8-2003

The Effect of Age and Advice Accuracy on Compliance with Decision Support

Susan E. Vallance
Embry-Riddle Aeronautical University - Daytona Beach

THE EFFECT OF AGE AND ADVICE ACCURACY ON COMPLIANCE WITH

DECISION SUPPORT

by

Susan E. Vallance

A Thesis Submitted to the

Department of Human Factors and Systems

in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Human Factors & Systems

Embry-Riddle Aeronautical University

Daytona Beach, Florida

August 2003

THE EFFECT OF AGE AND ADVICE ACCURACY ON COMPLIANCE WITH DECISION SUPPORT

by

Susan E. Vallance

This thesis was prepared under the direction of the candidate's thesis committee chair, Frances Greene, Ph.D., Department of Human Factors & Systems, and has been approved by the members of this thesis committee. It was submitted to the Department of Human Factors & Systems and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors & Systems.

THESIS COMMITTEE:

Frances Greene, Ph.D., Chair

Dennis Vincenzi, Ph.D., Member

Gerald Gamache, Ph.D., Member

MS HFS Program Coordinator

Department Chair, Department of Human Factors & Systems

ACKNOWLEDGEMENTS

The author would like to thank her committee members, Dr. Frances Greene, Dr. Dennis Vincenzi, and Dr. Gerald Gamache for their guidance, knowledge, and patience. She would especially like to thank Dr. Gamache for his continued support over the years and for his encouragement to explore the field of human factors. She also wishes to thank Dr. John Wise for his initial guidance at the start of her thesis work and for encouraging her to head for NASA and for trusting that she would eventually return to complete the thesis requirement.

Finally, the author wishes to gratefully thank her family and friends, especially Scott, Mom, Dad, Catherine, Murphy, George, Bugsy, Hedges, Abe, Lou, and Buckley for their endless support, patience, and love.

ABSTRACT

Author: Susan E. Vallance

Title: The Effects of Age and Advice Accuracy on Compliance with Decision Support

Institution: Embry-Riddle Aeronautical University

Degree: Master of Science in Human Factors & Systems

Year: 2003

This thesis was designed to determine whether age or the accuracy of advice provided significantly effects compliance with a computerized decision support assistant. 48 participants in two groups, aged 20-40 (younger adults) and 41-69 (older adults), performed a monitoring/vigilance task intended to be similar to screening baggage with an X-ray monitor. A decision support assistant was provided to assist participants in choosing one out of four gray circles that had the most contrast with the background screen. Compliance with the decision support assistant's advice was then assessed. Results indicated that the level of advice accuracy did have a significant effect on compliance with decision support. As the advice accuracy level decreased, compliance decreased for both age groups. Although previous literature indicates that older adults may have negative attitudes toward computers, no significance was found for age or the interaction effect of age and advice accuracy on compliance with decision support technology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Most employees in the workplace interact daily with technology, specifically computer technology, in fields such as transportation, power, manufacturing, health care, and the computer industry in general (Bunce & Sisa, 2002). Age is an increasingly important factor to consider when assessing interactions with computer technology. As technology advances rapidly, most assume that young and old users will maintain awareness and embrace these changes and innovations. However, age related differences have been observed regarding computer use and attitudes among older and younger adults (Hoot & Hayslip, 1983; Krauss & Hoyer, 1984; Ansley & Erber, 1988; Mackie & Wylie, 1988; Jay & Willis; 1992; Ryan, Szechtman & Bodkin, 1992; Dyck & Smither, 1994; Marquie, Thon & Baracat, 1994; Czaja & Sharit, 1998; Ellis & Allaire, 1999; Marquie & Huet, 2000; Jennings & Onwuegbuzie, 2001; Marquie, Jourdan-Boddaert & Huet, 2002).

Older adults, in particular, may encounter factors that interfere with their ability to learn and adjust to new and unfamiliar technology, computer technology in particular. In addition, age related changes in cognitive skills, abilities, and unwillingness to learn and use computer technology may play a role in making adaptation difficult for older adults (Ryan, et al., 1992; Czaja & Sharit, 1998; Marquie et al., 2002). Assumptions may be made that older adults are more resistant to and uncomfortable with new technology than younger adults (Hoot & Hayslip, 1983; Jennings & Onwuegbuzie, 2001). There is some possibility that younger adults may be inexperienced and possess negative computer attitudes, but it is easier to presume that they were introduced to computers at a very young age and will not have as many difficulties adapting. This puts older individuals at a disadvantage because designers often fail to consider them as potential users (Czaja & Sharit, 1998). Therefore, efforts must be made to encourage individuals of all ages to explore computer technology. The more experience users have with technology, the more positive attitudes they will possess regarding the technology (Krauss & Hoyer, 1984; Jay & Willis, 1992; Czaja & Sharit, 1998).

Research indicates that older adults possess more negative attitudes toward computers, and less experience with computers, however, this does not rule out the possibility for successful interactions with computer users, regardless of age. Computer technology, particularly artificial intelligence (AI), is consistently advancing and paving the way for supporting problem solving and cognitive activities, as well as the positive interaction between humans and computers. Decision Support Systems (DSS) assist users in searching for information and solving problems. They provide support during cognitive activities by monitoring, formulation, plan generation, and adaptation (Roth, et. al, 1987). Therefore, DSS could be a viable solution to introduce and assist those less experienced and comfortable with computers. Although the application of DSS may not always be successful, because it is often difficult to absolutely determine if a DSS will perform reliably and accurately. Users may be skeptical, still encounter difficulties, and may not respond as expected or intended even when interacting with a 100% reliable assistant (Wiegmann, Rich & Zhang, 2001).

Although DSS are designed with the best intentions to assist users and provide "expert" or "intelligent" advice, difficulties may be encountered due to the design of the DSS, its actual suitability for the task and its compatibility with the user. Considerations are necessary for individual differences, such as age, as well as the type of task that the

DSS is supporting. For example, age differences have been observed in relation to vigilance tasks (Thackray & Touchstone, 1981; Bunce & Sisa, 2002). The type of advice and the level of reliability or accuracy of the advice presented should also be addressed. If the DSS appears to be unreliable or inaccurate, responses from users may not be positive or cooperative. Disuse or misuse of DSS may occur due to these factors and may hinder the effectiveness and ultimately defeat the purpose of the problem solving or diagnostics support provided (Parasuraman & Riley, 1997; Wiegmann & Cristina, 2000; Wiegmann et al., 2001). However, if attempts are made to accurately and reliably design and present them, DSS may successfully provide insight into and assistance with difficult tasks encountered by users of all ages. This may then result in compliance with an effective and accurate DSS and positive user attitudes toward interaction with computers.

1.1 Statement of the Problem

Although research regarding computer attitudes, performance, and age differences is extensive, results have been mixed and often a complete range of ages groups is not sampled. Therefore, no definitive answers have been agreed upon regarding how age relates to computer attitudes and performance. Research is extremely limited regarding individual differences, specifically age, and the use of decision support systems (DSS) and automated diagnostic aids. Bunce & Sisa (2002) suggested based on their findings regarding age differences, vigilance task performance and perceived workload that it is vital to attempt to moderate age differences. One method may be to provide environmental support to possibly reduce demands on attention. Current agent supported computer work (ASCW) and artificial intelligence (AI) technology offers many options and tools for assisting users while navigating the World Wide Web or for providing environmental, problem solving and decision support. However, little research has been conducted to investigate how these tools and applications affect users when they do not perform as expected? If an individual is given the opportunity to complete a problem-solving task with an application characterized as an "intelligent" decision support assistant, will the user consider and comply with the assistant's advice based on their brief experience with the tool during a simulated task? If the computer occasionally or even frequently provides inaccurate assistance, will the user continue to comply with the supposed decision support assistant, based on the assumption that it is "intelligent?"

Based on a thorough literature search, it was determined that to date, there are no published studies which measure compliance of a wide age range of users with a decision support tool with varying levels of accuracy while performing a monitoring/vigilance task. This study attempted to determine the outcome of this scenario with adults in 2 age groups, younger adults, 20 to 40 years old and older adults, 41 to 69 years old, while performing a task involving determining the contrast of colored circles to a dark background. The task was similar to a monitoring/vigilance tasks performed by X-ray baggage screeners. For each decision, users were provided with the help of a decision support assistant known as CAL, which provided advice regarding which circle had the most contrast with the background at 2 random levels of accuracy (95-100% accurate and 65-70% inaccurate).

1.2 Review of the Literature

Interaction with computer technology is commonplace on the job. But what are the consequences if an adult at any age with varying computer experience, utilizing computer or monitoring technology, is presented with the opportunity to complete their daily tasks with the help of a computerized decision support assistant? Will their age play a role in whether or not the user complies with the DSS advice? What happens if the decision support assistant provides inaccurate advice? The purpose of this study is to address how age differences and the level of accuracy of advice provided by a simulated decision support assistant may effect compliance with the decision support assistant.

The following review of the literature will first discuss age differences and computer use. This section includes specific information regarding age differences and computer attitudes, age differences and the performance of vigilance tasks and age differences and reactions to computer based work errors. This is followed by a discussion of decision support systems (DSS), automated diagnostic aids and computer advice. Particular emphasis is placed on individual differences in the use of DSS, automated diagnostic aids and DSS reliability and accuracy, the perception of computer advice, and a possible application related to this study: decision support for X-ray passenger baggage screening. Compliance, the dependent variable for this study, is then discussed. And finally, conclusions based on the literature are drawn and hypotheses are provided.

1.2.1 Age Differences and Computer Use

Whether checking baggage at an airport, checking groceries, or checking bank account balances online, computers are unavoidable and indispensable. For many, it is difficult to recall a time when everyday life did not rely on these technological advances. On the other side of the spectrum, for some older adults this technology may not be so familiar. Those who may not have been in high school, or college, or even in the workplace when computers began to emerge significantly, may not have had the opportunity to acclimate themselves to computers. Some may simply not have the desire to work with these ever changing and advancing machines and accompanying technology.

Previous research has suggested that older adults are more negatively oriented to computers than younger adults (Hoot & Hayslip, 1983; Ansley & Erber, 1988; Ryan, Szechtman, & Bodkin, 1992; Dyck & Smither, 1994; Birdi & Zapf, 1997; Czaja & Sharit, 1998; Ellis & Allaire, 1999; Marquie, Bodddaert & Huet, 2002). This speculation may stem from the idea that older adults feel threatened by the demise of paper and pencil and the rise of computer based information-processing techniques (Ansely & Erber, 1988). Similar to decision-making research findings, if an individual, no matter the age, believes they cannot do something, they will not necessarily put forth the effort or try to succeed at the task (Marquie et al., 2002).

Computer anxiety may also be a factor that affects an older adult's computer skills and willingness to use computers (Dyck & Smither, 1994). Higher anxiety levels in younger adults may have negative effects on computer use and skills (Marcoulides, G.A., 1988). This is not to say that older adults in the world today are not proficient and positive regarding computers. Although interestingly enough, recent research indicates that computer attitudes are still low for older adults. Based on the reviewed sampling of mixed findings, for the purposes of this study, computer attitudes and experience are addressed specifically in relation to a user's age, while performing vigilance or monitoring tasks with the assistance of a decision support assistant with varying levels of decision support accuracy.

1.2.1.1 Age Differences and Computer Attitudes

Research regarding user attitudes and experience with computer technology in relation with age, has provided mixed results. Ansley & Erber conducted a study in 1988 to assess the relationship between older users and computer attitudes. Several aspects were addressed including older community living adults' attitudes toward computers and individual differences regarding these attitudes. The Wagman Cybernetics Attitude Scale (CAS) was used and responses from older adults were compared with scores from undergraduates who were assumed to be directly interacting with computers and computer technology. Sixty older adults ranging in age from 55-86 with a mean of 70.7 years old were interviewed. Findings indicated that the older participants in this study did not seem to have negative attitudes toward computer technology. They also did not seem hindered by task performance with computer presentation. Ansely & Erber (1998) concluded that no real differences were observed in computer attitudes between younger and older adults.

Based on findings from a computer experience questionnaire and computer anxiety and attitude scale, Dyck and Smither (1994), found that out of 219 young and 203 older adults, older adults had more positive computer attitudes than younger adults did. Although, the older adults did indicate they had less confidence in their ability to use computers and less experience than the younger adults. Dyck & Smither (1994) concluded that older adults' attitudes were more positive because of the type of experience they had with computers.

In a survey of six hundred twenty 18-70 year old office workers, regarding computer attitudes, computer use outside of the office and levels of computer training, Marquie, Thon, and Baracat (1994) found that computer experience was the most important factor that influenced computer attitudes. Unlike the previously mentioned research, older workers displayed greater negative attitudes and less knowledge about the use and utility of computers, and more fear regarding threats to employment. Older workers were less comfortable with computers and more sensitive to the less flexible operating procedures when performing computer tasks as opposed to paper based tasks.

In another study, Czaja & Sharit (1998) attempted to examine the influence of computer experience on the computer attitudes of older adults. This was an extension of prior research and examined a data entry, database inquiry, and accounts balancing task, all meant to resemble common tasks conducted daily in the workplace, each placing a different kind of demand on the individual. Czaja & Sharit (1998) conducted this research in an effort to understand whether characteristics of computer tasks influenced attitudes toward computers, with this understanding ultimately leading to more effective interface design and training. Their goal was to determine if computer experience causes a change in attitudes toward computers, if experience varies across attitude dimensions and if the effects vary with task characteristics and age. The study also considered computer attitudes and ratings of workload, stress, and arousal. Three hundred eighty four subjects aged 20-75, divided into three age groups, younger – 20-39, middle aged – 40-59, and older 60-75 adults participated in the study over the course of 5 days. Subjects were screened and trained and then participated in one the three simulated computer-based work tasks for 3 days. Throughout the course of the research, the Stress Arousal Checklist, NASA TLX and Attitudes Toward Computers Questionnaire (ATCQ) was administered.

Czaja & Sharit's (1998) found that computer attitudes are modifiable, and regardless of age or gender, positive attitudes resulted from direct experience with computers. They indicated that these findings parallel previous research and emphasize the importance of allowing computer users, novices especially, the chance to interact with computers. With more interaction and experience, one could assume that attitudes toward computers would become more positive. Findings also highlighted the importance of adequate training and the usability of interfaces.

Regarding age and computer attitudes, Czaja & Sharit (1998) age effects were observed for several attitude dimensions. Older adults reported feeling less competent and comfortable with computers, as well as that they had less control over the computers, similar to findings by Marquie et al., (1994). Czaja & Sharit (1998) also found prior experience with computers positively related to comfort, competence, and efficacy ratings with computers, while older participants demonstrated less prior experience than the younger participants did. Age effects also occurred even after controlling for experience, which suggests that other age related factors were related to computer attitude findings. This supported Czaja & Sharit's (1998) interest in assessing computer attitudes on a multidimensional level.

Jennings & Onwuegbuzie (2001) intended to replicate and extend past research related to computer attitudes to determine if associations with variables such as age, gender and attitudes demonstrated before the widespread use of graphical user interfaces (GUI) operating systems still existed. Three hundred fifty one undergraduate students in one of three age groups, 22 years or less, 23-29 years or 30 years or greater, participated. It was determined that age was significantly related to specific dimensions of computer attitude, but not consistently with each dimension. Similar to Czaja & Sharit (1998) and Loyd and Gessard (as cited in Jennings & Onwuegbuzie, 2001), the relationship of age to computer attitudes was significant, but inconsistent. The largest age group was the youngest and showed the lowest computer anxiety and highest computer confidence. This might be because the younger participants had exposure to computers throughout their educational experience, unlike the older participants. It is possible that participants 30 years and older were unfamiliar with computers, which may have led to the high computer anxiety and low computer confidence levels demonstrated. In regards to participants in the middle age group, medium to high levels of computer anxiety and confidence were observed. In conclusion, while Jennings & Onwuegbuzie (2001) found an association between computer attitudes and age, no clear linear trend was indicated.

Previous studies have suggested that difficulties and non-cognitive factors such as negative stereotypes related to age, attitudes, and the fear of computers and its consequences may cause difficulties experienced by older adults when mastering new information or technology (Marquie et al., 2002). According to Mackie & Wylie (1988), four factors affect user acceptance of technology. They include the user's acceptance of the technology and its purpose, the consistency of the technology's features with the users needs, experience with the technology, and available training or documentation support. In 1983, Hoot & Hayslip commented that the computer revolution was quickly passing older adults by and the growing computer industry exerted minimal effort to attempt to attract older adults' attention. They concluded that those in the computer industry might perceive older adults as incapable or simply unwilling to use computers.

In order for users to maximize the benefits of this technology, it is essential that users possess positive attitudes toward computer use in general (Jennings & Onwuegbuzie, 2001). Danowski and Sacks (1980) found that positive experiences with computer-mediated communication resulted in positive attitudes for older adults. The more experience users have with technology, the more positive attitudes they will have regarding the technology (Krauss & Hoyer, 1984; Jay & Willis, 1992; Czaja & Sharit, 1998). Ellis & Allaire (1999) recommend that when training older adults to use computers, focus should remain on increasing the user's level of knowledge and decreasing their level of anxiety related to computers. One previously successful method to achieve this knowledge and comfort level and attitude in general regarding computers is to provide hands on training for older adults (Jay & Willis, 1992; Czaja & Sharit, 1998; Ellis & Allaire, 1999).

1.2.1.2 Age Differences and Vigilance Tasks

Monitoring tasks such as X-ray baggage screening involve vigilance, although vigilance is not a variable formally assessed in the present study. Literature and research involving vigilance is extensive. Therefore, it is important and pertinent to briefly review some of the literature specifically related to age differences and vigilance tasks. Results of these studies may directly relate to the present study as participants' in this study will also vary in age and will be performing a simulated vigilance/monitoring task.

According to Bunce & Sisa (2002), evaluating age differences in the performance of demanding vigilance tasks for short durations is vital due to the safety critical aspects of monitoring in airports and other workplaces.

Bunce & Sisa's (2002) investigated the performance effects of age and the perceived level of workload during a 9 minute highly demanding vigilance task. A series of monochrome digits ranging from 0-9 were presented on the center of a PC screen. Reversing 30% of the pixels and defining the digit and its surround degraded digits. Sixty practice trials and 540 trials were administered requiring participants to respond to the specific target digit 0 by pressing the space bar. Distracter digits did not require a response. Participants were also expected to complete a card-sorting task, before completing the aforementioned vigilance task. Performance measures were taken and workload measures following all tasks. Twenty-six younger adults ranging in age from 16-35 years old and 24 older adults aged 45-65 years old participated in the study. Bunce & Sisa (2002) hypothesized that during the vigil older adults would have lower performance scores than the younger adults and older adults would perceive a greater increase in workload.

Bunce & Sisa's (2002) results indicated that for this sample of ages, no age differences existed in the ability to maintain a vigil. However, age differences related to perceived workload across the vigil did exist. They observed age related differences related to increased mental, temporal and physical demands and frustration. Mental and temporal demands differentiated younger and older adults throughout the study and were major sources of workload. Unexpected statistical significance was observed related to physical demands, but became non-significant when considering the relative importance of the source of workload in performing the task. Bunce & Sisa (2002) also observed higher levels of frustration in younger participants.

In summary, although performance across a vigil was equal for both age groups, older adults experienced greater demands on attentional resources during the vigil. Based on these findings, the assumption that attentional resources are limited, and that older adults' vigilance performance may suffer sooner than younger, these are important findings for monitoring and safety critical situations. For baggage screeners in particular, screeners may vary in age from 18 to over 60 years of age. Therefore, performance decrements with vigilance tasks whether age related or not could be very detrimental and could lead to greater problems especially if a suspicious or even worse, explosive item goes missed while scanning baggage.

Surwillo (as cited in Thackray & Touchstone, 1981) observed a greater performance decrement in sustained attention (the ability to maintain a high level of attention to a visual display over extended durations) with older adults during a vigilance task requiring minimum search. Thackray & Touchstone (1981) attempted to extend this research by experimenting to see if a task involving greater visual search would demonstrate even more sustained attention age differences. Forty five subjects aged 18-29, 40-50 and 60-70 years old were asked to monitor a display resembling an air traffic control radar display containing alphanumeric data blocks for 2 hours for occasional designated changes in the alphanumeric data. These changes were referred to as critical stimuli.

Based on other studies regarding selective attention, the ability to quickly detect relevant stimulus in the presence of irrelevant or competing stimuli, Thackray &

Touchstone (1981) hypothesized that the older the subject, the greater the mean detection time on the task. They also considered that previously observed performance differences related to age may only occur after an extended period of task performance, with impairment to performance occurring faster with older adults.

Thackray & Touchstone (1981) observed that older subjects exhibited longer average detection times and more errors of omission and commission while detecting critical altitude changes in the presence of many competing and similar stimuli. Initial age differences in detection time were not observed at first, but performance impairments were evident with the older group of adults as the task progressed in length. These findings agree with previous similar vigilance research comparing older and younger adults. Thackray & Touchstone (1981) concluded that age-related differences in selective attention are more pronounced after some time has elapsed during prolonged monitoring.

1.2.1.3 *Age Differences and Reactions to Computer Based Work Errors*

Birdi & Zapf (1997) conducted research to determine age-related differences in the affective and behavioral reactions of office workers who encounter problems while using computers. They proposed that older employees would demonstrate more negative emotion reactions to computer work errors and would be less likely to try to solve problems on their own without the use of support, than younger employees. They surmised that this would be due to older employees having less affective orientation to computers, a lack of experience with computers and the possibility that they would be more prone to errors in their computer work. Birdi & Zapf (1997) supposed that errors in computer work made by older adults might be a function of their negative affective orientation to computers. Regarding experience, Birdi & Zapf (1997) assumed, based on previous research, that older adults' lack of computer experience, knowledge, and familiarity as compared to younger adults may lead to negative affective reactions to dealing with computers and computer errors. Birdi & Zapf (1997) proposed if older adults make more errors than younger adults, this may explain these age differences in reactions to errors and strategies to dealing with them.

Birdi & Zapf utilized 134 subjects ranging in age from 19-55 years of age with 24 participants over the age of 40. A questionnaire regarding negative emotional reactions, attitudes to new technology, computer experience, demographics, attempts at self-correction for computer errors and sources of support for computer errors was administered. In addition, the number of errors encountered during the observation was recorded, negative emotional reactions, and attempts at self-correction.

Findings indicated that older adults did in fact display stronger negative reaction to errors in computer work, even after education, individual attitudes toward technology, computer experience and the total of errors made during a typical work task were controlled for. The questionnaire revealed that older adults were significantly less likely to try to solve the problem themselves. Finally, older adults referred more to manuals rather than asking colleagues or supervisors for support and assistance.

1.2.2 DSS, Automated Diagnostic Aids and Computer Advice

Bunce & Sisa's (2002) findings regarding age differences and vigilance task performance, and perceived workload suggest that it is vital to attempt to moderate age differences. One method may be to provide environmental support to help reduce demands on attentional resources. They proposed the introduction of intelligent interfaces for monitoring and vigilance tasks where the balance between man and machine control is considered a continuum. As task workload increases, automated systems will moderate the workload. This occurs if an optimum level of environmental support is provided and maintained around the middle of the active-passive continuum to facilitate mental operations related to the task then performance benefits may be observed, especially for older operators.

According to Roth, et al., (1987), when determining how to configure human and machine cognitive systems, it is vital to consider the machine's capabilities as extensions and expansions along dimensions of machine power. Machines are tools and people build and use them. Keeping this in mind, decision and problem solving support systems may be created as prostheses, which correct or remove deficiencies. Computational technology is developed as stand-alone or interactive experts that assist with problem solving at various levels. Decision support systems (DSS) are one example of the type of artificial intelligence computer applications or environmental support systems that can be implemented. DSS support solving problems and complex decision-making. They were originally defined by Gorry and Morton (as cited in Shim et al., 2002) as computer systems that dealt with problems at various stages. The stages can be defined as semi-structured or unstructured, routine and easy to solve, or new and difficult to solve. These computer systems are often designed to deal with specific portions of problems creating a human-machine, problem solving system. In the last two decades, research regarding DSS has expanded to group decision support systems and group support systems.

Executive information systems extend the use of DSS to the corporate level. Model management and knowledge-based decision support systems utilize artificial intelligence and expert system techniques to provide intelligent support for decision makers (Bonczek, R.H. et al., 1981; Courtney & Paradice, 1993, Shim et al., 2002).

Nuclear, aviation, manufacturing, and other industries currently DSS. Research related to Air Traffic Flow Management and decision support systems is roughly a decade old. DSS are also widespread on the World Wide Web delivering decision support information or tools to a variety of different users. Web based DSS has made decision relevant information and model driven DSS available to managers and staff in the workplace, no matter their geographic location.

Intelligent decision support designers may overestimate their ability to capture all of the relevant aspects of the problem-solving situation at hand in the behavior of the machine expert or assistant (Roth, 1987). Designers may fail to address or consider all of the anticipated and even expected aspects of the varied and complex operational situations the application is created for. Thus, the system cannot support or may even interfere with the user achieving their intended goals (Roth, 1987). This is often a problem with automation and support systems, similar to intelligent agents, intelligent help systems, and DSS. "There are several steps that can be taken to convert the power of the human machine expert into a more instrumental form. One is to build displays that provide a shared frame of reference for the person and machine" (Roth, p. 504, 1987).

1.2.2.1 Individual Differences and the Use of DSS

Findings from Czaja & Sharit's (1998) study provided interesting data related to initial computer task performance experiences that influence an individual's attitude and how attitude change is influenced by changes in performance as a function of task experience. They found that the nature of a computer task could affect both of these relationships. Changes in computer attitudes were not a function of performance level for the less cognitively demanding tasks such as data entry and database management. However, with the accounts balancing task, those who performed better had more positive computer attitudes than those who experienced initial difficulties with the task. Czaja & Sharit (1998) determined that this emphasizes the importance of matching the demands of a particular task with the cognitive skills of the user. If there is a mismatch, users may feel incompetent or incapable and negative attitudes toward the computer and interaction may increase and affect their potential willingness to work with the computer in the future.

These findings may also be applicable when considering use of a decision support system (DSS) or computer aided diagnosis system. Not only would it be optimal for the DSS technology to match up with the users cognitive abilities, but what if the DSS was the "incompetent" one? How would users react to a supposed "intelligent" decision making assistant that did not provide correct responses? How would this negative interaction affect an individual's computer attitudes and willingness to work with the computer in the future? Could this possibly also relate to the person's age, since previous research has indicate that older adults are less experienced with computers and therefore may be less apt to adapt to working with an accurate, let alone an inaccurate decision support assistant?

1.2.2.2 Automated Diagnostic Aids and DSS Reliability and Accuracy

Level of frustration and level of performance during initial interactions with technology may have an influence on attitude change. Therefore, it is essential to ensure that users are provided with adequate support during their interactions with technology. (Czaja & Sharit, 1998). Systems become increasingly more automated everyday. System operators process large amounts of intricate information that constantly grows in complexity. This growth has led to operators being removed from processes formally under their control and often times operators are incapable of visually assessing the state of the system or possible alternatives for action (Wiegmann, Rich & Zhang, 2001). This can often interfere with system control and diagnosis during failures, which leads to diagnostic and decision errors, often the leading cause of difficulties and even disasters. Efforts are made to decrease human error and its consequences by designing tools to improve diagnosis and decision-making. These efforts have included the development of computer decision-making and automated diagnostic aids.

Although designed with the best intentions, technology such as computer decision-making and automated diagnostic aids are not always perfect and often underutilized when they are less than 100% accurate. Disuse is often evident when an operator disagrees with an aid, even when the aid, although imperfect, is still statistically on average more accurate than an unaided diagnosis or decision (Parasuraman & Riley, 1997; Wiegmann & Cristina, 2000; Wiegmann et al., 2001, Wiegmann, 2002). The underestimation or misperception of an aid's reliability may result in distrust or disuse of automation. It's common that operators expect to interact with a perfect automated assistant, so if any errors do occur, operators focus too much on the errors, which leads to this underestimation of the reliability of the aid (Wiegmann et al., 2001, Wiegmann, 2002).

Wiegmann, Rich & Zhang (2001) examined the effects that different levels and changes in automation reliability had on users' "trust" in automated diagnostic aids. 47 participants were presented with 200 trials of a computer simulation task that involved diagnosing the validity of pump failures within a waste processing facility, using information provided only by an automated diagnostic aid. The likelihood of the aid presenting a correct diagnosis depended on the experimental condition the subject was assigned to. The users were not aware of this, but the aid was 60%, 80% or 100% reliable. In addition, the 60% reliable aid's accuracy increased to 80% half way through the trials, the 100% reliable aid's accuracy was reduced to 80% half way through the trials, and the 80% reliable aid's accuracy remained the same throughout the trials. Subjective perceived reliability of the aid and confidence ratings along with objective performance measures including concurrence with the aid's diagnosis and decision times were obtained.

Results obtained suggest that if automated diagnostic aid reliability differs initially by a magnitude of 20%, users will be sensitive (Wiegmann et al., 2001). Higher initial reliabilities were associated with higher agreement rates with the aids, higher confidence ratings in decisions on agreement trials, and higher estimates of aid reliability, as well as faster decision time on agreements. Performance over time led participants to exhibit sensitivity to changes in the aid's initial reliability. Group differences on all measures shifted as reliability shifted to 80%. No general contrast effects were observed, except for possibly subjective reliability estimates, which suggested that previous exposure to a particular aid reliability did not lead to the user's hypersensitivity to decreases or increases in reliability.

Wiegmann et al., (2001) also found that all 47 participants underestimated aid reliability in all conditions, reconfirming Wiegmann & Christina's 2000 findings that when not perfectly reliable, users underestimate true aid reliability. Although participants who interacted with the aid that was 80% correct throughout all of the trials, had more accurate user estimates of reliability, which suggest that multiple interactions with an aid may be needed to accurately calibrate trust in it. Thus, underestimates of reliability for aids with shifting reliabilities may have been because the user did not interact with the aid enough to come up with a more accurate estimation (Wiegmann et al., 2001). Participants' reliability estimates were lower, but correlated with observed agreement rates with the aid. This is possibly due to users expecting "perfect" automated assistance and when they do not receive it, the errors are more likely to be recalled. With that in mind, automation errors have a somewhat unjustified influence on users' aid reliability estimates and the use of imperfect aids, even when aided diagnosis is more accurate than unaided.

Because the users were provided with no information to base their diagnosis on, it may be best for the user to always agree with the aid to make the most correct decisions. Even under conditions of 100% aid reliability, periodic disagreement with the aids was observed (Wiegmann et al., 2001). Here, users may have felt that since the aids had been consistently correct for several trials, an incorrect diagnosis was soon to follow. Therefore, users may have even prematurely disagreed with the aid before the automation failure even happened. However, participants in this study continued to disagree with aids after multiple trials with the 100% accurate aid. Finally, a disassociation was observed between subjective reliability estimates and objective agreement rates, which suggested to the authors that they do not reflect the same fundamental constructs of automation "trust" and that reliability estimates may only appropriately be inferred via subjective measures. Based on their findings, Wiegmann et al., (2001) concluded that automation trust, utility, utilization strategy and reliance need to be operationally defined and measures of automation trust should be subjectively and objectively distinguished to effectively design automated aids that will optimally impact the performance and safety of systems.

Wiegmann (2002) continued his previous research by examining different types of utilization strategies used by users when automated diagnostic aids were less than 100% reliable, but the aid was more accurate than unaided diagnosis, and when the system provided no aid or insight to make their diagnosis. Fifty participants were presented with a set of 120 test trials to diagnosis the validity of system failures. For every trial, they were expected to use a diagnostic aid that they were not aware was only 80% reliable and were then asked to agree or disagree with the aid's advice. The frequency of these agreements and disagreements was assessed to determine the type of concurrence strategies used. When asked to determine system failures during the test trials, participants were presented with two aids of equal accuracy and were asked to pick one to use. After making a selection, an interim screen flashed and then the diagnostic aid's conclusion regarding the system failure was provided. The participant was then asked to choose whether they accepted or rejected the advice. They also indicated their confidence in their answer with a Likert rating scale (1 no confidence – 5 very confident). Feedback was provided regarding the correct diagnosis and a score was provided. Participants were also asked to rate the reliability of the diagnostic aids from 0%-100% reliable, following the completion of the test trials.

Results indicated that participants disagreed with the aids, even though the aids were more accurate than unaided diagnosis. Half of the participants disagreed with the aids before the aids made mistakes. Wiegmann (2002) surmised again that because the aids were correct several times, the participants expected an error and this might have led to premature disagreement with the aids. Only seven people disagreed with the aid on the first trial, which suggested the users trusted the aids initially. Both maximization and probability-matching utilization strategies were used as well. Participants that agreed with the aids most of the time and optimized their number of correct choices utilized the maximization strategy. Participants who used the probability-matching strategy had lower accuracy scores. Those who adopted maximization strategies seemed to trust aids initially and did not let failures affect them as much. Those who adopted probability-matching strategies seemed to have lower initial trust and adjusted to match the actual aid reliabilities (Wiegmann, 2002). Wiegmann (2002) concluded that differences in users' abilities to accurately calibrate trust in automation exist and may affect the type of strategy that is utilized when interacting with inaccurate automated diagnostic assistance. These strategies may not always result in optimal performance regarding the accuracy of diagnoses.

1.2.2.3 Perception of Computer Advice

Based on the increasing use of databases, decision support systems and knowledge systems and the reference to these and other computer applications as "intelligent," Waern & Ramberg (1996) conducted an experiment to determine possible differences between the perceptions of advice provided by computers versus humans. The research specifically concerned the relationship between the user's knowledge in the domain covered by the advice provided by the computer or human and how that information is perceived. Conclusions were drawn based on participants' ratings of their trust, in general and in relation to the advice, in the computer or person.

Thirty volunteers participated in both a problem-solving situation using a computer presented as an expert system providing advice and a problem-solving situation using paper and pencil and a human for advice. At the end of each situation, participants were asked to complete a questionnaire regarding their problem solving experience in both situations. Question categories included stress, confidence, control, understanding, motivation, efficiency, effort and trust.

Waern & Ramberg (1996) determined that general trust ratings were significantly different between humans and computers. On a scale from 1-10, trust in humans averaged 9.38 while trust in the computer expert average 7.55. However, trust in the computer expert, in relation to particular tasks, was not lower than trust in humans. In addition, whether participating in the computer or human problem-solving situation, correctness of the participant's response seemed to influence the individual's trust and self-confidence ratings. If the participant was correct, self-confidence was always higher than trust and when the problem was not solved, trust in the expert was higher.

One particularly noteworthy finding was that Waern & Ramberg's participants trusted incorrect advice from experts more than they did, if they made an error. Waern & Ramberg interpreted this to mean that the incorrect subjects could not understand that the expert's answer and explanation was incorrect. They emphasized that this could be detrimental in a real life problem solving situation or application in the workplace because individuals could not check a wrong answer if they were incorrect themselves. In addition, it was observed that subjects who thought they were correct reported decreased self-confidence levels when they got incorrect advice from an expert. This was assumed to indicate that an unexpected answer might make individuals question their own competence.

1.2.2.4 Possible Application: DSS for X-ray Passenger Baggage Screening

As computer technology advances, these innovations can be utilized to assist users in a multitude of ways. Young or old, novice or expert, all users may benefit from decision support and automated diagnostic aid technology. Computer technology may be utilized in the health industry to provide computer assisted health instruction, to monitor medication, and the control "intelligent" emergency response systems (Dyck and Smither, 1994). Decision support technology would be of use in all of these applications and may also be utilized specifically for computer aided diagnosis.

Following the tragic events of September 11, 2001, airport security has remained at heightened state to protect airline passengers, crew, and the world at large from any terror threats. A major part of this effort is the enhancement and increase in passenger baggage screening. Screeners all over the world, engage in vigilance monitoring tasks, with X-ray technology similar to that used in the medical field to screen for explosive and harmful objects and materials in travelers luggage. According to the United States' Transportation Safety Administration (TSA), to qualify for employment, baggage screeners' must possess the ability to lift a 50-pound bag, a high school diploma, American citizenship, and the ability to see color. Screeners must be 18 years of age and there is no age limit for employees. Although most have interacted with baggage screeners at some point since September 11, 2001, have many considered who these people are and how qualified and trained they may be to perform this important service. What kind of experience do these individuals have with computers and other technology? Are they competent and aware of what they are attempting to detect in the multitudes of carry on luggage they view day after day? Do the screeners simply rely on their own judgements or is there something more to those X-ray screens that aids in the screeners' decision as to which bags get pulled to be searched and which continue onto the aircraft? Finally, if decision support is provided to screeners monitoring the X-ray machines, is the information and accurate and will screens acknowledge and utilize the information provided?

Operator Assist is an example of an explosives and contraband detection option that was available on some Linescan X-ray security screening systems. It was designed with an algorithm to help the operator by highlighting objects of specific density & area (mass) and atomic number without stopping the belt. A red ellipse is drawn over objects of concern in real time (http://www.bombscan.com). However, as promising as this technology may sound, according to Eric Neiderman, TSA, use of Operator Assist was discontinued due to too many false alarms and errors made by Operator Assist and the subsequent distrust by users (personal communication, June 18, 2003).

1.2.2.5 Compliance

Gardner and Berry (1995) conducted three experiments regarding the introduction of advanced computer systems into the workplace. The three experiments assess the effects of various forms of computer-generated advice on concurrent and subsequent performance of participants controlling a simulated intensive care task. Qualitative advice, or non-specific, and quantitative, specific, advice was presented. Results indicated that users complied at different frequencies with the different kinds of advice being administered. Subjects complied 60% of the time with quantitative advice and 73% of the time with qualitative advice.

In a study conducted by Wise (1999), participants were also presented with qualitative and quantitative advice while controlling a simulated intensive care task. In addition to just assessing compliance rates, Wise also examined individual differences among users including trust, self-confidence and computer anxiety, and gender and past performance of users. As opposed to Gardner and Berry (1995), no significant differences were found in mean levels of compliance with qualitative and quantitative advice. Although, it was observed that ratings of trust, gender, previous performance and self-confidence significantly predicted compliance, while only trust significantly predicted compliance with qualitative advice. Age of the user was not assessed in this study. Wise concluded that DSS users display a great deal of variability in their frequency of use of DSS. This variability may be due to individual differences of the user. Wise's (1999) hypothesis that participants presented with qualitative advice would comply more than participants presented with quantitative or specific advice was disproved similar to Gardner & Berry's research (1995). This may be attributed to temporal differences in advice presentation (Wise, 1999).

1.3 Summary

This literature review demonstrates that although research regarding computer attitudes and age differences has produced mixed results, a majority of the findings indicate that older adults predominantly possess more negative computer attitudes, less experience and less confidence in interactions with computer technology. For example, Ryan, et al. (1992), Czaja & Sharit (1998), and Marquie et al., 2002 determined that age related changes in cognitive skills, abilities, or assumed or even actual unwillingness to learn and use new computer technology may also play a role in making adaptation difficult for older adults. While Hoot and Hayslip (1983) and Jennings and Onwuegbuzie (2001) indicated that assumptions may also be made that older adults are more resistant to and uncomfortable with new technology than younger adults are. Understanding the influence of individual characteristics such as age on computer attitudes is essential to developing effective methods to introduce computers to a wide range of users and to allow for adaptation to new technology and innovations.

Decision support research revealed that some form of environmental or decision support would be useful to moderate age differences related to vigilance task performance and perceived workload (Bunce & Sisa, 2002). They propose introducing intelligent interfaces for monitoring and vigilance tasks where the balance between man and machine control is considered a continuum. As task workload increases, automated systems could moderate the workload. Performance benefits could be observed for older operators if an optimum level of environmental support was provided and maintained around the middle of the active-passive continuum to facilitate task related mental operations. Notably, Bunce & Sisa's (2002) results indicated that with their sample of ages, no age differences existed in the ability to maintain a vigil.

Regarding individual differences and decision support, Czaja & Sharit's (1998) findings emphasized the importance of matching the demands of a particular task with the cognitive skills of the user. If there is a mismatch, users may feel incompetent or incapable and negative attitudes toward the computer and interaction may increase and effect their potential willingness to work with the computer in the future. This is applicable not only to task performance but also to decision support and computer aided diagnosis.

Findings showed that if automated diagnostic aid reliability differs initially by a magnitude of 20%, users will be sensitive (Wiegmann et al., 2001). Higher initial reliabilities were associated with higher agreement rates with the aids, confidence ratings in decisions on agreement trials, and estimates of aid reliability, as well as faster decision time on agreements. Performance over time exhibited user sensitivity to changes in the aid's initial reliability. Wiegmann et al., (2001) also found that all 47 participants underestimated aid reliability in all conditions, reconfirming Wiegmann & Christina's 2000 findings that when not perfectly reliable, users underestimate true aid reliability.

It is essential to develop a better understanding of the factors that affect a users trust in automation that is not 100% reliable. This investigation allows the possibility of designing interfaces to facilitate the users' calibration of their trust in automated aids. This would ultimately improve the performance of whatever system is being operated (Wiegmann et al., 2001). Age is one of these important factors that should be considered.

Regarding compliance with quantitative (specific) and qualitative (non-specific) advice, it was determined that users participating in Gardner and Berry's (1995) experiments to assess effects of DSS on the learning process complied with quantitative advice 60% of the time. They complied with quantitative advice 73% of the time. While Wise (1999) found no significant differences between compliance with qualitative and quantitative advice.

The literature review revealed no recent research related to age and the level of decision support advice accuracy. This study observed these variables to determine their effects on compliance with a DSS during the performance of a task. The task was intended to somewhat resemble a monitoring/vigilance task similar to X-ray passenger baggage screening for explosive objects and materials.

1.4 Statement of Hypothesis

It is expected that older adults will comply less with a computer based decision support assistant known as CAL, based on previous findings related to negative attitudes toward computers and lack of experience with computers & DSS. This may also occur because older adults' vigilance performance may suffer sooner due to less attentional resources. In addition, older workers may have stronger negative reactions to errors in computer work. Younger adults may be more open and familiar with DSS so they should comply more or simply be more open to computer experience. If an employee were to react negatively to decision support while screening baggage the consequences could be severe including frustration or even neglect for the important task they are performing because they have become to flustered to concentrate.

Compliance with CAL is expected to be greatest while interacting with CAL at an accuracy level of 95-100%. Subjects that interact with CAL at a 65-70% inaccuracy level are expected to display lower rates of compliance with CAL. It is also hypothesized that participants that interact with CAL at higher accuracy levels will have a higher compliance rate, possibly regardless of age.

METHOD

2.1 Participants

Forty-eight people participated in this study. Participants were separated into two equal groups based on age groups utilized in previous research: younger adults between the ages of 18 and 40 and older adults that were 41 and older. Actual ages for participants in this study range between the ages of 20 and 40 in the younger adult group and 41 to 69 in the older adults group. The average participant age was 42. The determination of 48 as an adequate sample size is based on previous literature reviewed in preparation for this study.

Selection of subjects was intended to mimic TSA baggage screener job requirements. TSA requires employees to be at least 18 years old, must be able to lift a 50lb bag and there is no top limit for age. Employees must also possess a high school diploma. Because these were the only criterion for selection, participants in this study were volunteers from a variety of locations.

2.2 Apparatus

The present study utilized a laptop computer equipped with Microsoft PowerPoint software. During the computerized portions of the experiment, the laptop monitor remained 5 degrees off vertical and the monitor remained no less than 20 and no greater than 30 inches from the participant's eyes throughout the PowerPoint task.

2.3 Design

This study was a 2x2 between subjects factorial design.

2.3.1 Independent Variable:

Advice accuracy was the first independent variable in this study. It was defined as the percentage of correct choices that are provided to the participant during the decision making task by the simulated decision support assistant known as CAL. Levels of accuracy of the tool utilized were programmed at two levels: 95-100% accurate and 65-70% accurate. Participants were randomly assigned to one of these conditions. When CAL was 95-100% accurate, CAL provided inaccurate decision making advice for 1 out of the 30 decision making trials. When CAL was 65-70% inaccurate, CAL provided inaccurate decision making advice for 20 out of the 30 decision making trials.

Age of the participant was the second independent variable. Subjects were assigned to one of two levels (younger (<40) and older adults (>41)) based on their reported age, before the experiment began.

2.3.3 Dependent Measures:

Compliance with the decision support assistant's advice was the dependent measure utilized in this study. This was measured by calculating the number of times that the participant agreed with CAL's choice, either because the participant and CAL made the same choice or because the participant chose to change their original choice to CAL's choice.

2.4 Procedure

The experimenter greeted participants and asked them to read and sign an informed consent form that included a brief summary of the study and the sequence of events they could expect. Participants were asked their age and whether they possessed a high school diploma. They were also asked to provide an email address or mailing address on this form. These will be utilized following the study to debrief participants on their individual performance and the results of the study if requested.

Participants were asked to take a seat directly in front of the laptop computer. The PowerPoint task was ready for use with the introductory screen visible on the laptop monitor. The laptop keyboard and mouse pad was completely covered with cardboard, except for a space revealing the SPACE BAR. The SPACE BAR was labeled and was the only part of the laptop that the participant used.

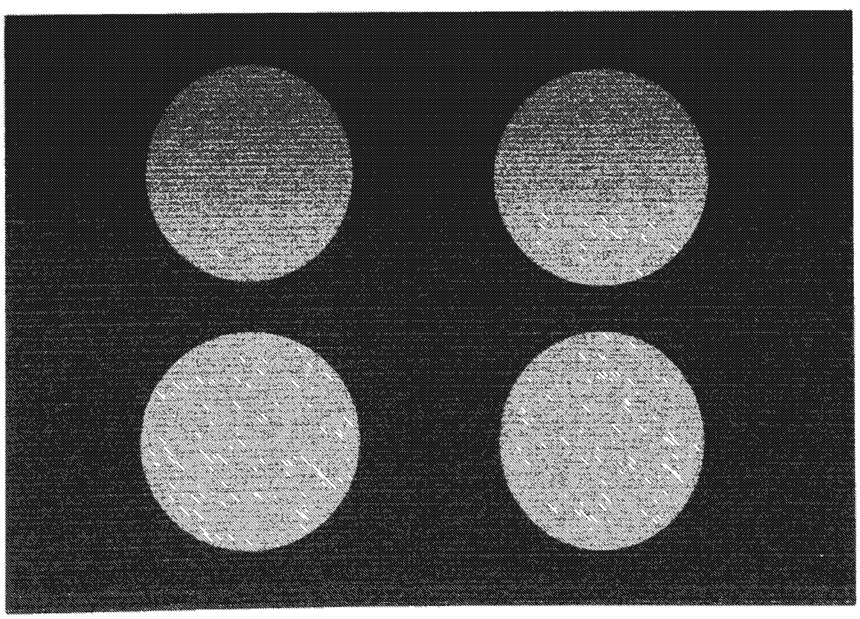
The experimenter reviewed the participant's age and randomly assigned them to one of the two levels of CAL based on their age. The laptop angle was also adjusted at this time. During this time, the participant reviewed the written instructions for the decision making task known as "An Introduction to CAL." The interaction with CAL was described as a quick activity to assess the use of an early phase prototype for a new decision support software package that can be used with Windows applications.

Next, the experimenter asked the participant to review the hand out next to the laptop containing the demographic questionnaire, answer sheets for the practice decision task and decision task, and the final questionnaire. The participant was also provided with a red pen to mark their answers on the answer sheet. They were then asked to ask any questions they had and to begin the practice session on the laptop when they were ready.

The practice decision-making session on the laptop consisted of nine introduction/ instructional PowerPoint slides and five practice decisions. The entire practice and simulated decision making task was conducted in PowerPoint. For each decision (whether during the practice or official decision making task), the participant was presented with a black box that contained four circles that were dark shades of gray that were barely distinguishable from the background. Two of the circles were always the same color and there was a slight difference between the other two, a just noticeable difference. The color of the circles was adjusted via luminance and color controls provided by PowerPoint.

Upon viewing the circles, the participant was asked to select one circle from the four that had the most contrast with the background. Contrast was defined in the instructions as how well a target stands out from its background (Sanders & McCormick, 1993). The location of the circle with the most contrast out of all four circles was randomly selected for each slide. Please refer to Figure 1 on page.

PRACTICE DECISION 1: Your Choice!



Which circle has the most contrast with the background? Indicate your choice in Box A on your answer sheet. Then press the SPACE BAR to advance

Figure 1. Practice Decision Task: The Participant's Choice

Once the participant made their choice, they were instructed to mark the corresponding circle on the answer sheet with the red pen provided. For example, during practice Decision 1, the participant marked a circle in Practice Decision 1 Box A. Please refer to Figure 2 below.

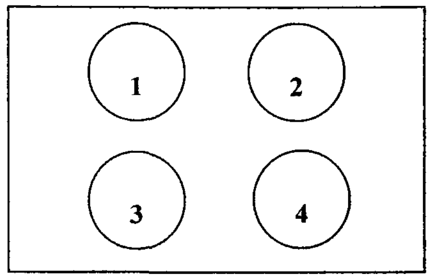
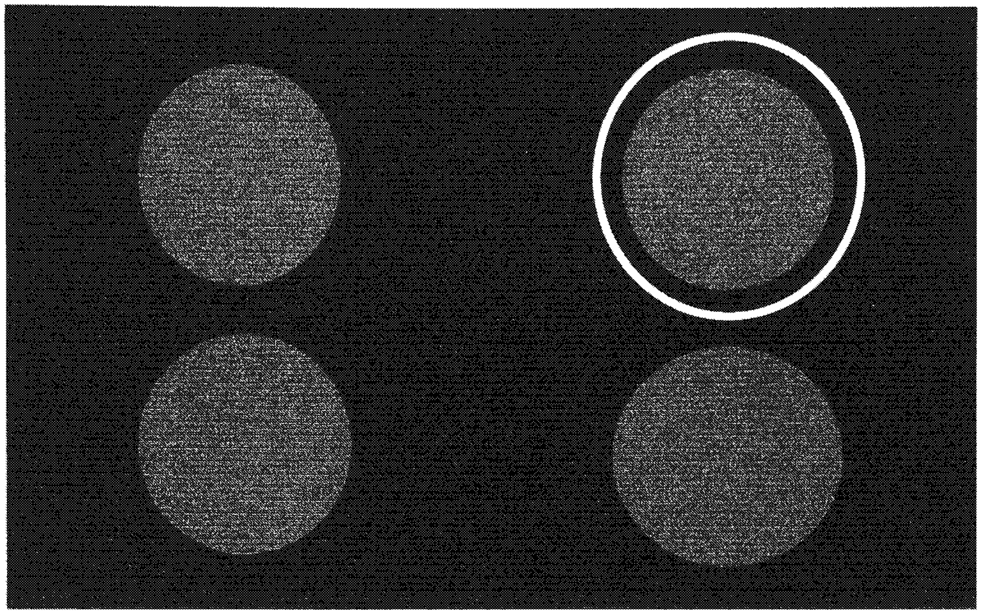


Figure 2. Practice Decision Task Answer Sheet Box A

The next screen on the laptop introduced CAL, the decision making assistant, who provided its own choice for the circle with the most contrast. Please refer to Figure 3 on page 37.

PRACTICE DECISION 1: CAL's Choice!



CAL determined Circle 2 has the most contrast. Indicate your final decision in Box B on your answer sheet. Then press the SPACE BAR to advance.

Figure 3. Practice Decision Task: CAL's Choice

After the participant viewed CAL's choice, they were instructed to return to Practice Decision 1 Box B and respond whether they agreed with CAL or would like to maintain their original choice. Please refer to Figure 4.

___ I agree with CAL. We have chosen the same circle.

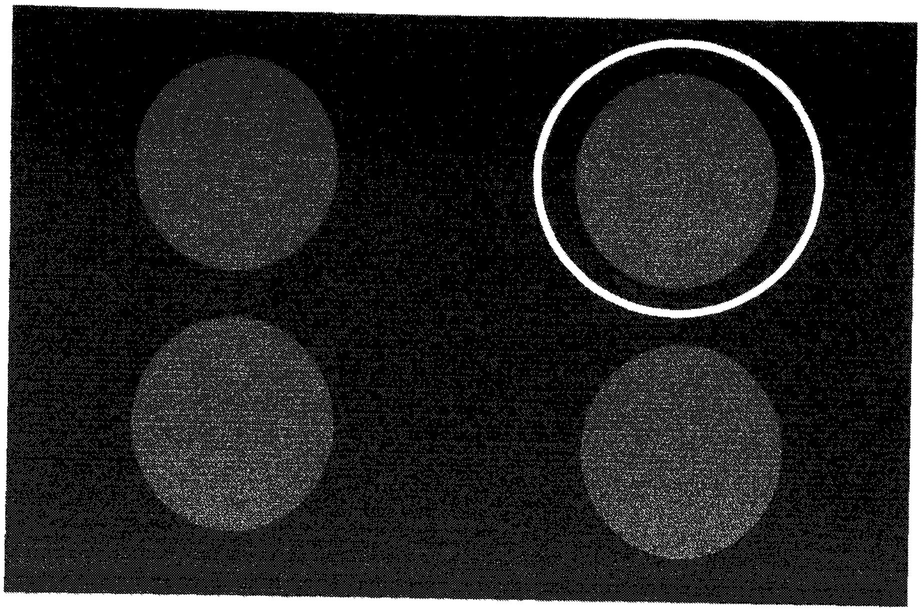
___ I agree with CAL's choice and want to change my decision.

___ I do not agree with CAL and do not want to change my decision.

Figure 4. Practice Decision Task Answer Sheet Box B

Finally, the correct answer was provided to the participant in the next slide. The participant completed 5 practice trials, 4 with CAL's choice being accurate and 1 with CAL inaccurate.

PRACTICE DECISION 1: The Correct Choice!



Circle 2 has the most contrast. Please press the SPACE BAR to advance to your next decision.

Figure 5. Practice Decision Task: Correct Choice

When the practice decisions were complete, the participant was given a chance to ask any final questions (none could be asked during the official decision making task) and then they were instructed to begin the 30 trial decision making task, which was set up exactly like the practice decision trials. The participants were randomly assigned to one of 2 levels of CAL's accuracy – 95-100% accurate and 65-70% accurate. Twelve younger adults were assigned to CAL at 95-100% accuracy and 12 to CAL at 65-70%

inaccuracy. Twelve older adults were assigned to CAL at 95-100% accuracy and 12 to CAL at 65-70% inaccuracy.

After all 30 trials were completed, the participant was asked to complete a final questionnaire at the end of the answer sheet that would be used for informational purposes regarding the subject's experience with CAL. Six questions were presented including modified versions of questions previously used by Wise (1999) and questions created by the experimenter. Please refer to Figure 6.

Final Questionnaire

Please complete the following questionnaire regarding your experience during this experiment:
1) Overall, how high was your self-confidence in the decisions you made?
Very Low Very High
1 2 3 4 5 6 7 8 9 10

2) Overall, how much did you trust CAL's decision making advice?
Very Low Very High
1 2 3 4 5 6 7 8 9 10

3) Please estimate the accuracy level of CAL's advice based on your experience.
Very Low Very High
0% 5% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

4) During the course of the experiment, there may have been times when you did not follow CAL's advice. Please elaborate on why you did not choose to follow CAL's advice in these situations.

5) Did you feel frustrated when CAL did not provide accurate advice? Yes or No

6) Do you feel that accurate decision support assistance may be useful in the workplace? Yes or No

Figure 6. Decision Task Final Questionnaire

Upon completion of the final questionnaire, the experimenter thanked the participant and collected their answer sheet. The participant was reminded that if they had provided an email or mailing address on the informed consent form, their individual results and a summary of the study's results would be sent to them at a later date. However, the individual's score was not actually used for this study, only the compliance score was utilized. The compliance score was determined by calculating the number of times the participant agreed with CAL's choice, either because the participant and CAL made the same choice or because the participant chose to change their original choice to CAL's choice.

RESULTS

A two-way analysis of variance (ANOVA) was calculated for effects of age and CAL's advice accuracy on compliance with CAL's decision support advice. All effects reported as significant in this study met a criterion of $p \leq .05$. A non-significant Levene's Test of Equality of Error Variance revealed homogeneity of variance at .538.

A little more than 85% of the variation in total compliance scores is explained by age, accuracy, and age/accuracy interaction ($R^2 = .856$), but the only significant main effect was for accuracy. Accuracy accounted for 85.5% of the variance ($R^2 = .855$). There was a significant main effect of CAL's advice accuracy on compliance, $F(1,44) = 258.603$, $p < .001$. This means that when age was ignored, CAL's advice accuracy level influenced compliance with CAL. The main effect of CAL's advice accuracy is shown graphically in Figure 7.

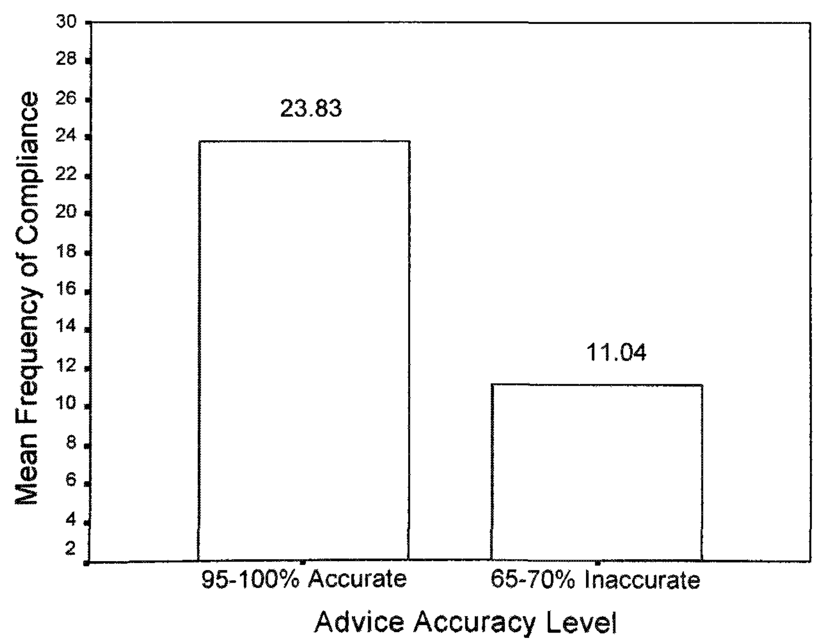


Figure 7. Main Effect of Advice Accuracy Level

There was a non-significant main effect of the age of the participant on the compliance with CAL's advice, $F(1,44) = 1.451$, $p = .235$. This means that overall when CAL's advice accuracy level is ignored, the age of the participant did not influence whether or not they complied with CAL's advice. The main effect of age is shown graphically in Figure 8.

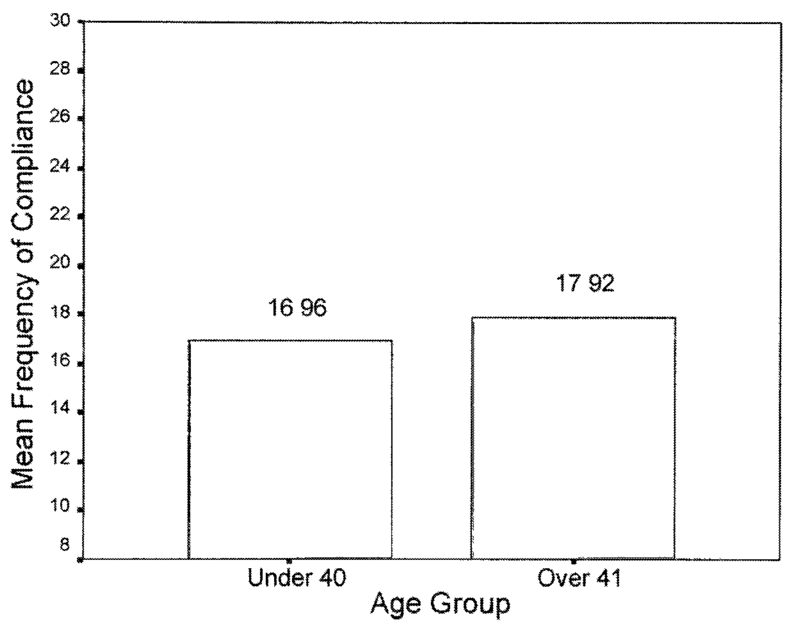


Figure 8. Main Effect of Age

There was a non-significant interaction effect between the age of the participant and the accuracy of CAL's advice, on the participant's compliance with CAL's decision support advice, $F(1,44) = 1.210$, $p = .277$. Compliance performance across age and CAL's advice accuracy is displayed in Figure 9 on page 43. Mean scores and complete ANOVA data for accuracy scores can also be viewed in Tables 5 and 6 in Appendix D.

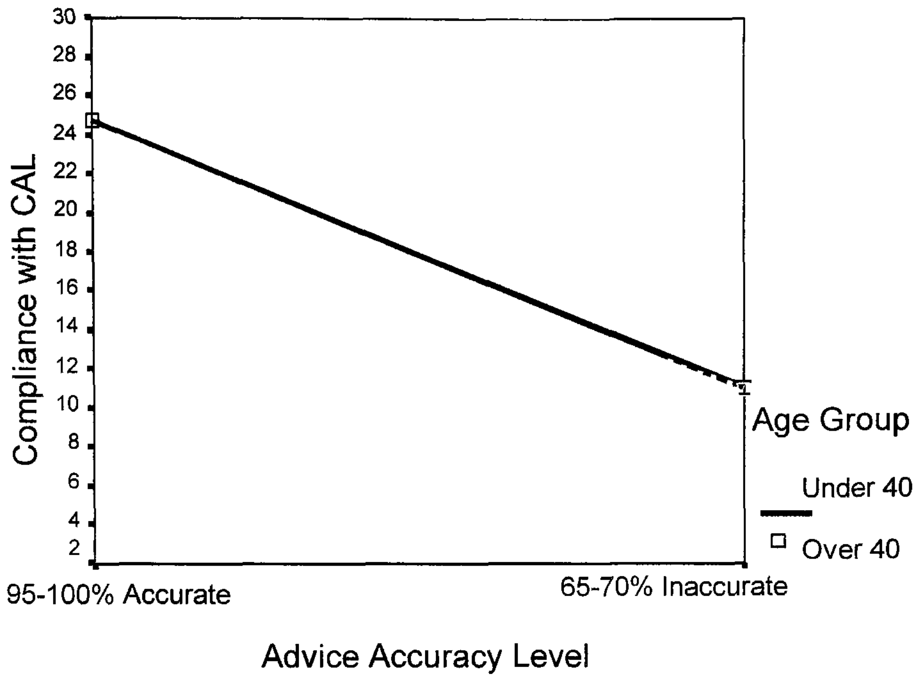


Figure 9: Interaction Effect of Age and Accuracy Level on Compliance

Figure 9 illustrates that regardless of age, as CAL's accuracy level decreased, compliance with CAL's advice decreased for both age groups in this study. The accuracy level effects appear to be very similar for both older and younger adults. The parallel lines indicate non-significance. However, towards the end of the plot, the lines do begin to intersect. This may indicate that if CAL's accuracy level began to decrease even further, a significant interaction between the effect of the participant's age and the accuracy of CAL's advice may have been observed.

Figures 10, 11, 12, and 13 graphically display the percentage of compliance as it relates to specific trials and the occurrence of the provision of inaccurate advice. Tables 10 and 11 assess the percentage of compliance specifically related to younger and older adults. Tables 12 and 13 assess the percentage of compliance related to the provision of

95-100% accurate advice and 65-70% inaccurate advice. The reference lines within all four of the graphs indicate decision trials where CAL provided inaccurate advice.

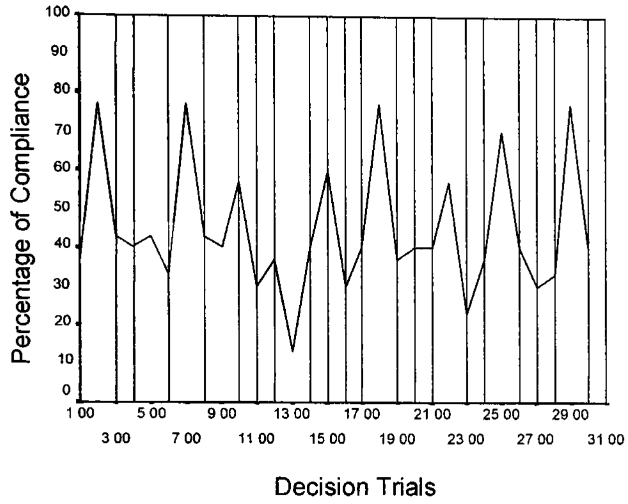


Figure 10. Percentage of Compliance – Younger Adults

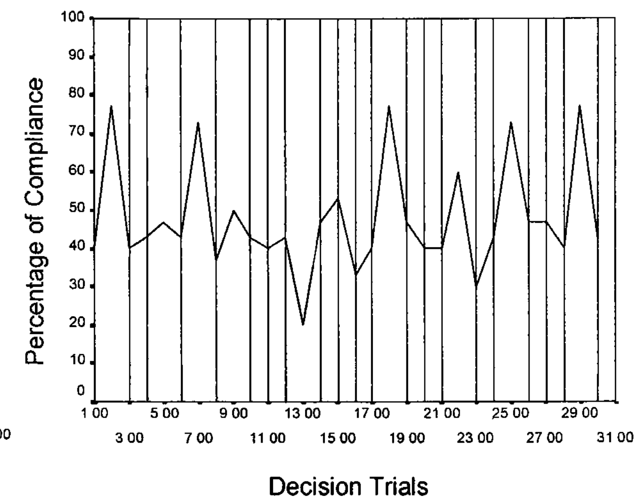


Figure 11. Percentage of Compliance – Older Adults

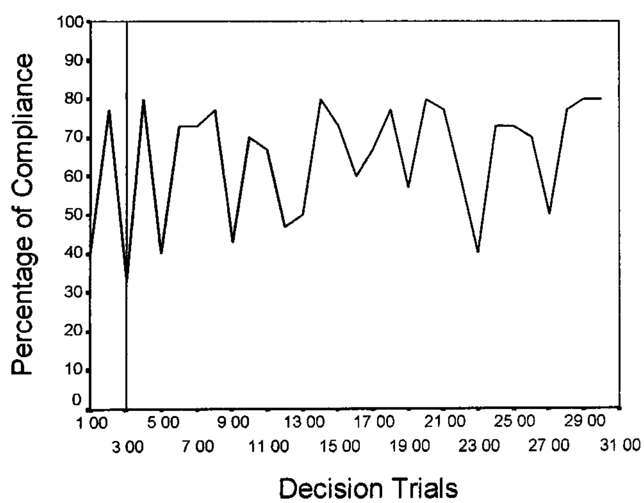


Figure 12. Percentage of Compliance - 95-100% Accuracy

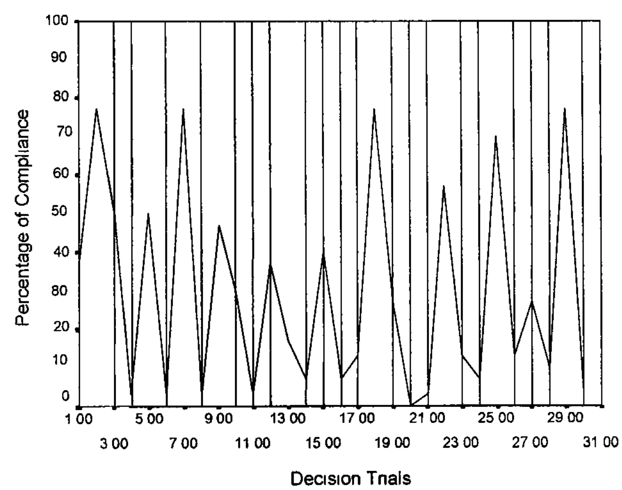


Figure 13. Percentage of Compliance – 65-70% Inaccurate

In addition to the ANOVA data obtained, the final questionnaire following the decision trials revealed no significant differences between younger and older adults' ratings of self-confidence in their own decisions, trust in CAL's advice, or estimates of CAL's accuracy level. Table 1 provides descriptive statistics related to self-confidence, trust, and accuracy estimates. Table 2 provides age-related t-scores for the final questionnaire data.

Table 1. Final Questionnaire Descriptive Statistics Related to Age

	Age	N	Mean	Standard Deviation	Standard Error Mean
Self-confidence	40 and under	24	8.7083	.9546	.1949
	41 +	24	8.2500	1.1887	.2426
Trust in CAL's Advice	40 and under	24	5.7083	1.6280	.3323
	41 +	24	5.3333	2.3713	.4840
CAL's Estimated Accuracy	40 and under	24	55.8333	22.0507	4.5011
	41 +	24	61.2500	24.7268	5.0473

Table 2. Age Related t-scores for Final Questionnaire Data

	t	df	Significance (2-tailed)	Mean Difference
Self-Confidence	1.473	46	.148	.4583
Trust in CAL's Advice	.639	46	.526	.3750
CAL's Estimated Accuracy Level	-.801	46	.427	-5.4167

The final questionnaire also provided data related to frustration with incorrect decision assistance and the possible provision of accurate decision support in the workplace. Five out of 24 participants under 40 and three out of 24 over 41 commented that they were frustrated when CAL provided incorrect decision assistance. Twenty-three out of 24 participants under 40 and all participants over 41 agreed that accurate decision support could be useful in the workplace.

Data from the final questionnaire also indicated no significance between participants' self-confidence when using the 95-100% accurate and 65-70% inaccurate version of CAL. However, trust in CAL's advice at the 95-100% accuracy level (M = 6.67, SD = 1.58) was significantly higher than trust in CAL's advice at the 65-70% inaccuracy level (M = 4.38, SD = 1.76), $t(46)$, $p < .001$. Estimates of CAL's advice accuracy at the 95-100% level (M = 72.08, SD = 21.46) were significantly higher than estimates of CAL's advice accuracy at the 65-70% inaccuracy level (M = 45, SD = 16.41), $t(46)$, $p < .001$. Please refer to Tables 3 and 4 below.

Table 3. Final Questionnaire Descriptive Statistics Related to Advice Accuracy

	Advice Accuracy	N	Mean	Standard Deviation	Standard Error Mean
Self-confidence	95-100% Accurate	24	8.3750	1.0555	.2155
	65-70% Inaccurate	24	8.5833	1.1389	.2325
Trust in CAL's advice	95-100% Accurate	24	6.6667	1.5788	.3223
	65-70% Inaccurate	24	4.3750	1.7647	.3602
CAL's Estimated Accuracy Level	95-100% Accurate	24	72.0833	21.4637	4.3812
	65-70% Inaccurate	24	45.0000	16.4184	3.3514

Table 4. Advice Accuracy Related t-scores for Final Questionnaire Data

	T	df	Significance (2-tailed)	Mean Difference
Self-confidence	-.657	46	.514	-.2083
Trust in CAL's Advice	4.741	46	.000	2.2917
CAL's Estimated Accuracy	4.910	46	.000	27.0833

Two out of 24 participants that used CAL at the 95-100% accuracy were frustrated when CAL provided incorrect advice and all 24 commented that accurate decision support could be useful in the workplace. Six out of the 24 participants that used CAL at the 65-70% inaccuracy level were frustrated when CAL provided incorrect decision support and 23 out of the 24 participants commented that accurate decision support could be useful in the workplace.

DISCUSSION

Previous research findings are mixed regarding older adults and computer attitudes. However, some studies have determined that older adults are more negatively oriented toward computers. Limited research is available related to decision support for monitoring or vigilance tasks and compliance with decision support at varying accuracy levels. Because previous research assessing the variables of age and decision support accuracy is limited, a heavy emphasis was placed on research related to age and computer attitudes when developing hypotheses related to user age and interaction with inaccurate decision support for this study. Previous studies have determined that factors including inexperience, anxiety, and negative attitudes may affect older adults' attitudes toward computers. These factors may interfere with their ability to learn and adjust to new and unfamiliar technology, computer technology in particular (Hoot & Hayslip, 1983; Ryan, et al., 1992; Czaja & Sharit, 1998; Marquie et al., 2002; Jennings & Onwuegbuzie, 2001).

Based on the findings highlighted in the literature review, it was hypothesized that older adults would comply less with CAL due to more negative attitudes toward computers and lack of experience with computers. It was hypothesized that this may also occur because older adults' vigilance performance may suffer sooner due to less attentional resources. Older adults were also expected to comply less due to their stronger negative reactions toward errors in computer work.

The results of this study surprisingly did not support this particular hypothesis. Regardless of advice accuracy level, age did not play a role in whether or not participants complied with decision support assistance in this study. Older adults (M = 17.92, SD =

7.40) complied with CAL slightly more than younger adults (*M* = 16.96, *SD* = 6.75) in both the 95-100% accurate condition and 65-70% inaccurate condition.

The specific occurrence and frequency of accurate and inaccurate advice should be addressed when considering interactions with decision support. These assessments may determine how much inaccurate advice effects compliance based on when and how inaccurate advice is provided. Figures 10, 11, 12, and 13 on page 44 of the Results section illustrate the effect that inaccurate advice for individual decision trials had on the percentage of compliance with CAL's advice for each decision trial. It also illustrates how previous errors effected compliance for subsequent decision trials. When considering compliance related to age in Figures 10 and 11, relatively similar patterns were observed for both younger and older adults. Especially low compliance percentages at trials 13 and 23 occurred with both groups. Overall it is apparent again that older adults complied slightly more than younger adults did. The graphs also show sharp increases in compliance levels between instances of the provision of inaccurate advice and decreases once CAL provided more inaccurate advice. This shows that participants did not necessarily lose all trust in the decision support after the provision of inaccurate advice, but it did cause lower levels of compliance.

Figures 12 and 13 illustrate how CAL's 95-100% accurate and 65-70% inaccurate advice levels effected compliance for specific trials. Differences in percentages of compliance for these two levels of accuracy were obviously different. The 95-100% accuracy graph shows more consistent compliance percentages as compared to the fluctuating and higher rates on the 65-70% accuracy graph. Compliance percentages for the 65-70% accuracy condition ranged all the way from zero to nearly 80%. It is interesting to note that for the 95-100% accuracy graph, the lowest compliance percentage is observed at decision trial 3, where CAL provides its only inaccurate advice for that condition. Compliance for the 95-100% accurate condition never drops below that point again, although it does get close. Percentage of compliance rates for the 65-70% accurate condition drop close to zero following the first error made by CAL at trial number 3. This graph clearly displays how quick participants seemed to be to not comply once inaccurate advice had been provided. The fairly consistent percentages of compliance for the 95-100% condition may be explained simply by the fact that no errors were made again following the one provision of inaccurate advice at trial 3. On the other hand, with 20 out of 30 responses from inaccurate CAL, it is not surprising that the percentages of compliance for the 65-70% condition fluctuated so much. The findings regarding the 65-70% level of inaccuracy supported Wiegmann et al., (2001) and Wiegmann & Christina's (2000) findings that when not perfectly reliable, users underestimate true aid reliability, and obviously compliance will decline. Although both accuracy conditions began to present CAL's inaccurate advice at trial three, in future research, it would be interesting to see how compliance percentages are effected exactly when inaccurate advice begins to be presented.

Subjective data obtained from three 10-point Likert scale questions from the final questionnaire determined that older adults had less self-confidence in their decisions (M = 8.25) than younger adults (M = 8.71), although the difference was not determined to be significant. No significant difference was observed between younger adults' (M = 5.71) and older adults (M = 5.33) trust in CAL's advice. This slightly lower self-confidence in decisions made by older adults while performing a monitoring task with computers is similar to previous findings related to computer attitudes and age. The younger adults' slightly higher rating of trust in decision support advice may also be attributed to previous findings that suggest younger adults are more familiar with computer technology.

Trends in previous research regarding age and computer experience often address older adults' negative attitudes and lower comfort levels with computer technology (Hoot & Hayslip, 1983; Jennings & Onwuegbuzie, 2001). Older adults that participated in these types of studies in 80's are more likely to have had no access to computers, which may explain their greater negative attitudes toward computers. In the twenty first century, it is a bit harder to believe that most people do not have some kind of access to or experience with computers. But this does not mean that there are still some older adults who do not consider themselves "computer people" by any means and were never exposed to computers.

Although the hypothesis that older adults would comply less with CAL was not supported, these findings may be due to older adults' determinations that the computer was automatically more correct. This may be explained by the aforementioned lower ratings of their own self-confidence in decision making and greater advice accuracy estimates. As mentioned in the original hypothesis, younger adults may be more familiar with computers and may be more skeptical and unwilling to always trust the computer's assistance. This may lead to the younger adults' slightly lower estimate of advice accuracy ($M = 55.83\%$) as compared to older adults ($M = 61.25\%$). Lower estimates of advice accuracy may also relate to previous findings that it is uncommon for operators to expect to interact with a perfect automated assistant. In general, if any errors do occur, operators may focus too much on those errors, leading to an underestimation of the aid's reliability (Wiegmann, Rich & Zhang, 2001).

Results of the present study supported the hypothesis that compliance with decision support advice from CAL would be highest with advice at a 95-100% accuracy level. Subjects that interacted with CAL at a 65-70% inaccuracy level would display lower rates of compliance with CAL. As CAL's level of accuracy decreased, compliance decreased for both age groups. The observed significant effect of accuracy level on compliance demonstrated that no matter the age, compliance declined if the computer did not consistently provide correct decision support advice.

Because the task performed in this study was simple, participants may have been more inclined to comply less as accuracy decreased. This may simply be because they were confident in their determination that the computer was making mistakes. However, this may not be the case in more complex tasks such as vigilant monitoring of baggage, where users may not be as quick or confident to doubt the computers' advice or question why some errors are being made. In addition, the consequences of a false alarm or a miss would be more costly with a more complex task such as baggage screening.

The final questionnaire provided some findings regarding the hypothesis that compliance increases with accuracy. The ANOVA data indicated that regardless of age, a significant main effect was observed for accuracy, $F(1,44) = 258.603$, $p < .001$. The final questionnaire data indicated a significant difference was determined between trust in CAL's advice and participant estimates of CAL's accuracy levels while using the 95-100% and 65-70% inaccurate versions of CAL. These subjective results indicate that as the computer became more inaccurate, the users became more confident, trusted the computer advice less, and became more frustrated with the inaccurate decision support. Interestingly, as CAL's accuracy level decreased, the participants' average estimate of CAL's accuracy ($M = 45\%$) was higher than the actual level. The incorrect estimated reliability of the CAL's advice for both conditions is similar to Wiegmann & Christina (2000) and Wiegmann, Rich & Zhang's (2001) findings that when not perfectly reliable, users underestimate true aid reliability. Participants may have felt that because the aid was consistently correct for several trials that an incorrect decision was soon to follow. Participants may have disagreed prematurely before an incorrect decision ever happened (Wiegmann, Rich & Zhang, 2001).

When participants utilized the 95-100% accurate version of CAL, 2 out of 24 participants indicated they were frustrated when CAL provided incorrect advice and all 24 felt that accurate decision support could be useful in the workplace. Six out of the 24 participants that used the 65-70% inaccurate version of CAL indicated they were frustrated when CAL provided incorrect decision support and 23 out of the 24 participants felt that accurate decision support could be useful in the workplace. A majority of the participants that commented that decision support could be useful in the workplace emphasized that the support must be accurate. One particular result of interactions with inaccurate advice, as demonstrated in this study, is frustration. Even when provided with 100% accurate advice, some individuals may become frustrated, which could interfere with concentration, mood and attitude changes, and possibly even errors. Although highly accurate systems may not alleviate all user frustration, it is essential for designers of decision support systems to address frustration and attempt to avoid it with the most accurate support possible.

Open ended, subjective comments were also provided by participants regarding why they did not always comply with CAL throughout the decision making task. Participants that utilized CAL at the 95-100% accuracy level indicated several reasons for noncompliance. Most simply stated they were more confident in their own responses than in the decision support advice. This important finding evokes the question, if users are more confident in their responses or abilities in general, than what is the purpose of decision support or even machines in general? Even at a high level of accuracy, a majority of the participants in the study indicated that they were still more confident in their own responses, as opposed to CAL's. One participant remarked that CAL was incorrect one time so during subsequent trials they did not trust CAL's advice. One participant simply indicated "I am not a computer person," while another stated that they knew the computer could be wrong, so they took a second look at their options and went with what they thought was best. Finally, one participant indicated that they trusted their decisions at first, but then began to trust CAL.

Participants that utilized CAL at the 65-70% inaccuracy level also provided subjective comments regarding why they did not always comply with CAL. Many indicated that they simply did not trust CAL, were more confident in their own responses, and that CAL was frequently and at times, obviously wrong. They therefore often went with their first instincts or impressions, which they considered usually right, so they did not change their answers to match CAL's. All of these comments emphasize the importance of encouraging the acceptance of computer technology, by attempting to increase positive experiences with computers, which may lead to more trust and positive attitudes regarding the systems.

When most participants did not comply with CAL, regardless of accuracy, they generally indicated this was because they had more confidence in their choices and not in CAL's decision advice. So, does this reflect the type of task and conditions that the participants worked under? Or on the other hand, did the users already possess these strong opinions about the reliability of computerized decision support or computers and machines in general? In the case of X-ray machines and baggage screening, X-ray machines, with or without decision support assistance, reveal to screeners what cannot be seen with the naked eye. Based on the participants' subjective comments regarding noncompliance with CAL, does this mean that these individuals' lack of confidence in CAL's decision support would also lead them to doubt the information presented to them by a computerized machine such as an X-ray machine? Whether dealing with X-ray machines, ATMs, or grocery scanners, these types of issues must be considered from the start. Whether people even realize they possess them or not, strong attitudes regarding compliance with non-humans or even other humans may create difficulties when implementing decision support for technology users with varying attitudes and experience levels.

No matter how confident participants may have been in CAL, the data obtained in this study does support the final hypothesis that when participants interacted with CAL at higher accuracy levels, higher compliance rates were observed, regardless of age. Accuracy had a significant effect on compliance. However age and the interaction effect of age and accuracy on compliance did not show significance. Although when viewing the interaction plot for compliance across age and accuracy level, it does appear that an interaction could be possible if CAL's accuracy level were to decrease further. This indicates that if the task was modified, a larger sample size was observed or if age and CAL's accuracy levels were broken down even further, a significant interaction may be observed between age, accuracy, and compliance. Specifically, if a more complex task than the decision making task performed in this study were presented, a different result may be observed and older adults may comply less with decision support, possibly because of their inexperience with complex computer tasks and feelings of inadequacy with computers. In addition, if more levels were added to the independent variables of age and advice accuracy, allowing for post hoc tests and more specific investigations, significance may be determined for the interaction effect of age and advice accuracy.

In regards to the provision of decision support assistance for baggage screeners, this study provides great insight into the effects of inaccurate advice on compliance with both inaccurate and accurate advice. As mentioned in the literature review, TSA previously utilized a baggage screening decision support tool known as Operator Assist. However, Operator Assist is no longer in use due to too many false alarms and misses and the subsequent distrust in the tool by screeners (Eric Neiderman, personal communication, June 18, 2003). The current research demonstrates that as accuracy levels of decision support or "intelligent" assistance begin to decline, user compliance with advice, no matter their age, will decline as well. This reinforces that fact that greater attempts must be made to design accurate decision support applications that will provide correct decision assistance to assure the compliance and use of the tools and most importantly insure the safety of those that are affected by the use of the tool.

RECOMMENDATIONS

The task utilized in this study was not complex, although it was intended to somewhat simulate the more complex task of monitoring X-ray baggage screening machines. Because the task was simple, participants in all age groups may have been more inclined to comply less as accuracy decreased simply because they could be certain the computer was making mistakes. This may not be the case in more complex tasks such as the vigilant monitoring of baggage where screeners may not be as inclined to doubt the computer's advice or question why errors are being made. It may be that the task in this study was not complex enough for participants to begin truly questioning their trust in their own responses or the computers. Older and less experienced participants may have felt more comfortable with the task because it did not directly replicate a more complex screening task. Therefore, age and accuracy did not show signs of significant interaction and effect compliance as expected.

Regardless of how simple the task and conditions may have been, this study provides excellent insight into how individuals aged 20-69 interact and comply with decision support at varying accuracy levels. In order to build upon the information derived from this study, it is suggested that the study be taken to the next level by utilizing a task that more closely replicates an actual baggage screening task. The task could be conducted similar to the present study with a decision support assistant with varying levels of accuracy to determine if the findings from this study carry over to a simulated baggage screening task. Follow on research should assess a larger sample size with additional age levels and levels of accuracy to ensure the possibility of post hoc tests and to increase the possibility of determining additional significant interactions and effects. Age could be grouped in smaller increments of 10 to 20 years (possibly 18-25, 26-35, 36-45, 46-55, etc) and advice accuracy could be broken down into smaller increments, as well (90-100% accurate, 70-80% accurate, 40-50% accurate, etc). Other independent variables could also be assessed including computer experience, trust, and even user personality characteristics, which may shed some light on how a person interacts with non-human assistants.

The current research reflects the importance of designing decision support or "intelligent" assistance that performs as expected and provides assistance to the user. It is human nature, no matter how technologically advanced a system may be, to have some doubt in a computerized system or frustration if errors are consistently or even inconsistently made. With this in mind, other considerations for further research may include assessing how varying levels of advice accuracy of decision support may lead to frustration or possibly excessive workload and how this may affect compliance or task performance in general.

As one participant in this study commented, the "buddy system" is always advantageous and there is value in providing accurate decision support in the workplace. Accurate decision support reinforces confidence and allows decision to be made easier and faster. Future research should be conducted to determine if an accurate decision support system really could be developed and utilized specifically for baggage screeners. Operator Assist failed to provide these screeners with accurate and useful decision support, but hopefully a similar new tool may be designed and successfully implemented. Implementing and successful use of this kind of tool requires evaluating the individual characteristics of the people who will be using these tools. Age is specifically important because it may reflect on individuals' computer attitudes and experience and predict how they will interact with a decision support assistant, especially if it does not always perform optimally. In a profession such as baggage screening, age is an important factor to consider because screeners' ages vary greatly.

Trends in the literature may seem obsolete to some regarding age and negative computer and technology attitudes, due to the increased accessibility of computers. But it is still an important factor to consider especially as it relates to compliance with decision support with varying advice accuracy levels. Therefore it is highly recommended that this research expand by using an actual simulated baggage-screening task. It is hoped that further and more specific conclusions related to baggage screening may be drawn regarding the effects of age, advice accuracy, and possibly even other individual characteristics on compliance with decision support systems.

REFERENCES

Ansley, J. & Erber, J. (1988). Computer interaction: Effects on attitudes and performance in older adults. *Educational Gerontology*, *14,* 107-119.

Birdi, K. & Zapf, D. (1997). Age differences in reactions to errors in computer based work. *Behaviour & Information Technology, 16,* 309-319.

Bonczek, R., Holsapple, C., Whinston, A. (1981). *Foundations of decision support systems.* New York: Academic Press.

Bunce, D. & Sisa, L. (2002). Age differences in perceived workload across a short vigil. *Ergonomics*, *45*, 949-960.

Courtney, J. & Paradice, D. (1993). Studies in managerial problem formulation systems. *Decision Support Systems*, *9,* 413-423.

Czaja, S., & Sharit, J. (1998). Age differences in attitudes towards computers. *Journal of Gerontology: Psychological Sciences*, *55B,* 329-340.

Danowski, J. & Sacks. W. (1980). Computer communication and the elderly. *Experimental Aging Research*, *6,* 125-135.

Dyck, J. & Smither, J. (1994). Age Differences in Computer Anxiety: The Role of Computer Experience, Gender and Education. *Journal of Educational Computing Research*, *10,* 239-248.

Ellis, R. & Allaire, J. (1999). Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety. *Human Factors*, *41,* 345-355.

Gardner, P. & Berry, D. (1995). The effect of different forms of advice on the control of a simulated complex system. *Applied Cognitive Psychology, 9,* S55-S79.

Hoot, J. & Hayslip, B. (1983). Microcomputers and the elderly: New directions for self-sufficiency and life-long learning. *Educational Gerontology*, 9, 493-499.

Jay, G. & Willis, S. (1992). Influence of direct computer experience on older adults' attitudes towards microcomputers. *Journal of Gerontology*, 47:250-257.

Jennings, S. & Onwuegbuzie, A. (2001). Computer attitudes as a function of age, gender, math attitude and developmental status. *Journal of Educational Computing Research*, 25(4): 367-384.

Krauss, I. & Hoyer, W. (1984). Technology and the older person: Age, sex and experience as moderators of attitudes towards computers. In P.K. Robinson, J. Livingston, & J.E. Birren (Eds.). *Aging and technological advances* (pp. 349-350). New York: Plenum Press.

Laguna, K. & Babcock, R. (1997). Computer anxiety in young and older adults: Implications for human-computer interactions in older populations. *Computers in Human Behavior, 13,* 317-326.

Mackie, R. & Wylie, C. (1988). Factors influencing acceptance of computer-based innovations. In M. Helander (Ed.) *Handbook of human-computer interaction* (pp. 1081-1106). New York: Elsevier Publishing Co.

Marcoulides, G. (1988). The relationship between computer anxiety and computer achievement. *Journal of Educational Computing Research, 4,* 151-158.

Marquie, J., Thon, B., & Baracat, B. (1994). Age influence on attitudes of office workers faced with new computerized technology. *Applied Ergonomics*, *25*, 130-142.

Marquie, J. & Huet, N. (2000). Age Differences in feeling-of-knowing and confidence judgments in two semantic knowledge domains. *Psychology and Aging*, *15*, 451-461.

Marquie, J., Jourdan-Boddaert, L. & Huet, N. (2002). Do older adults underestimate their actual computer knowledge. *Behaviour & Information Technology*, *21*, 273-280.

Operator Assist (OA) (n.d.). Retrieved June 23, 2003, from http://www.bombdetection.com/oa.shtml

Parasuraman, R. & Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Human Factors*, *39*, 230-253.

Roth, E., Bennett, K. & Woods, D. (1987). Human interaction with an "intelligent" machine. *International Journal of Man-Machine Studies*, 27, 479-525.

Ryan, E., Szechtman, B., and Bodkin, J. (1992). Attitudes toward younger and older adults learning to use computers. *Journal of Gerontology: Psychological Sciences*, *47*, 96-101.

Sanders, M. & McCormick, E. (1993). *Human Factors in Engineering and Design.* New York: McGraw-Hill.

Shim, J., Warkentin, M., Courtney, J., Power, D., Sharda, R., and Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems*, *33*, 111-126.

Thackray, R. & Touchstone, R. (1981). Age-related differences in complex monitoring performance. FAA-AM-81-12.

Waern, Y. & Ramberg, R. (1996). People's perception of human and computer advice. *Computers in Human Behavior, 12,* 17-27.

Wiegmann, D. & Cristina, F. (2000). Effects of feedback lag variability on the choice of an automated diagnostic aid: A preliminary predictive model. *Theoretical Issues in Ergonomic Science, 1,* 139-156.

Wiegmann, D., Rich, A. & Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. *Theoretical Issues in Ergonomic Science, 2,* 352-367.

Wiegmann, D. (2002). Agreeing with automated diagnostic aids: A study of users' concurrence strategies. *Human Factors, 44,* 44-50.

Wise, M. (1999). Individual Operator Compliance with a Decision-Support System (Doctoral dissertation, University of Central Florida, 1999). Dissertation Abstracts International, 60(7-B): 3603.

APPENDIX A: Informed Consent Form

The experiment in which you are about to participate in is designed to investigate the effects of the provision of decision support during monitoring or vigilance tasks, which require decision-making. Susan Vallance is conducting this study, in conjunction with the Human Factors and Systems Department at Embry-Riddle Aeronautical University in Daytona Beach, Florida to fulfill the requirements of the Master of Human Factors and Systems degree. Ultimately this experiment seeks to provide insight into the benefits of providing airport passenger baggage screeners with decision support assistance while scanning baggage.

You will be presented with a series of decision trials, which display 4 circles on a black background. You will be asked to determine which one of the 4 circles has the most contrast with the background. Contrast is defined as how well the circle stands out from its background. CAL, a simulated decision support assistant will provide its own advice for each decision. Nine introductory slides including detailed instructions for the task and five practice decision-making trials will be presented. You may ask questions during the practice trials but, not during the 30 official decision making trials. During the 30 decisions, you will be asked to make your own choice, CAL's choice will be displayed, and then you will be given an opportunity to change your response, if you feel it is necessary. Following the decision making task, you will be asked to complete a brief questionnaire. The entire experiment should take no longer than 30 minutes.

Any information you provide during this experiment will be held in strict confidence by the researchers. At no time will your name be reported along with your responses. All data will be reported in group form only. A final report of this experiment will be available at Embry-Riddle Aeronautical University's library once the experiment is complete. A summary of findings and your own individual results will also be sent to the email or mailing address that you provide below, once the report is completed.

Your participation in this study is voluntary. You are able to withdraw from the study at any time without penalty. You may also have your data removed from this research at any time.

I acknowledge that I have been informed of, and understand the nature and purpose of this study, and I freely consent to participate. I am also at least 18 years old and possess a high school diploma.

______________________________ __________

Signed Date

Please provide your email address and/or mailing address in the space below:

APPENDIX B: Demographic Questionnaire and Sample Answer Sheet

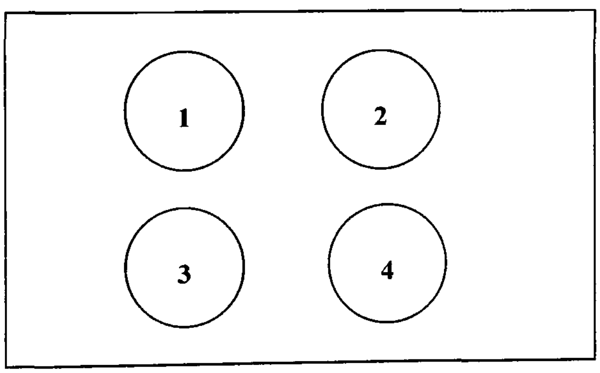
Name ________________________
Age ___
Level of Education (circle one)
High school diploma 1-2 years of college 3-4 years of college Graduate school
How often do you use computers? (circle one)
Every day Every 2-3 days Once a week Once a month Rarely ever
Are you familiar with Artificial Intelligence or Decision Support System Technology? Yes or No

Please review the instructions provided on the computer. Use the SPACE BAR to advance screens. Begin the practice session when you have completed a thorough review of the instructions. Consult the researcher with any questions you may have during the practice decision making task. A printed copy of the instructions is also available to you.

Decision Making Task Answer Sheet

Directions: Box A – Please place an X inside the circle you have chosen.
Box B – Please place an X next to your final decision.

Practice Decision 1
Box A



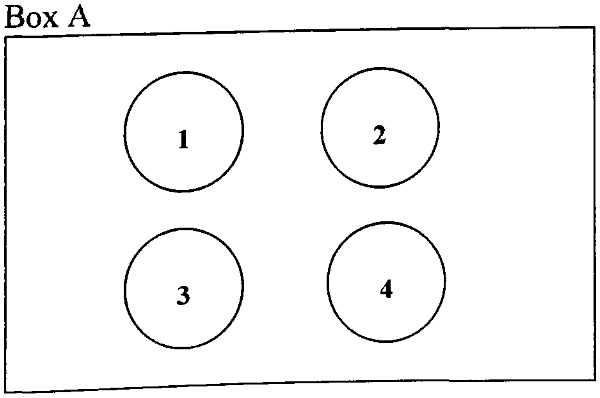
Box B

___ I agree with CAL. We have chosen the same circle.

___ I agree with CAL's choice and want to change my decision.

___ I do not agree with CAL and do not want to change my decision.

Decision 1
Box A



Box B

___ I agree with CAL. We have chosen the same circle.

___ I agree with CAL's choice and want to change my decision.

___ I do not agree with CAL and do not want to change my decision.

APPENDIX C: Instructions

The following task involves deciding which one of 4 circles presented appears to have the most contrast with the background.

You will complete this decision making task with the help of an "intelligent" decision support assistant known as CAL.

CAL has been programmed with an algorithm that can determine the contrast of objects presented to you on the screen.

Throughout this task, you will be expected to make 30 decisions.

Each decision involves determining which circle, out of the 4 presented, has the most contrast with the black background.

Contrast is defined as how well a circle stands out from its background.
The decision making task will proceed as follows:

1) For each of the 30 decisions, a box will be presented to you that contains 4 shaded circles.
2) Please look at all 4 circles and decide which circle appears to have the most contrast with the background.
3) On the paper answer sheet provided, please indicate which circle you have decided shows the most contrast with the background in Box A.
4) After you have written down your decision, please press the SPACE BAR to advance to the next screen.
5) CAL will then provide you with its choice for the circle with the most contrast.
6) Upon viewing CAL's choice, you will now have a chance to change your decision based on CAL's decision.
7) After you have considered CAL's decision, please return to your answer sheet and indicate whether you would like to maintain your decision or change it to match CAL's.
8) As a reminder, there is no guarantee that CAL's decision will be correct.
9) After recording your final decision, press the SPACE BAR to reveal the correct answer.
10) Once the correct choice is revealed, press the SPACE BAR to advance to your next decision.
11) Please continue until all 30 decisions have been made.

As a reminder, you are not required to choose the circle that CAL chooses.

CAL's contrast decision calculations are provided for assistance in making your decision. Remember, this is YOUR decision, whether based on CAL's calculations OR your own.

It is time for a practice session.

If you have any questions, please let your decision task supervisor know at this time.

When you are ready, please press the SPACE BAR to begin 5 practice decision tasks.

APPENDIX D: Tables

Table 5. Compliance Mean Scores by Age and CAL's Accuracy Level

Age Group	CAL's Accuracy Level	Mean	Standard Deviation	N
Under 40	95-100% Accurate	22.92	3.18	12
	65-70% Inaccurate	11.00	2.76	12
	Total	16.96	6.75	24
Over 40	95-100% Accurate	24.75	2.63	12
	65-70% Inaccurate	11.08	2.39	12
	Total	17.92	7.40	24
Total	95-100% Accurate	23.83	3.00	24
	65-70% Inaccurate	11.04	2.53	24
	Total	17.44	7.02	48

Table 6. ANOVA Tests of Between-Subjects Effects – Frequency of Compliance
Independent Variables: Age, Advice Accuracy Level

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared	Noncent. Parameter	Observed Power
Corrected Model	1983.729	3	661.243	87.088	.000	.856	261.264	1.000
Intercept	14595.188	1	14595.188	1922.240	.000	.978	1922.240	1.000
AGE	11.021	1	11.021	1.451	.235	.032	1.451	.218
ACCURACY	1963.521	1	1963.521	258.603	.000	.855	258.603	1.000
AGE * ACCURACY	9.188	1	9.188	1.210	.277	.027	1.210	.190
Error	334.083	44	7.593					
Total	16913.000	48						
Corrected Total	2317.813	47						

a Computed using alpha = .05
b R Squared = .856 (Adjusted R Squared = .846)

Table 7. Final Questionnaire Data

	Self Confidence in Decisions	Trust in CAL's Advice	CAL Accuracy Level Estimate
Under 40	M=8.708	M=5.708	M=55.83%
Over 40	M=8.25	M=5.333	M=61.25
CAL at 95-100%	M=8.375	M=6.67	M=68.333
CAL at 65-70%	M= 8.583	M=4.375	M=45