved procedures to train experienced decision makers.

Reference Notes

1. Ettenson, R. An empirical investigation of the characteristics and development of auditing expertise. Submitted as a Master's Thesis Proposal, Kansas State University, Manhattan Kansas, June 1980.

References

Anderson, N. H. Information integration theory: A brief survey. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), Contemporary developments in mathematical psychology, (Vol. 2). San Francisco: W. H. Freeman, 1974.

Brehmer, B. Effect of practice on utilization of nonlinear rules in inference tasks. Umea Psychology Reports, No. 113, 1977.

Brehmer, B. & Svensson, C. Learning to use functional rules in inference tasks. Scandinavian Journal of Psychology, 1976, 17, 313-319.

Bruno, K. J. The discrimination of assertions and implications: A training procedure for adults and adolesents. Journal of Educational Psychology, in press.

Bruno, K. J., & Harris, R. J. The effect of repetition on the discrimination of asserted and implied claims in advertising. Applied Psycholinguistics, in press.

Brunswick, E. Perception and the representative design of psychological experiments. (2nd ed.). Berkeley and Los Angeles: University of California Press, 1956.

Castellan, N. J. Jr. Multiple-cue probability learning with irrelevant cues. Organizational Behavior and Human Performance, 1973, 9, 16-29.

Clarke, G. R. The Study of the Soil in the Field. Oxford: Clarendon Press, 1936.

Dixon, P., & Just, M. Normalization of irrelevant dimensions in stimulus comparisons. Journal of Experimenal Psychology: Human Perception and Performance. 1978, 4(1), 36-46.

Edwards, W. Conservatism in human information processing. In B. Kleinmuntz (Ed.). Formal representation of human judgment. New York: Wiley, 1968.

Einhorn, H. J., & Hogarth, R. Unit weighting schemes for decision making. Organizational Behavior and Human Performance, 1975, 13, 171-192.

Einhorn, H. J., & Hogarth, R. Confidence in judgment: Persistence of the illusion of validity. Pschological Review, 1978, 85, 395-416.

Fischer, G. W., Edwards, W., & Kelly, C. W., III Decision theoretic aids for inference, evaluation, and decision making: A review of research and experience. (Incorporated Technical Report, 78-1-30). McLean, Va: Decisions and Designs, February, 1978.

Fischhoff, B. Perceived informativeness of facts. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 349-358.

Foss, J. E., Wright, W. R., & Coles, R. H. Testing the accuracy of field textures. Soil Science Society of America Proceedings, 1975, 59, 562-571.

Fromkin, H. L., Goldstein J. H., & Brock, T. C. The role of irrelevant derogation in vicarious agression catharsis: A field experiment. (Krannert Graduate School, Paper No. 447, Purdue University). Laffette, Ind.: Purdue University Press, 1974.

Gaeth, G. J., & Shanteau, J. Analysis of the soil survey laboratory data of some soils of Kansas: Validity of the texture classifications. Paperr presented at SCSS (Agronomy Scoiety) Meeting, Fort Collins, Colorado, 1979. (Also, Kansas State University Psychology Series Technical Report, 79-11).

Goodman, B., Fischhoff, B., Lichtenstein, S., & Slovic, P. The training of decision makers. (ARI TR-78-B3). Alexandra, Va: U.S. Army Research Institute for the Behavioral Sciences, March, 1978.

Hammond, K. R. Computer graphics as an aid to learning. Science, 1971, 172, 903-908.

Hammond K. R., Rohrbaugh, J., Mumpower, J., & Adelman, L. Social judgment theory: Applications in policy formation. In M. F. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes in applied settings. Academic press: New York, 1977.

Harris, R. J. The comprehension of pragmatic implications in advertising. Journal of Applied Psychology, 1977, 62, 603-608.

Hawando, T. A characterization of Crete, Smolan, and Geary edons of central Kansas. Unpublished Doctoral Dissertation, Kansas State University, Manhattan, Kansas, 1978.

Hodgson, J. M., Hollis, J. M., Jones, J. A., & Palmer, C. A comparison of field estimates and laboratory analyses of the silt and clay contents of some west midland soils. Journal of Soil Science, 1976, 27(3), 411-419.

Hoyer, W. J., Rebok, G. W., & Sved. S. M. Effects of varying irrelevant information on adult age differences in problem. Journal of Psychology, 1979, 34(4), 553-560.

Humphreys, P. Decision aids: Aiding decisions. In L. Sjoberg, T. Tyszka, & J. A. Wise (Eds.), Decision analyses and decision processes, (Vol. 1). Lund: Doxa, 1979.

Kahneman, D., & Tversky, A. On the pschology of prediction. Psychological Review, 1973, 80, 237-251.

Kahneman , D., & Tversky, A. Intuitive prediction: Biases and corrective procedures. Management Science, in press.

Krischer, J. P. The mathematics of cleft lip and palate treatment evaluation: Measuring the desirability of treatment outcomes. The Cleft Palate Journal, 1976, 13(2), 165-180.

Krueger, L. Effect of irrelevant surrounding material on speed of same-different judgment of two adjacent letters. *Journal of Experimental Psychology*, 1973, *98*, 252-259.

Layton, B. Perceptual noise and aging. *Psychological Bulletin*, 1975, *82*, 875-883.

Leenheer, L. D., Appellmans, F., & Caestecher, K. de. The objective value of evaluating texture in the field. *Transactions from the Fifth International Congress of Soil Science*, 1954, *4*, 377-382.

Lichtenstein, S., Earle, T. C., & Slovic, P. Cue utilization in a numerical task. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, *104*, 77-85.

Lichtenstein, S., Fischhoff, B. *Training for calibration*. (Tr-78-A32). Alexandria, Vir.: U.S. Army Research Institute for the Behavioral and Social Sciences, November, 1978. (DAHC 19-77-0-0019).

Lichtenstein, S., Fischhoff, B., & Phillips, L. C. Calibration of probabilities: The state of the art. In H. Jungermann & G. de Zeeuw (Eds.), *Proceedings of the Fifth Research Conference on Subjective Probability, Utility, and Decision Making*, 1977.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. Judged frequency of Lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, in press.

Moscovmitch, M., & Craik, F. I. M., Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. _Journal of Verbal Learning and Verbal Behavior_, 1976, _15_, 447-458.

Murphy, A. H., & Wrinkler, R. L. Can weather forecasters formulate reliablee probability forecasts of precipitation and tempetures? _National Weather Digest_, 1977, _2_(2), 2-9.

Norman, K. L. Dynamic processes in stimulus integration theory: The effects of feedback on averaging of motor movements. _Journal of Experimental Psychology_, 1974, _102_, 39-408. (a)

Norman, K. L. Rule learning in a stimulus integration task. _Journal of Experimental Psychology_, 1974, _103_, 941-947. (b)

Norman, K. L. Effect of feedback on the weights and subjective values in an information integration model. _Organization Behavior and Human Performance_, 1976, _17_, 367-387.

Pauker, S. G., Gorry, G. A., Kassirer, J. P., & Schwartz, W. B. Towards the simulation of clinical judgment: Taking a present illness by computer. _The American Journal of Medicine_, 1976, _60_(June), 981-996.

Phelps, R. H. _Expert livestock judgment: A descriptive analysis of the development of expertise_. Unpublished Doctoral Dissertation, Kansas State University, Manhattan, Kansas, 1977.

Rabbitt, P. M. A. Ignoring irrelevant information. British Journal of Psychology, 1964, 55, 403-414.

Rabbitt, P. M. A. An age decrement in the ability to ignore irrelevant information. Journal of Gerontology, 1965, 20, 233-238.

Russell, E. W. Soil Conditions and Plant Growth. (10th ed.). London: Longman, 1973.

Shanteau, J. Averaging versus multiplying combination rules of inference judgment. Acta Psychologica, 1975, 39, 83-89.

Shanteau, J. Information integration theory applied to training of experts. In J. F. Hair (Ed.), AIDs Proceedings, New Orleans, 1979.

Shanteau, J., & Gaeth, G. J. Accuracy of the field method of soil texture analysis: A psychological analysis. Journal of Soil and Water Conservation, 1979, under review. (Also, Kansas State University Psychology series technical Report, 79-1).

Slovic, P., & Fischhoff, B. On the psychology of experimental surprises. Journal of Experimental Psychology: Human Perception and Performance. 1977, 3, 544-551.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Troutman, C. M., & Shanteau, J. Inferences based on nondiagnostic information. Organizational Behavior and Human Performance, 1977, 19, 43-55.

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

Tversky, A., & Kahneman, D. Belief in the law of small numbers. Psychological Bulletin, 1971, 76, 105-110.

USDA Soil Survey Manual. Agricultural Handbook No. 18,. U.S. Department of Agricultural, 1951.

Wallsten, T. S. Bias in evaluating diagnostic information. Report No. 157, L. L. Thurstone Psychometric Laboratory, Chapel Hill, N.C., 1978.

Wallsten, T. S Processes and models to describe choice and inference. Report No. 158, L. L. Thurstone Psychometric Laboratory, Chapel Hill, 1979.

Williams, C. The effect of an irrelevant dimension on "same-different" judgments of multi-dimensional stimuli. Quartely Journal of Experimental Psychology, 1974, 26-31.

Youngblood, J., & Himmelfarb, S. The effects of prior neutral messages on resistance to evaluative communication. Psychonomic Science, 1972, 29(6A), 348-350.

Footnotes

[1]Personal communication with James Shanteau, May 25, 1980 concerning preliminary results of G. Nagy's Doctoral disseration, Kansas State University.

[2]The author wishes to sincerely thank Vernon Hamilton, of the USDA Soil Conservation Service, for the considerable effort he expended to provide four soils containing coarse fragments which ewere used in this experiment.

[3]The soils were analyzed by the front range laboratory, Fort Collins Colorado, 80522, using a pipette analysis as described in ASA Monograph #9, Pages 562-565. The organic material was destroyed using hydrogen peroxide. The technique is described in ASA Monograph #9, Pages 573-574. In addition, three of the filler soils were also analyzed by the National Soil Survey Laboratory in Lincoln, Nebraska. This was done to check the validity of the Front Range Laboratory, which is a private laboratory. The average difference between the analysis for sand, clay and silt was only 2%.

[4]The level of excessive moisture was selected with the help of Dr. O. W. Bidwell, Soil Scientist, Kansas State University . For each of the test soils the particular level of excessive moisture chosen was picked to relect a level above that normally used when soil is hand texture classified, but still be a level which might reasonably be encountered in some field situations.

[5]The silt estimates will not be considered in the analyses. The soil judges all determined the silt estimates from through subtraction of the sand and clay estimates from 100%. In additior, these is no accepted procedure for determining the amount of silt in the soil using the feel method. All of the judges reported calculating the percentage of silt by subtraction. Thus, the silt dependent measure is both statistically and psycholcgy dependent on the sand and clay estimates, and was felt tc be a relatively uninterpretable dependent measure.

[6]The %-of-sand, silt, and clay and the texture-classficiations were determined by Dr. O. W. Bidwell, soil scientist and prcfessor, Kansas State University. He is primarily responsitle for the training of the soil judges used as participants in the study. His validity has been previously tested against a laboratory analyses, and had been found to be quite high.

APPENDIX

Instructions and Protocols Used in Experiment

ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

Instructions for the Evaluation Sessions

In this session I would like you to determine the texture of a number of soil samples. You will be asked to texture a number of soils in various conditions. This is designed to simulate field situations. Do the best you can with each soil. You should assess the texture of each soil individually, and do not look back at your earlier responses.

You will record your responses to each soil in your booklet (here). Could you fill in the requested information on the second page (wait). Thank you. Do not turn the page until instructed to do so.

You will use your booklet to tell you which soil to texture. The container for each soil sample is numbered and you should match this number with the one in the upper right hand cornder of each page of the booklet. Once you find the soil sample whose number appears in the right hand corner you should write down the time you begin to texture it. There is no time limit, in fact you should take your time and be as careful as possible. After writing down the time, you texture the soil in the container. You may use the water bottle, acid bottle, and the texture triangle provided. I would like you to write in both an assessment of the texture classification of the soil and your estimates of the percentage of sand, clay, and silt. There will be spaces provided to write these responses in your booklet.

Also, for each soil you texture I would like you to indicate how sure you are that your texture classification is correct. We will be asking you to make a probability statement on how likely you think it is that you have made the correct texture classification for that soil. Could you turn the page please. To indicate how sure you are of your texture you will place a mark on the line you see here. The line is marked "Definitely Wrong" at the left end and "Definitely Correct" on the right end, as you can see. Because we are asking you to make a probability statement consider the

Definitely Wrong or left end to mean that there is 0% chance that your texture assessment is correct. The Definitely Correct or right end means that you are 100% certain that your texture assessment is correct. The mark in the middle corresponds to a 50% point.

To familiarize you with using this line could you put a mark where you would if you feel that your texture classification has a 75% chance of being correct. Good. Now could you turn the page please. On this line, put a mark where you would if you feel your texture classification has a 15% chance of being correct. Good.

Now turn the page and we will practice recording all of the information into the booklet. First, look in the upper right hand corner and see the soil as an example. Normally there would be a number there to tell you which soil to texture. Suppose you began to texture the soil at 8:35. Could you write it in. Now imagine you textured the soil and decided it was a silty clay loam, with 9% sand, 61% silt, and 30% clay. Could you record these responses in your booklet. Also, you feel that it has a 90% chance of being the correct texture. Could you mark the line as you would. Very good. Then you finished at 8:41. Could you record that.

When you have finished with a soil sample put the soil in the container marked with the same number which is behind the first container. Then you may use the sink and paper towels to clean up before you go on to the next sample.

Remember to texture the soil whose number is in the upper right hand corner of each page, to write in the starting and finishing time, and to be sure the percentages of sand, clay, and silt add up to 100%.

Do you have any questions? After you have finished with a soil sample and cleaned up move on to the next page in your booklet. Please do not look back at your earlier responses. Texture each soil individually please. If you have no questions you may begin by turning the page and texturing the soil whose number appears in the right hand corner. I will be here to answer any questions which might come up.

Cognitive Training for Excessive Moisture

I've been studying how soil scientists use the feel method of soil texture classification for several years now. I have found that the overall accuracy of texture judgments is generally quite high. However, when mistakes are made they seem often to be due to the presence of excessive moisture in the soil. Naturally, the moisture shouldn't really change the soil texture at all. What I would like to talk about today are some of the characteristics of excessively moist soil and some of the methods which might be used to texture these soils. I am specifically interested in discussing how you might go about accurately assessing the texture of excessively moist soils.

We will go over three things in this session. In the first part I will show you some evidence that excessive moisture even affects the accuracy of soil texture classifications made by experts. Second, we will discuss what excessive moisture is defined to be, and how you may go about recognizing it in the soil. And then third, we will discuss some ideas which you might find personally useful to help avoid problems in judgment caused by excessive moisture.

DEMONSTRATION OF THE PROBLEM WHICH EXCESSIVE MOISTURE CAUSES

Now we will dicsuss five pieces of evidence on the influence caused by excessive moisture in soil. Hopefully you will be convinced that moisture does indeed affect the accuracy of the soil texture judgments. Here is a list (point) of the five points; you can follow along on the list.

KSU has one of the best soil judging teams in the country. Yet, when the regional competition was held in Iowa last fall their performance was not as good as we had hoped. One obvious problem our team faced in Iowa was that the soil conditions there were very different from the soils they practiced on here in Kansas. One major difference was caused by the flooding of the soil pits. There had been heavy rains and many of the soils used in the competition were actually underwater. As you can imagine, these soils were excessively moist. Dr. Bidwell and I have discussed this problem and we think that, although there were a number of other factors which contributed to the difficulties the team had in Iowa, certainly excessively most soil was one of them. (Pause)

Recently I spoke to Dr. Holzhey who is the head of the National Soil Testing Lab in Lincoln. He felt that, in his experience, soils which are excessively moist appear to cause soil surveyors some difficulty. He suggested that soil scientists should take special care when judging the texture of a soil which is excessively moist. (Pause)

Also I've looked at published articles in soil science journals which dealt with this problem. I have found two articles in which moisture was considered as a possible source of errors. In one article the authors felt that mistakes in texture classifications were generally due to extraneous materials in the soil. Of course, excessive moisture is an extraneous material. In a similar study another group of researchers found that even when the soil surveyors tried to allow for the presence of extraneous materials, such as excessive moisture, it still caused errors in the texture classification judgments. (Pause)

I have just completed a survey of Kansas soil scientists asking them for their opinion about the adverse influence of excessive moisture. These soil scientists work for the USDA and are all quite experienced in determining the texture of soil in the field. I asked them to indicate how much interference was caused by excessive moisture, on a scale from 1 to 7. As you can see the interference of moisture was consistently rated as high (show table), with a mean of 5.2. Thus, these soil scientists definitely felt that moisture adversely influenced their professional judgments of the texture of soil in the field. (Pause)

EVIDENCE THAT EXCESSIVE MOISTURE AFFECTS THE ACCURACY OF TEXTURE JUDGMENTS

1. Experience of soil team in Iowa.
2. Dr. Holzhey's comments.
3. Research in soil science journals.
4. Survey of USDA soil scientists.
5. Analysis of the data in the Kansas Soil Description.

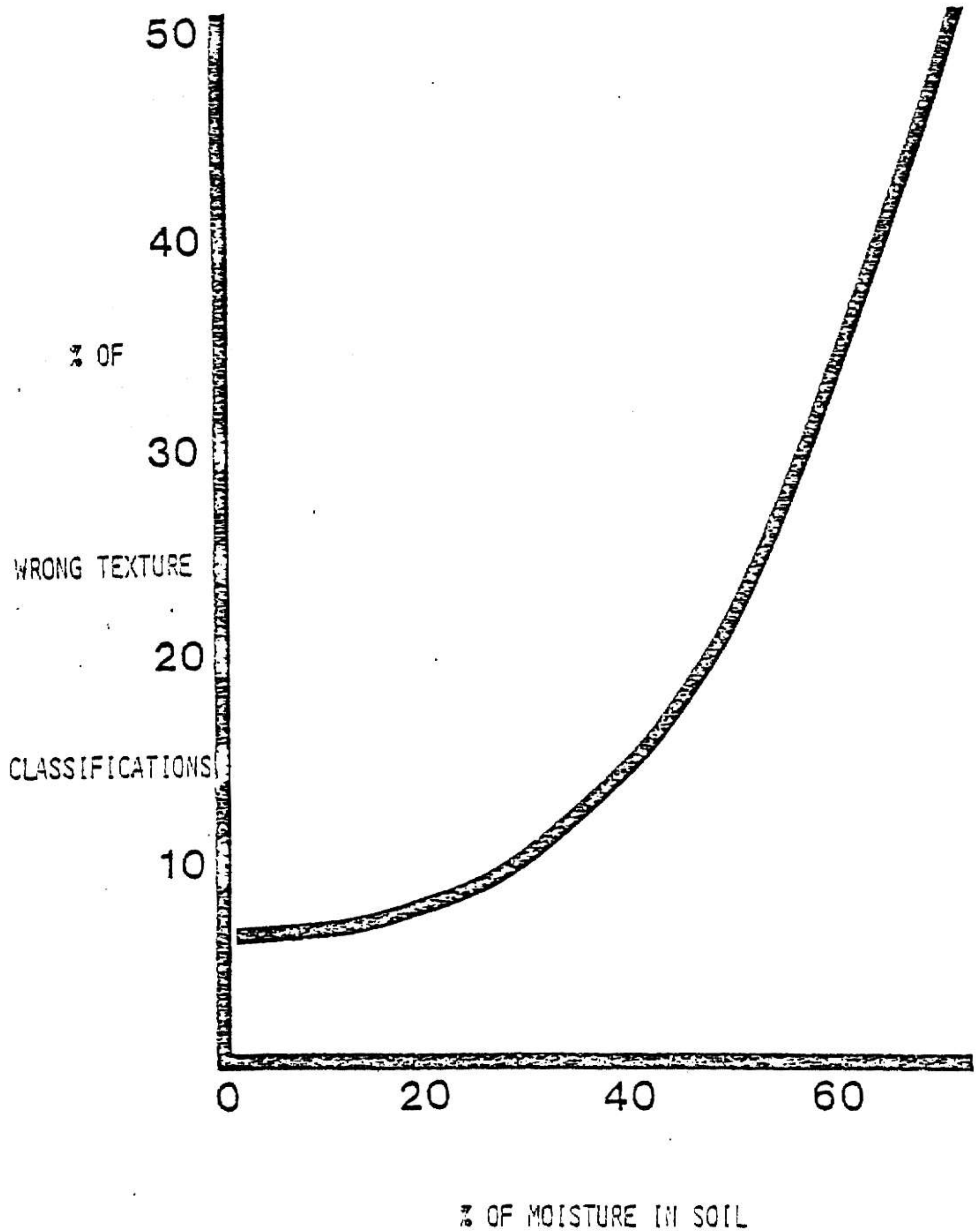
RATINGS[A] OF THE ADVERSE INFLUENCE OF EXCESSIVE MOISTURE

	Soil Scientist										
Factor	1	2	3	4	5	6	7	8	9	10	Avg.
Moisture	5	7	4	6	4	6	4	4	6	6	5.2

[A]Ratings were from 1 (little or no effect) to 7 (considerable or likely effect).

I have recently analyzed the data collected for the published Soil Survey Description of Kansas. As you may know, this contains information of both the surveyors' determination of the soil texture in the field and the lab analysis of each soil. I was interested in comparing the texture determined in the field with the texture analysis obtained from the lab. I have graphed a general representation of the relationship between the percentage of moisture and the number of errors (show graph). You can see that as the percentage of moisture goes up, so do the number of errors in the texture classifications. Roughly, when the moisture reaches a level of 60% the surveyors make a mistake over one-third of the time. You can imagine what might happen when the percentage of moisture gets higher than this. Does this graph make sense to you?

We have just discussed evidence showing that excessive moisture really does cause difficulties in making accurate soil texture assessments. Look over this list of the evidence which summarizes these points. I'd like to review them briefly. First, there is the experience of the soil judging team in Iowa. Remember they had considerable difficulty with soils which had excessive moisture in it. Then I indicated that Dr. Holzhey from the National lab has agreed that excessive moisture was a likely cause for errors in texture classifications, even by expert surveyors. We also talked about soil science research in which the authors indicated that they felt that excessive moisture might be a cause of some of the errors found in their studies. Next, experienced USDA soil scientists told me in my survey that interference from excessive moisture was quite high. Finally, we looked at a rough graph of data from the Kansas Soil Description and saw that as the percentage of moisture goes up so does the number of incorrect texture classifications. Do you have any questions about these points of evidence? Would you like me to go over any of them again?


50
40
% OF
30
WRONG TEXTURE
20
CLASSIFICATIONS
10
0
20
40
60
% OF MOISTURE IN SOIL

DEFINITION OF EXCESSIVELY MOIST SOIL AND WHAT TO LOOK FOR

As I am sure you already know, excessively mosit soil is soil which is wetter than field capacity. Just to refresh your memory, field capacity is the moisture level of the soil you typically achieve when you assess texture. Any soil sample which contains more moisture than field capacity is considered to be excessively moist. Does this agree with your understanding? (Yes-Good No-Can you accept this definition? It is the one currently used).

Although the definition of excessively moist soil is easily stated and understood there is more that I would like you to consider. The point I would like you to consider relates to the physical properties of the soil-moisture mixture. It is not unusual for there to be considerable differences in the amount of excess moisture in the soil. Some soils can have as much as 35% moisture in them. For this reason I would like you to specifically pay attention to the amount of moisture in the soil sample.

SEVEN SUGGESTIONS TO HELP REDUCE THE INTERFERENCE OF EXCESSIVE MOISTURE

I have seven suggestions which you may find useful in reducing the influence of moisture on your judgments of soil texture. I have listed them here (show list) on this chart. You can follow through the chart as we discuss each point.

The first step should be the removal of as much excess moisture as possible. You can get the moisture out of the way by letting it drain out. You must do this very carefully. It is possible to accidentally remove some of the soil with the water. There are two ways this can happen. First, if the soil is well-mixed then the clay and/or fine silt may be removed with the moisture. Of course this could result in a serious underestimation of the precentage of clay. Second, when the attempt is made to allow the water to run through your fingers it is possible the sand will sink through and not be picked up in your assessment of the texture. This would cause what could be a serious underestimation of the percentage of sand. It is probably best to remove only the moisture which definitely does not contain any of the soil particles. When in doubt, don't take it out.

Second, as we discussed before you should roughly try to estimate the amount of moisture in the soil sample you are texturing. This will help you to become aware of the moisture in the soil and to remember to watch out for the problems caused by the excessive moisture. We have seen earlier that when the percentage of moisture gets even moderately high the chances of making an error in the texture classification becomes quite high. Thus, you should assess the percentage of moisture in the soil and use that information to warn yourself that an error in the texture judgment is likely.

Next, after you have removed the excess moisture and estimated the amount of moisture you should roughly determine the condition of the soil-water mixture. Is the water and soil well mixed or settled? If it is well mixed it is likely to contain both clay and fine silt. However, if it is not mixed the moisture may have only fine clay in it. Thus, depending on the condition of the soil-water misture you may have to be suspicious of the clay and silt percentages or just the fine clay and keep this in mind.

Fourth, you should be especially careful and try to get a sample which is a good representation of the complete soil mixture. It is very important for the soil you texture to come from a good sample of the soil; one which contains the same proportion of the particles as in the original soil sample. You should look at the first sample you take and then use it to help you to decide how to take another one, if necessary. It may be necessary to try twice or more to get a representative sample. How difficult it is to get a good sample depends on the results of the previous step concerning the mixture.

Fifth on the list is a more general warning. You should make a special effort to assess the percentage of clay carefully. It seems likely that the excessive moisture will influence your ability to estimate the percentage of clay the most. Therefore it is especially important that you carefully evaluate the clay percentage. It might even be worthwhile to evaluate the percentage of clay twice; once in the beginning of your texture assessment and once at the end of your texturing procedure. Perhaps you might decide to take the average of these two values.

Sixth, you should also estimate the sand percentage very carefully for two reasons. If the sand is determined accurately then when you subtract to get silt the overall assessment will be more accurate. More importantly however, is the difficulty caused by the settling of the sand. Moisture is likely to allow the sand to settle to the bottom of the soil. Then, when you assess the soil it may be hard to find the correct amount of sand because you may get too much or too little from the bottom.

Finally, you should be careful that the large compacted particles of clay are adequately broken down. Sometimes particles of clay tend to stick together and appear to feel like sand. This can often cause trouble. If the soil has been excessively moist you may be tempted not to work it as much as you normally would. This may be no problem, but if the clay has not been broken down then you may tend to both underestimate the clay and overestimate the sand. I suggest that you make certain that the clay particles have been sufficiently broken down.

In summary I would like you to look over the list I gave you and have just discussed. Please read through them and summarize back to me what I have said.

SEVEN SUGGESTIONS TO HELP YOU TEXTURE SOILS CONTAINING EXCESSIVE MOISTURE

1. Remove the excess moisture.
2. Roughly estimate the amount of excessive moisture.
3. Determine the condition of the soil-moisture mixture.
4. Take a representative sample of the mixture.
5. Estimate the sand carefully.
6. Estimate the clay carefully.
7. Break down the compacted clay.

Perceptual Training for Excessive Moisture

In this session you'll actually be assessing the texture of soils which have excessive moisture in them. I would like you to concentrate on the things we talked about earlier in order to help you texture these soils. We will specifically be practicing the use of the seven suggestions I gave you in the last session. Feel free to ask questions at any time. Also, during this practice session feel free to redo any texture classification which you may be unsure of.

I'd like to remind you of my seven suggestions we talked about earlier. Here is that list. (Read through and remind for each point.)

Here is a soil to texture (#1). Could you determine the texture of it as carefully as possible. (Wait and record results; ___% sand, ___% clay, ______% silt, _________________ (texture), save soil). This soil was used in an earlier session and you assessed the texture of it then. When you saw this soil earlier you said then that it was a ____________ ________________ (use the worst response). This response you gave last _________ (day) agrees/disagrees with what you said before. Also, last time you said it had _______% sand, _____% clay, _____% silt. Your texture (if appropriate) and percentages have changed considerably. The only difference with the soil then was that it did not have any moisture in it when you judged it in the earlier session. The moisture in the soil had been removed then. What amount of moisture do you think is in the soil now. It is actually _____%. Let me say it again. The difference in what you said before and what you said today must be due to the moisture because it is the only thing different in the soils used in the two texture assessments. We will spend some time practicing with soil to see if we can help you with the excessive moisture problem.

What I would like you to do now is to look at the amount of moisture in this container (#2). I have added this amount of moisture to the same soil as you just textured. I would like you to texture this same soil with the amount of moisture I just showed you in it. Also, I would like you to follow the seven suggestions and carefully re-assess the texture. Remember to remove the excess moisture, think about the condition of the soil-moisture mixture, roughly estimate the percentage of moisture, and try to take a representative sample. Do you need to worry about breaking down the compacted clay particles?(Yes/No). Here is the soil with the moisture added (#3). (Record results; ___% sand, ___% clay, _____% silt, _______________). What amount of

(texture)

moisture do you think is in the soil? Actually, the soil had 30% moisture in it.

Now here is the same soil again with additional moisture in it (#4). Please retexture it and follow the list of seven suggestions. I would appreciate it if you would tell me out loud as you go through each of the seven things. (Record results; ___% sand, ___% clay, ____% silt, ______________). What amount

(texture)

of moisture do you think is in it? There actually is ___%. Now this soil has as much moisture as you should ever be faced with in the field. I would like you to go back to the first sample we worked with today and compare the current soil to that one. Remember that the only difference between the two soil samples is in the amount of moisture. Nothing else about the texture has been changed.

Here is another soil I would like you to texture (container #5). (Record results; ___% sand, ___% clay, ___% silt, _______________). It is not exces-

(texture)

sively moist. Now texture this soil (#6). It is the same soil with excessive moisture. (Record results; ___% sand, ____% clay, ___% silt, _________________).

(texture)

Could you tell me which of the seven suggestions you used and how (remind them of any they forget). What amount of moisture do you think there was in the soil? There was actually ___%.

Now texture this soil, again it was the same soil, but with even more moisture added (#7). You should practice using the seven suggestions again, but you don't need to say them aloud (Record results; ___% sand, ___% clay, ___% silt, _________________). What amount of moisture do you think there was in the soil?

(texture)

There actually was _____%.

Next, please look at the water in this container (#8). Now could you analyze the texture of this soil (#9). It has the same amount of moisture in it as you just saw added to it. (Record results; _____% sand, ____% clay, ___% silt, ______________________). Also, what amount of moisture do you think was in

(texture)

the soil? The actual amount was ____%.

Now I would like you to use the practice and experience you have gained from our work today to assess the texture of this soil, (#10). (Record results; ___% sand, ___% clay, ____% silt, ___________ (texture)). Actually you have seen this soil before when it had no moisture in it. When you judged the soil in the earlier session I had removed the moisture. And then you said it was a ___________ ____________, with _____% sand, ______% clay, ___% silt. (Use the best one this time). As you can see we have been able to practice and reduce the influence of the moisture considerably. It is very important that you try to remember what you did today when you texture soils in the future.

I would like to remind you of some of the things we practiced today in this session. These were: (repeat the summary from the verbal training, point). In this session we have stressed these aspects (point) of the problem of moisture and you have had a chance to practice using them. Do you have any questions? You will be asked to texture more soils in the next session. You should try and use the techniques we have practiced here when you assess the texture of these other soils.

Cognitive Training for Coarse Fragments

I've been studying how soil scientists use the feel method of soil texture classification for several years now. I have found that the overall accuracy of texture judgments is generally quite high. However, when mistakes are made they seem often to be due to the presence of coarse fragments in the soil. Naturally, the coarse fragments shouldn't really change the soil texture at all. What I would like to talk about today are some of the characteristics of soils containing coarse fragments and some of the methods which might be used to texture these soils. I am specifically interested in discussing how you might go about accurately assessing the texture of soils with coarse fragments.

We will go over three things in this session. In the first part I will show you some evidence that coarse fragments even affect the accuracy of soil texture classifications made by experts. Second, we will discuss what coarse fragments are defined to be and how you might go about recognizing them in the soil. And then third, we will discuss some ideas which you might find personally useful to help avoid problems in judgment caused by coarse fragments.

DEMONSTRATION OF THE PROBLEM WHICH IS CAUSED BY COARSE FRAGMENTS

Now we will discuss five pieces of evidence on the influence caused by coarse fragments in the soil. Hopefully you will be convinced that coarse fragments do indeed affect the accuracy of the soil texture judgments. Here (point) is a list of the five points; you can follow along on this list.

KSU has one of the best soil judging teams in the country. Yet, when the national competition was held in New Mexico last year their performance was not as good as we had hoped. One obvious problem our team faced in New Mexico was that the soil conditions there were very different from the soils they practiced on here in Kansas. One major difference was caused by the presence of coarse fragments in many of the soils used in the competition. As you can imagine, these soils were extremely difficult to texture accurately. Dr. Bidwell and I have discussed this problem and we think that, although there were a number of other factors which contributed to the difficulties the team had in New Mexico, certainly the coarse fragments in the soil was one of them. (Pause)

Recently I spoke to Dr. Holzhey who is the head of the National Soil Testing Lab in Lincoln. He felt that in his experience soils which contain fragments appear to cause soil surveyors some difficulty. He suggested that soil scientists should take special care when judging the texture of a soil which contains coarse fragments. (Pause)

Also I've looked at published articles in social science journals which dealt with this problem. I have found two articles in which coarse fragments were considered as a possible source of errors. In one article the authors felt that mistakes in texture classifications were generally due to the extraneous materials in the soil. Of course, coarse fragments are an extraneous material. In a similar study, another group of researchers found that even when the soil surveyors tried to allow for the presence of extraneous materials, such as coarse fragments, it still caused errors in the texture classification judgments. (Pause)

I have just completed a survey of Kansas soil scientists asking them for their opinion about the adverse influence of coarse fragments. These soil scientists work for the USDA and are all quite experienced in determining the texture of soil in the field. I asked them to indicate how much interference was caused by coarse fragments, on a scale from 1 to 7. As you can see the interference of coarse fragments was consistently rated as high (show table), with a mean of 4.4. Thus, these soil scientists definitely felt that coarse fragments adversely influenced their professional judgments of the texture of soil in the field. (Pause)

EVIDENCE THAT COARSE FRAGMENTS AFFECT THE ACCURACY OF TEXTURE JUDGMENTS

1. Experience of soil team in New Mexico.
2. Dr. Holzhey's comments.
3. Research in soil science journals.
4. Survey of USDA soil scientists.
5. Analysis of the data in the Kansas Soil Description.

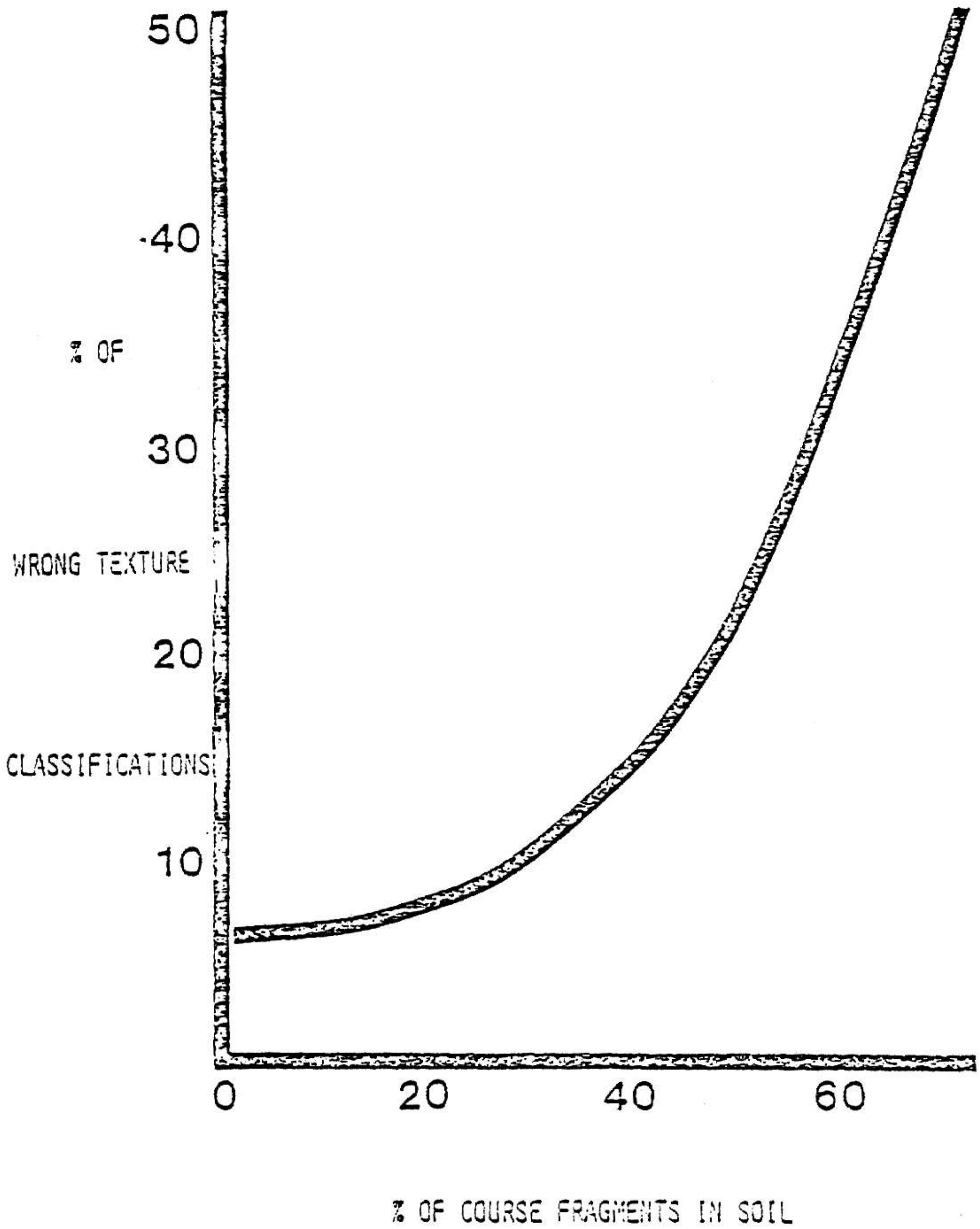
RATINGS[A] OF THE ADVERSE INFLUENCE OF COARSE FRAGMENTS

	Soil Scientist										
Factor	1	2	3	4	5	6	7	8	9	10	Avg.
Coarse Particles	2	7	1	4	3	6	1	7	6	7	4.4

[A]Ratings were from 1 (little or no effect) to 7 (considerable or likely effect).

I have recently analyzed the data collected for the published Soil Survey Description of Kansas. As you may know, this contains information of both the surveyor's determination of the soil texture in the field and the lab analysis of each soil. I was interested in comparing the texture determined in the field with the texture analysis obtained from the lab. I have graphed a general representation of the relationship between the percentage of coarse fragments and the number of errors (show graph). You can see that as the percentage of fragments go up, so do the number of errors in the texture classifications. Roughly, when the level of coarse fragments reaches a level of 60%, the surveyors make a mistake one-third of the time. You can imagine what might happen when the percentage of coarse fragments gets higher than this. Does this graph make sense to you?

We have just discussed evidence showing that coarse fragments really do cause difficulties in making accurate soil texture assessments. Look over this list of the evidence which summarizes these points (show list). I'd like to review them briefly. First, there is the experience of the soil judging team in New Mexico. Remember they had considerable difficulty with soils which had coarse fragments in them. Then I indicated that Dr. Holzhey from the National Lab had agreed that coarse fragments were a likely cause for errors in texture classifications, even by expert surveyors. We also talked about soil science research in which the authors indicated that they felt that coarse fragments might be a cause of some of the errors found in their studies. Next, experienced USDA soil scientists told me in my survey that interference from coarse fragments were quite high. Finally, we looked at a rough graph of data from the Kansas Soil Description and saw that as the percentage of coarse fragments goes up so does the number of incorrect texture classifications. Do you have any questions about these points of evidence? Would you like me to go over any of them again?


50
40
% OF
30
WRONG TEXTURE
20
CLASSIFICATIONS
10
0
20
40
60
% OF COURSE FRAGMENTS IN SOIL

DEFINITION OF COARSE FRAGMENTS AND WHAT TO LOOK FOR

As I am sure you already know, coarse fragments are any stony particles in the soil which are larger than 2mm in size. Just to refresh your memory as to how large 2mm is take a look at this sieve (show). It has 2 mm holes in it. Any fragment which will not pass through such a sieve is considered to be a coarse fragment. Any of the soil which will pass through the sieve is not considered to contain coarse fragments. Does this agree with your understanding? (Yes/good; No/ Can you accept this definition? It is the one currently used.)

Although the definition of coarse fragments is easily stated and understood there is more I would like you to consider. The point I would like you to consider relates to the physical properties of the soil sample. It is not unusual for there to be considerable differences in the amount of coarse fragments in the soil. Some soils can have as much as 35% coarse fragments in them. For this reason I would like you to specifically pay attention to the amount of fragments in the soil sample.

SEVEN SUGGESTIONS TO HELP REDUCE THE INTERFERENCE OF COARSE FRAGMENTS

I have seven suggestions which you may find useful in reducing the influence of fragments on your judgments of soil texture. I have listed them here (show list) on this chart. You can follow through the chart as we discuss each point.

The first step should be the removal of the very large fragments which are obviously larger than 2mm. You can get them out of the way just by physically pulling them out of the soil. You must do this very carefully. It is possible to accidentally remove soil with the coarse fragments. There are two ways this can happen. First, large sand which is nearly 2mm in size might be removed. Of course the result of this would be a serious underestimation of the percentage of sand. Second, clay might also be removed if it has not been broken down. This would cause what could be a serious underestimation of the percentage of clay. Thus it is probably best to remove only those coarse fragments which are definitely over 2mm in size. When in doubt don't take it out.

Second, as we discussed before you should roughly try to estimate the amount of coarse fragments in the soil sample you are texturing. This will help you to become aware of the coarse fragments in the soil and to remember to watch out for the problems caused by the coarse fragments. We have seen earlier that when the percentage of coarse fragments gets even moderately high the chances of making an error in the texture classification becomes quite high. Thus, you should assess the percentage of coarse fragments in the soil and use that information to warn yourself that an error in the texture judgment is likely.

Next, after you have removed the large coarse fragments and estimated the amount of the coarse fragments you should generally evaluate the size of the remaining coarse fragments. Are they close to 2mm in size, or are they larger? If they are nearly the size of sand, then you might get the two confused. Also, if the fragments are quite large it is possible that there are pieces of clay stuck into these large particles.

Fourth, you should attempt to determine the general shape of the particles. If they tend to be sharp they may stick out in your mind and be easier to ignore. However, if they are smooth and unobtrusive you may tend to get them confused with sand. This aspect is likely to vary from soil to soil. You should determine the shape of the coarse fragments and try to make whatever adjustments you think are needed.

Fifth on the list is a more general warning. You should make a special effort to assess the percentage of sand carefully. It seems likely that the coarse

SEVEN SUGGESTIONS TO HELP YOU TEXTURE SOILS CONTAINING COARSE FRAGMENTS

1. Remove the very large fragments.
2. Roughly estimate the amount of coarse fragments.
3. Evaluate the size of the remaining fragments.
4. Determine the shape of the fragments.
5. Estimate the sand carefully.
6. Estimate the clay carefully.
7. Break down the compacted clay.

fragments will influence your ability to estimate the sand the most. Therefore it is especially important that you carefully evaluate the sand percentage. It might even be worthwhile to evaluate the percentage of sand twice; once in the beginning of your texture assessment and once at the end of your texturing procedure. Perhaps you might decide to take the average of these two values.

Sixth, you should also estimate the clay percentage very carefully for two reasons. If the clay is determined accurately then when you subtract to get the silt the overall assessment will be more accurate. More importantly, however, is the difficulties which coarse fragments are likely to cause in your assessment of clay. Coarse fragments are likely to both reduce the length of the ribbon formed by the soil and also reduce the shine of the soil. Both these effects may make the determination of the clay percentage very difficult.

Finally, you should be careful that the large compacted particles of clay are adequately broken down. Sometimes particles of clay tend to stick together and appear to feel like sand. This can often cause trouble. If the soil has coarse fragments you may be tempted to not work the soil as much as you normally would. This may be no problem, but if the clay has not been broken down then you may tend to both underestimate the clay and overestimate the sand. I suggest that you make certain that the clay particles have been sufficiently broken down.

In summary, I would like you to look over the chart I gave you and have just discussed. Please read through them and summarize back to me what I have said.

Perceptual Training for Coarse Fragments

In this session, you'll actually be assessing the texture of soils which have coarse fragments in them. I would like you to concentrate on the things we talked about earlier in order to help you texture these soils. We will specifically be practicing the use of the seven suggestions I gave you in the last session. Feel free to ask questions at any time. Also, during this practice session feel free to redo any texture classification which you may be unsure of.

I'd like to remind you of my seven suggestions we talked about earlier. Here is that list. (Read through and remind for each point).

Here is a soil to texture (#1). Could you determine the texture of it as carefully as possible. (Wait and record results; ___% sand, ___% clay, ___% silt, _______________). This soil was used in an earlier session and you assessed
(texture)
the texture of it then. When you saw this soil earlier you said then that it was a ___________ ___________(use the worst response). The response you gave last ____ (day) agrees/disagrees with what you said today. Also, last time you said it had ___% sand, ___% clay, ___% silt. Your texture (if appropriate) and percentages have changed considerably. The only difference with the soil then was that it did not have any coarse fragments in it when you judged it in the earlier session. The coarse fragments in the soil now had been sieved out then. Also, the level of the coarse fragments is 10%. Let me say it again. The difference in what you said before and what you said today must be due to the coarse fragments because they are the only thing different in the soils used in the two texture assessments. We will spend some time practicing with soil to see if we can help you with the coarse fragment problem.

What I would like you to do now is feel the coarse fragments in this container (#2). These are the fragments which were originally in this soil. Remember the things we talked about as important aspects of coarse fragments? What is the shape of the particles; are they sharp or smooth? (underline). What is their size; do they feel near 2mm or are they quite large? (underline).

Next I will have you texture this same soil with these coarse fragments in it. I would like you to follow the seven suggestions and carefully re-assess the texture. Remember to remove the very large fragments, think about the size and shape we just talked about, and roughly estimate the percentage of coarse fragments in the soil sample. Do you need to worry about breaking down the compacted clay particles? Here is the soil with the coarse fragments added (#3). (Record results; ____% sand, ____% clay, ____% silt, _______________).
(texture

That soil, in fact, had 10% coarse fragments in it.

Now here is the same soil again with additional coarse fragments in it (#4). Please retexture it and follow the list of seven suggestions. I would appreciate it if you would tell me out loud as you go through each of the seven things. (Record results; ____% sand, _____% clay, ____% silt, __________________).
(texture)

What amount of coarse fragments do you think is in it? There actually is ____%. Now this soil has as much coarse fragments in it now as it did naturally. I would like you to go back to the first sample we worked with today and compare the current soil to that one. Remember that the only difference between the two soil samples is in the amount of coarse fragments. Nothing about the texture has been changed at all.

Here is another soil I would like you to texture (container #5). (Record results; ____% sand, ___% clay, ____% silt, _____________). There were no coarse
(texture)

fragments in the soil. Now texture this soil (#6). It is the same soil with the original coarse fragments left in. (Record results; ____% sand, ___% clay, ______% silt, _______ ___________). Could you tell me which of the seven sugges-
(texture)

tions you used and how. (Remind them of any they forget.) What amount of coarse fragments did you think there was in the soil? There was actually 15%. Now texture this soil; again it is the same, but with more of the original coarse fragments (#7). You should practice using the seven suggestions again, but you don't need to say them aloud (Record responses; ___% sand, __% clay, ___% silt, ___________ __________). What percentage of coarse fragments do you think were
(texture)

in the soil? There actually was 44%.

Next could you feel the coarse fragments in this container (#8). Now could you analyze the texture of this soil (container #9); it is the original soil for the same coarse fragments you just felt. (Record results; _____% sand, ____% clay, ___% silt, ______________). Also what was the percentage of coarse
(texture)

fragments? (____%). The actual amount of coarse fragments was 32%.

Now I could like you to use the practice and experience you have gained from our work today to assess the texture of this soil (_________________).
(texture)

(container #12). (Record results; ___% sand, ___% clay, ___% silt, ______________).
(texture)

Actually you have seen this soil before when it had no coarse fragments in it. For the soil you judged in the earlier session, I had sieved the coarse fragments out. And, when you say it earlier you said it was a ______________ ______________, with ___% sand, ___% clay, and ___% silt (use the closest one this time). As you can see, we have been able to practice and reduce the influehce of the coarse fragments considerably. It is very important that you try to remember what you did today when you texture soils in the future.

I would like to remind you of some of the things we practiced today in this session. These were: (repeat the summary from the verbal training and point). In this session we have stressed these aspects (point) of the problem of coarse fragments and you have had a chance to practice using them. Do you have any questions? You will be asked to texture more soils in the next session. You should try and use the techniques we have practiced here when you assess the texture of these other soils.

A COMPARISON OF LECTURE AND INTERACTIVE TRAINING DESIGNED TO REDUCE THE INFLUENCE OF INTERFERING MATERIALS: AN APPLICATION TO SOIL SCIENCE

by

GARY J. GAETH

B.S., Northern Michigan University, 1974
M.S., University of Wyoming, 1977

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER'S OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY
Manhattan, Kansas

ABSTRACT

This study was designed to test three general hypotheses concerning the influence of irrelevant info rmation on the judgments of tweleve experienced student soil judges, and the effectiveness of two training programs designed to reduce this influence. The first hypothesis predicted that material, technically irrelevant to soil judgments, would adversely influence these judgments. The second hypothesized that if the irrelevant materials did abversely influence the soil judgments then lecture and/or interactive training would reduce this influence. The third hypothesis predicted that reducing the influence of the irrelevant information through training (lecture or interactive) would led to improved accuracy relative to a performance criterion.

Twelve soil judges were used as participants. They attended five sessions; three evaluation sessions and two training sessions. The first, third and fifth sessions were evaluations, while the second session was lecture training, and the fourth session was interactive training. The evaluations each consisted of having the soil judges estimate the percentage of sand, clay, silt and assess the textural classification of 16 soils. These sessions were used both to determine the influence of the irrelevant inforamtion, and to aassess the impact of the training procedures. In addition, these responses were compared to a standard of performance defined by a professional soil

scientist and a valdity standard defined by a laboratory analysis of the soil. The performance and validity data allowed for a measurement of changes in accuracy (with respect to these standards). The two irrelevnat materials were excessive moisture and coarse fragments, both of which occur naturally in Kansas soils, and are defined by soil scientists to be irrelevant to soil judgment.

Analyses of the data from the evaluation sessions showed that both excessive moisture and coarse fragments influenced the judgments of the soil judges. Therefore, the irrelevant information contained in the excessive moisture and coarse was determined to be interfering. Consideration of the performance and validity data indicated that not only did irrelevant information influence judgments, but that the influence was adverse. That is, the interfering materials caused decreased convergence to the performance standardd. Thus, the first hypothesis was confirmed.

Following the first evaluation session the judges wer divided into two groups of six. One group was given lecture training designed to reduce the adverse influence of excessive moisture and the other group was given parallel lecture training designed to reduce the influence of coarse fragme nts. In general, both of the groups failed to show any impact of the lecture trainining. Some individual differences were found with respect to the impact of the lecture training.

Following the lecture training and the subsequent evaluation session (3), the same two groups of soil judges were given interactive training designed to reduced the influence of the interfering materials. The group that was originally given lecture insturctions for excessive moisture was also given interactive instructio ns for excessive moistuire. A parallel procedure was used for the coarse fragments group. The interactive training was found to significantly reduce the influence of the interfering materials (for both ting groups). In fact, the training for either of the two interfering materials generalized to the other interfering material. Consideration of the performance results showed that the interactive training improved performance through convergence to the standard. The validity restlts, although nonsignificant, showed a similar trend. Whether this improvement in accuracy was due entirely to the reduction of the influence of the interfering material is not known.