INDIVIDUAL DIFFERENCES, COGNITIVE ABILITIES, AND THE INTERPRETATION OF AUDITORY GRAPHS

Bruce N. Walker and Lisa M. Mauney

Sonification Lab, School of Psychology
Georgia Institute of Technology,
654 Cherry Street
Atlanta, GA 30332-0170
bruce.walker@psych.gatech.edu, kint16@yahoo.com

ABSTRACT

Auditory graphs exploit pattern recognition in the auditory system, but questions remain about the relationship between cognitive abilities, demographics, and sonification interpretation. Subjects completed a magnitude estimation task relating sound dimensions to data dimensions. Subjects also completed a working memory task (2-back task) and a spatial reasoning task (Raven's Progressive Matrices) to assess cognitive abilities. Demographics, such as gender, age, handedness, and musical experience, were also reported and included in the analysis.

A stepwise multiple regression analysis was performed to determine the relationship between the independent (cognitive abilities and demographics) and dependent (individual slopes and R-squared values) variables. The regression analysis indicates some support for most of the predictor variables, especially predicting R-squared values. The 2-back task does not seem to contribute significantly to the interpretation of sonifications and auditory graphs. However, Raven's and many of the demographic variables do show predictive value for interpretation of auditory graphs.

1. INTRODUCTION

Researchers often analyze large, multidimensional, and dynamic data sets. New sonification and auditory graphing techniques are being used to exploit the exceptional pattern recognition capabilities of the human auditory system [1]. The increasing prevalence of auditory graphing systems brings up many design issues. Auditory graph designers have three options to improve the effectiveness of their systems: (1) change the system (e.g., adding context), (2) change the human (e.g., training), and (3) use individual difference characteristics to select operators. The current study focuses on the possible demographics and cognitive abilities information needed for the selection of operators. For example, the military uses multi-modal watch stations, which contain both visual and auditory displays. The current study could lead the way to discovering more effective ways of selecting military personnel to work these stations, compared to the standard battery of aptitude tests given upon entering the military. Unfortunately, there remain many unanswered questions regarding the relationship between cognitive abilities, demographics, and the interpretation of sounds used to represent data. Limited research in this area has found some differences between groups of individuals, suggesting more investigation is required.

1.1. Individual Differences in Auditory Graph Interpretation

Walker found individual differences in the polarities of responses to data-to-sound mappings [1]. The polarity of a mapping describes how changes in the display dimension indicate changes in the data dimension. For example, as frequency increases so does the number of dollars in a given mapping; this represents a positive polarity. An example of a negative polarity would be as frequency increases, the number of dollars decreases. In some cases, a majority of listeners clearly prefers either a positive or negative polarity and in other cases the polarities are split within the data-to-display mappings.

There are also differences between groups of listeners. For example, Walker and Lane showed that there are differences in the way blind and sighted listeners respond to sonification mappings [2]. Neuhoff, Knight, and Wayand showed that in some cases musical experts and musical novices respond differently [3]. Such findings of individual differences have not been consistent or replicated. Further, it is not clear which other individual difference variables may play a role in auditory display interpretation. We contend that there may be cognitive differences that influence these types of tasks. The current study looks at a 2-back task as a measure of working memory, and the Raven's Progressive Matrices task as a measure of spatial abilities or general fluid intelligence. We describe these next.

1.2. Cognitive Abilities and Demographics in the Interpretation of Sound

One measure of cognitive abilities that is widely researched is the construct of working memory. Working memory was initially proposed by Baddeley and Hitch in 1974 and was developed by Baddeley in 1986 [4,5]. Working memory consists of temporary memory stores (i.e., visuo-spatial sketch pad and articulatory loop) for rehearsing stored information and a mechanism of central or executive attention that regulates the contents of the active portion of memory [4,5]. Working memory is distinct from short-term memory in that the working memory system is able to store and process information simultaneously, whereas short-term memory is just the temporary storage of information [5]. An n-back test was chosen for the current study because in order to interpret an auditory graph, one must store and process current and previously heard sounds. Although the magnitude estimation task is auditory in nature, a visual working memory measure was used. We chose to use this

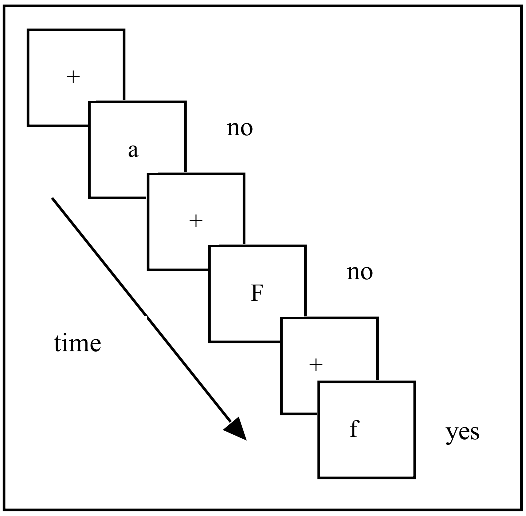

Figure 1: *Example of 2-back task. The 'yes' and 'no' beside the different blocks indicates the correct response that should be made upon seeing the letter.*

visual 2-back test because studies show that complex span tasks tend to load onto a single factor, regardless of the different stimuli used [7]. Performance on working memory measures are also correlated with performance on various higher-order cognitive tasks, such as reading comprehension, complex learning, and reasoning [4,6,7].

This correlation between working memory and reasoning is the reason we chose the Raven's Progressive Matrices task to further assess cognitive abilities; the Raven's is an observable measure of spatial reasoning [7]. Spatial reasoning and spatial cognition is related to a person's ability to mentally store information about the external world, visualize objects and scenes, perform tasks such as mental scanning or mental rotation, or coordinating motor activities and navigating through an environment [5]. We hypothesize that sound localization and the manipulation and interpretation of sounds are important characteristics of spatial reasoning, hence the use of the Raven's Progressive Matrices task. The visual nature of the Raven's should not interfere with its predictive power of auditory graph interpretation—the Raven's Progressive Matrices is correlated with pitch discrimination, which in turn is correlated with complex span tasks, such as the 2-back [7].

Despite the research that has been done on auditory graphs, there has not been any systematic evaluation of the relationship between a variety of variables and auditory graph interpretation. In addition to the two cognitive abilities, we also wanted to systematically study such demographics as gender, age, handedness, and musical experience, which we also included in our regression analysis.

1. PROCEDURE

1.1. Participants

Participants included 160 undergraduate students taking an introductory level psychology course at Georgia Institute of Technology (93 males, 67 females, mean age 19.9 years, age range 18-25). All received partial course credit for participating in this study.

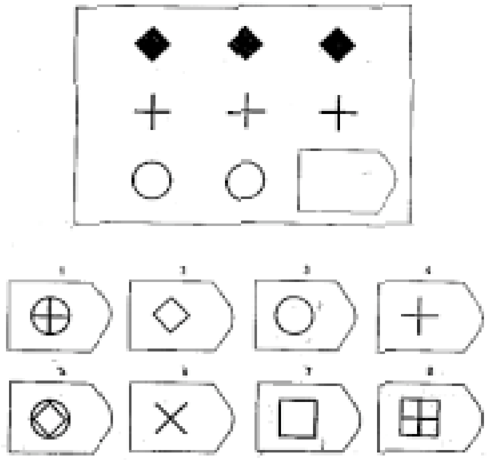

Figure 2: *Sample Raven's Matrix. One of the bottom eight pieces is the correct pattern to complete the top matrix.*

1.2. Demographics

The demographics that were collected on each subject were gender, age, handedness, and musical background. The gender and age of the participants were recorded upon arrival. The participants were then given an online version of the Edinburgh Handedness Test. Possible test scores range from -20 to 20, where negative is more left handed, positive is more right handed, and close to zero indicates the person is somewhat ambidextrous.

In keeping with more recent approaches to assessing musical experience (e.g., Neuhoff) [3], our assessment of musical experience consisted of the following four questions: (1) Have you ever played a musical instrument regularly? (2) How many years total did you play a musical instrument?; (3) At what age did you begin playing your first musical instrument?; (4) How many years of formal musical training have you had, either private or class instruction? However, to prevent multi-collinearity of the regression models, the only musical variable included in the multiple regression analysis was total years that the participants played a musical instrument.

1.3. Cognitive Abilities Measures

The measures of cognitive abilities included a working memory task and a spatial reasoning task.

The working memory task that was given to participants was an n-back task, specifically a 2-back task. The 2-back task was programmed in E-Prime and was run on a Windows-based PC. Participants had two blocks of practice trials before three blocks of actual trials. Each block consisted of letters being presented rapidly one at a time on the computer screen (letters were in white font and the background was gray). The stimuli consisted of upper and lowercase letters that had a "+" for a fixation point between the presentation of each letter. The participants were instructed to press the '1' key on the keypad if the current letter was the same as the letter presented 2 letters before, regardless of case type. Participants were instructed to press the '3' key on the keypad if the current letter was *not* the same as the letter presented 2 letters back, regardless of case type. See Figure 1

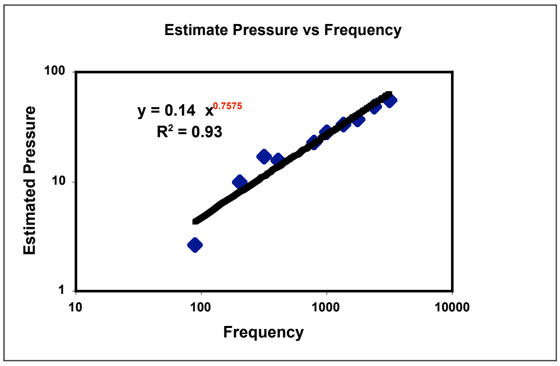

Figure 3: *Example of an aggregate group 'positive' polarity.*

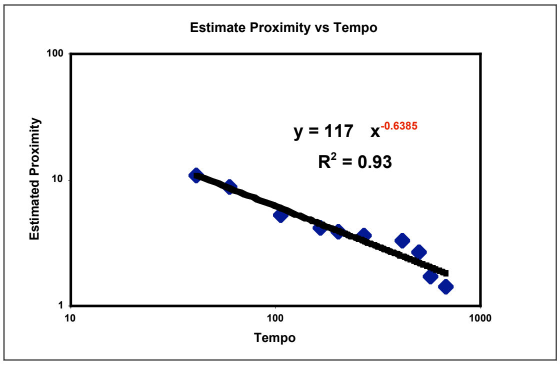

Figure 4: *Example of an aggregate group 'negative' polarity.*

for an example series. The experimenter stayed in the room during the practice sessions to make sure that the directions were clear to the participant. The experimenter was not present during the experimental trials. When the participant was finished with the task, the percentage of correct responses was recorded, since we were interested in participants' abilities to remember previous items.

The spatial reasoning task was measured by a computerized version of the Raven's Progressive Matrices task. This task was also programmed in E-Prime and run on a Windows-based PC. A series of pictorial matrices were shown on the computer screen to participants; the matrices were similar to a puzzle with a piece missing. The matrices were either all of one pattern with a piece missing from the bottom right corner or they were a 3x3 matrix of different pieces with the bottom right piece missing. In either case, the participants had a choice of six or eight pieces that would fill in the missing part of the puzzle. The figures often required mental rotation of objects to find what the next piece in the series would be. Participants were not allowed to use paper to work out any of the problems and had 15 minutes to finish the task. There were 36 puzzles total, but the participants were instructed to answer each question correctly before moving on to the next picture; accuracy was chosen over speed. E-prime allowed for several scores to be recorded, but we focused on the percentage correct (total correct divided by total number of matrices looked at).

2.4. Magnitude Estimation Task

To investigate the interpretation of sounds that represent data, a magnitude estimation task was performed. The details of the procedure are explained elsewhere by Walker [1]. However, here is a brief summary of the procedure.

Participants heard a series of sounds presented one at a time in random order and were asked to assign each sound a value that represented its magnitude for a particular data dimension. For example, "What 'Number of Dollars' does this sound seem to represent?" The participants gave each sound they heard a magnitude that could be any positive number, fraction, or decimal that they felt was appropriate. The three display dimensions were frequency, tempo, and spectral brightness (operationalized as the number of harmonics of the fundamental frequency). In Walker's previous study [1], only tempo and frequency was used, so brightness is a new display dimension added to the task. Also, a number of additional data dimensions were included in the current study. The complete list of data dimensions used in this study is: size, number of dollars, temperature, velocity, pressure, danger, mass, urgency, proximity, and attractiveness.

Participants completed three blocks of trials, one for each of the display dimensions. Each block consisted of sounds that varied in one of the three display dimensions and one data dimension. The data dimension that was associated with each block was randomly selected from the 10 choices of data dimensions. Each block presented 10 different sounds along that display dimension and each sound was presented twice, for a total of 20 sounds per block.

2. RESULTS

Several types of data analysis were performed on the data from this study. The group polarity and slope analysis, as described in full by Walker [1], will be briefly described next.

3.1. Slope and R-squared Analyses

For each participant, in each block of trials (i.e., each combination of display and data variables) the magnitude estimation slope was determined, as well as the R-squared value of that slope. These variables indicate how much change in the data dimension (e.g., temperature) is indicated by a given change in the display dimension (e.g., pitch). The R-squared value indicates how consistent a particular listener is in her response pattern. These two values (slope and R-squared) became the dependent measures used in the subsequent regression analyses. Individual slopes and R-squared values were calculated from the geometric means of the mapping of participants' answers against the actual stimulus values. Slopes that did not reach statistical significance were removed from further analysis ($r_{critical}$=|0.444|, p<0.05, 18 d.f.). Slopes that were positive and above 0.444 were classified as having a 'positive' polarity, those below -0.444 were classified as 'negative' polarity, and those that were not of statistical significance were classified as 'no' polarity. For each data-to-display pairing the results from each polarity were collapsed to create aggregate magnitude estimation slopes; where necessary, these groups were divided into positive and negative polarities. Figures 3 and 4 contain examples of what the overall group polarities look like from the study. In general, these overall slopes and R-squared values for the different data and display dimensions fell in line with previous studies [1,8].

3.2. Multiple Regression Analyses

A stepwise multiple regression was then performed to investigate the influence of the potential predictor variables (i.e., cognitive abilities and demographics) on the interpretation of auditory graphs, specifically the scaling of the individual mappings (slope values) and the R-squared values of those mappings. We were able to look at the detailed level of data and display dimension for aggregated slope and R-squared values, the results of which can be seen in Table 1.

The only correlation of the predictor variables that was statistically significant was between Ravens percent correct and 2-back percent correct (r=.214, p<.01). However, despite this significant correlation, regression collinearity analyses do not show any multi-collinearity with any of the variables; the regression models should therefore not be degenerative. The variables that made significant contributions (p<.05) to slope and R-squared values are shown in Table 1. The standardized beta (β) weights were reported in order to deal with differences in measurement units. The F-statistics and R^2 values for the significant predictors of R-squared are: frequency-velocity (F=17.279, R^2=.552), frequency-dollars (F=6.082, R^2=.303), frequency-urgency (F=4.925, R^2=.260), frequency-proximity (F=7.486, R^2=.348), tempo-attractiveness (F=6.713, R^2=.309), tempo-danger (F=5.809, R^2=.293), brightness-size (F=5.708, R^2=.290), brightness-dollars (F=16.186, R^2=.536). The F-statistics and R^2 values for the significant predictors of slope are: frequency-dollars (F=5.462, R^2=.281), frequency-mass (F=8.537, R^2=.379), tempo-danger (F=7.787, R^2=.357), tempo-mass (F=13.625, R^2=.493), brightness-dollars (F=21.289, R^2=.603), brightness-urgency (F=5.654, R^2=.288).

There is some support for several of the predictors, as seen in the table. One finding from the multiple regression analysis is that the 2-back task was not a uniquely significant contributor to the variance of either individual slopes or variance in R-squared values. That is, in this experiment the 2-back is not a good predictor of auditory graph interpretation. As the table indicates, the independent variables were better predictors of R-squared values than they were of individual slopes. The slope results are actually encouraging for display designers: Since individual slopes cannot be predicted by individual difference variables, it is implied that people can be trained to interpret auditory graphs for a particular slope, making design of graphs easier. Having predicted R-squared values is also good because individual differences can be used to select operators of auditory graphing systems. Raven's and gender seem to be good predictors of auditory graph interpretation, though certainly not universally effective. The measures of handedness and musical ability do not seem to be very significant contributors to the variance in sonification interpretation; their effects are sporadic. There are a number of points to consider in regard to these results, discussed next.

Display & Data Dimension	**2-back % Correct**	**Ravens % Correct**	**Gender**	**Age**	**Handedness**	**Total Years Played**
Frequency						
size						
temperature						
dollars			-.550 (R), -.530 (S)			
pressure						
velocity	.743 (R)					
urgency			-.510 (R)			
proximity		-.590 (R)				
attractiveness						
mass		.615 (S)				
danger						
Tempo						
size						
temperature						
dollars						
pressure						
velocity						
urgency						
proximity						
attractiveness	-.556 (R)					
mass			-.702 (S)			
danger			-.598 (S)	.542 (R)		
Brightness						
size		-.538 (R)				
temperature						
dollars	-.777 (S)		-.732 (R)			
pressure						
velocity						
urgency						
proximity		.536 (S)				
attractiveness						
mass						
danger						

Table 1: *Significant Predictors from Multiple Regression. Shows the cells that have significant beta (β) weights (p<.05) for individual slopes (indicated by an S) and for R-squared values (indicated by an R). Note that a significant beta weight indicates that the dependent variable is predicted by that independent variable, in that combination of data and display parameters.*

3. CONCLUSIONS

The results show that there is some support for the chosen predictor variables of cognitive abilities and demographics. However, these results are sporadic and not completely consistent. Based on the regression, there is a linkage between certain variables and interpretation of auditory displays. The story that results from this study is complex and no one single variable amongst the ones tested has ultimate predictive power. However, we can say that the presence of some significant predictors bodes well for this line of research into individual differences.

One thing we are attempting to predict is the interpretation of auditory graphs, but we may simply not have the right predictor variables. For example, the 2-back task did not produce the expected predictive power. However, there are many other tests that exist to measure working memory abilities, such as reading span and operation span [4,5,6]. Engle, a noted memory researcher, considers that n-back tests are not the best measures of working memory; he insists that measures such as operations span and reading span are better measures of the construct of working memory capacity [9]. This suggests that we may need to consider other predictor variables in the interpretation of auditory graphs.

In the current study, we are only studying magnitude estimation in order to look at the underlying psychoacoustic variables that play a role in the interpretation of auditory graphs. It remains to be seen if prediction in a more holistic or complete auditory task, such as Smith and Walker's point estimation task, may be more accurately predicted by these variables [10]. However, we caution that many questions remain and require further study.

In summary, some cognitive abilities, such as working memory, and demographics, such as gender, seem to be promising predictors of auditory graph interpretation. Implications of these findings are that the designer has the freedom to choose the slopes used by the system and that individual difference characteristics of people can be used to select operators of auditory graphs. The future development of more advanced auditory graphs and sonifications may depend on further research in the area of individual differences and general intelligence.

4. REFERENCES

[1] Walker, B.N. "Magnitude estimation of conceptual data dimensions for use in sonification", *Journal of Experimental Psychology: Applied*, 8, 2002.

[2] Walker, B.N. and D.M. Lane. "Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners," in *Proceedings of the International Conference on Auditory Display*. 2001. Espoo, Finland.

[3] Neuhoff, J.G., Knight, R., and Wayand, J. "Pitch change, sonification, and musical expertise: Which way is up?," in *Proceedings of the International Conference on Auditory Display*. 2002. Kyoto, Japan.

[4] Engle, R.W. "Working memory capacity as executive attention", Current Directions in *Psychological Science*, 11, 2002.

[5] Cornoldi, C. and Vecchi, T. *Visuo-spatial working memory and individual differences,* 2003, New York, NY: Psychology Press.

[6] Daneman, M. and Carpenter, P.A. "Individual differences in working memory and reading", *Journal of Verbal Learning and Verbal Behavior,* 19, 1980.

[7] Payne, T.W. "Working memory capacity and pitch discrimination", Doctoral Dissertation, 2003, Georgia Institute of Technology, Atlanta, GA.

[8] Walker, B.N. "Magnitude estimation of sound attributes for sonification, a study involving sighted college students", Unpublished raw data, 2002, Georgia Institute of Technology, Atlanta, GA.

[9] Engle, R.W. Personal correspondence, 2003.

[10] Smith, D.R. and Walker, B.N. "Tick-marks, axes, and labels: The effects of adding context to auditory graphs", in *Proceedings of the International Conference on Auditory Display.* 2002. Kyoto, Japan.