Holistic versus Decomposed Rating Scales: Which causes higher levels of cognitive load?

by

Alisa H. Watt

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science

In

Industrial & Organizational Psychology

Roseanne J. Foti, Chair

Robert J. Harvey

Helen J. Crawford

Jack W. Finney

May 7, 2003

Blacksburg, VA

Keywords: Decomposed, Holistic, Job Analysis, Cognitive load

Holistic versus Decomposed Rating Scales: Which causes higher levels of cognitive load?

Alisa H. Watt

Abstract

The purpose of this study was to help explore the assumptions in job analysis literature involving whether holistic or decomposed job analysis items lead to a higher level of cognitive load for raters. The main study, involving 303 undergraduate university students, was a 2 (Type of measure: holistic or decomposed) X 2 (Level of extraneous cognitive load: additional load or no additional load) within-subjects design. The 160 decomposed items analyzed in this study were pilot tested to ensure that they would correlate with the 17 holistic items. Under the additional cognitive load condition, participants memorized an 8-digit number, and then were asked to recall and recognize this number upon the completion of the rating task (this manipulation was performed for both the holistic and decomposed measures). Stability of ratings across conditions and interrater agreement were used as dependent measures. Results indicated that the holistic items (r=.74) had higher levels of stability across cognitive load conditions than did the decomposed items (r=.66). The levels of interrater agreement were not significantly different between three of the four conditions. In partial support of Butler and Harvey (1988), the level of interrater agreement for the Holistic additional cognitive load condition (r^*_{wg}=.33) was significantly lower than the interrater agreement for the remaining three conditions. The pattern of results supported prior research (Cornelius & Lyness, 1980; Lyness & Cornelius, 1982) indicating that, depending on the criteria being used, holistic items do not necessarily cause a higher level of cognitive load for raters than do decomposed items.

Acknowledgements

I would like to thank the following people for their support, assistance, and encouragement during the completion of this project:

Douglas Watt

Marguerite Watt

Dr. Roseanne J. Foti

Dr. Robert J. Harvey

Dr. Helen J. Crawford

Dr. Randolph L. Grayson

Dr. Larry D. Moore

Eran Hollander

Kristina A. Meacham

Eugene Kutcher

Holistic versus Decomposed Rating Scales: Which causes higher levels of cognitive load?

Throughout the job analysis literature, researchers have speculated that the success or failure of holistic and decomposed measures has been because of the levels of cognitive load or cognitive demands experienced by the rater at the time of the decision. Fischer (1977) described holistic measures as those methods that require the decision-maker to directly assign over-all values to each outcome, while he described decomposed methods as those that divide the evaluation task into a set of simple subtasks to which the decision-maker must assign values. It has been assumed that the higher the levels of cognitive load being experienced by the rater, the more inaccurate his or her ratings (on either holistic or decomposed scales) then become (Morgeson & Campion, 1997). The speculation that excessive cognitive load caused by the rating scale itself affects the rater's performance level is problematic as it has been used to explain too much. This speculation has been used to explain the success of both holistic and decomposed measures. However, it has also been used to explain the failure of both holistic and decomposed measures. It would seem that no one really knows whether holistic measures or decomposed measures (or both) cause excessive cognitive load for the rater.

Currently, job analysts use decomposed measures (i.e., the Position Analysis Questionnaire, or PAQ) or holistic measures (i.e., questionnaires from the Occupational Information Network, or O*NET) in order to provide information about jobs. As both types of measures have costs associated with them (both monetary and time related), it is important to determine which type of measure results in less cognitive load (therefore making the ratings more accurate and more effective). This study is now examining one aspect of the holistic versus decomposed items debate: cognitive overload. As one recurring theme in the job analysis literature is to blame poor performance of holistic and decomposed measures on cognitive overload, this study is now looking at this aspect of the literature empirically. This study will shed some light on this applied issue with theoretical underpinnings. Once this is known, job analysts will be able to make a more informed decision as to which type of rating scale to use.

Holistic versus Decomposed Measures: Utility

In the past, researchers (Butler & Harvey, 1988; Morera & Budescu, 2001; Ravinder, 1992) have compared the utility of holistic measures versus decomposed measures in a variety of contexts such as decision-making and job analysis.

Decision Making Literature. The principle of "divide and conquer", which suggests that complex decisions should be broken into smaller, more manageable parts, and that the decisions based on these parts should then be aggregated to form an overall decision value (Morera & Budescu, 2001), has been applied to this comparison. In general, the literature (Fischer, 1977; Morera & Budescu, 1998) tends to agree that decomposed measures (which follow the "divide and conquer" principle) are more accurate than holistic measures.

For example, Armstrong, Denniston, and Gordon (1975) had 151 participants respond to questions with known answers. One group responded to a direct (holistic) form of the question (i.e., How many families were living in the United States in 1970?). The other group responded to the decomposed version (What was the population of the U.S. in 1970? How many people were there in the average family living in the U.S. in 1970?). Armstrong et al. found that participants using the decomposed measures, rather than those using the holistic versions, made more correct estimates for 12 out of 13 questions. They also found that answers obtained from decomposed measures were less likely to have large errors.

Ravinder (1992) compared the amount of random error associated with holistic versus decomposed evaluations. One group (N = 14) of participants evaluated hypothetical job offers that were described with three attributes (city, annual salary, and type of work) while the other group (N = 14) evaluated offers described by five attributes (city, annual salary, type of work, annual raises, and job security). Both groups of participants evaluated the offers holistically and through a decomposed function. Two weeks later, the participants evaluated the same offers both holistically and in a decomposed manner (this time the order of the type of measure was reversed). Ravinder concluded that the random error associated with holistic evaluations significantly increased as the number of attributes being considered increased from three to five ($F = 2.85$, $p = .00$). This result was not found for the decomposed evaluations. He also concluded that for both the three-attribute case as well as for the five-attribute case, the random error variance for the decomposed evaluations ($M = .039$ for the three attributes, $M = .024$ for five attributes) was smaller than the random error variance of the holistic evaluations ($M = .054$ for three attributes, $M = .083$ for five attributes) in both cases. When looking at the stability of each measure across time, Ravinder examined the correlations of both measures across the two sessions. He found that the decomposed measure had a higher level of temporal stability ($r = .935$ for three attributes, $r = .970$ for five attributes) than the holistic measure ($r = .861$ for three attributes, $r = .830$ for five attributes). Morera and Budescu (2001) also found that decompositional judgments were more consistent than holistic judgments.

Job Analysis Literature. The comparison of holistic and decomposed ratings is also found in the job analysis literature (Butler & Harvey, 1988; Cornelius & Lyness, 1980; Lyness & Cornelius, 1982, Sanchez & Levine, 1994). Butler and Harvey (1988) compared holistic versus decomposed measures in an attempt to find a way to lower the cost and intrusiveness of the Position Analysis Questionnaire (PAQ) (Mecham, McCormick, & Jeanneret, 1977). Three groups of raters (3 professional job analysts familiar with the PAQ, 9 graduate students familiar with the PAQ, and 63 undergraduates unfamiliar with the PAQ) made holistic ratings of the PAQ dimensions for four jobs found to be familiar to undergraduates in a pilot study (Secretary, homemaker, restaurant food server, and fast-food counter clerk). These holistic ratings were then compared to the professional analysts' decomposed ratings on those same four jobs. Butler and Harvey based their analysis of the quality of the holistic ratings on levels of interrater agreement and levels of convergence between the holistic ratings and the scores of the professionals' decomposed ratings. They found that all three groups of raters had low quality holistic ratings when compared with the ratings made from decomposed measures. For the holistic ratings, the average interrater agreement across jobs was $r = .26$ for the undergraduates, $r = .43$ for the graduate students, and $r = .49$ for the professional raters. This is in comparison to the professional rater's level of interrater agreement ($r = .87$ at the item level and $r = .86$ at the dimension level) for the decomposed ratings. Butler and Harvey speculated that under the holistic judgment condition, each of the three groups of raters experienced cognitive overload and therefore produced lower quality levels of ratings. They then concluded that if this were the case, holistic ratings should not be applied to job analysis measures.

Not all researchers agree, however, that decomposed measures always outperform holistic measures. Cornelius and Lyness (1980) had 115 job incumbents evaluate their own jobs using 13 job analysis scales on two occasions (the first session was between 3 to 8 weeks prior to the second session). They compared three judgment strategies: a decomposed strategy that used an overall algorithm to determine the overall rating, a decomposed strategy that was followed by the participant making an overall holistic estimate, and a holistic strategy. When using interrater agreement as criterion, the results indicated that the decomposed-algorithm strategy (r=.54) and the holistic strategy (r=.39) were equally effective, while the decomposed-holistic strategy

(r=.30) was ineffective. When using intrarater reliability across time as a criterion, the results again indicated that the decomposed-algorithm strategy (r=.82) and the holistic strategy (r=.80) were equally effective, while the decomposed-holistic strategy (r=.64) was ineffective. Cornelius and Lyness speculated that the decomposed-holistic strategy caused the subjects to experience cognitive overload while the other two strategies did not. Cornelius and Lyness did note that the decomposed-algorithm strategy was superior to the holistic strategy for tasks involving high levels of information rather than low levels of information.

Lyness and Cornelius (1982) compared the same three judgment strategies as Cornelius and Lyness (1980). However, Lyness and Cornelius (1982) compared participants' ratings (N = 270) formed from these strategies in terms of intrarater reliability across testing occasions, intrarater convergence across rating methods, and interrater agreement, using both correlations and mean absolute deviations as dependent measures. In this study, the results showed that the decomposed-algorithm strategy was superior to the other two when the mean absolute deviation was the criterion. When evaluating the strategies with this dependent measure, the decomposed-algorithm strategy resulted in better reliability over time, convergence across rating methods, and interrater agreement than the other two strategies. When the standard deviations of the ratings under each strategy were compared, those for the holistic ratings were found to be the largest, while those for the decomposed-algorithm strategy were found to be the smallest. When basing the results on correlations, however, the holistic strategy was as effective as the decomposed-algorithm judgment approach.

Sanchez and Levine (1994) had incumbents of three different jobs (N=86) analyze their jobs using either decomposed or holistic ratings. Half of the participants received inferential training during which the researchers explained that raters are prone to inferential errors.

Strategies were given to these participants to help them to avoid making these types of errors. Interrater agreement, correlation with experts' consensus ratings, and the root mean square deviation from the experts' ratings were used as dependent variables. Sanchez and Levine found that decomposed type items helped improve participants' ratings on those items for jobs where the tasks were less complex than the job as a whole. For jobs where the tasks were more or equally complex as the job as a whole, the decomposed items did not help to improve ratings. For example, for firefighters, interrater agreement was higher when ratings were made on the decomposed instrument (r=.50) as opposed to the holistic (r=.36). However, the opposite was found for mail adjusters and police sergeants. Sanchez and Levine also noted that for highly educated incumbents, the holistic measure produced better ratings than the decomposed measure. In addition, the results from this study suggested that the decomposed items were not as effective as holistic for supervisory or managerial type jobs.

Holistic versus Decomposed Measures: Effectiveness

Depending on the criterion being used, some studies comparing holistic and decomposed measures found that holistic measures are actually more effective than decomposed measures. However, as mentioned earlier, some researchers (Butler & Harvey, 1988; Cornelius & Lyness, 1980) have proposed that ratings based on holistic measures may be inferior because they cause the subject to be overloaded with information. Morgeson and Campion (1997), in a review of studies detailing a number of social and cognitive sources of potential inaccuracy in job analysis, found that people's limited information-processing capabilities could lead to inaccuracy when performing job analyses. As making judgments about an entire job (holistic judgments) is more complex than making judgments about specific tasks (decomposed judgments), one would assume that decomposed judgments would be more accurate than holistic because less recall and integration of information is required.

However, Morgeson and Campion also mentioned that information overload can occur as a result of attempting to analyze large amounts of information or information that is complex. As decomposed measures require judgments about specific tasks, they generally require more judgments to be made in order to arrive at the same overall judgment that a holistic measure would provide. Therefore, as decomposed measures usually require the rater to consider a larger number of items than do holistic measures, it is possible that raters could be experiencing a higher degree of information overload as a result of decomposed measures.

The decision-making literature and the job analysis literature provide strong examples of the debate over which type of measure leads to the rater experiencing a higher degree of cognitive load. In some respects, the debate would seem clear cut and many would assume that ratings formed from decomposed measures are superior to those formed from holistic measures. After all, it has been found that the random error variance of holistic measures is greater than the random error variance of decomposed measures and decomposed measures are found to be both more consistent and more accurate. However, as Morgeson and Campion (1997) pointed out, a conflict in the theories does exist. It can be considered that decomposed measures and holistic measures are both equally ineffective in that they can both cause excessive cognitive load, and that they are both equally effective depending on the criteria on which effectiveness is based. While many studies have speculated that one type of measure causes more cognitive load than the other, this conflict has not yet been resolved empirically. In order to more fully explore these issues, one must turn to literature that addresses cognitive load theory.

Cognitive Load Theory

According to Paas, van Merrienboer, and Adam (1994), the term cognitive load represents the demands that performing a particular task impose on the cognitive system. This multidimensional construct consists of causal factors (the task's environmental characteristics, subject characteristics and the interaction of the two), and assessment factors (mental load, effort, and performance). Task characteristics include structure, time pressure, type of reward system, and novelty. The task's environmental demands include noise and temperature. Subject characteristics include cognitive capabilities, prior knowledge, and style. The subject and the task can interact and such things as the subject's internal motivation, arousal, and the internal criteria of optimal performance can also affect the cognitive system.

Assessment factors include mental load, effort, and performance. Mental load is assumed to be intrinsic within a task. This is the task's difficulty level and is imposed by the task itself or by environmental demands (Paas et al., 1994). Mental effort, on the other hand, is the amount of resources the subject actually dedicates towards a task. Task performance is the third dimension of measurement for cognitive load.

Sweller (1994) considered cognitive load as a multidimensional construct. He suggested that cognitive load is a combination of at least two separate factors: extraneous cognitive load and intrinsic cognitive load (as mentioned in Paas et al., 1994). Instructional method, materials, and additional tasks are all things that can impose extraneous cognitive load. Intrinsic cognitive load, however, is determined by the information itself. That is, the amount of element interactivity, or the amount of elements that must be considered simultaneously in order for the information to make sense, determines the amount of intrinsic cognitive load a piece of information possesses (see Kirschner, 2002; Pollock, Chandler, & Sweller, 2001 for similar explanations).

Intrinsic cognitive load closely matches Paas et al.'s concept of mental load. Extraneous cognitive load is similar to Paas et al.'s concepts of causal factors (the task's environmental characteristics, subject characteristics, and the interaction between the two) and the assessment factor of mental effort. Extraneous cognitive load only interferes in a subject's understanding when there is high intrinsic cognitive load. If a person only has a maximum amount of mental effort that can be utilized at one time, that amount of mental effort must be directed to the critical task as well as any other factor. Therefore, when the difficulty level of the task is high, other factors such as additional tasks, time pressure and the subject's cognitive capabilities have a greater effect on the outcome of task performance as each of these additional factors detract from the amount of mental effort that can be applied to the intrinsic cognitive load of the critical task.

Based on Sweller's (1994) explanation, one would assume that holistic measures would have a higher level of intrinsic cognitive load because more elements must be considered simultaneously in order for this type of judgment to be made. On the other hand, decomposed measures would have a lower level of intrinsic cognitive load because judgments are being formed on a single dimension at a time. However, as mentioned earlier, Morgeson and Campion (1997) stated that information overload (this could be considered to be intrinsic cognitive load) might also be caused by exposure to large amounts of information. Therefore, according to Paas et al.'s (1994) and Sweller's (1994) conceptualizations of cognitive load theory, the performance of an additional task would have a larger negative effect on holistic ratings than it would on decomposed ratings, but according to Morgeson and Campion's (1997) statement, the opposite will be true.

Manipulating Cognitive Load

The preceding argument shows that if a task's level of intrinsic cognitive load is high, then the degree of extraneous cognitive load will have a large effect on task performance. Extraneous cognitive load can be manipulated in a variety of ways. For example, Silvera (2000) examined the effects of cognitive load on strategic self-handicapping. Participants were given a series of practice tests that increased in difficulty. Before being given the actual test, participants were given a choice as to whether they preferred a relatively simple test, matched in difficulty to the most difficult practice test on which they had correctly answered every item, or whether they preferred a more difficult test, matched in difficulty to a practice test on which they incorrectly answered one or two items. Cognitive load was manipulated through time pressures placed on the participant during the time of their choice. Cognitively busy participants were asked to make the decision quickly while participants not under the cognitive load condition were asked to carefully consider their choice. It was found that the cognitively busy participants (61.29%) were more likely than non-busy participants (33.33%) to choose the easy task that has a higher level of success probability.

Van Knippenberg, Dijksterhuis, and Vermeulen (1999) also manipulated cognitive load through time pressure. In their study, they explored the effects of cognitive load and stereotypes on judgments of a defendant's guilt, the severity of punishment, and memory of information of an earlier presentation of a crime. Here, cognitive load was manipulated through the speed at which the materials were presented to the participants. Participants could either read the information at a self-paced rate (by pushing the 'enter' button for more information) or at a fixed rate of 8 s per screen (which was established in pilot testing to be enough time for the information to be scanned very quickly). It was found that cognitive load did not affect estimates of a defendant's guilt when the participant had been presented with a positive stereotype. However, when presented with a negative stereotype, higher estimates of guilt were made when the participant was under the cognitive load condition (M=64.4) than under the no load condition (M=44.6). All proposed prison sentences were higher under the cognitive load condition when the participant was presented with a negative stereotype. Under the negative stereotype condition, higher cognitive load (M=3.60) enhanced memory for incriminating evidence compared to low load (M=2.88). Under the positive stereotype condition, higher cognitive load (M=2.44) led to enhanced memory for exonerating evidence compared to low load (M=1.89).

Pontari and Schlenker (2000) explored the effects of cognitive load on self-presentation. In order to manipulate cognitive load, they provided those participants in the high-load condition with an 8-digit number 15 s before the start of the experimental task. The task required the participant to answer interview questions in such a way as to present a certain type of impression to the interviewer (the participant was asked to portray, through his/her answers, demeanor, etc, the impression that the participant was an extrovert or an introvert). The participants were told to try to memorize the 8-digit number because they would be asked to recall it upon the completion of their task. The participants who correctly recalled the 8-digit number upon completion of the task were said to have maintained a higher level of cognitive load throughout the task than those participants who did not have to perform this additional task. Of the participants, 71% recalled the 8-digit number perfectly. Pontari and Schlenker found that under the cognitive load condition, incongruent self-presentations were less effective than when the participant was not under the cognitive load condition.

Pendry and Macrae (1999) studied the effects of cognitive load on perceived group variability. Participants were asked to form an impression of either a homogenous or heterogeneous group after being presented with 30 traits that were said to describe the group. Participants in the high-load condition were also given 25 s to memorize an 8-digit number before the start of the experimental task; participants were asked to maintain this number throughout the task and recall it once the task was completed. All of the sample recalled the 8-digit number perfectly at the conclusion of the focal task.

Gilbert and Osborne (1989) conducted four studies during which participants were asked to watch a videotape of a female speaking with a stranger. Half of the participants were told that the female was discussing anxiety-inducing topics, while the other half were told that the female was discussing commonplace topics. After the video, participants were asked to rate the female's trait anxiety. Those participants in the cognitive load manipulation were presented with an 8-digit number for 25 s prior to watching the video. After the rating task, the participants then recalled the 8-digit number. All of the sample recalled the number perfectly. The results indicated that while under the cognitive load condition, participants did not use contextual information to correct the stereotypes they had formed of the female whom they were observing.

Blessum, Lord, and Sia (1998) had participants read two descriptions of and form opinions of two men, both homosexual, but one more atypical than the other. Under the cognitive load manipulation, participants were asked to memorize the 8-digit number prior to reading the description of the man and were asked to recall that number once they had finished reading the description. Participants were then asked to check off activities that they would be willing to perform in order to help each of the men. Under control conditions, participants were less willing to help an atypical homosexual male than a typical homosexual male. When participants were under the cognitive load condition, this effect did not occur, indicating that the participants had relied more on heuristics than on the actual descriptions when forming impressions of the male. Of the participants, 94% correctly recalled 6 digits of the 8-digit number.

Harris and Perkins (1995) asked participants to perform the NASA Moon Landing Survival Task. This task requires participants to rank a list of 15 items in their importance to the survival of a person who had to travel 200 miles on the moon to a rendezvous point. Once the participants had completed this task alone, they were then given false information about their future partner (the participants were either told that their future partner had a high or low grade point average). Participants in the cognitive load manipulation were then given the 8-digit number for 25 s to memorize. Those participants then rehearsed that number throughout their interaction with the partner (during this interaction, the two participants completed the NASA task once more, as a team. Once the NASA task was completed, those in the cognitive load manipulation were asked to recall the 8-digit number and to recognize that number out of a list of 10 similar numbers. Participants were then asked about the interaction with and impressions of their partner. Of the participants, 56.6% correctly recalled the 8-digit number, while 100% correctly recognized the 8-digit number. In a later paper, Georgesen and Harris used a list of 8 distracters that were more similar than the ones used in Harris and Perkins (1995). The recognition rates were just as high with the more difficult distracters as they were using the original set of numbers (M. Harris, personal communication, March 10, 2003). The task of memorizing and later recalling an 8-digit number has effectively been used as a way to manipulate extraneous cognitive load in studies where heuristics, stereotypes, and impression formation are being studied. In the past, researchers have used this task when testing undergraduate students (presumably with a range of different IQ values). Based on past literature, the recall task, when used with the recognition task, effectively enables the participants to be under an additional source of cognitive load, while also enabling a higher inclusion rate of participants in analyses. The experimenter can be sure that the rater has maintained a level of extraneous cognitive load by checking the accuracy of the recalled number and the correctness of the recognized number upon completion of the critical task.

Effect of Cognitive Load: The Accuracy Criterion Debate

To test the effect of cognitive load, researchers often examine a participant's score on a test under varying levels of cognitive load and judge that score based on the participant's level of accuracy. Kruglanski (1989) discussed the problem of operationalizing accuracy within social perception and cognition literature. Accuracy can be defined as a correspondence between a judgment and a criterion. This would correspond to Pollock et al.'s (2001) study where participants answered written and practical tests in addition to the cognitive load measure. The scores on the written and practical tests were compared to real scores; the answers were either correct or incorrect.

Kruglanski (1989) also defined accuracy as a consensus, or interpersonal agreement between judges. This is the definition of accuracy that is often used in the job analysis literature (Butler & Harvey, 1988). A participant's judgments are considered accurate if they agree with the judgments of other participants. When participants are rating their own job, however, there is neither a true, real score; nor is there necessarily agreement between judges (as participants who have the same title might have different focuses within their roles). For example, if two secretaries from two different companies are rating the amount of written comprehension that is necessary in order to perform their duties, the ratings may be different. One Secretary might be required to integrate large amounts of written information in order to create memos, while the other's duties might involve more dictation, etc.

In the realm of job analysis literature, there is much debate as to what qualifies as an acceptable accuracy criterion. This debate is exemplified through a series of recent articles (Harvey & Wilson, 2000; Morgeson & Campion, 2000; Sanchez & Levine, 2000). Sanchez and Levine (2000) posited that accuracy based on the convergence with a "true score" (similar to Kruglanski, 1989 and Pollock, et al. 2001) is not practical in job analysis. In this regard, they stated that, if a "proxy criterion itself is conjectural and therefore subject to revision, the same judgment that previously deviated from the criterion could later concur with the revised criterion and thus be deemed correct" (p.813). Sanchez and Levine instead believe that the consequences of judgment (the inferences one is able to make based on the job analysis) should be used in order to determine the quality of a judgment.

Morgeson and Campion (2000) detailed several current conceptualizations of acceptable accuracy criteria. Among these, interrater agreement and convergence with a true score were mentioned. Morgeson and Campion mentioned interrater agreement as a part of a multidimensional consideration of job analysis accuracy. As one aspect of this multidimensional idea, interrater agreement can help to establish a level of accuracy, but as it is only one index in a multidimensional idea, interrater agreement can also fail to establish accuracy on its own. In regards to using convergence with a "true score" as an accuracy criterion, Morgeson and Campion agreed with Sanchez and Levine in that a "true score" does not necessarily exist in job analysis. As Morgeson and Campion stated, "if jobs change over time, then the true score model would not appear to apply" (p. 822).

Harvey and Wilson (2000) disagreed with many aspects of both Sanchez and Levine's arguments and with Morgeson and Campion's arguments. Harvey and Wilson believe that there is a "potential to achieve a gold standard of accuracy" (p.838). They posited that the accuracy of job analysis ratings can be verified by reviewing position ratings made by independent, job-knowledgeable subject matter experts. They contend that this type of criterion should be used in order to determine accuracy of job analysis ratings, rather than to follow Sanchez and Levine's argument of using consequential validity as an accuracy criterion.

The accuracy criterion debate might seem to present a problem when attempting to determine whether holistic or decomposed measures cause the same or differing amounts of cognitive load for a rater. As will be pointed out again later, however, the current study is not attempting to demonstrate whether the ratings made by the raters are accurate representations of a certain position. Rather, the current study is attempting to evaluate the differences between the types of ratings made on holistic and decomposed type items when a rater is and is not experiencing an additional source of cognitive load.

Present Study

The present study followed some of the same procedures as were performed in the work of Butler and Harvey (1988) and Gibson (2001). Butler and Harvey compared the levels of interrater agreement within three groups of participants' holistic ratings on the PAQ (PAQ unfamiliar undergraduates, PAQ familiar graduate students, and professional job analysts who were familiar with the PAQ). The interrater agreements within the undergraduate group, within the graduate group, and within the professional group were all compared to one another.

Butler and Harvey concluded that the holistic ratings were inferior to decomposed ratings because the holistic measures led to low levels of interrater agreement for each group and that the holistic ratings for each group had low amounts of convergence with the professionals' decomposed ratings of the PAQ. Butler and Harvey speculated that these results were due to the holistic ratings causing the participants to experience higher levels of cognitive load.

There are some issues with this speculation, however. First, the studies of Cornelius and Lyness (1980) and Lyness and Cornelius (1982) point out that the criterion on which the judgment is based can affect whether decomposed or holistic measures are seen as equally effective or if one is seen as superior to the other.

Second, it has been assumed by the literature that decomposed measures provide the most accurate judgments. For example, this has been assumed in the work of Butler and Harvey (1988), as the three groups of participants were only required to make holistic judgments based on the PAQ. These judgments were then compared to expert ratings on the decomposed version of the instrument in order to judge whether holistic ratings were indeed more effective. In addition, Gibson (2001) had three groups of participants (job incumbents, graduate students and professors in Industrial/Organizational Psychology, and job analysis-naïve undergraduates) rate holistic items from the O*NET GWA questionnaire and decomposed items from the CMQ. Accuracy was determined by comparing the ratings on the holistic and decomposed items to an external criterion set of job analysis data. This external criterion set consisted of ratings from the CMQ.

However, as Morgeson and Campion (1997) pointed out, decomposed measures can lead to cognitive overload because a larger number of judgments about a variety of tasks must be made in order to judge the same overall category found within a holistic scale. Therefore, convergence with decomposed scores may not be the best criterion on which to judge the effectiveness of holistic judgments, as the decomposed ratings may also be flawed due to cognitive overload.

The present study proposed to examine the assumptions that either holistic judgments or decomposed judgments lead to higher levels of cognitive load. Research (Pendry & Macrae, 1999; Pontari & Schlenker, 2000) has shown that while experiencing a low level of cognitive load; it is possible to engage in more than one task simultaneously without the additional tasks affecting performance levels of the critical task. However, if a high level of cognitive load is being experienced due to a particular task, then the addition of another task will lead to an altered performance on both tasks.

This study had participants make both holistic and decomposed ratings on the position of Secretary. These tasks were done with and without the presence of an additional task (this additional task will increase the participants' level of extraneous cognitive load). The holistic ratings made while performing no other task (HolNo) were compared to the holistic ratings made while performing an additional task (HolL). The same was done for the decomposed judgments (DecNo and DecL). If the judgments made when performing the additional tasks were not consistent with the judgments made while performing the decision task alone, the degree of inconsistency would show the extent to which that type of judgment causes cognitive overload.

As mentioned earlier, the purpose of this study was not to determine the accuracy level of the participants' ratings. Instead, the purpose of this paper was to determine the degree of inconsistency in ratings when the raters were using different types of items and are under different levels of cognitive load. Therefore, the criteria used in this study were interrater agreement and levels of stability across cognitive load conditions. The differences in the degree to which ratings were consistent and inconsistent within an undergraduate sample, a graduate student sample, and an incumbent sample (Secretaries) will be explored. Rather than solely utilize a criteria that was based on decomposed items, holistic and decomposed ratings were compared across these types of raters. This unique contribution to the literature is in an attempt to give the holistic items a "fighting chance" at being found more effective. In this manner, the following hypotheses were proposed:

In order to examine the propositions made by Morgeson and Campion (1997): that information overload can occur when the amount of information to be considered is very large (as in the case of decomposed measures); and that information overload can occur when the information to be considered is very complex (as in the case of holistic measures); the following two hypotheses (1A-1B) were proposed.

Hypothesis 1A. The level of stability between holistic ratings made alone (HolNo) and decomposed ratings made alone (DecNo) will be higher than the level of stability between HolNo and holistic ratings made while under an additional cognitive load task (HolL).

Hypothesis 1B. The level stability between HolNo and DecNo will be different from the level of stability between DecNo and decomposed ratings made while under an additional cognitive load task (DecL).

The logic of the preceding hypotheses can be explained through the following scenario. If, for example, holistic ratings actually do cause more cognitive load than decomposed measures, then (1A) the stability between HolNo and DecNo would be greater than the stability between HolNo and HolL; (1B) the stability between DecNo and DecL would be greater than HolNo and DecNo.

Hypothesis 2. Ravinder (1992) found that decomposed measures provide more stable evaluations across time than holistic measures. If holistic measures are less apt to have temporal stability, it can be hypothesized that the addition of extraneous cognitive load will be more likely to have a larger negative effect on holistic measures than on decomposed measures. Therefore, it is hypothesized that the level of stability will be lower between HolNo and HolL than between the decomposed judgments.

Hypothesis 3A. Based on Butler and Harvey (1988), it was hypothesized that the average interrater agreement will be lower for HolNo than for DecNo.

Hypothesis 3B. The average interrater agreement will be lower for HolL than for DecL.

Hypothesis 3C. The average interrater agreements will be lower for the HolL than for the HolNo and DecNo conditions.

Hypothesis 3D. The average interrater agreement will be lower for the DecL condition than for the DecNo condition and the average interrater agreement for the DecL condition will be different than the agreement for the HolNo condition.

Method

Participants

Based on a power analysis, it was determined that 250 participants would be necessary in order to have a significant amount of power to detect a small effect size. Data was collected from 314 undergraduate students obtained from the psychology subject pool and an introductory computer science course at Virginia Tech. A total of 11 undergraduate students were excluded from the analyses because they did not successfully perform the cognitive load manipulation check. Thus, a total of 303 undergraduates were included in the analyses. The participants received extra credit points in exchange for their participation.

In the past, studies involving cognitive load (Gilbert & Osborne, 1989; Harris & Perkins, 1995; Pendry & Macrae, 1999; Pontari & Schlenker, 2000) as well as studies involving job analysis instruments (Butler & Harvey, 1988; Lyness & Cornelius, 1982) have used undergraduate students as participants. While undergraduates were naïve to the job analysis instruments utilized here, this study examined the effect of cognitive load on ratings, not at how accurate those ratings were in comparison to professional job analysts. Therefore, consistent with Sanchez & Levine (2000) and Morgeson & Campion (2000), the levels of interrater agreement and stability across cognitive load conditions among the undergraduate sample are acceptable criteria to demonstrate how cognitive load affects the ratings made on holistic and decomposed measures.

A problem occurred during data collection where the last 80 responses of the two decomposed sections were not recorded for 253 of the undergraduates. Therefore, a sample of 50 undergraduates (included in the total N=303) was collected with all of the data being recorded in all of the sections.

In addition to the undergraduate sample, 10 graduate students (ranging in experience as graduate students from 1 year to 4 years) in Industrial/ Organizational Psychology and 8 people in the position of Secretary (at Virginia Tech) also participated in this study.[1] Similarly to Butler and Harvey (1988), the graduate student sample was compared to the results of the undergraduate sample. There were no problems during the data collection of these samples.

Research design

The study was a 2 (Type of measure: holistic or decomposed) X 2 (Level of extraneous cognitive load: additional load or no additional load) within-subjects design.

Dependent Variables.

Interrater agreement and stability across cognitive load and measure type conditions were used to examine the changes that occurred across the different types of measures and different amounts of load.

Cognitive Load Manipulation

For the cognitive load independent variable, an 8-digit number was presented to the participant for 25 s (as done in Pendry & Macrae, 1999). Following their performance on the rating task (holistic or decomposed measure), the participants were asked to recall the 8-digit number. After they had recalled the number, they were presented with 10 different 8-digit numbers from which to choose the one they had memorized. The numbers used during the first presentation of the cognitive load manipulation were the same as those used by Harris and Perkins (1995). A variation of those numbers was used during the second presentation of this manipulation.

Measure-Type Manipulation

The Occupational Information Network (O*NET) has replaced the Dictionary of Occupational Titles (DOT). As the DOT's replacement, the O*NET is affecting job analysis in many ways. This new system uses a variety of different holistic scales in order to collect job analysis data from incumbents and store this information within its electronic system (Peterson et al., 2001). For this study, 33 items taken from both the O*NET GWA and Knowledge questionnaires were used as the holistic measure.

Pilot testing: Creating the decomposed measure. To test the hypotheses, it was necessary to find decomposed and holistic items that were highly correlated. For example, Gibson (2001) used the O*NET GWA questionnaire as the holistic measure and the Common Metric Questionnaire, or CMQ (Harvey, 1993), as the decomposed measure. However, items on these two scales are not correlated very strongly. Gibson's (2001) correlations between dimensions on the O*NET GWA questionnaire and the CMQ ranged from -.099 to .160. As the hypotheses of this study depended on the holistic and decomposed items used being highly correlated, holistic items were taken from the O*NET GWA and Knowledge questionnaires, while decomposed measures were developed from the items taken from those questionnaires. The decomposed items were developed from the Handbook of Human Abilities: Definitions, Measurements, and Job Task Requirements (Fleishman & Reilly, 1992), task inventories from various positions contained within the O*NET database, and the Job Element Inventory (JEI; Cornelius & Hakel, 1978).

In order to be sure that the holistic and decomposed items were highly correlated, a pilot test was conducted. Participants (N=50) in the pilot test received various O*NET holistic items as well as a number of related decomposed items. The participants matched the holistic items to the decomposed items. In a way similar to how the BARS format was developed, an 80% rule was utilized. That is, 80% of the sample had to agree that a certain decomposed item matched a holistic item in order for the items to be used in this study.

The experimental study consisted of 33 holistic (see Appendix A) and 242 decomposed items (see Appendix B). Of those, 17 holistic items and 160 decomposed items were actually used in the analyses. The starred items in Appendix A and Appendix B indicate those items used in the analyses. The remaining items were included as distracters. Figure 1 depicts an example of the Likert type scale (ranging from 1 = Not important to 5 = Extremely important) that participants used to rate both the holistic and decomposed items.

Reading Span and Digit Span.

Participants performed both a modified reading span task (Daneman & Carpenter, 1980) and a modified digit span task. The reading span task consisted of presenting participants with a series of sentences, one sentence at a time. Participants were asked to read each sentence quietly to themselves and then to hit the continue button at the bottom of the screen to move on to the next screen (only one sentence was presented on each screen). At the end of each series of sentences, participants were asked to recall the last word of each sentence. When the participants were asked to recall the last word of each sentence, a screen was presented that had the same number of blocks as there were words to recall. Participants typed the words into the blocks, one word per block. Participants were presented with 5 sets each of 2 sentences, 3 sentences, 4 sentences, and 5 sentences. Participants were then presented with 3 sets of 6 sentences. In total, the participants were presented with 88 sentences. In order to compute reading span, the number of correct responses for each sentence set was calculated. To determine their reading span level, participants had to correctly recall all of the words in 3 out of 5 sets for each series of sentences. For example, in order to have a reading span of 4, the participant must have correctly recalled the words in 3 out of 5 sets where 2 sentences were presented, 3 out of 5 sets where 3 sentences were presented, and 3 out of 5 sets where 4 sentences were presented. Possible reading span scores ranged from 2 to 6.

This modified version of the reading span task differs from Daneman and Carpenter's task, in that the task used here was presented online on the computer. Daneman and Carpenter's task is presented to one participant at a time. The participant was presented with index cards, on which the sentences were written. Participants were presented with one sentence at a time, and were asked to read each sentence out loud. At the completion of each sentence set, the participant was presented with a blank card. At that time, the participant was asked to recall the last word of each sentence. No cues were provided to the participant indicating the number of words that should be recalled at the end of each set.

Daneman and Carpenter (1980) found that their reading span task correlated with verbal SAT scores ($r = .59$, $p<.01$). In particular, the reading span task was correlated strongly with those items in the verbal SAT section used to test comprehension – specifically those items involving fact questions ($r = .72$, $p<.01$) and those items involving pronominal reference questions ($r = .90$, $p<.01$). As the reading span task used in this study was a modified version of Daneman and Carpenter's task, it can not be assumed that the reading spans computed on this sample would correlate with verbal SAT scores.

Digit span was measured through the use of both a forward and a backward digit span task that was modified in presentation format from the WAIS-III. The forward digit span task required participants to recall a series of numbers in the order they were read, while the backward digit span task required participants to recall a series of numbers in the opposite order in which they were read. Participants listened to each series of numbers as they were read aloud by the experimenter, 1 s per digit. Once the experimenter finished reading the set of numbers, the participants then typed the numbers into boxes provided on the screen. The number of boxes on the screen corresponded with the number of digits the participant was being asked to recall. Possible total digit span scores ranged from 0 to 30.

These modified digit span tasks differed from the WAIS-III digit span tasks in the following ways. As a part of the WAIS-III, the digits are read out loud by an experimenter. The participant then recalls the number out loud. Cues are not given to the participant as to how many numbers should be recalled at any time. The WAIS-III correlates .88 with the Stanford-Binet-IV. The test-retest reliability for the forward and backward digit span tasks is in the .70s

(Conoley, Impara, & Murphy, 1995). Again, as the digit span tasks utilized in this study were modified versions of the WAIS-III digit span tasks, one cannot assume that the results for this sample would correlate as well with the Stanford-Binet-IV. Also, one cannot assume that the results for this sample would have a similar test-retest reliability as the WAIS-III task.

For the undergraduate sample, the means (with standard deviations in parentheses) for the reading span task and the total digit span task were 3.78 (1.27) and 18.82 (4.50) respectively. The correlation between the reading span scores and total digit span scores was .14 (p=.05). For the graduate sample, the means (with standard deviations in parentheses) for the reading span task and the total digit span task were 3.85 (.94) and 22.70 (4.11) respectively. The correlation between the reading span scores and total digit span scores was .50 (p=.05).

Instrument Presentation.

Following the logic of Gibson (2001), in both conditions (no additional cognitive load and additional cognitive load), participants rated the holistic items first. Specifically, this was to prevent the decomposed items from causing priming effects. In all cases, participants received the no additional cognitive load condition followed by the additional cognitive load condition.

The holistic and decomposed items in the two load conditions were the same, but they were presented in a different order for each condition. It was assumed that the participants would not remember their responses to each individual item from time one to time two for various reasons. First, each time the participant encountered the holistic and decomposed item sets; the individual items were in a different order. The participants were not asked to remember their ratings of each of the items as they performed the first sets of the holistic and decomposed tasks. As they were not asked to remember these ratings, it is not likely that the participants would make an effort to rehearse or maintain this information in memory. Without rehearsal, information in short-term memory will decay (Reitman, 1974). Also, participants were not given the time, nor were they requested to rehearse their responses to any of the holistic and decomposed items. Of the 33 holistic items, only 17 of these items were actually included in the analyses. Of the 242 decomposed items, only 160 of these items were included in the analyses. The remaining items were used as distracters. It is assumed that because these remaining items were less related to the position of Secretary, they were more salient for the raters. If a participant were to remember his/her response to an individual item, he/she would be more likely to remember their response to a distracter item, rather than a focal item.

Second, proactive interference was expected to occur. That is, once the participant is rating the decomposed task for the second time, it is expected that memories of material seen earlier (holistic items) will be confused with other similar material (many of the items are similar in nature as the many of the items involved Secretarial tasks). Therefore, the participants would not necessarily remember the exact rating he or she had given on an identical item at an earlier time. Also, retroactive interference was expected to occur. That is, once the participant was rating the holistic task for the second time, the memory of the decomposed items (which were similar in nature to the holistic items) should have interfered with the memory of how the participant rated the original holistic item.

Thirdly, there were 242 decomposed items. As decomposed items do not need to be deeply processed in order to rate the item, and as this was a large number of similar items, it is unlikely that the participants remembered their exact ratings on each of the individual items due to the large amount of items that were included. There were also 33 holistic items. As holistic items require a deeper level of processing in order to be answered, there was a greater concern about whether the participant would remember his/her rating on a specific holistic item.

However, there was an interval of approximately 30-40 minutes between the participants' first viewing of the holistic items and the second viewing. Once the holistic items were viewed for the first time, 242 similar decomposed items, a reading span task, and finally a digit span task followed these items.

For the larger undergraduate sample (N = 253), 10 holistic and 57 decomposed items were used in the analysis. For the smaller undergraduate sample (N = 50) and for the graduate student sample (N=10), 17 holistic and 160 decomposed items were used in the analyses.

Procedure

Participants were videotaped during the experimental session to ensure that they did not have any writing utensils or any other means of recording any part of the study while participating. Partitions were placed around the computers so that the participants could not see one another's responses. The study was conducted on the Internet in a computer lab. Participants accessed the website on which the study was located in order to perform the study. Groups of 1-19 participants were tested at the same time.

Participants were first presented with an overall description of the tasks that they would be performing during the study. Next, they received a description of the position of Secretary (see Appendix C). The position of Secretary was chosen because it has been shown that undergraduate students are familiar with this position (Butler & Harvey, 1988). Participants were then given a description of the holistic items (see Appendix D) (as was done in Gibson, 2001) and then rated the 33 holistic items. Once that was completed, participants again read the same description of the Secretary position, and were then given a description of the decomposed items (see Appendix E) (as was done in Gibson, 2001). Participants then rated the 242 decomposed items.

For both the holistic and decomposed items, the items were listed on one scrollable screen. In order to continue to the next task, the participant had to click on the ‘continue’ button at the bottom of the screen. If the participant did not rate an item, a message prompted the participant to return to that item and rate it. The participant could not continue to the next screen until all of the items had been rated.

In order to assess each participant’s level of working memory capacity, participants were then presented with an online (modified) version of Daneman and Carpenter’s (1980) reading span task and an online (modified) version of the Wechsler Adult Intelligence Scale Third Edition (WAIS-III) forward and backward digit span tasks.

Upon completion of the reading span and digit span tasks, participants were once again presented with the description of the Secretary position. The participants then went through the cognitive load manipulation, followed by the second holistic rating task (this time the items were presented in a different order).

Participants were then shown the description of the Secretary position. They again were presented with the cognitive manipulation where they were presented with a new 8-digit number, followed by the second decomposed rating task (in a different order than it had originally been presented). After completing the second decomposed rating task, the participants were asked to recall and then recognize the 8-digit number.

At this point, two questions regarding the participants’ knowledge about the position of Secretary and other, related positions were asked. The participants were asked if they had ever worked at any of 17 positions related to the job of Secretary (according to the O*NET database). They were then asked the degree to which they were confident that they knew the types of duties a Secretary does and does not perform. The participants then saw the final screen, which thanked them for their participation and reminded them to hand in the informed consent form.

Manipulation check

Under the higher cognitive load condition, those participants who correctly recalled the 8-digit number at the end of the task or who could correctly recognize that number from the list of 10 numbers for both the Holistic and the Decomposed additional cognitive load conditions were considered as having experienced the manipulation, and were included in the analyses. Participants who could either correctly recall or recognize the number only in the Holistic load or only in the Decomposed additional cognitive load conditions were not included.

Results

Pilot Study

Percent agreement was used in order to test that 80% of the participants agree that a certain decomposed item matches to a certain holistic item. The percent agreement index was computed by:

P% = (Number of agreements/ Number of agreements + number of disagreements) X 100%. Values of p% range from 0% to 100% (Suen & Ary, 1989). The results of the pilot study were used in order to create the decomposed items in the study. 17 holistic items and 160 decomposed items were chosen as focal items for this study. There were at least 4 decomposed items associated with each holistic item. Table 1 shows the percent agreement from participants in the pilot study between all of the holistic and decomposed items used in the analyses. Tables 2-18 show (for the undergraduate sample in the main study) the correlations between all of the holistic and decomposed items used in the analyses as well as the means and standard deviations of the participants' ratings for each item. Table 19 shows the average correlations for the items in each item set (for example, all items associated with item set A: HolNo, HolL, DecNo, and DecL conditions). This table indicates that the correlations between items in this study were much higher than in previous research. The correlations in this study ranged from .28 to .62 while correlations between decomposed and holistic items used in Gibson (2001) ranged from -.99 to .16.

Manipulation Check: Cognitive Load

Previous researchers who have used recall of an 8-digit number as a measure of cognitive load (Gilbert & Osborne, 1989; Pendry & Macrae, 1999; Pontari & Schlenker, 2000) have agreed that this manipulation worked if the participants can correctly recall from memory the 8-digit number at the end of the experimental task. In addition to the free recall task, Harris and Perkins (1995) also included a recognition task (conducted after the free recall task) that had participants identify the 8-digit number from a list of similar 8-digit numbers. Therefore, based on the previous research, this study considered the cognitive load manipulation to have had the desired effect if the participants could either recall the 8-digit number at the conclusion of the experimental task or if they could identify that number from a list of similar 8-digit numbers. In order to be included, the participant had to correctly perform the cognitive load manipulation for both the holistic and the decomposed measures. Of the 314 participants in the sample, 35 participants (11.1% of the sample) could not correctly recall either of the 8-digit numbers, but recognized both of the 8-digit numbers correctly. 102 of the participants (32.5% of the sample) recalled at least one of the 8-digit numbers correctly and also correctly recognized both of the numbers. 166 of the participants (52.9% of the sample) correctly recalled both of the 8-digit numbers and also correctly recognized both of the numbers. Of the 314 participants in the sample, 11 participants (3.5% of the sample) were excluded from the analyses because they failed to correctly recognize the numbers and also failed to correctly recall the 8-digit numbers. All participants who incorrectly recognized the 8-digit numbers also incorrectly recalled the 8-digit numbers.

Those participants for whom all of the data was recorded were not significantly different from the larger sample. For the stability hypotheses, the correlations between the conditions were computed for both the larger and smaller undergraduate samples, converted into *z*-scores, and then compared. The correlations between the HolNo and the DecNo conditions were not significantly different between the larger undergraduate sample (r=.62) and the smaller undergraduate sample (r=.72) (z_{test}(300)=.48, p=*ns*). The correlations between the DecNo and the DecL conditions were not significantly different between the larger undergraduate sample (r=.92) and the smaller undergraduate sample (r=.94) two undergraduate samples (z_{test}(300)=.06, p=*ns*). The correlations between the HolNo and the HolL conditions were not significantly different between the larger undergraduate sample (r=.79) and the smaller undergraduate sample (r=.71) (z_{test}(300)=1.09, p=*ns*). Therefore, both samples were included in the stability analyses.

The levels of interrater agreement for the smaller sample were also not significantly different from those of the larger sample. There were no significant differences between the two samples in levels of interrater agreement for the HolNo condition (t(25)=.22, p=*ns*), the HolL condition (t(25)=.62, p=*ns*), the DecNo condition (t(215)=.78, p=*ns*), or the DecL condition (t(215)=.53, p=*ns*).

The analyses conducted to test the hypotheses of this study included both undergraduate samples. Therefore, the analyses are based on 303 undergraduate students (50 of those had all of the data recorded) and 10 graduate students in Industrial/Organizational Psychology.

Stability across cognitive load conditions and Interrater agreement

Stability across cognitive load conditions was tested with the Pearson's *r* statistic where $r_{AB} = \sigma^2_t/\sigma^2_x$. Here, σ^2_t is the true score variance, and σ^2_x is the total variance of Test A or Test B. Pearson's *r* statistic is a measure of reliability when the assumptions of parallel tests are met. The items used in the experimental conditions were determined by the percent agreement found from the results of the pilot test; therefore, the measures used satisfied these assumptions. Therefore, for the purposes of this study, the Pearson's *r* statistic is an estimate of the proportion of true variance across tests for the raters (Suen & Ary, 1989).

Interrater agreement was tested with $r^*_{wg(J)}$, where $r^*_{wg(J)} = 1-$ average s^2_x/ s^2_{EU}. In this equation, average s^2_x = obtained average variance of the items in the scale, and s^2_{EU} = the variance under a uniform distribution of responses. This version of the r_{wg} formula should be used, as it is invariant with respect to the number of items, thereby allowing for meaningful comparisons between scales with differing numbers of items (as is the case with the decomposed and holistic scales being used in the current study) (Lindell, Brandt, & Whitney, 1999).

Stability Hypotheses.

In order to test those hypotheses that involved stability across cognitive load conditions (Hypotheses 1A, 1B, and 2), a test score for each condition was created for each participant. To compute this score, for each participant, the average rating across all of the items in each condition was computed. That is, each participant received one score for the HolNo, HolL, DecNo, and DecL conditions, for a total of four scores per participant. Table 20 lists the means and standard deviations for the test scores for each condition across all of the undergraduate participants.

Covariates. Partial correlations between time 1 and time 2 for each item in the stability hypotheses were conducted in order to determine if confidence in knowledge of the position of

Secretary, digit span, and reading span should be used as covariates in the analyses. For example, for Hypothesis 1A, partial correlations were conducted between the test scores for the HolNo and HolL conditions using no covariates, knowledge confidence as a covariate, reading span as a covariate, digit span as a covariate, and all of these tests as covariates together. These correlations were then transformed into *z*-scores and compared. It was found that for all comparisons, the *z*-scores were equivalent, regardless of whether or not covariates were used. The *z*-scores were also equivalent regardless of which covariates were used. For example, when examining the correlation between the HolNo and the HolL conditions under the different patterns of covariates, it was found that all of the *z*-scores were equal to .95.

It was decided to report analyses for Hypothesis 1A, 1B, and 2 using both reading span and digit span together as covariates in order to utilize all of the information that was collected. Digit span and reading span were measures of the participants' level of working memory capacity. As working memory capacity would influence the participants' ability to rate two tasks simultaneously, these two measures were used as covariates. Confidence in the knowledge of the position of Secretary was not included as a covariate, as it was not a published test. Therefore, for these three hypotheses, all of the correlations conducted between conditions were partial correlations with both reading span and digit span acting as covariates.

In order to test Hypotheses 1A-2, partial correlations were computed between the different conditions for each of the hypotheses. For example, for Hypothesis 1A, correlations were computed between the HolNo and HolL conditions and between the HolNo and DecNo conditions. These correlations were then transformed into *z*-scores and then compared using the Hotelling-Williams test or a *z*-test.

Table 21 lists the results of the Hotelling-Williams tests conducted for Hypotheses 1A-1B. The Hotelling-Williams test is used to compare z-scores where the sample correlations are not independent because they (a) are computed on the same sample of subjects and (b) are computed using a common variable. The formula for the Hotelling-Williams test is as follows:

$$t_{(n-3)df} \sim (r_{yx} - r_{yz}) \sqrt{[(n-1)(1 + r_{xz})] / [2(n-1)/(n-3) \,|R|\, + \bar{r}^2 (1 - r_{xz})^3]}$$

where $\bar{r} = (r_{yx} + r_{yz}) / 2$, and $|R| = 1 - r^2_{yx} - r^2_{yz} - r^2_{xz} + 2\, r_{yx}\, r_{yz}\, r_{xz}$.

Hypothesis 1A predicted that the level of stability between the Holistic no additional cognitive load (HolNo) and the Decomposed no additional cognitive load (DecNo) conditions would be higher than the level of stability between the HolNo and the Holistic load (HolL) conditions. The Hotelling-Williams test was conducted on the *z* scores from the undergraduate sample in order to test this hypothesis. The stability between the HolNo and the DecNo conditions ($r = .55$) was not significantly different from the stability between the HolNo and the HolL conditions ($r = .74$) ($t(300) = 1.33$, $p = ns$).

Hypothesis 1B predicted that the level of stability between the HolNo and the DecNo conditions would be different from the level of stability between the DecNo and the DecL conditions. The Hotelling-Williams test was conducted on the z scores from the undergraduate sample and found that the stability between the DecNo and the DecL conditions ($r = .66$) was not significantly different than the reliability between the HolNo and the DecNo conditions ($r = .55$) ($t(300) = 1.25$, $p = ns$).

Hypothesis 2 predicted that the level of stability between the DecNo and the DecL conditions would be higher than the reliability between the HolNo and the HolL conditions. A z-test was conducted to analyze this hypothesis. The formula used for the z-test was: $z_{test} = (z_r - z_p) / \sqrt{(1/(n-3))}$. This formula is used when a z-score is being compared to a constant. As the data in this study was collected from only one sample, this test seemed more appropriate than the test between two independent samples. The results of this test indicated that for the undergraduate sample, the stability between the DecNo and the DecL conditions (r = .66) was lower than that of the HolNo and the HolL conditions (r = .74) (z_{test}(300)= 2.72, p=.05). Table 22 lists the results of the z-test conducted to analyze Hypothesis 2.

As Hypothesis 2 examined the levels of stability across the holistic measure to the levels of stability across the decomposed measure, this hypothesis was also analyzed in the same manner as done in Ravinder (1992). For each participant, the scores in the HolL condition were partially correlated with the scores in the HolNo condition. The scores in the DecL condition were partially correlated with the scores in the DecNo condition. That is, for each of the 303 participants, there were now two correlations (one between the two holistic conditions and one between the two decomposed conditions). These correlations were then transformed into z-scores. The average z-score across the participants was computed for both the holistic and decomposed conditions. These two z-scores were then compared using the same z-test as above (comparing a correlation to a constant). When computed this way, the decomposed condition (r=1.405) had a higher level of stability across load conditions than the holistic measure (r=.06) (z_{test}(300)= 14.56, p=.05). A possible explanation for this could be that because the holistic condition had fewer items than the decomposed condition, that there was less freedom in the holistic condition for items to not be consistent. For the decomposed condition, more items could be inconsistent (in ratings across load conditions) while not affecting the overall correlation between the two load conditions for the decomposed measure.

When comparing the expert graduate student sample to that of the undergraduate student sample, it was found that there were significant differences. The following formula was used to test the differences between these two independent samples:

$$z_{test} = (z_1 - z_2) / \surd((1/(n_1 - 3)) + (1/(n_2 - 3)))$$

Table 23 shows the results of the tests between the *z* scores conducted in order to determine the differences between the graduate and undergraduate sample. There were no significant differences in the levels of stability between the HolNo and the HolL conditions for the graduate student sample ($r = .88$) than for the same condition with the undergraduate sample ($r = .74$) ($z_{test}(310)=1.23$, $p=ns$). The levels of stability under the DecNo and the DecL conditions for the graduate student sample ($r = .98$) were higher than between those conditions for the undergraduate student sample ($r = .66$) ($z_{test}(310)=3.94$, $p=.05$). The levels of stability between the HolNo and the DecNo conditions for the graduate students ($r = .89$) were higher than the stability between those conditions for the undergraduate samples ($r = .55$) ($z_{test}(310)=2.10$, $p=.05$).

Interrater Agreement Hypotheses.

In order to find the interrater agreement for each condition, the variance for each item (all holistic and all decomposed items being considered in the analyses) was first calculated. The formula for r_{wg} was then calculated for each condition utilizing all of the items in that condition (17 items for the holistic conditions and 160 items for the decomposed condition). A repeated measures analysis of variance was conducted in order to determine if there were any main effects or an interaction among the levels of interrater agreement for the different conditions. This test was conducted in order to determine if there was an interaction between the type of measure and extraneous cognitive load. The results indicated that there was a main effect of type of measure

(F (1,16)=5.93, p= .03) and also a main effect of cognitive load (F (1,16)=18.48, p=. 00), but that there was no significant interaction between the two variables (see Table 24).

Table 25 indicates the level of interrater agreement (r^*_{wg}) for each condition and the results of the one-way analyses of variance tests conducted to examine Hypotheses 3A-3D. One-way analyses of variance tests are usually computed on conditions consisting of independent samples. However, as the data in this study was only collected from one sample, the design was completely within-subjects, and there were unequal amounts of data in each cell, there was no within-subjects analysis that would allow the comparison between the levels of interrater agreement in each of the four conditions. Therefore, one-way analyses of variance were conducted in order to explore the relationships in the levels of interrater agreement for each of the four conditions.

Hypothesis 3A predicted that the average interrater agreement for the HolNo condition would be lower than the average interrater agreement for the DecNo condition.
A one-way analysis of variance was conducted on the data from the undergraduate sample and found that there were no significant differences between the levels of interrater agreement for the HolNo condition (r^*_{wg}= .59) than for the DecNo condition (r^*_{wg}= .56) (F(1,175)=.261, p = ns).

Hypothesis 3B predicted that the level of interrater agreement would be lower for the HolL condition than for the DecL condition. A one-way analysis of variance was conducted on the data from the undergraduate sample and supported this hypothesis. The level of interrater agreement for the DecL condition (r^*_{wg}= .54) was higher than the level of interrater agreement for the HolL condition (r^*_{wg}= .33) (F(1,175)=7.44, p=.01).

Hypothesis 3C predicted that the level of interrater agreement for the HolL condition would be lower than the level of interrater agreement for the HolNo condition. A one-way analysis of variance was conducted on the data from the undergraduate sample and indicated that the agreement in the HolNo condition (r^*_{wg}= .59) was higher than the agreement in the HolL condition (r^*_{wg}= .33) ($F(1,32)$= 10.74, p = .00). Hypothesis 3C also predicted that the level of interrater agreement for the HolL condition would be lower than the level of agreement in the DecNo condition. A one-way analysis of variance was conducted and indicated that the level of agreement in the DecNo condition (r^*_{wg}= .56) was higher than the agreement in the HolL condition (r^*_{wg}= .33) ($F(1,175)$=10.04, p = .00).

Hypothesis 3D predicted that the level of interrater agreement for the DecL condition would be lower than the level of agreement for the DecNo condition. A one-way analysis of variance was conducted on the data from the undergraduate sample and indicated that there were no significant differences between agreement levels for the DecNo condition (r^*_{wg}= .56) and agreement levels for the DecL condition (r^*_{wg}= .54) ($F(1,318)$= .51, p=*ns*). Hypothesis 3D also predicted that the level of interrater agreement for the DecL condition would be different than the agreement levels for the HolNo condition. A one-way analysis of variance was conducted and indicated that there was no significant difference between the levels of agreement in the DecL condition (r^*_{wg}= .54) and the HolNo condition (r^*_{wg}= .59) ($F(1,175)$=.65, p=*ns*).

The undergraduate sample was compared to the graduate student sample. Table 26 indicates the level of r_{wg} for both the graduate and undergraduate samples and the results of the statistical tests conducted comparing these two groups. One-way analyses of variance indicated that the expert graduate student sample (r^*_{wg}= .52) did not have significantly different levels of interrater agreement than the undergraduate sample (r^*_{wg}= .33) in the HolL condition ($F(1,32)$=3.38, p=*ns*). There were significant differences between the undergraduates (r^*_{wg}= .56) and the graduate students (r^*_{wg}= .67) in the levels of interrater agreement in the DecNo condition

($F(1,318)=9.98$ $p=.00$). There were also significant differences between the undergraduates ($r^*_{wg}=.54$) and the graduate students ($r^*_{wg}=.69$) in the DecL condition ($F(1,318)=23.01$, $p=.00$). There were no significant differences between the graduate students ($r^*_{wg}=.63$) and the undergraduate students ($r^*_{wg}=.59$) in the HolNo condition ($F(1,32)=.18$, $p=ns$).

Supplemental Analyses

Accuracy. Two forms of accuracy were computed on the data from the undergraduate sample as supplementary analyses. In order to look at the overall level of accuracy for this sample, as well as the accuracy for each of the four conditions, the graduate students' scores were used as the "true score". Two of the four components (Elevation and Stereotype Accuracy) of Cronbach's (1955) accuracy components were computed. It was not possible to compute the remaining two components (Differential Elevation and Differential Accuracy) because only one position was being rated in this study. Elevation measures the degree to which a rater's average observed rating equals the average of the target scores. It is computed as $E^2 = (x_. - t_.)^2$. Stereotype Accuracy (Hauenstein & Alexander, 1991) was also computed, where $SA^2 = 1/n \sum [(x_j - x_.) - (t_j - t_.)]^2$. Here, x_j and t_j = observed rating and target score on dimension j, and $x_.$ and $t_.$ = observed mean rating and mean target score over all dimensions. Stereotype Accuracy measures the degree to which a rater's observed ratings and the target ratings are correlated and the degree to which a rater's variance in ratings equal the variance of the target scores. Larger levels of elevation and stereotype accuracy indicate lower levels of accuracy because these components are partitions of the overall inaccuracy score (Butler & Harvey, 1988).

A post-hoc repeated –measures analysis of variance indicated that for elevation, there was a main effect of measure ($F(1,302) = 149.54$, $p= .00$), a main effect of cognitive load ($F(1,302) = 99.04$, $p= .00$), and an interaction between measure and cognitive load ($F(1,302) =$

44.06, p= .00) (see Table 27). As the repeated-measures analysis of variance indicated that there were main effects of measure and cognitive load, as well as an interaction, further analyses were conducted in order to test the differences between the four different conditions. A paired-sample t-test was used to analyze the differences between the four conditions. While a post-hoc test would normally be conducted, in this case there were no between subjects factors. Therefore, post-hoc analyses could not be conducted in the repeated-measures analysis of variance. For the undergraduate sample, the DecNo condition showed the greatest amount of inaccuracy in terms of elevation (E^2 = .79), followed by the HolL condition (E^2 = .33), the DecL condition (E^2 = .22), and the HolNo condition (E^2 = .14). All of the conditions were significantly different from one another (see Table 28).

A post-hoc repeated-measures analysis of variance indicated that in terms of stereotype accuracy, there was a significant main effect of measure ($F(1,302)$ = 165.08, p= .00), a main effect of cognitive load ($F(1,302)$ = 189.28, p= .00), and an interaction between measure and cognitive load ($F(1,302)$ = 822.67, p= .00) (see Table 29). As the repeated-measures analysis of variance indicated that there were main effects of measure, cognitive load, as well as an interaction between these two variables, further analyses were conducted. Again, a paired-sample t-test was used to analyze the differences between the four conditions. While a post-hoc test would normally be conducted, in this case there were no between subjects factors. Therefore, post-hoc analyses could not be conducted within the repeated-measures analysis of variance. The results of the paired-samples t-test indicated that in terms of stereotype accuracy, the DecNo condition again had the greatest amount of inaccuracy (SA^2= 77.13), followed by the HolL condition (SA^2= 55.10) and the DecL condition (SA^2= 56.62). The HolNo condition showed the least amount of inaccuracy (SA^2= 13.37). The HolL and the DecL conditions were not significantly different from one another in terms of stereotype accuracy. All other conditions were significantly different from one another (see Table 30).

Discussion

This study explored empirically a theoretical question, the answer to which has been assumed (but not before tested) in the job analysis literature. Authors in the job analysis literature, such as Butler & Harvey (1988) posit that decomposed type items are superior to holistic type items, while others contend that there are some benefits to utilizing holistic questions (Cornelius & Lyness, 1980; Lyness & Cornelius, 1982, Peterson, et al., 2001). Specifically, this study endeavored to explore which type of item results in the rater experiencing a higher degree of cognitive load.

As mentioned by Paas et al. (1994), levels of cognitive load experienced by a rater can be elevated by several things and can affect raters in many ways. Extraneous cognitive load can be problematic for a rater when there is a high level of intrinsic cognitive load associated with the task (Sweller, 1994). While it is harder to manipulate aspects of the rater (such as internal motivation, ability, etc.) researchers and practitioners do have the ability to choose whether to utilize holistic or decomposed items, thereby gaining some way of influencing the amount of cognitive load a rater is experiencing.

There are differing beliefs as to which type of measure is associated with a higher level of intrinsic cognitive load. In regards to holistic items, Sweller's (1994) explanation indicated that the level of cognitive load intrinsic to a task could be elevated in relation to the number of elements that must be mentally integrated in order to form one comprehensive answer to a question. Holistic job analysis instruments tend to have fewer items, but each item requires a higher level of mental integration in order to respond appropriately. The rater must incorporate memories of multiple past activities in order to answer an item.

While the complexity of each holistic item may lead to an elevated level of cognitive load, this argument does not necessarily indicate that decomposed instruments are the better choice. Decomposed job analysis instruments tend to be much longer than holistic instruments. For example, the Position Analysis Questionnaire (a decomposed instrument) consists of 189 items, nearly five times the amount of items within the O*NET's General Work Activity questionnaire (41 items). While the individual items in a decomposed instrument may require the rater to perform less mental integration in order to rate a specific item, the sheer length of decomposed instruments may cause the rater to experience an elevated level of cognitive load. As Morgeson and Campion (1997) indicated, information overload can be a result of exposure to large amounts of information at one time.

It is important to learn whether holistic or decomposed items lead to higher levels of cognitive load for a rater. As assumed by Morgeson and Campion (1997), the more cognitively taxed a rater is, the less accurate his/her ratings become. It would seem beneficial for practitioners and researchers to be aware of empirical evidence that helped to indicate which type of job analysis items are more mentally taxing to perform. In this regard, the current study was conducted.

Hypotheses 1A-1B. The purpose of Hypotheses 1A and 1B was to examine whether information overload occurred in holistic type measures (due to the complexity of the questions) or in decomposed type measures (due to the length of the questions) (Morgeson & Campion, 1997). There were no significant differences in the level of stability between the Holistic no additional cognitive load and the Decomposed no additional cognitive load conditions and between the Holistic load and Holistic no additional cognitive load conditions. There were no significant differences in the stability between the Decomposed load and Decomposed no additional cognitive load conditions and the stability between the Holistic no additional cognitive load and the Decomposed no additional cognitive load conditions.

Hypothesis 2. The purpose of Hypothesis 2 was to determine whether holistic or decomposed items had a higher level of temporal stability (Ravinder, 1992). The results of Hypotheses 1A and 1B indicate that the level of stability in the no additional cognitive load condition (for both the holistic and decomposed measure) was not significantly different from the stability across the Holistic conditions (additional cognitive load and no additional cognitive load) and was also not significantly different than the stability across the Decomposed conditions (additional cognitive load and no additional cognitive load). However, the stability between the Decomposed load and the Decomposed no additional cognitive load conditions was significantly lower than the stability between the Holistic load and Holistic no additional cognitive load condition.

Hypotheses 3A-3D. The purpose of Hypotheses 3A through 3D was to test which condition led to the highest level of interrater agreement. It was found that there were no significant differences between the levels of interrater agreement in the Decomposed additional cognitive load, Decomposed no additional cognitive load, and Holistic no additional cognitive load conditions. However, the level of interrater agreement for each of these conditions was higher than the level of interrater agreement in the Holistic additional cognitive load condition. It was found that both the type of measure and cognitive load variables resulted in main effects for the interrater agreement hypotheses.

Comparing undergraduates to graduate students. It was found that the expert graduate students did not have significantly different levels of stability in ratings than the undergraduate students between the Holistic no additional cognitive load/Holistic additional cognitive load conditions. The graduate students did have a higher level of stability in ratings between the Holistic no additional cognitive load/ Decomposed no additional cognitive load conditions, and between the Decomposed no additional cognitive load/ Decomposed additional cognitive load conditions. The graduate students also had higher levels of interrater agreement than the undergraduate students in both of the Decomposed conditions.

Accuracy. When using the graduate students scores as the "true" scores, it was found that the Decomposed no additional cognitive load condition produced the lowest level of accuracy (according to both elevation and stereotype accuracy), followed by the Holistic additional cognitive load and the Decomposed additional cognitive load conditions. The Holistic no additional cognitive load condition showed the highest level of accuracy (according to both elevation and stereotype accuracy).

Again, the purpose of Hypothesis 1A and 1B was to examine whether information overload occurred in holistic type measures (due to the complexity of the questions) or in decomposed type measures (due to the length of the questions) (Morgeson & Campion, 1997). The results of Hypothesis 1A and 1B indicated that holistic ratings do not necessarily lead to a higher level of cognitive load for the rater. In actuality, the results of this study showed that the level of stability between the Holistic no additional cognitive load and the Decomposed no additional cognitive load condition was not significantly different than the level of stability between the two Holistic conditions. The level of stability between the two decomposed conditions was also not significantly different than the stability in ratings made between the Holistic no additional cognitive load and the Decomposed no additional cognitive load condition.

Hypothesis 2 was proposed in order to determine whether holistic or decomposed items had a higher level of temporal stability. The results in this study did not replicate the findings of Ravinder (1992), as here, the decomposed items had a lower level of temporal stability than the holistic items. It is interesting to note that the expert graduate students had a higher level of stability between the two holistic conditions and between the two decomposed conditions than the undergraduate sample did. This indicates that when a rater is familiar with the instrument, temporal stability between the load conditions can be improved.

The results of Hypothesis 2, combined with Hypotheses 1A and 1B indicate that holistic items are not necessarily inferior to decomposed items, as past literature has assumed (Butler & Harvey, 1988). In fact, the holistic items result in higher levels of stability across conditions than the decomposed items. In addition, the Holistic no additional cognitive load condition was found to have the highest level of accuracy when being compared to the graduate students' true scores. These results support the findings of Cornelius and Lyness (1980) in that when intrarater reliability across time was used as a criterion, the results indicated that the holistic strategy was found to be equally effective as the decomposed strategy. In the current study, the stability hypotheses indicate that the holistic items were either seen as equally effective as or superior to the decomposed items. The current study also supports the finding of Sanchez and Levine (1994), in that for educated raters, performance on holistic items may be superior to performance on decomposed items.

As mentioned earlier, Hypotheses 3A through 3D were proposed in order to test which condition led to the highest level of interrater agreement. Butler and Harvey (1988) found that holistic measures had lower levels of interrater agreement than decomposed measures, presumably because the holistic measures led to a higher level of cognitive load than the decomposed. However, all of the participants did not perform on both holistic and decomposed measures in their study.

Hypotheses 3A through 3D indicate that the levels of interrater agreement were not significantly different between three of the four conditions. The Holistic additional cognitive load condition was the only condition that differed from (was significantly lower than) the others. These hypotheses could be interpreted as indicating that the holistic measure does not lead to a higher degree of cognitive load for the rater. When under the no load condition, levels of interrater agreement for the holistic measure were not significantly different from interrater agreement levels for either of the two decomposed measure conditions.

The results of the interrater agreement hypotheses indicate that the holistic and decomposed measures differentially affect the rater when under the additional cognitive load condition. It could be that raters are using heuristics and stereotypes when performing both the holistic and the decomposed rating tasks. Similar to Harris and Perkins (1995), performing an additional task while rating on the decomposed measure may force raters to rely more heavily on stereotypes and engage in biased information processing. This is supported by the fact that when rating on the decomposed measure, the amount of variance in responses was smaller in the additional cognitive load condition than in the no additional cognitive load condition. Also, when an additional task was performed in conjunction with the decomposed measure, levels of accuracy increased (both for stereotype accuracy and elevation). For the holistic measure, however, variance in responses increased as an additional task was performed and levels of accuracy decreased with the addition of another task. It would seem that for the holistic measure, raters are utilizing fewer shared stereotypes when another task is being performed simultaneously.

Graduate students had significantly higher levels of interrater agreement than the undergraduates in the Decomposed no additional cognitive load and the Decomposed additional cognitive load conditions. In addition to being more familiar with the types of items contained in holistic and decomposed measures, the graduate students would be more likely to share the same stereotypes and heuristics involving the position being rated. This increased level of shared stereotypes may have led to the higher levels of interrater agreement in the decomposed conditions. Again, as mentioned earlier, there may be something intrinsically different about the holistic type items that do not enable a rater to engage in the stereotypes that the rater may normally utilize. While the graduate students did have higher levels of interrater agreement than the undergraduate students, the levels of interrater agreement for the graduate students were not as high as normally found in the literature. However, the levels of interrater agreement found in this study were similar to Gibson (2001), and therefore were still utilized as points of comparison for the undergraduate students.

Hypotheses 3A-3D would suggest that for the undergraduate students, under the no additional cognitive load condition, interrater agreement is not significantly different for either the Holistic or Decomposed conditions. However, when additional cognitive load is present, the interrater agreement for the Holistic condition is lower than for the Decomposed condition. This would imply an interaction between the measure conditions and the load conditions. The results from the within-subjects analysis of variance did indicate a main effect of both measure and load conditions, but the interaction was not significant.

It is interesting to note that while the Holistic additional cognitive load condition was the only condition that was significantly lower than the others in terms of interrater agreement, the highest level of stability was found across the holistic measure. The results would have appeared to be more logical if the highest level of stability had been found across the decomposed measure, as both load conditions in this measure had similar levels of interrater agreement. It could be that the interrater agreement for the Holistic no additional cognitive load was overestimated, as this condition had a smaller range than the other conditions, which affects the interrater agreement index, r^*_{wg}. When there is a restricted range, r^*_{wg} tends to overestimate the level of interrater agreement (Harvey & Hollander, 2002). If the two holistic conditions in reality have a more similar level of interrater agreement (if the interrater agreement for the Holistic no additional cognitive load is actually lower than reported by r^*_{wg}), then the higher level of stability found between these two conditions would be more logical.

The interaction between measure type and level of cognitive load that has been suggested by the findings of the current study (for the interrater agreement hypotheses, although the interaction itself was found non-significant) supports the cognitive literature in a larger sense. As previously demonstrated (Pendry & Macrae, 1999; Pontari & Schlenker, 2000), if one is experiencing a lower level of cognitive load, it is possible to perform more than one task concurrently. If one is experiencing a higher level of cognitive load, then the performance on the critical task will be altered. The current findings suggest that the addition of an extraneous source of cognitive load will differentially affect holistic and decomposed measures, depending upon the criterion that is being examined. The level of stability was higher between the two holistic conditions than between the two decomposed conditions. In addition, the level of stability was higher between the two holistic conditions than between the two conditions with no additional cognitive load. A different pattern is found with interrater agreement, however, suggesting that the holistic items lead to a higher level of cognitive load than the decomposed. Under interrater agreement, there was only a decrement in performance for the Holistic additional cognitive load condition. This argument is not supported, though, by the fact that the Holistic no additional cognitive load condition actually had the highest level of accuracy when comparing stereotype accuracy and elevation across conditions. It is interesting to note that Decomposed no additional cognitive load had the lowest level of accuracy in terms of both elevation and stereotype accuracy.

The results of this study indicate that the holistic condition may not have induced a higher degree of cognitive load for the rater, but that the holistic condition did prevent the raters from utilizing stereotypes and heuristics when the rater was given an additional task to perform. Contrary to this, it would seem that for the decomposed measure, as an additional source of cognitive load was placed on the rater, raters were more likely to engage in stereotype use (as evidenced by a decreased amount of variance in responding, and higher levels of accuracy as cognitive load was increased). This is useful information for researchers and practitioners to have when designing tests. If one wants to help ensure that the raters will not be engaging in heuristics when responding, then perhaps the holistic items would be a better choice. This study demonstrates (similar to Sanchez & Levine, 1994) that perhaps the context in which the ratings are being used determines whether single (global) items or decomposed items are more useful.

Limitations. There were a few limitations to this study. First, a problem occurred during data collection: the last 80 responses for both of the decomposed sections were not recorded for a large portion of my undergraduate sample. This limited several analyses with that sample to a subset of the original number of items that had been planned to analyze. Another sample of 50 undergraduates was collected, however, and for the smaller sample all responses were recorded and analyzed.

Originally, the intention had been to compare the undergraduate data against both expert graduate students and Secretary incumbents. Unfortunately, it was not possible to gather enough Secretaries in order to make worthwhile comparisons. Therefore, no comparisons were made between the undergraduate and Secretary samples. While data was collected from only 10 graduate students (which is not a significantly larger amount of data than the amount collected from Secretarial participants), the data from the graduate students was used in order to provide some point of comparison to the undergraduates as was done in Gibson (2001).

Another possible limitation was the length of the decomposed set of items. While some may argue that the length of an actual decomposed job analysis instrument may sometimes be longer than the 242 items used in this study, other studies have actually used decomposed instruments with fewer items. For example, Butler and Harvey (1988) used a form of the PAQ that consisted of 189 items as the decomposed measure in their study. The Job Element Inventory (JEI), another decomposed job analysis instrument consists of 153 items. While these example used fewer items than the current study, other decomposed job analysis instruments do use a larger number than what was used in this study. As mentioned earlier in the paper, alternative explanations for the current findings are possible. It has been demonstrated (Butler & Harvey, 1988; Cornelius & Lyness, 1980; Lyness & Cornelius, 1982) that different patterns of findings can occur depending on the criterion used to determine the amount of cognitive load that is intrinsic to holistic and decomposed items. If different criterion were used, then perhaps the patterns demonstrated in this study may have been different.

Another limitation is that the cognitive load manipulation might not have been a concurrent task for the participants. As no data was collected on how often, or if, the participant was rehearsing the 8-digit number during the holistic and decomposed tasks, one cannot say with certainty that the participants were performing both tasks simultaneously. It is possible that the participants memorized the number during the 25 s presentation, performed the rating task, and then recalled and recognized the number. However, as the cognitive load manipulation differentially affected the levels of interrater agreement for the holistic and decomposed measures, one might assume that the participants were rehearsing the 8-digit number while rating the decomposed and holistic items. In the future, data should be collected regarding whether or not the participant rehearsed the number during the focal task.

Another possible limitation of this study was that the cognitive load manipulation was always presented second rather than being counterbalanced. This manipulation was always presented second because it was thought that the addition of a second task in conjunction with having already performed the rating tasks once would cause a higher level of cognitive load (as the rater may have been fatigued from the amount of tasks they had performed thus far). Also, it was thought that by presenting this manipulation second, there would be a decrement in the raters' ability to recall the responses made during the first presentation of the holistic and decomposed tasks.

By having this manipulation second, however, a confound could have been introduced into the study. It could be that the results of this study were due to rater fatigue, rather than a significant effect of the cognitive load manipulation. A way to test this would be to collect more data during which the cognitive load manipulation was presented first, thereby allowing the opportunity to test to see if the cognitive load manipulation actually did cause higher levels of cognitive load for the raters.

Contributions. This study does contribute to the body of job analysis literature. First, this study was an attempt to demonstrate, rather than assume, whether holistic or decomposed items cause higher levels of cognitive load for a rater. As this study was a laboratory study, the results may not be widely generalizable. However, this study makes a significant contribution in that it examines levels of stability in ratings across conditions and interrater agreement. Because these two dependent measures were studied using the data from the same sample, there were no between-subjects factors (thereby making it difficult to conduct certain analyses that may have been deemed more appropriate). This study examined one aspect, cognitive load, of a debate over what types of job analysis items are more effective.

The results (found using stability across cognitive load conditions and interrater agreement as criterion) do not support the ideas of many researchers (Butler & Harvey, 1988; Ravinder, 1992) who believe that holistic items do lead to a higher level of cognitive load than decomposed items. This study used both holistic and decomposed measures as true scores, rather than only comparing both types of ratings to decomposed measures as has been done in the past. When comparing both types of ratings to a decomposed criterion, it would seem logical that the decomposed measure would be more similar to the criterion than the holistic measure. However, the results of this study demonstrate that when given a fair shot of succeeding, the holistic measures do not perform as badly in comparison to the decomposed measures as some may assume.

This study also contributes to the literature because the decomposed items used here were in many cases highly correlated with the holistic items. While there were some items that did not correlate well, the range of average correlations between item types was an improvement over past studies that have tried to compare holistic and decomposed items (see Gibson, 2001).

Comparisons between the undergraduate sample and the expert graduate sample suggest that it is possible to have improved performance on both holistic and decomposed items if the rater is familiar with the instrument. Graduate students, familiar with the instruments used in this study, had higher levels of stability than undergraduates across all conditions and also higher levels of interrater agreement for the two decomposed conditions. This is consistent with other research that has indicated that inferential training can help to improve ratings on both holistic and decomposed items (Sanchez & Levine, 1994). The results from this study indicate that the differences found between holistic and decomposed measures in past research (Butler & Harvey, 1988) may not be due to the holistic measures causing a higher degree of cognitive load for the rater. Further research (where both the holistic and decomposed measures have a chance of being shown as accurate and useful) is needed to more fully explore the differences between the levels of cognitive load that decomposed and holistic job analysis items induce in naïve raters. As incumbents are being utilized more often as sources for job analysis ratings (as the O*NET is currently doing), it is important to discover exactly how these different item types affect the cognitive processes of naïve raters.

References

Armstrong, J.S., Denniston, W.B., & Gordon, M.M. (1975). The use of the decomposition principle in making judgments. *Organizational Behavior and Human Performance, 14*, 257-263.

Blessum, K.A., Lord, C.G., & Sia, T.L. (1998). Cognitive load and positive mood reduce typicality effects in attitude-behavior consistency. *Personality and Social Psychology Bulletin, 24, (5)*, 496-504.

Butler, S.K., & Harvey, R.J. (1988). A comparison of holistic versus decomposed rating of Position Analysis Questionnaire work dimensions. *Personnel Psychology, 41,* 761-771.

Conoley, J.C., Impara, J.C., & Murphy, L.L. (Eds.) (1995). *The twelfth mental measurements yearbook*. Oxford: Oxford University.

Cornelius, E.T., & Hakel, M.D. (1978). *A study to develop an improved enlisted performance evaluation system for the U.S. Coast Guard*. Washington, DC: Department of Transportation, United States Coast Guard.

Cornelius, E.T., & Lyness, K.S. (1980). A comparison of holistic and decomposed judgment strategies in job analyses by job incumbents. *Journal of Applied Psychology, 65, (2)*, 155-163.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Fink, A., & Neubauer, A.C. (2001). Speed of information processing, psychometric intelligence, and time estimation as an index of cognitive load. *Personality and Individual Differences, 30, (6)*, 1009-1021.

Fischer, G.W. (1977). Convergent validation of decomposed multi-attribute utility assessment procedures for risky and riskless decisions. *Organizational Behavior and Human Performance, 18,* 295-315.

Fleishman, E.A., & Reilly, M.E. (1992). *Handbook of Human Abilities: Definitions, measurements, and job task requirements*. Palo Alto, CA: Consulting Psychologists Press, Inc.

Gibson, S.G. (2001). *Call me old fashioned – Is my job analysis accurate or not?* Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.

Gilbert, D.T. & Osborne, R.E. (1989). Thinking backward: Some curable and incurable consequences of cognitive busyness. *Journal of Personality and Social Psychology, 57, (6)*, 940-949.

Harris, M.J. & Perkins, R. (1995). Effects of distraction on interpersonal expectancy effects: A social interaction test of the cognitive busyness hypothesis. *Social Cognition, 13, (2)*, 163-182.

Hauenstein, N.M.A, & Alexander, R.A. (1991). Rating ability in performance judgments: The joint influence of implicit theories and intelligence. *Organizational Behavior and Human Decision Processes, 50*, 300-323.

Harvey, R.J. (1993). *Research monograph: Development of the Common-Metric Questionnaire (CMQ)*. Personnel Systems and Technologies Corporation & Virginia Polytechnic Institute and State University.

Harvey, R.J., & Hollander, E. (2002). Assessing interrater agreement in the O*NET. In M.A. Wilson (Chair), *The O*NET: Mend it or end it?* Symposium presented at the Annual Conference of the Society for Industrial and Organizational Psychology, Toronto.

Harvey, R.J., & Wilson, M.A. (2000). Yes Virginia, there is an objective reality in job analysis. *Journal of Organizational Behavior, 21*, 829-854.

Kirschner, P.A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction, 12, (1)*, 1-10.

Kruglanski, A.W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106, (3)*, 395-409.

Lindell, M.K., Brandt, C.J., & Whitney, D.J. (1999). A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement, 23, (2)*, 127-135.

Lyness, K.S., & Cornelius, E.T. (1982). A comparison of holistic and decomposed judgment strategies in a performance rating simulation. *Organizational Behavior and Human Performance, 29*, 21-38.

Mecham, R.C., McCormick, E.J., & Jeanneret, P.R. (1977). *Position Analysis Questionnaire user manual (System II.)*. West Lafayette, IN: Purdue University Bookstore.

Morera, O.F., & Budescu, D.V. (1998). A psychometric analysis of the "Divide and Conquer" principle in multicriteria decision-making. *Organizational Behavior and Human Decision Processes, 75*, (3), 187-206.

Morera, O.F., & Budescu, D.V. (2001). Random error reduction in analytic hierarchies: A comparison of holistic and decompositional decision strategies. *Journal of Behavioral Decision Making, 14*, 223-242.

Morgeson, F.P., & Campion, M.A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology, 82*, (5), 627-655.

Morgeson, F.P., & Campion, M.A. (2000). Accuracy in job analysis: toward an inference-based model. *Journal of Organizational Behavior, 21*, 819-827.

Paas, F.G., van Merrienboer, J.J.G., & Adam, J.J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*, 419-430.

Pendry, L.F., & Macrae, C.N. (1999). Cognitive load and person memory: The role of perceived group variability. *European Journal of Social Psychology, 29*, 925-942.

Peterson, N.G., Mumford, M.D., Borman, W.C., Jeanneret, P.R., Fleishman, E.A., Levin, K.Y., et al. (2001). Understanding work using the Occupational Information Network (O*NET): Implications for practice and research. *Personnel Psychology, 54,* 451-492.

Pollock, E., Chandler, P., & Sweller, J. (2001). Assimilating complex information. *Learning and Instruction, 12, (1)*, 61-86.

Pontari, B.A., & Schlenker, B.R. (2000). The influence of cognitive load on self-presentation: Can cognitive busyness help as well as harm social performance? *Journal of Personality and Social Psychology, 78, (6)*, 1092-1108.

Ravinder, H.V. (1992). Random error in holistic evaluations and additive decompositions of multiattribute utility – an empirical comparison. *Journal of Behavioral Decision Making, 5* 155-167.

Reitman, J.S. (1974). Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior, 13(4)*, 365-377.

Sanchez, J.I., & Levine, E.L. (1994). The impact of raters' cognition on judgment accuracy: An extension to the job analysis domain. *Journal of Business and Psychology, 9, (1)*, 47-57.

Sanchez, J.I., & Levine, E.L. (2000). Accuracy or consequential validity: which is the better standard for job analysis data? *Journal of Organizational Behavior, 21*, 809-818.

Silvera, D.H. (2000). The effects of cognitive load on strategic self-handicapping. *British Journal of Social Psychology, 39*, 65-72.

Suen, H.K. & Ary, D. (1989). *Analyzing Quantitative Behavioral Observation Data.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

Van Knippenberg, A., Dijksterhuis, A., & Vermeulen, D. (1999). Judgment and memory of a criminal act: The effects of stereotypes and cognitive load. *European Journal of Social Psychology, 29*, 191-201.

Appendix A.

All holistic items presented in focal study.

How **important** is JUDGING THE QUALITIES OF OBJECTS, SERVICES, OR PEOPLE to the performance of the position of Secretary?
(Judging the qualities of objects, services, or people – Assessing the value, importance, or quality of things or people.)

How **important** is MEDICINE AND DENTRISTY to the performance of the position of Secretary?
(Medicine and dentistry – Knowledge of the information and techniques needed to diagnose and treat human injuries, diseases, and deformities. This includes symptoms, treatment alternatives, drug properties and interactions, and preventive health-care measures)

* How **important** is TELECOMMUNICATIONS to the performance of the position of Secretary?
(Telecommunications – Knowledge of transmission, broadcasting, switching, control, and operation of telecommunication systems.)

How **important** is PROCESSING INFORMATION to the performance of the position of Secretary?
(Processing information – Compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data.)

How **important** is HISTORY AND ARCHEOLOGY to the performance of the position of Secretary?
(History and archeology – Knowledge of historical events and their causes, indicators, and effects on civilizations and cultures)

* How **important** is COMMUNICATING WITH PEOPLE OUTSIDE THE ORGANIZATION to the performance of the position of Secretary?
(Communicating with people outside the organization – Communicating with people outside the organization, representing the organization to customers, the public, government, and other external sources. This information can be exchanged in person, in writing, by telephone, or in e-mail.)

How **important** is BUILDING AND CONSTRUCTION to the performance of the position of Secretary?
(Building and construction – Knowledge of materials, methods, and the tools involved in the construction or repair of houses, buildings, or other structures such as highways and roads.)

* How **important** is SCHEDULING WORK AND ACTIVITIES to the performance of the position of Secretary?
(**Scheduling work and activities** – Scheduling events, programs, and activities,

as well as the work of others.)

* How **important** is MAKING DECISIONS AND SOLVING PROBLEMS to the performance of the position of Secretary?
(Making decisions and Solving Problems – Analyzing information and evaluating results to choose the best solution and solve problems.)

How **important** is PHYSICS to the performance of the position of Secretary?
(Physics – Knowledge and prediction of physical principles, laws, their interrelationships, and applications to understanding fluid, material, and atmospheric dynamics, and mechanical, electrical, atomic, and sub-atomic structures and processes.)

* How **important** is ECONOMICS AND ACCOUNTING to the performance of the position of Secretary?
(Economics and Accounting – Knowledge of economic and accounting principles and practices, the financial markets, banking and the analysis and reporting of financial data.)

How **important** is INTERPRETING THE MEANING OF INFORMATION FOR OTHERS to the performance of the position of Secretary?
(Interpreting the meaning of information for others – Translating or explaining what information means and how it can be used.)

* How **important** is OPERATING VEHICLES, MECHANIZED DEVICES, OR EQUIPMENT to the performance of the position of Secretary?
(**Operating vehicles, mechanized devices, or equipment** – Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft.)

* How **important** is ESTABLISHING AND MAINTAINING INTERPERSONAL RELATIONSHIPS to the performance of the position of Secretary?
(Establishing and Maintaining Interpersonal Relationships – Developing constructive and cooperative working relationships with others, and maintaining them over time.)

How **important** is PRODUCTION AND PROCESSING to the performance of the position of Secretary?
(Production and processing – Knowledge of raw materials, production processes, quality control, costs, and other techniques for maximizing the effective manufacture and distribution of goods)

* How **important** is MATHEMATICS to the performance of the position of Secretary?
(**Mathematics** – Knowledge of arithmetic, algebra, geometry, calculus, statistics, and their applications.)

* How **important** is CLERICAL KNOWLEDGE to the performance of the position of Secretary?
(Clerical knowledge– Knowledge of administrative and clerical procedures and systems such as word processing, managing files and records, stenography and transcription,

designing forms, and other office procedures and terminology.)

How **important** is BIOLOGY to the performance of the position of Secretary?
(Biology – Knowledge of plant and animal organisms, their tissues, cells, functions, interdependencies, and interactions with each other and the environment.)

* How **important** is FINE ARTS KNOWLEDGE to the performance of the position of Secretary?
(Fine Arts – Knowledge of the theory and techniques required to compose, produce, and perform works of music, dance, visual arts, drama, and sculpture.)

How **important** is REPAIRING AND MAINTAINING ELECTRONIC EQUIPMENT to the performance of the position of Secretary?
(Repairing and maintaining electronic equipment – Servicing, repairing, calibrating, regulating, fine-tuning or testing machines, devices, and equipment that operate primarily on the basis of electrical or electronic (not mechanical) principles.)

. How **important** is TRANSPORTATION to the performance of the position of Secretary?
(Transportation – Knowledge of principles and methods for moving people or goods by air, rail, sea, or road, including the relative costs and benefits.)

* How **important** is WORKING WITH COMPUTERS to the performance of the position of Secretary?
(Working with Computers – Using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information.)

How **important** is GEOGRAPHY to the performance of the position of Secretary?
(Geography – Knowledge of principles and methods for describing the features of land, sea, and air masses, including their physical characteristics, locations, interrelationships, and distribution of plant, animal, and human life.)

* How **important** is HANDLING AND MOVING OBJECTS to the performance of the position of Secretary?
(**Handling and Moving objects** – Using hands and arms in handling, installing, positioning, and moving materials, and manipulating things.)

* How **important** is knowledge of FOREIGN LANGUAGE to the performance of the position of Secretary?
(Foreign Language – Knowledge of the structure and content of a foreign (non-English) language including the meaning and spelling of words, rules of composition and grammar, and pronunciation.)

How **important** is CHEMISTRY to the performance of the position of Secretary?
(Chemistry – Knowledge of the chemical composition, structure, and properties of substances and of the chemical processes and transformation that they undergo. This includes uses of chemicals and their interactions, danger signs, production techniques, and disposal methods.)

* How **important** is CUSTOMER AND PERSONAL SERVICE to the performance of the position of Secretary?
(Customer and Personal Service – Knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction.)

How **important** is UPDATING AND USING RELEVANT KNOWLEDGE to the performance of the position of Secretary?
(Updating and using relevant knowledge – Keeping up-to-date technically and applying new knowledge to your job.)

How **important** is PUBLIC SAFETY AND SECURITY to the performance of the position of Secretary?
(Public safety and security – Knowledge of relevant equipment, policies, procedures, and strategies to promote effective local, state, or national security operations for the protection of people, data, property, and institutions)

* How **important** is the ENGLISH LANGUAGE to the performance of the position of Secretary?
(English Language – Knowledge of the structure and content of the English language including the meaning and spelling of words, rules of composition, and grammar.)

How **important** is MECHANICAL to the performance of the position of Secretary?
(Mechanical – Knowledge of machines and tools, including their designs, uses, repair, and maintenance)

* How **important** is FOOD PRODUCTION to the performance of the position of Secretary?
(**Food Production** – Knowledge of techniques and equipment for planting, growing, and harvesting food products (both plant and animal) for consumption, including storage/handling techniques.)

* How **important** is COMMUNICATING WITH SUPERVISORS, PEERS, OR SUBORDINATES to the performance of the position of Secretary?
(Communicating with Supervisors, Peers, or Subordinates – Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person.)

* . Item used in analyses.

Appendix B

All decomposed items presented in the focal study

*1 input information in prescribed formats such as reports, mailing lists, accounts, payrolls, staff records, or scientific information
*2 fly a jet
*3 contact supervisors as part of the job (examples: foremen, first-level supervisors)
*4 understand oral instructions
5 determine whether it is more cost effective to move goods by train or by plane
*6 write a newspaper article in a foreign language
*7 write software for keeping track of parts in inventory
8 understand physics principles
*9 schedule meetings with other people
10 place password protection on classified computer files
*11 navigate a speedboat
*12 provide information to assist clients
*13 good interpersonal skills when dealing with customers or clients
14 set up mail-merge functions in a computer program to enable multiple letters to be directed to Individuals in a number of locations
15 take a class in US history
*16 perform duties such as data entry
*17 place telephone callers on hold
*18 write an English language review of a book written in a foreign language
19 use a common household bug spray
*20 determine the meal selection for a cafeteria
21 perform open heart surgery
22 verify information contained in a report
*23 print out address labels using a computer
24 replace a valve on a steam pipe
*25 teach a college English class
*26 move light objects on occasion
27 use a seatbelt
28 make repairs by removing and replacing circuit boards
29 teach local history to school children
*30 use English words or sentences in speaking so others can understand
*31 arranging for the payment of dividends, and observing all legal requirements
32 determine which type of packing material is most efficient to use to ship a product
*33 paints scenic backgrounds
*34 file letters alphabetically
35 fill a tooth cavity
*36 discuss issues and problems with others in the organization
37 overhaul an airplane jet engine
*38 possess strong problem-solving skills

39 determine which tools are needed in order to build a roof
*40 ability to interact in a friendly manner with others in the organization
*41 add, subtract, multiply, divide, and manipulate numbers quickly and accurately
*42 drive a car
43 command a military operation
*44 maintain good working relationships with almost all coworkers and clients
*45 prepare the work schedule for subordinates
*46 supervise and coordinate activities of staff
47 code data to be used in research analysis
*48 address envelopes or prepare envelope labels, using typewriter or computer
49 organize a sobriety checkpoint with local police officers
*50 compose music
51 use complex test equipment to calibrate electronic equipment
*52 butchers animals
53 control air traffic at a busy airport
*54 navigate a helicopter
*55 knowledge of calculus
*56 organize information to be mailed out
*57 check messages sent to and from computers
*58 play a minor role in a local theater play
*59 organize business itineraries, travel arrangements, conferences, meetings and social functions
60 determine the quality of education in the local school system
61 knowledge of the French revolutionary war
*62 drive an 18-wheel tractor-trailer
63 know the capital of the United States
*64 use keyboard devices (examples: calculators, typewriters, computers)
*65 place, move, or assemble objects
66 determine which type of material is most efficient when constructing a road
*67 set up spreadsheets
*68 knowledge of arithmetic
*69 keep superiors informed of any potential problems
*70 organize a work schedule that is repetitive and easy to plan
*71 washes, peels, cutes, and seeds fruits and vegetables
72 inspect a building site for safety violations
73 determine the value of property lost in a fire
*74 understand someone with a heavy foreign accent
*75 maneuver water craft
*76 able to solve problems
77 determine the costs of flying to a conference across the country
*78 ability to assess the needs of customers
*79 knowledge of how to apply various types of mathematics
*80 cook food
81 apply a new technique for installing carpet
*82 deal in a tactful manner with others outside the organization

*83 connect/disconnect telephone calls
*84 exchange greetings with a coworker
85 put a computer back into its packing materials
*86 operate a bulldozer
*87 dispatch outgoing mail
88 learn new techniques for combining chemicals
*89 choreograph dance routines
90 know the distribution of plant life in Africa
91 use a crowbar to pry open a box
*92 settle disputes among others
*93 supervise people who are not employees
*94 maintain records
*95 gathers and arranges material to be typed, following instructions
*96 maintain calendar and coordinates conferences and meetings
97 use a band-aid
*98 cut and de-bone meat prior to cooking
99 diagnose a patient with cancer
*100 transfer telephone callers from one extension to another
*101 learn new computer programs to perform work more effectively
*102 derive a complex mathematical equation
*103 report the results of a sales meeting to a supervisor
*104 conduct research on the internet
105 understand how to repair a jackhammer
*106 handle and address customer complaints
107 manage an international shipping company distribution center
*108 fill out income tax returns
*109 operate a forklift
110 develop a map of the world showing mountains, deserts, and rivers
111 interpret how foreign tax laws apply to U.S. exports
*112 file papers and documents so that they can be easily found when needed
113 build a high-rise office tower
*114 use word-processing equipment
*115 stores completed documents on computer had drives or data storage mediums, such as a disk
*116 good spelling skills
117 predict whether the interaction of two chemicals will be dangerous
*118 add two numbers
119 understand how to properly maintain a power saw
120 understand what to feed different wild animals at a rescue center
*121 say "please" and "thank you" in a foreign language
*122 prepare master copies
123 explain how the American Disabilities Act can affect an employee's income
*124 keep a major corporation's financial records
*125 develop a new, world-wide telecommunications network
*126 feed domestic animals

127 know which areas of the country contain a higher population of a certain animal
*128 ability to maintain working relationships over time
129 relocate rescued animals into the wild
*130 ability to interact in a friendly manner with those outside of the organization
131 repair electronic equipment
*132 act in a play
133 keep current on changes in maintenance procedures for repairing sports cars
*134 knowledge of geometry
*135 understand information spoken in English
136 Identify Turkey on a world map
*137 type and re-arrange information, such as highlighting parts of the text, moving paragraphs from one page to another, And putting information into columns.
*138 good knowledge of the spoken English language
*139 resolve conflicting findings
140 isolate and identify a new virus
*141 paints murals
*142 print out letters using a computer
*143 ability to develop working relationships with others
*144 ask directions in a foreign city
145 set a broken arm
*146 make appointments using a predetermined schedule
*147 knowledge of statistics
148 replace the filters in a furnace
*149 handle money
*150 enter employee information into a computer database
*151 design an artistic display for a major trade show
*152 contact executives and other high officials as part of the job (examples: corporation vice-presidents, government officials)
*153 carve meat
*154 ability to work with computers
*155 contact outside professionals as part of job (examples: doctors, lawyers, professors, engineers, consultants)
*156 act as an advisor to superiors
*157 knowledge of algebra
*158 perform basic math, such as finding percentages and taking square roots
*159 run a 100,000 acre farm
160 predict whether the contents in a pipe will explode under different weather conditions
*161 load boxes
162 ride a train to work
163 use knows to adjust a television picture
*164 move equipment and supplies
165 steer a large freighter though a busy harbor
*166 refer customers to appropriate contacts, either in the organization or elsewhere
167 calculate the adjustments for insurance claims
*168 prepare financial statements and budgets

*169 seasons food according to recipes or personal judgment
*170 write an essay in a foreign language
*171 type 30 words per minute
*172 apply general rules to specific problems and come up with logical answers
173 interpret a blood pressure reading
*174 retrieve data from computers
175 design a cleaner burning gasoline engine
*176 good knowledge of the written English language
177 choose the proper type of wood for adding a deck onto a house
*178 perform tasks such as bookkeeping
*179 roasts meats, fish, vegetables, and other foods
*180 ability to quickly analyze and resolve specific problems
181 learn information related to a complex and rapidly changing technology
*182 dial a phone
*183 carry out tasks such as word processing
184 understand what types of fertilizer are needed in order to grow better crops
185 compile data for a complex scientific report
*186 checks completed work for spelling, grammar, punctuation, and format
*187 edit a feature article in a local newspaper
188 determine what effects the destruction of the Berlin wall had on German culture
*189 operate telephone switchboards and consoles
*190 operate a commercial fishing boat
*191 resolve discrepancies on accounting records
192 develop a safe commercial cleaner
*193 set up mail-merge functions in a computer program to enable multiple letters to be personalized
*194 set up a new computer system for a large multinational company
*195 assist other staff in the company with their questions
196 investigate the effects of pollution on marine plants and animals
*197 clean food preparation equipment
198 repair an air conditioner
*199 contact special interest groups (examples: property owners, clubs, governmental inspectors or regulators)
*200 tell a story in a foreign language
*201 arrange appointments for callers/customers
*202 quickly carry out arm-hand movements
*203 hover a helicopter in strong wind
*204 constructs artistic forms from metal or stone
*205 draw up plans of action
*206 observes and tests food to determine that it is cooked
*207 work with those outside of the organization
208 determine the age of bones for placing them in fossil history
209 send out accounts
210 determine which distribution company to utilize in order to ship a product most cost effectively
*211 bakes bread, rolls, cakes, and pastry
212 supervise an appliance assembly line

*213 develops drawings for use in publications and exhibits
214 determine the best way to dispose of hazardous chemicals
*215 ability to evaluate customer satisfaction
216 keep up with price changes in a small retail store
217 determine whether to remove a tree that has been damaged
218 establish the value of a recently discovered art work
*219 design a set to be used for a play
*220 contact middle management or staff personnel as part of the job (examples: department heads, administrators)
*221 keep supervisor informed
222 determine the quality of service received at a restaurant
223 tabulate the costs of parcel deliveries
*224 store data on computers
*225 sketches artwork
226 calculate water pressure through a pipe
*227 attend a popular music concert
*228 sort incoming mail
*229 contact outside specialists as part of job (examples: technicians, photographers, law enforcement agents)
230 interpret a complex experiment in physics for general audiences
*231 strong editing skills
*232 answer billing questions from customers
233 fix a plumbing leak in the ceiling
234 explain how employment laws can affect a small company
235 use the proper concentration of chlorine to purify a water source
*236 represent the company in dealings with other companies, banks, and shareholders
*237 encourage the efforts of others
*238 make alterations to information already stored on a computer
*239 understand written sentences and paragraphs
*240 key in codes necessary to process information
*241 able to help customers in a friendly and efficient manner
242 perform tasks such as banking

* . Item used in analyses.

Appendix C

Description of Secretary position

Secretaries perform routine clerical and administrative functions including, but not limited to: drafting correspondence, scheduling appointments, organizing and maintaining paper and electronic files, or providing information to callers.

Appendix D

Description of Holistic items

Holistic measures are those methods that require the decision-maker to directly assign over-all values to each outcome. An example of a holistic question is:

How important is THINKING CREATIVELY to the position of Secretary?

(Thinking Creatively: Developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions.)

Appendix E

Description of the Decomposed measures

Decomposed measures are those methods that divide the evaluation task into a set of simple subtasks to which the decision-maker must assign values. Decomposed questions involve smaller amounts of information than holistic questions. An example of a decomposed question is:

How important is it that a Secretary be able to schedule appointments?

Footnotes

1. As of April 18, 2003, I was only able to collect data from 8 Secretaries. As mentioned in the addendum to my proposal, I am defending my results without the data from the Secretarial sample.

Table 1

Percent agreement from pilot test for each item

Item	Percent Agreement
A. *English Language – Knowledge of the structure and content of the English language including the meaning and spelling of words, rules of composition, and grammar*.	
A1. checks completed work for spelling, grammar, punctuation, and format	84.0
A2. strong editing skills	80.0
A3. good spelling skills	80.0
A4. good knowledge of the spoken English language	96.0
A5. edit a feature article in a local newspaper	80.0
A6. teach a college English class	92.0
A7. understand written sentences and paragraphs	80.0
A8. understand information spoken in English	88.0
A9. use English words or sentences in speaking so others can understand	92.0
A10. understand oral instructions	80.0
A11. good knowledge of the written English language	88.0
B. *Foreign Language – Knowledge of the structure and content of a foreign (non-English) language including the meaning and spelling of words, rules of composition and grammar, and pronunciation.*	
B1. write an essay in a foreign language	96.0
B2. write a newspaper article in a foreign language	100.0
B3. tell a story in a foreign language	100.0
B4. understand someone with a heavy foreign accent	80.0
B5. say “please” and “thank you” in a foreign language	84.0

Item	Percent Agreement
B6. ask directions in a foreign city	84.0
B7. write an English language review of a book written in a foreign language	80.0
C. *Clerical – Knowledge of administrative and clerical procedures and systems such as word processing, managing files and records, stenography and transcription, designing forms, and other office procedures and terminology.*	
C1. sort incoming mail	84.0
C2. type and re-arrange information, such as highlighting parts of text, moving paragraphs from one page to another, and putting information into columns	80.0
C3. perform duties such as data entry	80.0
C4. type 30 words per minute	80.0
C5. organize information to be mailed out	80.0
C6. dispatch outgoing mail	80.0
C7. file papers and documents so that they can be easily found when needed	92.0
C8. prepare master copies	84.0
C9. gathers and arranges material to be typed, following instructions	84.0
C10. carry out tasks such as word processing	80.0
C11. address envelopes or prepare envelope labels, using typewriter or computer	80.0
C12. input information in prescribed formats such as reports, mailing lists, accounts, payrolls, and staff records or scientific information	80.0
C13. maintain records	88.0
C14. file letters alphabetically	80.0
C15. perform tasks such as bookkeeping	80.0
C16. use word processing equipment	80.0
D. *Scheduling work and activities – Scheduling events, programs, and activities, as well as the work of others*	
D1. prepare the work schedule for subordinates	80.0
D2. maintain calendar and coordinates conferences and meetings	80.0
D3. draw up plans of action	80.0

Item	Percent Agreement
D4. organize a work schedule that is repetitive and easy to plan	80.0
D5. supervise and coordinate activities of staff	84.0
D6. schedule meetings with other people	80.0
D7. arrange appointments for callers/customers	80.0
D8. organize business itineraries, travel arrangements, conferences, meetings, and social functions	80.0
D9. make appointments using a predetermined schedule	84.0
E. *Telecommunications – Knowledge of transmission, broadcasting, switching, control, and operation of telecommunication systems.*	
E1. operate telephone switchboards and consoles	88.0
E2. dial a phone	88.0
E3. transfer telephone callers from one extension to another	88.0
E4. connect/disconnect telephone calls	84.0
E5. place telephone callers on hold	84.0
E6. develop a new, worldwide telecommunications network	80.0
F. *Customer and Personal Service – Knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction.*	
F1. able to help customers in a friendly and efficient manner	84.0
F2. handle and address customer complaints	84.0
F3. ability to assess the needs of customers	80.0
F4. good interpersonal skills when dealing with customers or clients	80.0
F5. answer billing questions from customers	88.0
F6. ability to evaluate customer satisfaction	80.0
F7. provide information to assist clients	80.0
F8. refer customers to appropriate contacts, either in the organization or elsewhere	80.0

Item	Percent Agreement
G. *Communication with people outside the organization – Communicating with people outside the organization, representing the organization to customers, the public, government, and other external sources. This information can exchanged in person, in writing, by telephone, or in e-mail.*	
G1. work with those outside of the organization	84.0
G2. contact outside specialists as part of job (examples: technicians, photographers, law enforcement agents)	80.0
G3. contact special interest groups (examples: property owners, clubs, governmental inspectors or regulators)	88.0
G4. ability to interact in a friendly manner with those outside of the organization	88.0
G5. supervise people who are not employees	80.0
G6. represent the company in dealings with other companies, banks, and shareholders	80.0
G7. contact outside professionals as part of job (examples: doctors, lawyers, professors, engineers, consultants)	80.0
G8. deal in a tactful manner with others outside the organization	80.0
H. *Communicating with supervisors, peers, or subordinates – Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person.*	
H1. act as an advisor to superiors	80.0
H2. report the results of a sales meeting to a supervisor	80.0
H3. contact executives and other high officials as part of the job (examples: corporation vice-presidents, government officials)	92.0
H4. assist other staff in the company with their questions	80.0
H5. contact middle management or staff personnel as part of the job (examples: department heads, administrators)	88.0
H6. keep superiors informed of any potential problems	80.0
H7. keep supervisor informed	84.0
H8. discuss issues and problems with others in the organization	84.0
H9. contact supervisors as part of the job (examples: foremen, first-level supervisors)	84.0
I. *Establishing and maintaining interpersonal relationships – Developing constructive and cooperative working relationships with others, and maintaining them over time.*	
I1. ability to develop working relationships with others	80.0

I2. maintain good working relationships with almost all coworkers and clients	80.0
I3. ability to maintain working relationships over time	80.0
I4. ability to interact in a friendly manner with others in the organization	80.0
I5. settle disputes among others	80.0
I6. encourage the efforts of others	80.0
I7. exchange greetings with a coworker	80.0

J. *Mathematics – Knowledge of arithmetic, algebra, geometry, calculus, statistics, and their applications*

J1. knowledge of calculus	88.0
J2. add two numbers	80.0
J3. knowledge of geometry	100.0
J4. derive a complex mathematical equation	100.0
J5. knowledge of arithmetic	96.0
J6. add, subtract, multiply, divide, and manipulate numbers quickly and accurately	88.0
J7. perform basic math, such as finding percentages and taking square roots	92.0
J8. knowledge of algebra	100.0
J9. knowledge of statistics	84.0
J10. knowledge of how to apply various types of mathematics	84.0

K. *Economics and Accounting – Knowledge of economic and accounting principles and practices, the financial markets, banking and the analysis and reporting of financial data.*

K1. send out accounts	80.0
K2. handle money	80.0
K3. prepare financial statements and budgets	84.0
K4. keep a major corporation's financial records	80.0
K5. fill out income tax returns	80.0
K6. resolve discrepancies on accounting records	84.0
K7. arranging for the payment of dividends, and observing all legal requirements	80.0

Item	Percent Agreement
L. *Working with computers – Using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information.*	
L1. print out address labels using a computer	80.0
L2. ability to work with computers	96.0
L3. print out letters using a computer	80.0
L4. set up mail-merge functions in a computer program to enable multiple letters to be personalized	80.0
L5. store data on computers	92.0
L6. set up spreadsheets	80.0
L7. enter employee information into a computer database	80.0
L8. check messages sent to and from computers	80.0
L9. retrieve data from computers	92.0
L10. write software for keeping track of parts in inventory	88.0
L11. make alterations to information already sorted on a computer	84.0
L12. use keyboard devices (examples: calculators, typewriters, computers)	80.0
L13. set up a new computer system for a large multinational company	88.0
L14. conduct research on the internet	84.0
L15. stores completed documents on computer hard drives or data storage mediums, such as a disk	80.0
L16. learn new computer programs to perform work more effectively	80.0
L17. set up mail-merge functions in a computer program to enable multiple letters to be directed to individuals in a number of locations	80.0
L18. key in codes necessary to process information	80.0
M. *Making decisions and solving problems – Analyzing information and evaluating results to choose the best solution and solve problems.*	
M1. resolve conflicting findings	80.0
M2. ability to quickly analyze and resolve specific problems	80.0
M3. possess strong problem-solving skills	88.0
M4. apply general rules to specific problems and come up with logical answers	84.0
M5. able to solve problems	80.0

Item	Percent Agreement
N. *Handling and moving objects – Using hands and arms in handling, installing, positioning, and moving materials, and manipulating things.*	
N1. place, move, or assemble objects	80.0
N2. quickly carry out arm-hand movements	80.0
N3. load boxes	92.0
N4. move light objects on occasion	84.0
N5. move equipment and supplies	80.0
O. *Operating vehicles, mechanized devices, or equipment – Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or watercraft.*	
O1. drive a car	84.0
O2. operate a commercial fishing boat	80.0
O3. navigate a helicopter	88.0
O4. operate a bulldozer	80.0
O5. hover a helicopter in strong wind	92.0
O6. fly a jet	92.0
O7.navigate a speedboat	96.0
O8.drive an 18-wheel tractor-trailer	96.0
O9. maneuver water craft	88.0
O10. operate a forklift	84.0
P. *Fine arts – Knowledge of the theory and techniques required to compose, produce, and perform works of music, dance, visual arts, drama, and sculpture.*	
P1. compose music	96.0
P2. develops drawings for use in publications and exhibits	84.0
P3. paints scenic backgrounds	96.0
P4. design a set to be used for a play	80.0
P5. constructs artistic forms from metal or stone	80.0
P6. choreograph dance routines	92.0

Item	Percent Agreement
P7. sketches artwork	96.0
P8. paints murals	96.0
P9. act in a play	84.0
P10. attend a popular music concert	80.0
P11. play a minor role in a local theater play	80.0
P12. design an artistic display for a major trade show	84.0
Q. *Food production – Knowledge of techniques and equipment for planting, growing, and harvesting food products, (both plant and animal) for consumption, including storage/handling techniques.*	
Q1. determine the meal selection for a cafeteria	80.0
Q2. run a 100,000 acre farm	80.0
Q3. roasts meats, fish, vegetables, and other foods	88.0
Q4. feed domestic animals	80.0
Q5. clean food preparation equipment	80.0
Q6. washes, peels, cuts, and seeds fruits and vegetables	92.0
Q7. carve meat	96.0
Q8. observes and tests food to determine that it is cooked	92.0
Q9. cut and de-bone meat prior to cooking	84.0
Q10.bakes bread, rolls, cakes, and pastry	88.0
Q11. butchers animals	84.0
Q12. cook food	80.0
Q13. seasons food according to recipes or personal judgment	80.0

Table 2

Correlation matrix for Item A across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	A1N	A2N	A3N	A4N	A5N	A6N	A7N	A8N	A9N	A10N	A11N
AN	4.73	.60	0.16	0.19	.39**	.56**	0	0.06	0.09	.45**	.44**	.35**	.32*
A1N	4.55	.78		.60**	.47**	.42**	0.11	0.07	.77**	0.18	.32*	.33*	.58**
A2N	4.16	1.05			.31*	0.26	0.24	0.25	.38**	0.05	.30*	0.23	.32*
A3N	4.39	0.83				.40**	0.19	0.05	.45**	.33**	.20**	.27**	.36*
A4N	4.61	0.70					0.06	0	.48**	.62**	.43**	.31**	.50**
A5N	1.54	0.98						.28*	0.14	0.02	0.02	0.12	0.14
A6N	1.08	0.34							0.04	-0.06	-0.04	-0.06	0.05
A7N	4.56	0.73								0.13	.34*	.28*	.56**
A8N	4.55	0.87									.41**	.28**	0.13
A9N	4.59	0.67										.30**	.44**
A10N	4.56	0.81											0.23
A11N	4.58	0.75											

	AL	A1L	A2L	A3L	A4L	A5L	A6L	A7L	A8L	A9L	A10L	A11L
AN	0.07	.34**	0.09	.35**	.46**	0.03	0.08	.37**	.42**	0.26	.27**	.48**
A1N	.46**	.86**	.75**	.71**	.66**	0.15	-0.02	.68**	.82**	.67**	.37**	.53**
A2N	0.27	.50**	.60**	.50**	.36*	0.21	0.13	.42**	.51**	.47**	.32*	0.24
A3N	.15*	.53**	.49**	.68**	.34**	0.07	-0.03	.37**	.28**	.64**	.42**	.34**
A4N	0.02	.51**	.44**	.53**	.69**	0.09	-0.02	.56**	.56**	.67**	.40**	.65**
A5N	0.04	0.19	0.25	0.19	0.13	0.86**	0.21	0.15	0.17	0.13	0.08	0.07
A6N	.12*	0.03	0.18	0.03	-0.01	0.24**	0.86**	-0.1	-0.04	0.06	-0.02	-0.03
A7N	.52**	.80**	.60**	.67**	.67**	0.17	-0.03	.73**	.76**	.72**	.34*	.55**
A8N	0.06	.31**	0.11	.34**	.46**	0.05	-0.01	.37**	.41**	0.18	.37**	.58**
A9N	.12*	.26**	0.23	.30**	.42**	0.06	-0.08	.34**	.38**	.46**	.30**	.41**
A10N	.17**	.36**	0.12	.35**	.30**	-0.2	-0.2	.31**	.23**	.33*	.45**	.25**
A11N	.39**	.54**	.42**	.50**	.45**	0.18	0.01	.43**	.60**	.49**	0.26	.39**

	M	SD	A1L	A2L	A3L	A4L	A5L	A6L	A7L	A8L	A9L	A10L	A11L
AL	2.73	1.34	0.09	.46**	.15**	0.05	.17**	-0.03	-0.01	0.11	.67**	.24**	-0.03
A1L	4.51	0.78		.78**	.76**	.55**	.20**	0.01	.56**	.45**	.79**	.49**	.49**
A2L	4.16	1.05			.71**	.67**	0.26	0.08	.67**	.65**	.67**	.40**	.45**
A3L	4.39	0.83				.55**	.15*	0.02	.57**	.46**	.74**	.52**	.50**
A4L	4.61	0.70					0.04	-0.03	.64**	.67**	.81**	.51**	.74**
A5L	1.54	0.98						0.26	0.06	0.08	0.22	0.07	0.03
A6L	1.08	0.34							-0.07	0	-0.03	-0.1	0.03
A7L	4.56	0.73								.48**	.84**	.47**	.56**
A8L	4.55	0.87									.74**	.32**	.58**
A9L	4.59	0.67										.46**	.59**
A10L	4.56	0.81											.38**
A11L	4.58	0.75											

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 3

Correlation matrix for Item B across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	B1N	B2N	B3N	B4N	B5N	B6N	B7N	BL	B1L	B2L	B3L	B4L	B5L	B6L	B7L
BN	2.76	1.13	.58**	.37**	.42**	.35**	.51**	.33**	.36**	.32**	.55**	.32**	.32**	.35**	.45**	.29*	.29**
B1N	1.25	0.66		.22	.23	.21	.51**	.41**	.40**	.68**	.69**	.57**	.46*	.17	.53**	.26	.73**
B2N	1.33	0.69			.24	.20**	.33**	.18**	.57**	.18**	.20	.47**	.43**	.22**	.28**	.12	.43**
B3N	1.10	0.36				.29*	.42**	.35*	.35*	.34*	.65**	.42**	.86**	.34*	.43**	.11	.54**
B4N	3.18	1.23					.50**	.25**	.17**	.08	.22	.20**	.24**	.68**	.49**	.14	.21**
B5N	2.53	1.27						.39**	.35**	.22**	.58**	.38**	.40**	.46**	.79**	.41**	.41**
B6N	1.50	0.87							.27**	.20**	.47**	.31**	.35**	.23**	.43**	.66**	.35**
B7N	1.22	0.60								.21**	.28	.68**	.63**	.19**	.31**	-.09	.56**
BL	1.49	.85									.62**	.24**	.15*	.06	.21**	.28*	.35**
B1L	1.22	.67										.35*	.66**	.41**	.54**	.33*	.79**
B2L	1.20	.58											.75**	.14*	.36**	.03	.57**
B3L	1.17	.51												.19**	.32**	.09	.60**

	M	SD	B1N	B2N	B3N	B4N	B5N	B6N	B7N	BL	B1L	B2L	B3L	B4L	B5L	B6L	B7L
B4L	3.10	1.26													.47**	.28*	.16**
B5L	2.50	1.36														.41**	.38**
B6L	1.33	.65															.17
B7L	1.31	.85															

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 4

Correlation matrix for Item C across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	C1N	C2N	C3N	C4N	C5N	C6N	C7N	C8N	C9N	C10N	C11N
CN	4.80	.59	0.22	.23**	.31**	0.08	.33**	.16**	.32**	.20**	.30**	.31*	.33**
C1N	4.25	.77		0.25	0.2	.35*	.65**	.49**	0.26	.50**	.45**	.47**	.72**
C2N	4.43	.82			.41**	0.26	.51**	.50**	.47**	.57**	.54**	.48**	.51**
C3N	4.39	.83				0.04	.43**	.36**	.40**	.39**	.43**	0.01	.44**
C4N	4.12	.99					0.24	0.2	0.05	0.27	0.27	.33*	.30*
C5N	4.39	.80						.53**	.52**	.52**	.57**	.61**	.64**
C6N	3.90	1.13							.46**	.51**	.44**	.34*	.64**
C7N	4.67	.68								..40**	.58**	.54**	.52**
C8N	4.07	1.04									.46**	.38**	.49**
C9N	4.52	.72										.44**	.52**
C10N	4.71	.54											.57**
C11N	4.32	.95											

	C12N	C13N	C14N	C15N	C16N	CL	C1L	C2L	C3L	C4L	C5L	C6L
CN	.28**	.25**	.23**	0.12	.33**	0.01	.19**	.30**	.29**	.16**	.27**	.26**
C1N	.28*	.36*	.43**	0.12	0.02	0.24	.51**	.46**	.51**	.42**	.55**	.53**
C2N	.31**	.46**	.48**	.48**	.42**	0.06	.56**	.69**	.47**	.38**	.58**	.53**
C3N	.39**	.43**	.37**	-0.02	.37**	0.06	.46**	.42**	.64**	.29**	.48**	.42**
C4N	0.22	.40**	0.22	.38**	0.06	0.13	.37**	0.19	0.26	.84**	.29*	.31*
C5N	.32**	.47**	.55**	.30*	.38**	0.07	.58**	.53**	.54**	.34**	.65**	.57**
C6N	.32**	.47**	.55**	.30*	.38**	0.07	.58**	.53**	.54**	.34**	.65**	.57**
C7N	.25**	.48**	.45**	.35*	.45**	0.11	.49**	.43**	.48**	.29**	.52**	.48**

C8N	.30**	.43**	.34**	0.23	0.22	0.01	.55**	.58**	.49**	.27**	.56**	.59**
C9N	.33**	.57**	.50**	.32*	.37**	.20**	.46**	.51**	.50**	.34**	.55**	.45**
C10N	0.22	0.18	.67**	.40**	.47**	.60**	.45**	.60**	.53**	0.26	.54**	.42**
C11N	.33**	.46**	.53**	0.22	.37**	0.06	.57**	.54**	.53**	.35**	.59**	.51**

	C7L	C8L	C9L	C10L	C11L	C12L	C13L	C14L	C15L	C16L
CN	.36**	.26**	.22**	0.25	.28**	.24**	0.11	.28**	0.09	0.19
C1N	.39**	.43**	.50**	.35*	.49**	.46**	0.19	.32*	.29*	.39**
C2N	.42**	.53**	.38**	.73**	.56**	.40**	.51**	.49**	.44**	.52**
C3N	.36**	.39**	.34**	0.21	.45**	.35**	0.24	.40**	-0.02	0.13
C4N	0.2	0.19	-0.01	.36**	.31*	0.07	0.21	0.21	.49**	0.24
C5N	.61**	.46**	.44**	.47**	.67**	.55**	0.25	.55**	0.23	.42**
C6N	.61**	.46**	.44**	.47**	.67**	.55**	0.25	.55**	0.23	.42**
C7N	.57**	.39**	.41**	.30*	.57**	.38**	0.13	.52**	.33*	.30*
C8N	.40**	.69**	.41**	.33*	.54**	.45**	.30*	.45**	.37**	0.22
C9N	.49**	.43**	.43**	.44**	.56**	.42**	0.26	.53**	.32*	.50**
C10N	.64**	0.17	.39**	.55**	.72**	.29*	.29*	.53**	.49**	.67**
C11N	.52**	.44**	.35**	.49**	.69**	.52**	0.26	.51**	0.24	.49**

	M	SD	C13N	C14N	C15N	C16N	CL	C1L	C2L	C3L	C4L	C5L	C6L
C12N	4.67	.65	.32**	.32**	.31*	.34**	-0.04	.31**	.37**	.36**	.23**	.35**	.29**
C13N	4.52	.75		.44**	.46**	.41**	0.06	.46**	.47**	.48**	.28**	.46**	.43**
C14N	4.47	.80			.43**	.36**	0.1	.48**	.47**	.46**	.36**	.54**	.48**
C15N	3.90	1.35				.29*	.34*	.40**	.43**	0.27	.39**	.36**	.40**
C16N	4.66	.73					0.08	.30**	.39**	.39**	.23**	.37**	.33**
CL	1.83	1.37						0.05	0.06	.12*	0.04	0.03	0.09
C1L	4.10	1.01							.57**	.60**	.47**	.75**	.76**

C2L	4.38	.88	.61**	.38**	.63**	.62**
C3L	4.36	.80		.36**	.58**	.61**
C4L	4.13	1.03			.42**	.41**
C5L	4.19	.89				.73**
C6L	3.97	1.12				
C7L	4.57	.77				
C8L	3.90	1.13				

	C7L	C8L	C9L	C10L	C11L	C12L	C13L	C14L	C15L	C16L
C12N	.28**	.34**	.28**	0.13	.32**	.33**	0.22	.28**	.39**	0.03
C13N	.43**	.40**	.30**	.54**	.47**	.37**	.47**	.44**	.45**	0.26
C14N	.50**	.37**	.40**	.40**	.55**	.42**	.32*	.70**	.40**	.60**
C15N	.36*	0.25	0.2	.35*	.29*	0.18	.39**	.32*	.55**	.30*
C16N	.41**	.27**	.27**	.41**	.43**	.29**	0.26	.38**	0.17	.64**
CL	0.08	0.05	.12*	.42**	.12*	.24**	0.24	0.1	.31*	.45**
C1L	.55**	.55**	.41**	.49**	.63**	.53**	.45**	.60**	.50**	.40**
C2L	.52**	.61**	.45**	.64**	.71**	.51**	.41**	.55**	.36**	.68**

C3L	.50**	.49**	.43**	.49**	.60**	.51**	.43**	.55**	.34*	.62**
C4L	.36**	.35**	.31**	.38**	.41**	.30**	0.21	.41**	.45**	0.28
C5L	.59**	.52**	.48**	.50**	.70**	.62**	.47**	.67**	.40**	.58**
C6L	.54**	.60**	.52**	.59**	.68**	.63**	.50**	.61**	.50**	.45**
C7L		.46**	.47**	.60**	.65**	.52**	.39**	.60**	0.25	.66**
C8L			.55**	.37**	.53**	.49**	.29*	.43**	0.18	0.22

	M	SD	C10L	C11L	C12L	C13L	C14L	C15L	C16L
C9L	4.19	1.04	.33*	.53**	.49**	0.12	.41**	0.25	.43**
C10L	4.57	.64		.63**	.30*	.56**	.43**	.42**	.71**
C11L	4.32	.89			.62**	.28*	.67**	0.23	.72**
C12L	4.17	1.06				0.21	.49**	0.2	.33*
C13L	4.27	.90					.29*	.37**	.39**
C14L	4.36	.93						0.22	.43**
C15L	3.76	1.24							.37**
C16L	4.57	.64							

**. Correlation is significant at the 0.01 level.
*. Correlation is significant at the 0.05 level.
0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 5

Correlation matrix for Item D across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	D1N	D2N	D3N	D4N	D5N	D6N	D7N	D8N	D9N	DL	D1L	D2L	D3L	D4L	D5L	D6L	D7L	D8L	D9L
DN	4.58	.75	.14*	.34**	.13	.26**	.17**	.45**	.07	.27**	.24**	.07	-.05	.20**	.04	.21	.19**	.30**	.23**	.22**	-.02
D1N	3.45	1.26		.18**	.28	.54**	.47**	.16**	-.01	.08	.15**	.19**	.51**	.17**	.31**	.53**	.35**	.19**	.04	.11*	.32*

D2N	4.56	.72			.24	.37**	.18**	.51**	.09	.58**	.62**	.02	.29*	.48*	.02	.43**	.24**	.53**	.56**	.44**	.42**
D3N	2.24	1.27				.41**	.45**	.06	.22	-.04	.26	.20	.08	.19	.64**	.35*	.35*	.26	.20	.16	.19
D4N	3.61	1.20					.40**	.29**	.18	.36**	.33**	.13*	.49**	.35**	.28**	.73**	.33**	.31**	.34**	.40**	.22
D5N	2.82	1.30						.11	-.10	.20**	.13*	.16**	.25	.13*	.42**	.45**	.48**	.15*	.09	.18**	.08
D6N	4.74	.59							.07	.42**	.41**	.11	.15	.41**	.01	.25	.18**	.46**	.40**	.35**	.46**
D7N	4.43	.92								-.08	.17	.31*	.25	.41**	.04	.13	.10	.38**	.35*	.13	.20
D8N	4.28	.96									.47**	.04	.16	.39**	.06	.22	.22**	.40**	.36**	.46**	.23
D9N	4.35	.89										.06	.24	.40**	.06	.46**	.23**	.46**	.58**	.32**	.44**
DL	3.11	1.13											.32*	.12*	.31**	.23	.22**	.07	.04	.11*	.32*
D1L	3.63	1.15												.28*	.28*	.47**	.34*	.23	.30*	.42**	.21
D2L	4.03	1.02													.13*	.33*	.26*	.56**	.53**	.51**	.60**
	M	SD	D1N	D2N	D3N	D4N	D5N	D6N	D7N	D8N	D9N	DL	D1L	D2L	D3L	D4L	D5L	D6L	D7L	D8L	D9L
D3L	2.10	1.21														.45**	.42**	.10	.09	.20**	.14
D4L	3.82	1.29															.39**	.38**	.38**	.38**	.29*
D5L	2.85	1.33																.19**	.22**	.36**	.27
D6L	4.49	.83																	.68**	.47**	.46**
D7L	4.08	.75																		.44**	.44**

D8L	4.22	1.08	.43**
D9L	4.08	1.01	

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 6

Correlation matrix for Item E across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	E1N	E2N	E3N	E4N	E5N	E6N	EL	E1L	E2L	E3L	E4L	E5L	E6L
EN	3.52	1.33	.45**	.02	.05	.12*	.05	.15**	.24**	.48**	.05	.12*	.25	.06	.15**
E1N	3.69	1.48		.26	.18	.33*	.26	.09	.43**	.82**	.31*	.45**	.29*	.29**	.12
E2N	4.76	.59			.62**	.56**	.55**	-.19	.10	.21	.68**	.80**	.77**	.69**	-.08
E3N	4.45	.85				.63**	.42**	-.04	-.03	.09	.53**	.63**	.57**	.57**	-.01
E4N	4.27	1.02					.53**	-.06	-.02	.23	.50**	.62**	.75**	.60**	.02
E5N	4.17	1.12						-.01	-.05	.12	.43**	.48**	.59**	.59**	.06
E6N	1.20	.70							.20**	.19	-.17**	-.02	-.09	-.03	.61**
EL	1.81	1.21								.50**	-.12*	.01	.26	.05	.17**
E1L	3.75	1.44									.34*	.44**	.29*	.27	.00
E2L	4.64	.72										.64**	.87**	.55**	-.09
E3L	4.42	.85											.82**	.69**	.00
E4L	4.51	.76												.86**	-.03
E5L	4.24	1.01													.08
E6L	1.27	.80													

**. Correlation is significant at the 0.01 level. *. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 7

Correlation matrix for Item F across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	F1N	F2N	F3N	F4N	F5N	F6N	F7N	F8N	FL	F1L	F2L	F3L	F4L	F5L	F6L	F7L	F8L
FN	4.2	.98	-.10	.45**	.45**	.53**	.10	.33*	.43**	-.11	.08	.07	.42**	.43**	.33**	.44**	.19	.34**	.31**
F1N	4.5	.97		.57**	.47**	.34*	.52**	.46**	.18	.51**	.66**	.69**	.65**	.73**	.49**	.62**	.49**	.63**	.40**
F2N	3.8	1.1			.65**	.45**	.49**	.69**	.43**	.27	.19**	.57**	.61**	.59**	.49**	.60**	.59**	.48**	.50**
F3N	4.0	1.1				.55**	.36**	.63**	.50**	.20	.19**	.53**	.50**	.59**	.47**	.61**	.47**	.50**	.39**
F4N	4.7	.68					.31*	.42**	.60**	.19	.15*	.51**	.38**	.42**	.51**	.41**	.43**	.45**	.47**
F5N	3.1	1.2						.46**	.17	.18	.44**	.47**	.61**	.62**	.32*	.59**	.49**	.51**	.32*
F6N	3.4	1.1							.61**	.16	.46**	.57**	.61**	.72**	.39**	.68**	.63**	.57**	.42**
F7N	4.3	.88								.10	.22**	.33*	.31**	.35**	.35**	.37**	.31*	.46**	.37**
F8N	4.2	.99									.55**	.51**	.42**	.46**	.43**	.28*	.24	.46**	.30*
FL	3.4	1.1										.69**	.22**	.23**	.13*	.27**	.48**	.21**	.13*
F1L	4.1	1.0											.59**	.75**	.64**	.63**	.59**	.79**	.66**

	M	SD	F1N	F2N	F3N	F4N	F5N	F6N	F7N	F8N	FL	F1L	F2L	F3L	F4L	F5L	F6L	F7L	F8L
F2L	3.7	1.2												.61**	.52**	.66**	.63**	.58**	.54**
F3L	3.8	1.1													.52**	.74**	.75**	.57**	.50**
F4L	4.2	.97														.53**	.49**	.65**	.67**
F5L	3.6	1.1															.63**	.55**	.51**
F6L	3.2	1.2																.60**	.49**
F7L	3.9	.99																	.64**

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 8

Correlation matrix for Item G across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	G1N	G2N	G3N	G4N	G5N	G6N	G7N	G8N	GL	G1L	G2L	G3L	G4L	G5L	G6L	G7L	G8L
GN	4.39	.92	.39**	.19	.12	.37**	.16**	.36**	.24**	.34**	.07	.35**	.24**	.33*	.35**	.37**	.31*	.31*	.31**
G1N	3.45	1.39		.57**	.27	.72**	.20	.43**	.47**	.37**	.57**	.73**	.61**	.61**	.56**	.36**	.27	.63**	.57**
G2N	3.73	1.17			.62**	.53**	.14	.45**	.62**	.30*	.61**	.53**	.59**	.69**	.59**	.40**	.47**	.70**	.45**
G3N	3.33	1.21				.44**	.10	.51**	.53**	.37**	.31*	.37**	.46**	.60**	.42**	.34*	.51**	.48**	.29*
G4N	4.03	1.07					.23**	.31*	.48**	.61**	.00	.60**	.55**	.46**	.69**	.25	.21	.57**	.54**
G5N	2.15	1.12						.08	.21**	.22**	.03	.30**	.21**	.21	.26**	.78**	.17	.10	.28**
G6N	2.82	1.48							.45**	.40**	.44**	.42**	.30*	.56**	.41**	.35*	.82**	.47**	.44**
G7N	3.76	1.17								.46**	.05	.58**	.63**	.52**	.53**	.28*	.39**	.54**	.48**
G8N	3.95	1.17									.01	.58**	.42**	.39**	.64**	.44**	.37**	.51**	.58**
GL	1.66	1.27										.08	.05	.48**	.07	.43**	.34*	.62**	.03
G1L	3.70	1.17											.62**	.67**	.73**	.41**	.34*	.74**	.66**

	M	SD	G1N	G2N	G3N	G4N	G5N	G6N	G7N	G8N	GL	G1L	G2L	G3L	G4L	G5L	G6L	G7L	G8L
G2L	3.62	1.25												.48**	.57**	.31*	.18	.53**	.51**
G3L	3.39	1.30													.61**	.48**	.53**	.72**	.60**
G4L	4.06	1.02														.43**	.42**	.60**	.70**
G5L	2.27	1.37															.38**	.37**	.52**
G6L	2.96	1.52																.40**	.41**
G7L	3.78	1.10																	.61**
G8L	3.73	1.27																	

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 9

Correlation matrix for Item H across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	H1	H2N	H3N	H4N	H5N	H6N	H7N	H8N	H9N	HL	H1L	H2L	H3L	H4L	H5L	H6L	H7L	H8L	H9L
HN	4.8	.49	.09	.15*	.24**	.02	.13	.30**	.22	.29**	.42**	.04	.15**	-.01	.27**	.18**	.34**	.30**	.01	.20	.31**
H1N	2.4	1.3		.33**	.26**	.54**	.28*	.26**	.16	.34**	.20**	.09	.71**	.38**	.28**	.42**	.30**	.28**	.19	.38**	.26**
H2N	3.5	1.2			.44**	.14	.33*	.42**	.50**	.45**	.24**	.11	.38**	.55**	.42**	.48**	.49**	.46**	.45**	.40**	.38**
H3N	4.0	1.2				.42**	.63**	.27**	.44**	.37**	.42**	.12*	.32**	.33*	.64**	.38**	.58**	.39**	.51**	.45**	.55**
H4N	3.5	1.1					.27	.45**	.31*	.47**	.27	.23	.58**	.27	.43**	.61**	.37**	.35*	.27	.63**	.36**
H5N	4.2	1.1						.37**	.58**	.18	.48**	.53**	.31*	.17	.51**	.39**	.53**	.44**	.54**	.35*	.57**
H6N	4.2	.92							.59**	.44**	.37**	.08	.34**	.40**	.35**	.40**	.46**	.59**	.45**	.65**	.38**
H7N	4.6	.76								.39**	.50**	.64**	.33*	.28*	.57**	.33*	.70**	.68**	.66**	.46**	.58**
H8N	3.5	1.2									.35**	.02	.40**	.35*	.44**	.48**	.43**	.44**	.29*	.53**	.39**
H9N	4.3	.83										.05	.26**	.34*	.48**	.32**	.48**	.40**	.56**	.42**	.50**
HL	1.8	1.4											.11	.26	.16**	.09	.14*	.11	.57**	.39**	.14*

	M	SD	H2N	H3N	H4N	H5N	H6N	H7N	H8N	H9N	HL	H1L	H2L	H3L	H4L	H5L	H6L	H7L	H8L	H9L
H1L	2.6	1.3											.35*	.40**	.52**	.42**	.40**	.39**	.53**	.36**
H2L	3.1	1.5												.40**	.33*	.32*	.39**	.37**	.40**	.36*
H3L	3.8	1.2													.50**	.69**	.39**	.69**	.51**	.73**
H4L	3.3	1.1														.54**	.48**	.50**	.53**	.50**
H5L	3.9	1.0															.62**	.83**	.63**	.76**
H6L	3.8	1.1																.76**	.61**	.57**
H7L	4.2	.95																	.50**	.68**
H8L	3.2	1.3																		.49**
H9L	3.9	1.1																		

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 10

Correlation matrix for Item I across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	I1N	I2N	I3N	I4N	I5N	I6N	I7N	IL	I1L	I2L	I3L	I4L	I5L	I6L	I7L
IN	4.14	.90	.50**	.51**	.48**	.46**	.32**	.24	.43**	.14*	.47**	.45**	.45**	.30*	.37**	.31*	.30*
I1N	4.12	.96		.60**	.70**	.51**	.37**	.39**	.55**	.17**	.66**	.67**	.84**	.61**	.37**	.59**	.63**
I2N	4.40	.83			.58**	.57**	.36**	.23	.53**	.18**	.58**	.66**	.56**	.56**	.35**	.53**	.43**
I3N	3.96	1.11				.53**	.33**	.36**	.49**	.08	.64**	.60**	.47**	.25	.32**	.41**	.37**
I4N	4.54	.74					.23**	.17	.44**	.16**	.45**	.45**	.62**	.51**	.26**	.33*	.34*
I5N	2.44	1.07						.48**	.44**	.15**	.35**	.37**	.34*	.32*	.68**	.78**	.25
I6N	2.92	1.26							.44**	.23	.24	.18	.28*	.13	.55**	.61**	.48**
I7N	3.77	1.08								.13*	.52**	.56**	.63**	.43**	.42**	.63**	.67**
IL	3.02	1.18									.17**	.21**	.84**	.73**	.25**	.48**	.56**
I1L	3.89	1.09										.76**	.79**	.66**	.45**	.51**	.55**
I2L	4.06	.98											.72**	.82**	.44**	.45**	.44**
I3L	3.94	1.12												.78**	.43**	.54**	.57**

	M	SD	I1N	I2N	I3N	I4N	I5N	I6N	I7N	IL	I1L	I2L	I3L	I4L	I5L	I6L	I7L
I4L	3.94	1.16													.37**	.44**	.46**
I5L	2.55	1.15														.75**	.47**
I6L	3.24	1.29															.58**
I7L	3.57	1.19															

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 11

Correlation matrix for Item J across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	J1N	J2N	J3N	J4N	J5N	J6N	J7N	J8N	J9N	J10N
JN	2.84	1.01	.54**	.47**	.46**	.40**	.44**	.54**	.51**	.54**	.47**	.57**
J1N	1.74	.92		.35**	.57**	.55**	.33**	.47**	.43**	.61**	.54**	.52**
J2N	3.95	1.19			.36**	.20**	.63**	.63**	.68**	.47**	.46**	.54**
J3N	1.60	.80				.51**	.37**	.40**	.48**	.70**	.48**	.53**
J4N	1.34	.70					.29**	.32**	.40**	.56**	.50**	.47**
J5N	3.68	1.21						.63**	.68**	.51**	.43**	.62**
J6N	3.75	1.12							.65**	.53**	.45**	.67**
J7N	3.31	1.33								.62**	.53**	.63**
J8N	2.22	1.22									.64**	.67**
J9N	2.09	1.11										.53**
J10N	2.86	1.24										
JL	1.44	.85										
J1L	1.73	1.00										

	JL	J1L	J2L	J3L	J4L	J5L	J6L	J7L	J8L	J9L	J10L
JN	.19**	.29*	.48**	.37**	.32**	.50**	.56**	.52**	.54**	.51**	.61**
J1N	.20**	.57**	.39**	.50**	.51**	.40**	.47**	.42**	.65**	.60**	.60**
J2N	.16**	.41**	.81**	.40**	.20**	.63**	.63**	.68**	.47**	.46**	.54**
J3N	.13*	.63**	.41**	.71**	.38**	.45**	.50**	.45**	.62**	.48**	.51**
J4N	.14*	.73**	.28**	.49**	.32**	.32**	.41**	.34**	.52**	.51**	.56**
J5N	.17**	.45**	.67**	.43**	.27**	.75**	.67**	.67**	.55**	.47**	.68**
J6N	.25**	.52**	.66**	.54**	.29**	.62**	.79**	.69**	.59**	.49**	.72**
J7N	.15*	.48**	.73**	.47**	.32**	.72**	.76**	.80**	.65**	.59**	.66**
J8N	.22**	.57**	.54**	.51**	.42**	.60**	.62**	.59**	.83**	.65**	.68**
J9N	.17**	.56**	.47**	.38**	.41**	.50**	.54**	.53**	.62**	.81**	.57**
J10N	.24**	.40**	.60**	.34*	.36**	.64**	.73**	.65**	.67**	.56**	.60**
JL		.43**	.16**	.44**	.30**	.16**	.22**	.17**	.24**	.18**	.67**
J1L			.38**	.73**	.61**	.43**	.54**	.47**	.73**	0.25	.66**

	M	SD	J3L	J4L	J5L	J6L	J7L	J8L	J9L	J10L
J2L	3.84	1.23	.45**	.21**	.70**	.74**	.77**	.58**	.49**	.51**
J3L	1.63	.96		.57**	.57**	.55**	.54**	.66**	.37**	.57**
J4L	1.30	.71			.33**	.36**	.30**	.44**	.46**	.45**
J5L	3.04	1.30				.75**	.77**	.66**	.57**	.71**
J6L	3.42	1.23					.80**	.68**	.58**	.74**
J7L	3.22	1.27						.67**	.57**	.62**
J8L	2.38	1.24							.65**	.76**
J9L	2.12	1.11								.40**
J10L	2.59	1.19								

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 12

Correlation matrix for Item K across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	K1N	K2N	K3N	K4N	K5N	K6N	K7N	KL	K1L	K2L	K3L	K4L	K5L	K6L	K7L
KN	2.85	1.15	.12*	.45**	.42**	.35**	.46**	.35*	.49**	.32**	.35**	.47**	.49**	.39**	.42**	.38**	.41**
K1N	3.20	1.28		.12*	-.14	-.03	.00	-.16	.03	.06	-.11*	.15**	.03	.02	.03	-.00	-.13
K2N	3.20	1.28			.50**	.44**	.46**	.51**	.41**	.37**	.40**	.80**	.54**	.40**	.48**	.47**	.40**
K3N	2.70	1.32				.46**	.34*	.66**	.42**	.56**	.49**	.52**	.68**	.45**	.46**	.58**	.60**
K4N	2.27	1.31					.32**	.44**	.42**	.21**	.43**	.42**	.57**	.71**	.43**	.48**	.35*
K5N	3.25	1.57						.48**	.39**	.28**	.38**	.52**	.45**	.22	.66**	.40**	.47**
K6N	2.42	1.32							.28*	.60**	.40**	.42**	.48**	.13	.44**	.46**	.54**
K7N	2.43	1.27								.34**	.40**	.42**	.48**	.13	.44**	.46**	.54**
KL	2.12	1.05									.25**	.36**	.32**	.39**	.31**	.26**	.56**
K1L	3.45	1.22										.44**	.51**	.28*	.39**	.53**	.54**
K2L	2.54	1.26											.67**	.31*	.61**	.52**	.36**
K3L	2.60	1.29												.52**	.62**	.60**	.51**

	M	SD	KN	K1N	K2N	K3N	K4N	K5N	K6N	K7N	KL	K1L	K2L	K3L	K4L	K5L	K6L	K7L
K4L	2.84	1.42														.19	$.50^{**}$	$.43^{**}$
K5L	2.15	1.24															$.51^{**}$	.34*
K6L	2.64	1.28																$.56^{**}$
K7L	2.12	1.26																

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 13

Correlation matrix for Item L across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	L1N	L2N	L3N	L4N	L5N	L6N	L7N	L8N	L9N	L10N	L11N
LN	4.31	.92	.27**	.42**	.27**	.34*	.40**	.23**	.33**	.22**	0.2	.19**	0.2
L1N	4.04	1.04		.25**	.53**	.42**	0.27	.40**	.45**	.38**	.33*	0.07	.33*
L2N	4.52	.77			.39**	.30*	.43**	.33**	.49**	.36**	.29*	0.06	0.25
L3N	4.49	.81				.28*	0.23	.40**	.50**	.44**	.55**	0	.42**
L4N	3.90	.98					.45**	.48**	.64**	.42**	.42**	.46**	.45**
L5N	4.53	.61						0.15	.51**	.33*	.36**	.29*	.39**

v

	L12N	L13N	L14N	L15N	L16N	L17N	L18N	LL	L1L	L2L	L3L	L4L
LN	.30**	0.07	.23**	.25**	.29**	.34**	0.22	.35**	.21**	.39**	.27**	.29*
L1N	.37**	0.19	.26**	.42**	.30**	.27**	0.27	.27**	.62**	.25**	.53**	.53**
L2N	.43**	-0.13	.24**	.41**	.31**	.23**	0.27	.26**	.35**	.55**	.37**	.30*
L3N	.46**	0.07	.27**	.59**	.31**	.17**	0.09	.40**	.58**	.37**	.66**	.35*
L4N	0.24	.29*	.28*	.54**	.53**	.58**	.70**	.38**	.47**	.38**	.55**	.64**
L5N	.45**	0.24	0.1	0.28	.53**	.30*	.44**	.66**	0.28	.64**	.38**	.47**

v

	L5L	L6L	L7L	L8L	L9L	L10L	L11L	L12L	L13L	L14L	L15L	L16L
LN	.34**	.28**	0.24	.32**	0.18	.17**	.19**	.31*	0.18	.27**	.28**	.28**
L1N	.44**	.43**	.41**	.43**	0.2	.15**	.38**	.63**	0.16	.30*	.41**	.32**
L2N	.50**	.39**	0.27	.39**	.33*	.12*	.33**	.29*	-0.23	0.07	.37**	.40**
L3N	.51**	.50**	.36*	.44**	0.21	0.09	.47**	.50**	0.21	0.23	.45**	.40**
L4N	.56**	.46**	.65**	.57**	.50**	.33*	.33*	.45**	0.19	.41**	.55**	0.27
L5N	.67**	0.21	.48**	.48**	.53**	0.21	.47**	.49**	0.27	0.23	.58**	.56**

	L17L	L18L
LN	.29**	.21**
L1N	.32**	.31**
L2N	.23**	0.27
L3N	.30**	.31**
L4N	.72**	.60**
L5N	.62**	.49**

	M	SD	L7N	L8N	L9N	L10N	L11N	L12N	L13N	L14N	L15N	L16N	L17N
L6N	4.22	.96	.50**	.29**	0.09	0.05	0.04	.30**	0.19	.27**	.47**	.33**	.23**
L7N	4.31	.87		.41**	.48**	.12*	.48**	.45**	0.25	.32**	.53**	.37**	.29**
L8N	4.17	1.04			.35*	.14*	.47**	.32**	0.27	.25**	.46**	.41**	.33**
L9N	4.31	.86				.34*	.61**	.28*	0.27	.33*	0.27	.34*	0.18
L10N	4.74	.59					0.25	-0.03	.48**	.25**	0.01	0.08	.28**
L11N	2.92	1.26						.34*	.35*	0.21	.30*	.42**	0.08
L12N	4.66	1.17							-0.01	.23**	.31**	.31**	.13*
L13N	1.57	1.04								0.06	0.1	0.24	0.26
L14N	2.96	1.27									.21**	.31**	.26**
L15N	4.50	.79										.39**	.43**
L16N	1.34	.70											.37**
L17N	3.41	1.30											
L18N	3.94	1.16											
LL	4.06	.98											
L1L	4.13	1.03											

	L18N	LL	L1L	L2L	L3L	L4L	L5L	L6L	L7L	L8L	L9L
L6N	.34*	.24**	.43**	.29**	.47**	0.25	.45**	.71**	.31*	.39**	0.06
L7N	.53**	.28**	.53**	.45**	.59**	.68**	.65**	.51**	.66**	.54**	.66**
L8N	.30*	.32**	.39**	.35**	.42**	.53**	.48**	.38**	.44**	.63**	.52**
L9N	.40**	.34*	0.16	.38**	.32*	.39**	.52**	.52**	.47**	.54**	.64**
L10N	.31*	-0.02	0.07	0.08	0.02	.30*	0.07	0.1	0.26	.14*	.36**
L11N	.54**	0.26	0.24	.47**	.43**	.43**	.59**	0.26	.46**	.52**	.61**
L12N	0.15	.32**	.41**	.48**	.45**	.29*	.46**	.29**	.39**	.37**	.31*
L13N	0.24	0.11	-0.21	0.22	0.07	.29*	.31*	0.17	.29*	.32*	.31*
L14N	.32*	.23**	.31**	.31**	.24**	0.26	.31**	.29**	.28*	.30**	0.23
L15N	.39**	.34**	.48**	.42**	.57**	.45**	.58**	.48**	.45**	.43**	.48**
L16N	.44**	.28**	.32**	.41**	.41**	.50**	.47**	.33**	.47**	.49**	.55**
L17N	0.26	.18**	.24**	.25**	.21**	.62**	.29**	.24**	0.28	.35**	.32*
L18N		.28*	0.26	.44**	.45**	.53**	.58**	.44**	.58**	.55**	.60**
LL			.26**	.33**	.34**	.32*	.33**	.31**	.37**	.40**	.34*
L1L				.34**	.66**	.48**	.53**	.46**	.49**	.48**	.29*

	L10L	L11L	L12L	L13L	L14L	L15L	L16L	L17L	L18L
L6N	.12*	.39**	.29*	-0.04	.32*	.41**	.34**	.32**	.29**
L7N	.16**	.53**	.60**	0.16	.34*	.52**	.47**	.35**	.41**
L8N	.14*	.41**	.50**	.29*	0.14	.47**	.39**	.36**	.33**
L9N	0.25	.50**	.38**	0.28	.42**	.51**	.65**	.43**	.43**
L10N	.51**	.14*	0.26	.37**	.34*	0.03	0.09	.22**	.15*
L11N	0.13	.54**	.55**	0.26	.29*	.37**	.47**	.51**	.34*
L12N	0.07	.33**	.57**	0.03	0.17	.35**	.30**	.23**	.27**
L13N	.32*	.28*	0.16	.52**	.45**	0.17	0.21	0.27	0.26
L14N	.25**	.30**	0.15	0.07	.73**	.24**	.32**	.23**	.29**
L15N	0.04	.49**	.44**	.28*	0.04	.62**	.47**	.36**	.40**
L16N	0.1	.44**	.42**	0.24	.29*	.44**	.57**	.34**	.40**
L17N	.29**	.24**	0.2	0.22	0.27	.38**	.28**	.63**	.35**
L18N	0.26	.37**	.46**	0.11	.41**	.52**	.47**	.56**	.54**
LL	0.02	.27**	.47**	0.21	0.25	.33**	.31**	.20**	.27**
L1L	.12*	.50**	.54**	0.16	0.09	.52**	.37**	.40**	.36**

	M	SD	L3L	L4L	L5L	L6L	L7L	L8L	L9L	L10L	L11L	L12L	L13L
L2L	4.41	.84	.43**	.34*	.55**	.37**	.49**	.44**	.61**	0.1	.45**	.44**	.29*
L3L	4.44	.84		.51**	.68**	.53**	.68**	.55**	.46**	0.08	.61**	.73**	0.11
L4L	4.12	1.03			.63**	.31*	.59**	.62**	.61**	0.23	.35*	.56**	0.25
L5L	4.37	.79				.51**	.63**	.66**	.77**	.13*	.72**	.74**	0.27
L6L	4.12	.98					.45**	.48**	0.26	.14*	.51**	.34*	0.15
L7L	4.35	.72						.57**	.66**	0.08	.46**	.62**	.28*
L8L	4.08	1.01							.70**	.17**	.61**	.51**	0.27
L9L	4.47	.70								0.23	.61**	.56**	.30*
L10L	1.67	1.15									.17**	0.19	.38**
L11L	3.97	1.16										.49**	.30*
L12L	4.61	.63											0.14
L13L	1.41	.96											
L14L	3.00	1.44											
L15L	4.19	.99											

	L14L	L15L	L16L	L17L	L18L
L2L	0.15	.40**	.51**	.34**	.37**
L3L	0.11	.57**	.47**	.42**	.44**
L4L	.35*	.53**	.49**	.84**	.46**
L5L	.33*	.67**	.61**	.43**	.48**
L6L	.34*	.47**	.47**	.36**	.38**
L7L	.39**	.66**	.42**	.60**	.54**
L8L	.34*	.63**	.57**	.42**	.48**
L9L	.30*	.64**	.63**	.63**	.52**
L10L	0.13	0.04	0.1	.27**	.21**
L11L	0.27	.60**	.64**	.44**	.48**
L12L	0.22	.52**	.41**	.56**	.33*
L13L	0.17	0.23	0.22	0.26	.38**
L14L		0.21	.30*	.28*	.33*
L15L			.54**	.45**	.44**

	M	SD	L17L	L18L
L16L	3.75	1.17	.36**	.48**
L17L	3.71	1.31		.36**
L18L	3.25	1.39		

** . Correlation is significant at the 0.01 level. *. Correlation is significant at the 0.05 level. 0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 14

Correlation matrix for Item M across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	MN	M1N	M2N	M3N	M4N	M5N	ML	M1L	M2L	M3L	M4L	M5L
MN	3.25	1.05		.32**	.59**	.56**	.50**	.48**	.28**	.26	.48**	.51**	.45**	.54**
M1N	2.79	1.26			.71**	.49**	.56**	.51**	.38**	.68**	.66**	.56**	.58**	.56**
M2N	3.18	1.20				.67**	.81**	.77**	.76**	.69**	.78**	.80**	.71**	.75**
M3N	3.23	1.16					.58**	.71**	.36**	.37**	.61**	.69**	.59**	.71**
M4N	2.94	1.17						.57**	.64**	.63**	.67**	.67**	.56**	.68**
M5N	3.50	1.07							.39**	.56**	.64**	.67**	.62**	.68**
ML	3.96	1.03								.49**	.39**	.40**	.42**	.40**
M1L	2.63	1.30									.64**	.63**	.58**	.59**
M2L	3.02	1.20										.77**	.76**	.72**
M3L	2.87	1.24											.79**	.73**
M4L	3.12	1.20												.73**
M5L	3.23	1.15												

**. Correlation is significant at the 0.01 level.
*. Correlation is significant at the 0.05 level.
0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 15

Correlation matrix for Item N across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	NN	N1N	N2N	N3N	N4N	N5N	NL	N1L	N2L	N3L	N4L	N5L
NN	2.38	1.10		.52**	.38**	.31**	.32**	.44**	.29**	.46**	.29**	.34**	.41**	.45**
N1N	2.34	1.17			.45**	.37**	.41**	.70**	.30**	.68**	.44**	.41**	.52**	.44**
N2N	2.53	1.39				.29*	.35*	.49**	.57**	.40**	.84**	.35*	.43**	.37**
N3N	1.66	.90					.39**	.80**	.31**	.63**	.34**	.63**	.58**	.61**
N4N	2.75	1.21						.59**	.05	.42**	.37**	.37**	.65**	.63**
N5N	1.96	1.04							.72**	.71**	.63**	.70**	.73**	.68**
NL	1.96	1.06								.55**	.15**	.24**	.75**	.66**
N1L	1.88	1.05									.51**	.59**	.56**	.38**
N2L	2.23	1.27										.37**	.62**	.57**
N3L	1.66	.91											.51**	.36**
N4L	2.33	1.16												.63**
N5L	1.88	1.13												

**. Correlation is significant at the 0.01 level. *. Correlation is significant at the 0.05 level. 0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 16

Correlation matrix for Item O across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	ON	O1N	O2N	O3N	O4N	O5N	O6N	O7N	O8N	O9N	O10N
ON	1.39	.73		.36**	~	0.07	0.08	~	.19**	.16**	0.05	.13*	.12*
O1N	2.02	1.21			~	0.09	.19**	~	.15**	.25**	0.09	.20**	.20**
O2N	1.00	.00				~	~	~	~	~	~	~	~
O3N	1.03	.26					.67**	~	.64**	.66**	.90**	.60**	.71**
O4N	1.03	.30						~	.51**	.84**	.76**	.87**	.83**
O5N	1.00	.00							~	~	~	~	~
O6N	1.06	.37								.59**	.61**	.62**	.61**
O7N	1.06	.35									.66**	.80**	.84**
O8N	1.01	.23										.63**	.71**
O9N	1.05	.36											.92**
O10N	1.04	.32											
OL	3.68	1.40											
O1L	1.58	.96											

	OL	O1L	O2L	O3L	O4L	O5L	O6L	O7L	O8L	O9L	O10L
ON	0.02	.30**	.12*	0.08	0.06	~	0.11	0.07	0.05	~	.34*
O1N	0.01	.70**	0.07	.14*	.11*	~	0.11	.16**	0.09	~	0.12
O2N	~	~	~	~	~	~	~	~	~	~	~
O3N	0.02	.24**	.20**	.64**	.92**	~	.84**	.88**	.90**	~	~
O4N	0.02	0.23	.20**	.92**	.73**	~	.63**	.70**	.76**	~	-0.02
O5N	~	~	~	~	~	~	~	~	~	~	~
O6N	-0.01	.27**	0.01	.52**	.67**	~	.79**	.66**	.61**	~	-0.02
O7N	0.05	.30**	0.1	.47**	.68**	~	.68**	.72**	.66**	~	-0.02
O8N	0.01	.21**	-0.01	.58**	.97**	~	.83**	.87**	1.00**	~	~
O9N	-0.02	.22**	-0.02	.45**	.65**	~	.69**	.62**	.63**	~	0.21
O10N	-0.01	.25**	-0.02	.61**	.77**	~	.73**	.77**	.71**	~	.30*
OL		0	0.03	-0.05	0	~	0.03	0	0.01	~	.58**
O1L			0.05	.18**	.24**	~	.25**	.29**	.21**	~	0.04

	M	SD	O3L	O4L	O5L	O6L	O7L	O8L	O9L	O10L
O2L	1.04	.31	.20**	-0.01	~	0.02	0.03	-0.01	~	-0.02
O3L	1.04	.40		.70**	~	.60**	.75**	.58**	~	~
O4L	1.02	.24			~	.86**	.95**	.97**	~	~
O5L	1.00	.00				~	~	~	~	~
O6L	1.03	.27					.82**	.83**	~	~
O7L	1.02	.26						.87**	~	~
O8L	1.01	.23							~	~
O9L	1.00	.00								~

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

~. Correlation could not be computed because at least one of the variables is constant.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 17

Correlation matrix for Item P across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	P1N	P2N	P3N	P4N	P5N	P6N	P7N	P8N	P9N	P10N	P11N
PN	1.29	.57	.21**	0.26	.33**	0.2	.21**	.52**	.28**	0.07	-0.07	.17**	.22**
P1N	1.05	.31		.32*	.36**	1.00**	.73**	.43**	0.09	.69**	-0.02	.68**	.32**
P2N	1.29	.76			.51**	.32*	~	0.1	.51**	~	-0.06	.60**	.46**
P3N	2.74	1.35				-0.02	.59**	-0.03	.41**	.43**	-0.02	.40**	.31**
P4N	1.04	.28					~	.43**	-0.02	~	-0.02	-0.03	.70**
P5N	1.05	.31						~	.21**	.75**	~	.65**	.43**
P6N	1.06	.31							-0.03	~	-0.03	-0.04	.29*
P7N	1.03	.18								0.09	-0.02	.24**	.24**
P8N	1.04	.35									~	.70**	.26**
P9N	1.02	.14										-0.03	-0.03
P10N	1.07	.34											.24**
P11N	1.17	.58											
PL	3.81	1.42											

	PL	P1L	P2L	P3L	P4L	P5L	P6L	P7L	P8L	P9L	P10L	P11L
PN	0.07	0.2	-0.09	.21**	0.1	~	.17**	0.06	.14*	-0.07	0.06	.24**
P1N	-0.06	1.00**	-0.03	.40**	.47**	~	.59**	.13*	.67**	-0.02	.65**	.44**
P2N	.51**	.32*	.29*	~	.32*	~	.32*	-0.06	.32*	-0.06	~	.46**
P3N	-0.02	-0.02	-0.03	.54**	.29**	~	.45**	.27**	.26**	-0.02	.41**	.27**
P4N	0.2	1.00**	-0.03	~	1.00**	~	1.00**	-0.02	-0.02	-0.02	~	.70**
P5N	-0.03	~	~	.58**	.47**	~	.70**	.37**	.75**	~	.78**	.44**
P6N	0.27	.43**	-0.03	~	.43**	~	.43**	-0.03	-0.03	-0.03	~`	.29*
P7N	-0.03	-0.02	-0.03	.19**	0.01	~	.24**	.19**	0.1	-0.02	0.04	.15**
P8N	-0.05	~	~	.69**	.55**	~	.73**	.34**	.34**	~	.95**	.26**
P9N	0.2	-0.02	-0.03	~	-0.02	~	-0.02	-0.02	-0.02	1.00**	~	-0.03
P10N	-0.02	-0.03	-0.03	.38**	.38**	~	.45**	.15**	.80**	-0.03	.64**	.31**
P11N	0.04	.70**	.66**	.39**	.29**	~	.55**	.51**	.21**	-0.03	.24**	.69**
PL		0.2	0.17	-0.01	-0.07	~	-0.05	-0.01	-0.06	0.2	-0.04	-0.01

	M	SD	P2L	P3L	P4L	P5L	P6L	P7L	P8L	P9L	P10L	P11L
P1L	1.02	.14	-0.03	~	1.00**	~	1.00**	-0.02	-0.02	-0.02	~	.70**
P2L	1.08	.44		~	-0.03	~	-0.03	-0.03	-0.03	-0.03	~	.66**
P3L	1.05	.35			.37**	~	.72**	.47**	.47**	~	.66**	.18**
P4L	1.08	.50				~	.43**	.15**	.46**	-0.02	.55**	.38**
P5L	1.00	.00					~	~	~	~	~	~
P6L	1.05	.28						.64**	.59**	-0.02	.70**	.33**
P7L	1.04	.27							.28**	-0.02	.30**	.14*
P8L	1.05	.29								-0.02	.80**	.31**
P9L	1.02	.14									~	-0.03
P10L	1.04	.35										.30**
P11L	1.13	.49										

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

~. Correlation could not be computed because at least one of the variables is constant.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 18

Correlation matrix for Item Q across cognitive load conditions (HolNo, HolL, DecNo, DecL)

N=303	M	SD	Q1N	Q2N	Q3N	Q4N	Q5N	Q6N	Q7N	Q8N	Q9N	Q10N	Q11N
QN	1.19	.50	.40**	.40**	-0.03	.35**	.34**	0.2	.34**	0.22	.33**	.46**	.34**
Q1N	1.24	.65		.40**	0.02	.43**	.55**	0.16	.37**	0.05	.38**	.45**	.49**
Q2N	1.04	.31			~	.41**	.41**	~	.49**	~	.41**	~	.68**
Q3N	1.12	.59				0.05	0.02	0.24	.31*	0.11	.97**	0.03	~
Q4N	1.15	.53					.55**	.32*	.42**	0.27	.33**	87**	.44**
Q5N	1.13	.57						0.1	.49**	0.04	.58**	.95**	.61**
Q6N	1.06	.20							.81**	.86**	0.1	.31*	~
Q7N	1.06	.38								.40**	.49**	0.18	.66**
Q8N	1.10	.46									0.04	.33*	~
Q9N	1.12	.54										0.04	.52**
Q10N	1.10	.46											~
Q11N	1.03	.30											
Q12N	1.19	.63											
Q13N	1.16	.64											

	Q12N	Q13N	QL	Q1L	Q2L	Q3L	Q4L	Q5L	Q6L	Q7L	Q8L
QN	.31**	0.18	.35**	0.28	~	.30**	.26**	.42**	.40**	.44**	.38**
Q1N	.41**	0.19	.25**	0.22	~	.34**	.44**	.48**	.38**	.35**	.35**
Q2N	.44**	~	.18**	~	~	.51**	.51**	~	.61**	.68**	.54**
Q3N	.95**	.90**	0.13	0.07	~	.91**	-0.03	0.04	0.13	~	0.13
Q4N	.39**	.41**	.26**	0.03	~	.41**	.60**	.85**	.54**	.50**	.55**
Q5N	.60**	.44**	.29**	-0.05	~	.50**	.66**	1.00**	.45**	.45**	.44**
Q6N	.46**	0.2	.63**	0.24	~	-0.04	.57**	0.15	.81**	~	.81**
Q7N	.48**	0.27	.17**	.32*	~	.50**	.59**	0.2	.51**	.62**	.45**
Q8N	.35*	0.08	.55**	0.1	~	-0.04	.90**	0.06	.85**	~	.85**
Q9N	.66**	.88**	.19**	-0.05	~	.66**	.53**	0.05	.51**	.43**	.43**
Q10N	0.08	.42**	.91**	-0.06	~	-0.04	.28*	.95**	.40**	~	.40**
Q11N	.54**	~	.21**	~	~	.65**	.75**	~	.50**	.75**	.44**
Q12N		.86**	.24**	0.16	~	.69**	.59**	0.03	.49**	.42**	.45**
Q13N			.45**	0.04	~	.82**	-0.04	.45**	0.11	~	0.11

	Q9L	Q10L	Q11L	Q12L	Q13L
QN	-0.06	.31**	.40**	.27**	-0.09
Q1N	.38**	.26**	.46**	.34**	0.01
Q2N	~	.51**	.71**	.40**	~
Q3N	0.21	.91**	-0.03	.86**	.87**
Q4N	0.18	.37**	.49**	.44**	-0.01
Q5N	0.22	.42**	.59**	.51**	0.02
Q6N	.57**	0.1	-0.04	.52**	0.08
Q7N	.70**	.41**	.64**	.52**	0.12
Q8N	.28*	0.19	-0.03	.49**	0.02
Q9N	0.22	.58**	.48**	.63**	.90**
Q10N	.28*	0.04	.90**	0.12	0.02
Q11N	~	.54**	.90**	0.12	0.02
Q12N	0.19	.58**	.44**	.72**	.78**
Q13N	0.19	-0.02	.41**	.76**	.78**

	M	SD	Q1L	Q2L	Q3L	Q4L	Q5L	Q6L	Q7L	Q8L	Q9L	Q10L	Q11L
QL	1.86	.99	0.12	~	.18**	.23**	.83**	.22**	.22**	.16**	.36*	.17**	.19**
Q1L	1.08	.27		~	-0.05	-0.04	-0.05	-0.06	~	-0.06	-0.04	-0.05	-0.04
Q2L	1.0	.00			~	~	~	~	~	~	~	~	~
Q3L	1.07	.40				.62**	-0.03	.51**	.68**	.58**	-0.02	.86**	.60**
Q4L	1.09	.47					-0.03	.55**	.61**	.58**	-0.02	.52**	.62**
Q5L	1.08	.44						0.2	~	0.2	.30*	-0.03	.94**
Q6L	1.10	.40							.60**	.80**	.70**	.44**	.53**
Q7L	1.02	.25								.69**	~	.64**	.78**
Q8L	1.09	.43									.70**	.48**	.44**
Q9L	1.02	.14										-0.02	-0.02
Q10L	1.05	.37											.53**
Q11L	1.03	.27											
Q12L	1.13	.55											
Q13L	1.14	.63											

	Q12L	Q13L
QL	.14*	-0.09
Q1L	0.12	-0.06
Q2L	~	~
Q3L	.82**	.84**
Q4L	.62**	-0.03
Q5L	0.01	0.03
Q6L	.59**	0.12
Q7L	.53**	~
Q8L	.63**	0.12
Q9L	0.16	0.2
Q10L	.66**	.84**
Q11L	.51**	-0.03
Q12L		.79**

** . Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

~. Correlation could not be computed because at least one of the variables is constant.

0. Correlation was less than or equal to .004 but greater than 0 or Correlation was less than or equal to -.004.

Table 19

Average correlations (across both measure and level of cognitive load) for each item set

Variable	*r*
A	.35
B	.39
C	.41
D	.28
E	.33
F	.47
G	.44
H	.41
I	.48
J	.52
K	.41
L	.38
M	.61
N	.50
O	.48
P	.30
Q	.62

Table 20

Means and standard deviations for participant item ratings in each of the conditions before being correlated for stability hypotheses

Variable	N items	*M*	*SD*	*Min*	*Max*
HolNo	17	3.29	.37	2.00	4.24
HolL	17	2.66	.50	1.40	4.24
DecNo	160	3.22	.53	1.89	4.86
DecL	160	2.95	.46	1.65	3.93

Table 21

Results from the Hotelling-Williams tests conducted for Hypotheses 1A and 1B

Source	*r*	*df*	*t*	*p*
HolNo-DecNo	.55			
HolNo-HolL	.74	300	1.33	*ns*
DecNo-DecL	.66			
HolNo-DecNo	.55	300	1.25	*ns*

Table 22

Results from the z-test conducted for Hypothesis 2

Source	*r*	*df*	*z*	*p*
DecNo-DecL	.66			
HolNo-HolL	.74	300	2.72	.05

Table 23

Comparison between undergraduates and graduate students for levels of stability between conditions

Source	N	*r*	*df*	*z*	*p*
HolNo-HolL			310	1.23	*ns*
Grad	10	.88			
Undergrad	303	.74			
HolNo-DecNo			310	3.94	.05
Grad	10	.89			
Undergrad	303	.55			
DecNo-DecL			310	2.10	.05
Grad	10	.98			
Undergrad	303	.66			

Table 24

Results from the repeated measures analysis of variance for conditions with Interrater agreement

Source	*df*	*MS*	*F*	*p*
Measure (M)	1, 16	.397	5.93	.03
Cognitive Load (CL)	1, 16	.706	18.48	.00
M X CL	1, 16	-.005	.95	.37

Table 25

Results from the one-way analyses of variance tests conducted to test Hypotheses 3A-3D (Interrater agreement)

Source	N	r_{wg}	*df*	*MS*	*F*	*p*
HolNo	17	.5942				
DecNo	160	.5583	1, 175	.02	.26	.61
DecL	160	.5353				
HolL	17	.3344	1, 175	.621	7.44	.01
HolNo	17	.5942				
HolL	17	.3344	1, 32	.574	10.74	.00
DecNo	160	.5583				
HolL	17	.3344	1, 175	.770	10.04	.00
HolNo	17	.5942				
DecL	160	.5353	1, 175	.053	.65	.42
DecNo	160	.5583				
DecL	160	.5353	1, 318	.042	.51	.48

Table 26

Results of the one-way analyses of variance conducted to test differences between undergraduate and graduate students for interrater agreement

Source	N	r_{wg}	*df*	*MS*	*F*	*p*
HolNo			1, 32	.014	.18	.68
Grad	10	.6349				
Undergrad	303	.5942				
HolL			1, 32	.298	3.38	.07
Grad	10	.5216				
Undergrad	303	.3344				
DecNo			1, 318	.920	9.98	.00
Grad	10	.6655				
Undergrad	303	.5583				
DecL			1, 318	1.940	23.01	.00
Grad	10	.6910				
Undergrad	303	.5353				

Table 27

Repeated measures analysis of variance for levels of elevation for the measure type and cognitive load conditions

Source	*df*	*MS*	*F*	*p*
Measure (M)	1, 302	22.12	149.54	.00
Cognitive Load (CL)	1, 302	11.22	99.04	.00
M X CL	1, 302	44.06	197.85	.00

Table 28

Mean levels of elevation (E^2) and results of pair-wise t-tests for all conditions

Condition	N	(E^2)	*t*	*df*	*p*
HolNo	17	.14	9.56	302	.00
HolL	17	.33			
HolNo	17	.14	14.78	302	.00
DecNo	160	.79			
HolNo	17	.14	4.94	302	.00
DecL	160	.22			
HolL	17	.33	12.04	302	.00
DecNo	160	.79			
HolL	17	.33	4.95	302	.00
DecL	160	.22			
DecNo	160	.79	13.42	302	.00
DecL	160	.22			

Table 29

Repeated measures analysis of variance for levels of stereotype accuracy for the measure type and cognitive load conditions

Source	*df*	*MS*	*F*	*p*
Measure (M)	1, 302	322891.86	165.08	.00
Cognitive Load (CL)	1, 302	34114.96	189.28	.00
M X CL	1, 302	293381.39	822.67	.00

Table 30

Mean levels of stereotype accuracy (SA^2) and results of pair-wise t-tests for all conditions

Condition	N	(SA^2)	*t*	*df*	*p*
HolNo	17	13.37	30.90	302	.00
HolL	17	55.10			
HolNo	17	13.37	33.29	302	.00
DecNo	160	77.13			
HolNo	17	13.37	17.94	302	.00
DecL	160	56.62			
HolL	17	55.10	7.65	302	.00
DecNo	160	77.13			
HolL	17	55.10	.45	302	.65
DecL	160	56.62			
DecNo	160	77.13	15.64	302	.00
DecL	160	56.62			

Figure Captions

Figure 1. Example of scale that was used with both the holistic and decomposed measures.

Figure 1.

How <u>important</u> is SCHEDULING WORK AND ACTIVITIES to the performance of *your current job*?

Not Important*	Somewhat Important	Important	Very Important	Extremely Important
1	2	3	4	5

VITA

Alisa H. Watt
4 Bainbridge Court
Hamilton Square, NJ 08690
alwatt@vt.edu

EDUCATION

MS in I/O Psychology, Virginia Tech, May 2003
BA in Psychology, Lafayette College, May 2001

WORK HISTORY WITHIN EDUCATIONAL SETTINGS

January 2002-Present
Graduate Assistant for the Minority Academic Opportunity Programs Office, the Graduate School of Virginia Tech, and the Office of Multicultural Affairs

- Organized programs designed to retain minority students on campus
- Develop and maintain a peer mentoring program for undergraduate students
- Help to organize information that is used to recruit minority students to campus

August 2001- January 2002
Graduate Teaching Assistant

- Teach 3 sections of the recitation portion of the Introduction to Psychology course
- Plan and grade over 100 assignments each week
- Responsible for all lectures

Presentations

- Hollander, E., Watt, A.H., & McKinney, A.P. (2003). NBADS: Further support to its advantages over other formats. Presented at SIOP 2003.

- Watt, A.H., & Shaw, J. (2000). The role of power in perceptions of sexual harassment. Presented at the Eastern Psychological Conference 2000.