

The Personal Equation of Interaction for Interface Learning: Predicting the Performance of Visual Analysis Through the Assessment of Individual Differences

by

Tera Marie Greensmith

B.A., University of North Carolina at Charlotte, 2008
A.A., Central Piedmont Community College, 2005

Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

School of Interactive Arts and Technology
Faculty of Communication, Art and Technology

© Tera Marie Greensmith 2015

SIMON FRASER UNIVERSITY

Fall 2015

All rights reserved.

However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for Fair Dealing. Therefore, limited reproduction of this work for the purposes of private study, research, education, satire, parody, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name: Tera Greensmith
Degree: Doctor of Philosophy
Title: The Personal Equation of Interaction for Interface Learning: Predicting the performance of visual analysis through the assessment of individual differences

Examining Committee:

Dr. Wolfgang Stuerzlinger
Chair
Professor

Dr. Brian Fisher
Senior Supervisor
Associate Professor

Dr. Steve DiPaola
Supervisor
Professor

Dr. Halil Erhan
Supervisor
Associate Professor

Dr. Lyn Bartram
Internal Examiner
Professor

Dr. Giuseppe Carenini
External Examiner
Associate Professor
Computer Science
University of British Columbia

Date Defended/Approved: October 20, 2015

Ethics Statement



The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

- a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

- b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

- c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

- d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

The Personal Equation of Interaction (PEI) for Interface Learning is a short self-report psychometric measure which predicts reasoning outcomes of interface learning such as accurate target identification and insights garnered through and inferred from learning interaction. By predicting outcomes, we consider why some interfaces are more appropriate than others, provide a tool for intuitive interface design, and advance the pursuit and design of interface individuation. Through study designs which use comparative interfaces and simple but imperative tasks to any interface learning, such as target identification and inferential learning, we evaluate the accuracy of analysts and how it is impacted by graphical representation. By using psychometric items culled from normed trait assessment, we have created a measure which predicts accuracy and learning, called the Personal Equation of Interaction. This prediction tool can be used in a variety of ways, including as a function or equation that puts a number on the association between analyst and interface. We also use the PEI to build profiles of analyst expert cohorts and discuss how its use might impact Visual Analytics.

Keywords: cognitive science; visual analytics; data visualization; interface learning; individual differences; laboratory studies

Dedication

For my wife Penelope, who persists in believing that I can do anything.

Table of Contents

Approval.....	ii
Ethics Statement.....	iii
Abstract.....	iv
Dedication.....	v
Table of Contents.....	vi
List of Tables.....	ix
List of Figures.....	x
List of Acronyms.....	xii
Glossary.....	xiii
Chapter 1. Introduction	1
1.1. Research Questions	1
1.2. Why Is The Study Of Reasoning Important To Visual Analytics?	2
1.3. How Do We Use Individual Differences To Predict Reasoning Outcomes?	5
1.3.1. How Can Reasoning Be Studied?	6
1.4. How Do We Use Individual Differences To Study And Predict Visual Analytics Reasoning Outcomes?	12
1.4.1. The Non-Mystery of the Assumed Innate.....	13
1.5. What Is A Personal Equation And How Can It Be Used?.....	16
1.6. Summary.....	20
1.7. References.....	22
Chapter 2. Using personality factors to predict interface learning performance.....	26
2.1. Overview	26
2.2 Abstract.....	26
2.2. Introduction.....	27
2.3. Comparative Study	31
2.3.1. Interfaces.....	31
2.3.2. Psychometric Measures	34
2.3.3. Protocol	36
2.3.4. Results	37
Descriptive Results	37
Inferential task performance	39
Procedural task performance.....	39
2.4. Discussion	43
2.5. Conclusion.....	46
2.6. References.....	47
Chapter 3. Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction.	49
3.1. Overview	49
3.2. Abstract	49

3.3.	Introduction.....	50
	The first question	54
	The second question	54
3.4.	Comparative Studies	55
	3.4.1. Interfaces.....	55
	3.4.2. Psychometric measures	56
	3.4.3. Participants.....	58
	3.4.4. Study Protocol	59
3.5.	Results	60
	3.5.1. Completion times and personality factors	61
	3.5.2. Task Errors and Personality Factors.....	62
	3.5.3. Insight generation and personality factors.....	65
3.6.	Discussion	66
3.7.	Conclusion.....	73
3.8.	References	74

Chapter 4.	The personal equation of complex individual cognition during visual interface interaction	76
4.1.	Overview	76
4.2.	Abstract	76
4.3.	Introduction.....	77
	4.3.1. The Challenge of Complex Cognition	78
	4.3.2. The Sensemaking Loop.....	79
	4.3.3. The Absence of Precedent	84
4.4.	Individual Differences	86
	4.4.1. The Human Cognition Model	87
	4.4.2. The Personal Equation of Interaction (PEI).....	90
	The Personal Equation of Interaction Defined	91
	Research in the Personal Equation of Interaction	95
	Results of Study 1.....	97
	Results of Study 2.....	97
	Discussion of Results	98
4.5.	Conclusion.....	98
4.6.	References	99

Chapter 5.	Impact of personality factors on interface interaction and the development of user profiles: Next steps in the personal equation of interaction	103
5.1.	Overview	103
5.2.	Abstract	103
5.3.	Introduction to comparative studies	104
5.4.	Comparative studies	111
	5.4.1. Interfaces.....	111
	5.4.2. Psychometric measures	112
	5.4.3. Participants.....	113
	5.4.4. Study Protocols	114
5.5.	Results	121
	5.5.1. Completion times and personality factors	121

5.5.2. Task errors and personality factors.....	123
5.6. Introduction to an early profile of efficient users	125
5.7. Personality factors as predictors.....	128
5.8. Early profile of the efficient procedural learner.....	135
5.9. Discussion.....	136
5.10. Conclusion.....	142
5.11. References.....	143
Chapter 6. Discussion	147
6.1. Chapter 2.....	148
6.2. Chapter 3.....	151
6.3. Chapter 4.....	156
6.4. Chapter 5.....	157
6.5. References.....	159
Chapter 7. Conclusion.....	161
7.1.1. Ever onward and upward.....	164
7.1.2. Implications	165

List of Tables

Table 1. Items in the short measure.	40
Table 2. PCA components of the short measure	41

List of Figures

Figure 1. The interaction between human cognition and the interface. <i>source:</i> Green, Ribarsky, and Fisher (2008)	3
Figure 2. Example of the relationship between the PEI predicted values and the actuals.	19
Figure 3. The NCBI Interface.....	33
Figure 4. The main view of GVIS.....	34
Figure 5. MDS clusters in the 9-item measure.....	42
Figure 6. The NCBI MapViewer.....	56
Figure 7. The MainView of GVis.....	58
Figure 8. Correlations of GVis total completion times (in seconds) across procedural tasks and Locus of Control, Extraversion and Neuroticism scores.....	63
Figure 9. Correlations of Map Views Total completion times (in seconds) across procedural tasks and locus of control, extraversion, and neuroticism scores	64
Figure 10. Insights and Locus of control.....	68
Figure 11. Insights and Extraversion	69
Figure 12. Insights and Neuroticism score.....	69
Figure 13. Summary of findings.....	72
Figure 14. The sensemaking loop. From Russell and Card, 1993.	81
Figure 16. The human cognition model (Green & Ribarsky, 2008).	89
Figure 17. Items in the 9-item short measure. Adapted from Green, Jeong and Fisher, 2010.	96
Figure 18. Summary of Findings from Study 2.....	98
Figure 19: The NCBI Viewer.....	109
Figure 20: The main view of GVis.....	110
Figure 21. GVIS Correlations across Procedural Tasks completion times and personality traits.....	116
Figure 22. Correlation across tasks in MapViewer between completion times and personality traits.....	117
Figure 23. Scatter Plot overlay of Correlation of Generated Insights and Locus of Control	118
Figure 24. Scatterplot overlay of generated insights and Big Five Extraversion.....	119
Figure 25. Scatterplot overlay of correlation between generated insights and Big Five Neuroticism.	120

Figure 26. A summary of chapter findings.	126
Figure 27. A summary of the correlations between all aspects of the Personal Equation.....	126
Figure 28 . A summary of the user profile of a more efficient procedural learning in both interfaces.....	137

List of Acronyms

GOMS	Goals, Operators, Method, and Selection rules
GUI	Graphical User Interface
GVis	Genomic Visualization
IPIP	International Personality Item Pool
LOC	Rotter's Locus of Control
PEI	Personal Equation of Interaction
SFU	Simon Fraser University

Glossary

Term	Definition
individual difference	variance between participants in any measurable aspect of a cognitive study.
Inferential learning	A method of learning that uses the cognitive integration of pertinent information
interface learning	Acquiring the ability to use an interface through the use of reasoning and reasoning outcomes. This study evaluates 2 types of interface learning: procedural and inferential.
Personal Equation of Interaction	a function (or an equation) that allows us to use what we know about individual differences to predict interface learning outcomes.
procedural learning	The knowing-how to do something, the acquisition of knowledge about how to do or complete a task or goal
Psychometric	A measurement of a psychological construct or attribute.
Reasoning	The process of thinking about things or concepts in order to reach an cognitive outcome, such as a decision or judgement
visual analytics	the science of analytical reasoning supported by visual interfaces
Visualization	The graphical representation of data or information which uses interaction.

Chapter 1. Introduction

OUR information-rich age has blessed us with data in such abundance that it often defies description. Organizing, representing, and manipulating these data has engendered new domains of expertise in data mining and visualization. For data analysts, whose tasks are to evaluate complicated situations in real time and make evidence-based decisions in a timely manner, such as air traffic controllers, emergency response teams, and weather warning systems, the retrieval of information can involve a chaotic confluence of data, timing, and analytical cognition.

In this work, we will answer a central question, the pertinence of which may not be readily apparent now, but it will be by the end of this chapter. This research question can be broken down into several sub-questions, each of which will be addressed in this section. We will also address key concepts or drivers that motivated the research that is discussed in Chapter 2 through Chapter 5 before our final thoughts are provided in Chapter 6.

1.1. Research Questions

What is the Personal Equation of Interaction for Interface Learning? In other words, what personal equation of individual differences predicts reasoning outcomes during visual analytics interface learning?

This central question can be broken down into the following sub-questions:

1. Why is an analysis of reasoning important to visual analytics?
2. How do we use individual differences to predict reasoning outcomes?

3. How can reasoning be studied?
4. How do we use individual differences to study and predict reasoning outcomes?
5. What is a personal equation and how can it be used?

We will now discuss these questions one by one.

1.2. Why Is The Study Of Reasoning Important To Visual Analytics?

Visual analytics uses visual artifacts to aid in the solutions to “wicked problems” (Thomas & Cook, 2005). By definition, these problems require iterations of reasoning and reasoning outcomes before a solution can be built. In previous work (2008), we examined in depth the interaction between the visual analytics interface and reasoning outcomes. Based on the evaluation of the several interfaces and the types of analysis they were built to support, we presented a limited process model of the interaction of human cognition and the interface during visual analytics analysis. Several types of reasoning outcomes seemed common to the interactions with all of the visual analytics interfaces evaluated: information discovery, search by example or pattern, new knowledge creation, and the generation and analysis of hypotheses (please see Figure 1).

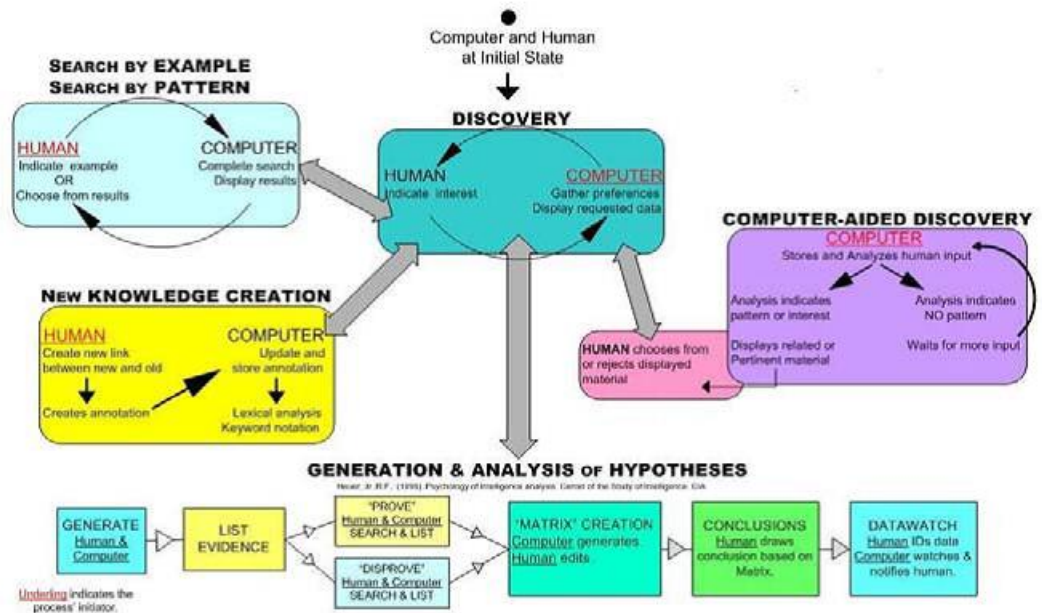


Figure 1. The interaction between human cognition and the interface. *source:* Green, Ribarsky, and Fisher (2008)

In addition to this early effort to diagram the process of complex interface interaction, several previous studies have attempted to study the interaction of reasoning and the visual analytics interface (whether the researchers realized that they were studying reasoning or not). For example, Ware, Neufeld, and Bartram (1999) as well as Bartram and Yao (2008) used animation to show how a reasoner might infer causality (a type of judgment) through the animation of contextual cues. Saraiya et al. (2006) conducted a longitudinal case study of hypothesis generation, which, by design, involves a wide variety of reasoning types, such as classification and satisficing (which are rule-based reasoning, usually deduction). Sensemaking, which is the label given to early generalizations about presented data or systems of data and is useful in hypothesis generation (Pirolli & Card, 1999), is a form of abduction. General inferences about the

meaning of the relationships between visualized data (e.g., those evaluated by Saraiya et al.) are an information integration cognitive process that uses induction.

The work of Correa, Chan, and Ma (2009) is part of a considerable body of literature that evaluates uncertainty in visualized data validity and relationships. The study of uncertainty visualization is, among other things, a study of analytical cognition. Much of the reasoning concerns the validity of data, its relationships, or its pertinence to the problem at hand. Correa et al. (2009) evaluated the problem of uncertainty in data validity by multiple transformations of the data so that uncertainty measures could be aggregated. These aggregations were visualized through scatterplots and correlational matrices, which we judged to be preferable in helping an analyst reason through uncertainty. Uncertainty can be an aspect of data, the noise in the data, the classification of data types, and so on.

In recent work (Greensmith, 2016), we explored uncertainty in classification types. We built composite glyphs, the categories of which were ambiguously defined (e.g., there was no determining attribute for either category, and Category A attributes could be also be found on Category B exemplars). This allowed us to evaluate multiple types of categorization and the effects of individual differences on the accuracy of these categorical decisions.

Irani and Ware (2000, 2003) also approached reasoning by using glyphs, positing that composite glyphs built with Biederman's geons (1987) would allow the analyst to deduce the correct classification more intuitively, especially in complex concepts. They also found that these composite glyphs improved long-term memory by improving recall.

Some researchers of visual analytics have studied cognition holistically. Kazancioglu, Platts, and Caldwell (2005) evaluated the visualizing of strategic decision making by focusing on the formation of strategies and the role that visual artifacts play in this process. Strategy, much like hypothesis generation, involves more than one type of reasoning. Deduction (e.g., rule-based methods, such as target identification), induction (e.g., information integration), and abduction (often referred to as sensemaking) all play

a role. In Chapter 5, we discuss in detail these three types of reasoning and the manner in which they affect visual analysis.

In Chapters 2, 3, and 5, we study the interface design structures commonly used as visual artifacts in reasoning. We comparing the performances of the participants in a GOMS-based interface and a data visualization that were equipped with the interaction paradigms of zooming, brushing, linking, and glyphs, and we evaluated the interface learning as either categorization tasks of classification (a rule-based reasoning) or inference (an information-integration reasoning).

Heer and Agrawala (2008) argued that the abduction of sensemaking could be a social process. They explored the relevance of collaborative cognition concepts, such as peer production and the cost of intelligence. This research was one of several lines of inquiry in which visual analytics researchers sought to overlay established theories of cognition on visual analysis with varying degrees of success. Another example is Meyer (2010), which applied theories of perceptual understanding (including Pylshyn's FINST, 2007) to visual analytics.

Although it may not be explicitly defined, the role of reasoning is well recognized in the visual analytics literature. It has always been studied through reasoning outcomes, however, and it is often based on natural observations or simple evaluations. The research that we describe in the following chapters differs from its contemporaries in that the focus is on reasoning and on predicting reasoning outcomes (visual analytics as the applied domain) instead of visual analytics interfaces and their use in secondary or ancillary discussion. A knowledge domain requires the study of both form and function and of both theory and practice.

1.3. How Do We Use Individual Differences To Predict Reasoning Outcomes?

Visual analytics is the science of analytical reasoning supported by visual interfaces (Cook & Thomas, 2005). We have established that reasoning is generally

considered an asset in the process of visual analytics. However, what makes analytical reasoning unique? In a recent paper (Green & Maciejewski, 2013), we proffered the following definition:

Analytical reasoning can be defined in a variety of ways. In addition to the Kantian idea of analytical reasoning as an evaluation of the validity or virtue of the proposition itself, we will also consider analytical reasoning as a determination about the value of given associations between concepts or statements. Note that except determinations about validity, no other outcomes are required in analytical reasoning. This is important because it highlights a core characteristic: reasoning has little or no explicit observable behavior. Reasoning is usually not defined as the outcome; it is defined as how the outcome is made possible. This may not be explicitly stated, but it is a common assumption in the psychology of reasoning literature. Because reasoning and the cognitive processes it informs are so closely interrelated, they are often studied together.

Johnson-Laird (1978, 1980) studied mental models through the decisions that participants make about formal syllogisms through deductive reasoning. His research demonstrated that these models are used to make decisions and solve problems, but a model or a system of mental models can also be used to make a variety of decisions or create multiple problem solutions. That is, a model is not the decision or the problem's solution; it is how the decision or solution is reached. Johnson-Laird postulated this (Johnson-Laird & Shafir, 1993), arguing that reasoning and decision-making inform each other, but the two are separate cognitive processes (p. 4)

In other words, analytical reasoning is a decision-making process. The decision is the degree of confidence in the validity of the visual analytics process for the problem at hand. Hence, analytical reasoning might be seen as an umbrella decision process that aids the reasoning outcomes of the types of reasoning discussed in the previous section by assuring a sufficient degree of confidence in the reasoning outcomes.

1.3.1. How Can Reasoning Be Studied?

As a field of study, reasoning is challenging because it is involved in many types of cognitive tasks. Many cognitive tasks common to visual analysis and its outcomes are closely associated with reasoning, such as judgment (Tversky & Kahneman, 1983; Piaget, 2002) and decision-making (e.g., Legrenzi, Girotto, & Johnson-Laird, 1993; Evans, Over, & Manktelow, 1993). With the possible exception of stimuli detection, reasoning is a handy toolbox for almost every aspect of human cognition. Indeed, there

are so many reasoning outcomes that we often give distinct labels to groups of reasoning-dependent outcomes, such as classification, target identification, and satisficing. These labels tend to describe the outcome of the reasoning process, not the reasoning process itself.

There is a reason for the emphasis on reasoning outcomes. For the researcher, reasoning tends to be an unseen process. It is a difficult research topic because no method to quantify and evaluate this cognition is readily apparent. However, as mentioned earlier, this has not stopped theorists such as Johnson-Laird, Stanovich, and Gigerenzer from tackling reasoning from the standpoint of evaluating reasoning outcomes. In each case, the research focused on small problems, the solutions of which would allow the researcher to make defensible assumptions about how the solution was reached (i.e., the reasoning). By choosing a problem or task that clearly requires the use of one type of reasoning (such as deduction) over any other, researchers can constrain their tasks to evaluating the reasoning outcomes and inferring insights about the reasoning itself. The study protocols of Johnson-Laird, for example, included deductive syllogisms, which could be difficult to do correctly. Syllogisms are based in formal logic, and their use in reasoning research tends to be normative. Johnson-Laird used the responses to each task as the basis for his “mental models,” which are a theory of reasoning cognition. In one paper (1978), he described his problems as follows:

Each subject was asked to make a deduction from the 27 pairs of premises that are shown in Table 1 with their valid conclusions italicized. The problems were presented with a sensible content of a sort unlikely to predispose subjects toward a particular conclusion. Hence, a typical pair was:

*None of the musicians is an inventor.
All of the inventors are professors.*

In another paper that explained the relevance of his research for cognitive science, he described how a tasks' “atmosphere” or phrasing could control outcomes. The use of negative aggregates such as “none” elicited different deduction than positive ones elicited (e.g., “all” and “always”) (1980):

One datum that is difficult to reconcile with the effect is that certain premises from which a valid conclusion can be drawn tend to be judged not to imply any conclusion. Here is an example:

*Some of the beekeepers are artists.
None of the chemists is a beekeeper.*

When such premises were presented in one experiment, 12 out of 20 subjects declared that there was no valid conclusion that could be drawn from them (see Johnson-Laird & Steedman, 1978). In fact, there is a valid conclusion:

Some of the artists are not chemists.

Moreover, it is entirely congruent with the atmosphere effect, particularly because the first premise is particular, and negative because the second premises negative. Only 2 of the 20 subjects drew this conclusion. Such findings require at the very least some modification of the atmosphere hypothesis.

Stanovich, whose broad interests in reasoning tended to focus on cognitive ability and its effects on rational outcomes, administered a variety of tasks to his participants. One syllogistic task, which was first used by Markovitz and Nantel (1989), asked the participants to make decisions about the validity of eight syllogisms that “followed logically but were unbelievable”:

*Premise 1: All things that are smoked are good for the health.
Premise 2: Cigarettes are smoked.*

Conclusion: Cigarettes are good for the health.

Other commonly used tasks which Stanovich employed were Wason’s Selection Task (which is discussed in detail later in this section) and statistical reasoning tasks similar to those in Fong, Krantz, and Nisbett (1986):

Probabilistic-Structure 1: At Stanbrook University, the Housing Office determines which of the 10,000 students enrolled will be allowed to live on campus the following year. At Stanbrook, the dormitory facilities are excellent, so there is always great demand for on-campus housing. Unfortunately, there are only enough on-campus spaces for 5,000 students. The Housing Office determines who will get to live on campus by having a Housing Draw every year: every student picks a number out of a box over a 3-day period. These numbers range from 1 to 10,000. If the number is 5,000 or under, the student gets to live on campus. If the number is over 5,000, the student will not be able to live on campus. On the first day of the draw, Joe talks to five people who have picked a

number. Of these, four people got low numbers. Because of this, Joe suspects that the numbers in the box were not properly mixed, and that the early numbers are more favorable. He rushes over to the Housing Draw and picks a number. He gets a low number. He later talks to four people who drew their numbers on the second or third day of the draw. Three got high numbers. Joe says to himself, "I'm glad that I picked when I did, because it looks like I was right that the numbers were not properly mixed." What do you think of Joe's reasoning? Explain.

Scenario-based laboratory tasks such as this one are not uncommon, and they involve the use of a story to elicit a response or series of responses. Researchers have used many such tasks to evaluate aspects of reasoning. For example, Tversky and Kahneman (1981) used several versions of the "disease problem". They described two scenarios and asked the participants to choose one:

Problem 1 [$N = 152$]: Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

If Program A is adopted, 200 people will be saved. [72 percent]

If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved. [28 percent]

Which of the two programs would you favor?

Most of the participants in this study picked Program A. When the problem was presented in other ways and used different scenarios, the participants changed their minds and chose Program B. Tversky and Kahneman demonstrated that context was highly pertinent when an analyst must reason through two similar but different choices. Reasoning outcomes, they argued, depend not only on the problem but also on how it is described or presented.

The reasoning research reviewed so far has been highly textual or reading based, but that is not always the case. Sometimes visual elements are also used. Cherubini, a student of Johnson-Laird, used both syllogisms and graphical card tasks to

move Johnson-Laird's models into fast heuristics. As we discuss in Chapter 5, his research demonstrated that after being exposed to a logical proposition as few as three times, his participants used inference to infer a deductive rule that they then used to solve similar problems (See Chapter 5 for a detailed description of this experiment.)

Cox and Griggs (1981) used an adaptation of the Wason Card Selection Task (1966), which was termed the "Drinking Age Problem". Wason used letters and numbers on his cards, but found consistently poor performance in an abstract task. The Drinking Age Problem applied the same task, but it was framed in a situation to which the participants could relate. Stanovich (1999, p. 128) described the task as follows:

When testing the rule "if a person is drinking beer, then the person must be over 19 years of age," and when given the four cards beer, Cole, 22 and 16 to represent P, not P, Q and not Q respectively, performance is markedly superior to that on the abstract selection task.

Moving away from deductive and inferential syllogisms in the conclusion, Gigerenzer studied inference and induction by using the participants' *a posteriori* knowledge of the world to make quick decisions about the size of European cities (1996):

Which city has a larger population? (a) Hamburg (b) Cologne.

Gigerenzer postulated that these one-decision tasks were an example of bounded rationality that was based on limited information rules. Instead of acquiring knowledge about the task through syllogisms, scenarios, or graphical representations, his participants based their decisions on their own knowledge.

In these and similar cases, the typical approach was to develop a narrow laboratory task so that the research could reasonably be considered to have content validity. Yamauchi and Markman (2000), for example, designed studies that created a series of tasks in two ambiguous categories, which were learned through exemplars. The participant then had to make a single decision about each presented visual artifact. In this task, the reasoning that classifies a visual artifact into one of two categories is

different from the reasoning that uses category labels to infer the characteristics of the category. While complex, these tasks were carefully constrained.

For a visual analytics researcher, the problem with this approach, of course, is that there is little in its scope or practice of visual analytics that could be considered narrow. Its questions are complex, and its interfaces are interactive. Furthermore, it can reasonably be asserted that visual analytics requires the analyst to use more than one type of reasoning in any given analytics session (e.g., the multi-stage process of hypothesis generation).

In this research program, we chose a blended approach. We used real-life visual analytics interfaces as part of our stimulus set. Both interfaces used throughout this program were expert systems built on the same dataset. The first interface was a Graphical User Interface (GUI) called MapViewer that continues to be used by genomic researchers across the United States. By using real-life interfaces, we were able to create a holistic environment that closely mimicked the problems that analysts would face. This improved the overall usefulness of the measured outcomes.

However, although we used a real world interface, we paired the GUI with relatively straightforward single and multiple-point decision tasks. The study tasks asked participants either to find a specific item in the interface (target identification) or to make some inference about the category by integrating information that is more complex. Target identification involves deductive or top-down reasoning, while information integration requires induction or bottom-up reasoning. Instead of using traditional reasoning tasks, we borrowed concepts from other interface evaluations, which regularly asked participants to find something or use some aspect of the interface in order to evaluate the interface's learnability (e.g. Irani & Ware, 2000, 2003). Because these laboratory tasks were undertaken on real world interfaces, the participants were able to reason more effectively with a visualized interface, but had an easier time finding solutions in a known interaction paradigm, such as a GUI interface.

We have adapted reasoning research protocols so that they can be directly applied to visual analytics research. In Greensmith (2016), we used composite glyphs to study two distinct types of visual analytics categorization. Classification and attribute

inference performance were predicted by measured psychometric differences. We isolated a predictive measure that could predict both types of categorization. While the studies were conducted online, the tasks were performed in a laboratory. No previous research has reported the use of a similar composite glyph in the real world. Although the tasks were administered online, we used careful adaptations of the category-attribute schedules in previous reasoning research (e.g. Yamauchi & Markman, 2000).

1.4. How Do We Use Individual Differences To Study And Predict Visual Analytics Reasoning Outcomes?

When the goal is to study a phenomenon through prediction, the obvious concern is that the researcher will never find a quantifier that can reasonably be expected to predict the phenomenon. It helps if the predictor is reliable and can be generalized to the target population.

How does the researcher stumble upon predictors? In cognitive research, a common place to start is to search individual differences. *In study of cognitive psychology, an individual difference is the variance between participants in any measurable aspect of a cognitive study.* Not surprisingly, individual difference is a broad topic in the cognition literature. Individual differences have been used to predict everything from affect (Gross & John, 2003; Nowicki & Duke, 1994), social categorization (Moscowitz, 1993), and rationality (Stanovich, 1999) to computer skills (Harrison & Rainer, 1992). A thorough literature review on the effects of individual differences is beyond the scope of the current work.

This program of research focuses on a specific type of variance in the search for predictors: inherent differences in personality and learning style. By inherent, we mean predictors that the participant could be said to have had since birth. These predictors are not particularly malleable; they can be shaped by early human experience, but by adulthood, they are considered stable. For this reason, they are called *traits* (Rotter, 1966). In other words, the analyst brings the same set of individual differences to every visual analytics task. These differences cannot be “changed” or diminished, so they are ideal for the task of prediction. Once measured, they can be used to generalize

expectations to other analysts with the same set of differences. What makes the study of these differences even more fascinating for the task at hand is that these differences can interact with each other, trending together systemically and allowing for the development of expert profiles. This interaction of differences can be seen in the profiles of chosen professions. Persons who self-select to acquire a set of job skills often have similar trait profiles (Blau, Super, & Brady, 1993; Gambles, Wilkinson, & Dissanayake, 2003; Warbah et al., 2007; Rose et al, 1982; Schroeder, Broach & Young, 1993). In Chapter 5, we develop a trait profile of excellent procedural learners and find that these superior performers tend to be moody extraverts who dislike novel environments and new experiences.

1.4.1. The Non-Mystery of the Assumed Innate

At this point in the discussion of inherent traits as predictors of visual analytics cognition, there is usually a critic who insists that superior analytical performance has nothing to do the analyst's personal proclivities but is somehow a pure function of cognitive ability and intuitive interfaces. This view is bemusing because cognitive ability itself is a stable trait-like individual difference (Roberts et al., 2007; Ackerman, 2003; Plomin, 1999). Some analysts have better cognitive abilities than others and are better able to make sense of the complex visual analytics environment. Is not this the very reason to build "intuitive interfaces" in the first place, as artifacts that support cognitive shortcomings in the analyst (Green & Ribarsky, 2008)? Is it really so strange that some analysts might be born with natural abilities that others do not have?

Thousands of psychometric measures of inherent differences have been reported in the psychological literature. Each measure was carefully evaluated statistically and compared to measures of similar constructs. From the near-beginning, when Sir Francis Galton (1869) declared in defense of early Darwinian eugenics that "man's natural abilities are derived by inheritance" (1896, p. 1) to the present day and Pinker's emphatically defended innate characteristics (2003) against Locke's *tabula rasa* and ghosts in machines, researchers have striven to isolate and measure the many aspects of unseen human personality, cognitive styles, and reasoning. They have particularly sought to demonstrate the ways in which such aspects affect outcomes

whether sociological, academic, or cognitive. With reference to this established corpus, our current research is one applied example of the statistical association of psychometrics with observed outcomes.

This current research is not a philosophical discussion of nature versus nurture. Nor does it grind the axe of personality's role in, frankly, much of anything at all. We found very early in our research that innate personality traits were linked with learning performance (Green & Najarian, 2007). This somewhat crude seminal work found correlations between the Big Five personality traits and the participants' ability to build complex 3D structures from pictorial instructions. It seemed only natural to learn what else personality traits could predict and to transfer that predictive power into the complexity of visual analytics. It is not the personality traits themselves that are the fascination. It is the statistical power in the systemic assembly of self-report psychometrics that can and does predict complex cognitive outcomes. If we could find similar predictive power in the color of a participant's hair or the time of day, we would abandon the current approach for one that was computationally easier.

To be completely frank, in some ways this current study of inherent analyst characteristics could be seen as a reaction to the overwhelming view in visual analytics that the human is an information processor. Even a superficial study of human reasoning, including the role of Piagetian adaptation and accommodation, the rule mechanization of Cherubini, or even the satisficing elimination of Gerginzer, reveals the rigidly hierarchical and overly simplistic assumptions of information processing theory. Human cognition concerns the use of tools. It is not one cognitive tool, but many cognitive tools used in combinations, in varying orders of engagement depending on problem and context, in varying degrees of rationality depending on the tool user, and in the face of cognitive fallacies, adverse affect, priming and contextual errors, including illusion. All interact with and are bounded by innate traits that are consistent enough to be predicted. In view of this complexity, the present research program may indeed be seen as one small assertion that human cognition is not a computer.

The choice of psychometrics in this research protocol was not random. Of the thousands of measures available, several stood out early and often as likely predictors.

Constructs such as locus of control have been known since the 1960s, and they have already been debated, reviled, and then eventually accepted by the scientific community (Marsh & Richards, 1986; Anderson, 1977). Furthermore, some of our earliest laboratory research on the role of affect in learners showed that the Big Five Personality was influential in cognitive outcomes (Green & Fisher, 2010). Moreover, research such as Judge et al. (2002) demonstrated correlations between locus of control, the Big Five Personality Index, and other measures. These inherent constructs, for whatever reason, tended to appear as predictors in study after study. As demonstrated in Chapters 2, 3, and 5, the present research found a high level of inter-correlation between psychometrics. Extraversion items consistently correlated with neuroticism items. Locus of control items consistently correlated with neuroticism and tolerance of ambiguity items to the degree that the process of isolating predictors could isolate a very short measure with a high degree of prediction, as shown in Chapter 2.

In this research, isolating psychometric items as part of a predictive measure follows the basic rules of psychometric assessment development. If possible (always, in the case of the current work), previously used and normed items (e.g., survey questions) are used. The self-reported item responses are evaluated statistically for their relationships to each other and to the desired cognitive outcome. Items that might be related but are not strong predictors are eliminated. The remaining items are examined by using established protocols of factor reduction, focusing on items that are the best predictors of outcomes. These items are “spun” around each other, to reveal those that clump together because of similarities in their variances. These subgroups of items are then evaluated as both whole measures and separate measures to determine the best representatives of the group. When the best items are isolated, the newly minted psychometric measure is evaluated. Its predictive power is tested holistically, and its internal strength is measured through goodness-of-fit tests, such as the Kolmogorov-Smirnov. A detailed description of this process of assessment creation is provided in Chapter 2.

1.5. What Is A Personal Equation And How Can It Be Used?

We have saved the definition of a personal equation of interaction for discussion in the conclusion of this chapter, for it depends heavily on everything that has been previously discussed. So far, we have discussed the need for the understanding of individual differences and how they affect reasoning outcomes. We have discussed the ubiquity of reasoning in visual analytics and in cognition in general. However, how do we take a measure of reasoning and use it to predict visual analytics outcomes?

Earlier in this chapter, we briefly touched on the many kinds of reasoning tasks that visual analytics can employ. Reasoning, of course, is an internal cognitive process and can only be studied by measuring its outcomes. In this line of research, these outcomes are accuracy ratings and timing, which are measures of how effectively and efficiently the reasoner was able to use of the visual interface to achieve an analytical outcome. These measures comprise half the equation—the measure of reasoning outcomes.

The other half of the equation is the measure of personal traits (e.g., psychometrics) and how they affect the way the analyst discovers, uses, and synthesizes data in the quest for accurate reasoning outcomes. In this research, before the participants perform the task, they are administered a battery of psychometrics that have been normed as predictive of the desired reasoning outcome. This battery could contain items from any of the individual differences that we have already discussed. In Chapters 2, 3, and 5, Rotter's Locus of Control items and those of the Big Five attributes of Neuroticism and Extraversion are identified as useful predictors of the reasoning outcomes. The analyst's unique score on psychometrics is then used to predict the likelihood of a desirable outcome.

In other words, the Personal Equation of Interaction for Interface Learning is a function (or an equation) that allows us to use what we know about individual differences to predict interface learning outcomes. The psychometrics which predict interface learning accuracy and speed will be to some degree unique to this particular task. Although similar cognitive tasks might overlap, the types of individual differences that

predict this type of reasoning are likely to deviate slightly from the differences that predict categorization outcomes, for example.

This is the reason that some of the literature on the Personal Equation of Interaction (PEI) refers to it as a matrix of functions (see Chapter 4). Each function will uniquely predict one type of cognitive task more accurately than the other functions in the PEI. There will likely be overlap, but each type of cognition in visual analytics must be studied separately and that knowledge is added to the entire PEI as a research program.

In a simple example, we will use a three-item measure that was isolated in Greensmith (2016), which predicts the classification of composite glyph performance. The three items are taken from the Rotter Locus of Control (1966) and the Index of Learning Styles (1988). Combined into a whole measure, the three items account for 38% of the variance in the participants' classification scores, which indicates that the PEI for classification is moderately predictive.

We could use the PEI in several ways, but for the purposes of illustration, we will employ an early test of the relationship between the Personal Equation and its power to predict classification. We evaluated the strength (or importance) of each item in the Personal Equation by ranking the beta coefficients. The strongest item was Locus of Control 13, which asked the participants to choose which statement best described themselves:

- *When I make plans, I am almost certain that I can make them work.*
- *It is not always wise to plan too far ahead because many things turn out to be a matter of good or bad fortune anyhow.*

The second most important variable was the ninth item of the Locus of Control.

- *I have often found that what is going to happen will happen.*
- *Trusting fate has never turned out as well for me as making a decision to take a definite course of action.*

The last item was the first item on the Sequential/Global continuum of the Index of Learning Styles, which is a measure of whether participants use a top-down organizational style or a bottom-up, sequential style:

- *I tend to understand the details of a subject but may be fuzzy about its overall structure.*
- *I tend to understand the overall structure but may be fuzzy about its details.*

The scores for each of these three items were entered into an Enter-method Chi-based regression model. The regression model statistically scored each participant with a probability score that reflects a prediction of how accurate the participant will be in the performance of the task. This prediction can be reflected in the Standardized Predictive Classification Value (ZPRED) for each participant. The ZPRED is the predicted accuracy computation for the participant. In other words, the model used what it had learned about each participant's individual differences to predict the likely accuracy score for each participant. The ZPRED is not the predicted score itself, but it is a weight that will allow us to find the predicted score.

We then graphed the relationship between the model computation (ZPRED) and the actuals to find an equation that would reasonably predict actual reasoning outcomes. In Figure 2, each dot represents the relationship between each participant actual accuracy score (the Y axis) and the accuracy score that the model predicted for the participant (the ZPRED on the X axis).

The solid line in Figure 2 reflects the relationship between the participant's accuracy score and the model's computation. By graphing the relationship, we can find the function that allows us to predict a participant's accuracy score from the ZPRED.

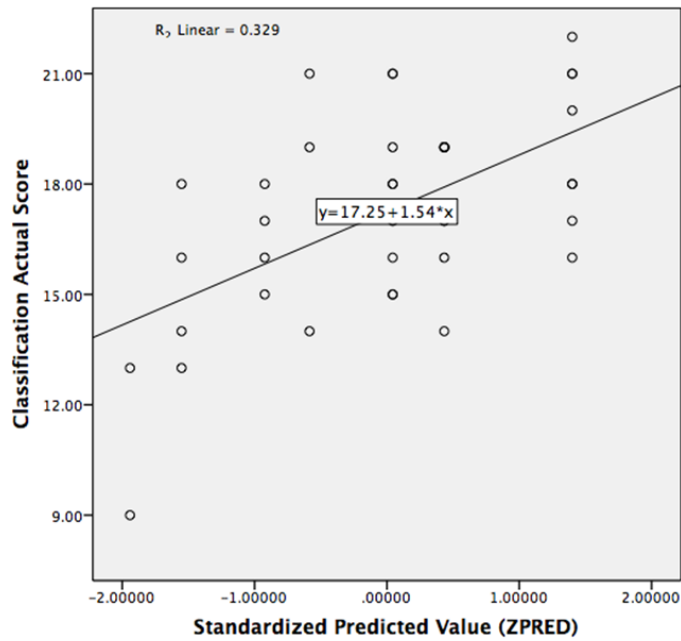


Figure 2. Example of the relationship between the PEI predicted values and the actuals.

As Figure 2 shows, this method uses a very simple equation to predict a participant's actual accuracy:

$$f(\text{Number Correctly Classified}) = 17.25 + 1.54(\text{ZPRED})$$

This function is an expression of the Y intercept of the line, which is generically expressed as $y = mx + b$. Put another way, the function allows us to derive the classification accuracy for each participant by using this Personal Equation of Interaction.

As an example, for Participant 1, whose ZPRED is .044, the PEI yields

$$f(\text{Number Correctly Classified}) = 17.25 + 1.54(.044)$$

$$f(\text{Number Correctly Classified}) = 17.32$$

In this case, the model predicted a rounded accuracy score of 17, which was indeed the actual accuracy score for Participant 1.

Obviously, running the model on what is essentially the model's training set is hardly a validation of the model. In Greensmith (2016), we validated the PEI by using it to predict performance in a replication of the study. Although for some purposes, this replication may suffice as validation, we realize that norming the PEI for classification will require a replication with a large participant set, preferably using stimulus sets that are similar but not exactly the same as the training set. However, this example suffices to demonstrate how the PEI could work.

1.6. Summary

The foregoing discussion has answered each sub-question. Let us review the research question:

What is the Personal Equation of Interaction for Interface Learning?

In the following chapters, we will create a series of study protocols that will use real-life visual analytics interfaces and laboratory tasks to evaluate interface learning through the reasoning outcomes of target identification and information integration, by both procedural and inference learning.

We will predict the outcomes of these visual analytics tasks by using self-report measures of individual differences to create a model or short measure that uniquely predicts these reasoning outcomes. We will also use the PEI for Interface Learning as a foundation for the PEI functional matrix.

Is the PEI useful in practice? This is still an open question, but a few applications seem promising. As the interface interaction becomes more intimate, the knowledge of why participants prefer one interaction style or paradigm will inform the design for the target users. This works in multiple directions. The pre-design understanding of the analysts' PEI profile allows the designer to choose displays and interaction schemas

that are the best suited for the expert cohort. For example, some research suggested that a majority of schoolteachers have a preference for learning through reading texts and organizing data in a bottom-up, sequential fashion (Felder & Silverman, 1988). With this knowledge, any expert educational system would not be heavily pictorial or follow Schniederman's mantra (1966), which begins with a top-down overview. Instead, the data would be organized in a topical, step-by-step manner using textual pointers and paragraphs organized into readable information. The PEI does not show that one type of knowledge representation is always better than another is; it selects the best type of representation based on the target analyst audience. In an expert cohort that is highly visual and spatially coherent, the schoolteacher's very wordy expert system would be ineffective. The visual-spatial analyst might be able to adapt to the schoolteacher's interface, but insights would be synthesized slowly, and some pertinent information might be missed entirely. The PEI contributes to the domain of interface design by using the PEI to define the best interface that a visual-spatial expert would likely be able to accomplish key hypothesis-generation tasks the most effectively through interactive top-down paradigms, such as zooming, bushing, and linking. The best visual analytics interface is one that the analyst can easily learn and make sense of and one in which the data are organized in a way that the analyst understands best.

In reverse, the PEI contributes in a very real way with expert system recommendations. Knowing what we do about schoolteachers in the United States, we can choose tools and interfaces that are the best adapted to their organizational proclivities. Knowing *why* our analysts seem to see insights in one interface more easily than in another aids not only the visual analytics process for the user but prevents blaming the interface without cause. Schoolteachers may not work well with an interactive pictorial representation, but that does not mean that the interface is a failure; it is simply being used to support the wrong type of analyst.

And lastly, and to us most importantly, the PEI contributes to the realm of reasoning research by studying the 'black box' reasoning processes by predicting their outcomes. The intellectual beauty of understanding why some people reason better than others provokes a re-evaluation the manner in which reasoning is studied. The PEI tells us why one reasoner is better than another during task, and provides impetus for

research that compares the ‘why’ of each reasoning outcome with the ‘why’ of other reasoning outcomes. In the research that follows, we do this comparison in order to find a group of psychometrics that predict all of the primary task outcomes of the study. But in ongoing research, the PEI has become a benchmark for describing how reasoning tasks relate to each other. For example, locus of control explains both bottom-up and top-down categorization, but with different emphases that seem to have something to do with how comfortable the analyst is with the unknown. Knowing *what* and *when* have been the traditional domains of reasoning research; we intend to also know the *why*.

In the following chapters, we describe the boundaries of the PEI for Interface Learning. In Chapter 4, we review the theory of the PEI. In Chapter 2, we define the PEI task protocol of using common interface learning tasks and an administered battery of psychometrics. In this chapter, we will also use statistics to isolate the PEI. In Chapters 3 and 5, we demonstrate the reliability of Locus of Control, Extraversion, and Neuroticism as useful PEI constructs in predicting learning outcomes. We explore the PEI in Chapter 5 by using what we know to build the PEI’s first profile, which is a description of individual differences in a superior procedural learner. We offer concluding thoughts in Chapter 6.

1.7. References

- Ackerman, P. L. (2003). Cognitive ability and non-ability trait determinants of expertise. *Educational Researcher*, 32(8), 15–20.
- Anderson, C. R. (1977). Locus of control, coping behaviors, and performance in a stress setting: a longitudinal study. *Journal of Applied Psychology*, 62(4), 446.
- Bartram, L., & Yao, M. (2008) Animating causal displays. IEEE-VGTC Symposium on Visualization 2008 (Vol. 27, pp. 751–758). Presented at the Eurographics, Crete, Greece: Blackwell Publishing.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94(2), 115.
- Blau, T. H., Super, J. T., & Brady, L. (1993). The MMPI good cop/bad cop profile in identifying dysfunctional law enforcement personnel. *Journal of Police and Criminal Psychology*, 9(1), 2–4.
- Cherubini, P. & Mazzocco, A. (2004). From models to rules: Mechanization of reasoning as a way to cope with cognitive overloading in combinatorial problems. *Acta Psychologica*, 243, pp. 223–243.

- Cook, K. A., & Thomas, J. J. (2005). *Illuminating the path: The research and development agenda for visual analytics* (No. PNNL-SA-45230). Richland, WA (US): Pacific Northwest National Laboratory (PNNL).
- Correa, C., Chan, Y. H., & Ma, K. L. (2009, October). A framework for uncertainty-aware visual analytics. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium* (pp. 51–58). IEEE.
- Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition, 10*(5), 496–502.
- Evans, J. S. B., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition, 49*(1), 165–187.
- Evans, J. S. B., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition, 49*(1), 165–187.
- Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles. *Engineering Education, April*, 674–681.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18*(3), 253–292.
- Galton, F. (1869). *Hereditary genius*. Macmillan and Company.
- Gambles, M., Wilkinson, S. M., & Dissanayake, C. (2003). What are you like? A personality profile of cancer and palliative care nurses in the United Kingdom. *Cancer Nursing, 26*(2), 97–104.
- Green, T. M., & Maciejewski, R. (2013, January). A role for reasoning in visual analytics. In *System Sciences (HICSS), 2013 46th Hawaii International Conference* (pp. 1495–1504). IEEE.
- Green, T. M., & Ribarsky, W. (2008, April). Using a human cognition model in the creation of collaborative knowledge visualizations. In *SPIE Defense and Security Symposium* (pp. 69830C-69830C). International Society for Optics and Photonics.
- Green, T. M., Ribarsky, W., & Fisher, B. (2008, October). Visual analytics for complex concepts using a human cognition model. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium* (pp. 91–98).
- Green, T. M., & Najarian, K. (2007). *Is emotion regulation a predictor of learning performance and associated cortical activity?* Poster session presented at the North Carolina Cognition Conference, February 2007. Chapel Hill, NC, USA.
- Greensmith, T. (2016, under review). The personal equation of interaction for categorization. *Proceedings on Hawaii International Conferences on System Sciences*, January, 2016. Kauai.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348.
- Harrison, A. W., & Rainer Jr, R. K. (1992). The influence of individual differences on skill in end-user computing. *Journal of Management Information Systems, 93–111*.

- Irani, P., & Ware, C. (2000). Diagrams based on structural object perception. Paper presented at the Working Conference on Advanced Visual Interfaces, Palermo, Italy.
- Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, 4(1), 71–115.
- Johnson-Laird, P. N., & Shafir, E. (1993). The interaction between reasoning and decision making: An introduction. *Cognition*, 49(1), 1–9.
- Johnson-Laird, P. N., & Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology*, 10(1), 64–99.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, 83(3), 693–710.
- Kazancioglu, E., Platts, K., & Caldwell, P. (2005). Visualization and visual modeling for strategic analysis and problem-solving. In *IEEE International Conference on Information Visualization* (pp. 61–69). Minneapolis MN: IEEE-Wiley.
- Legrenzi, P., Girotto, V., & Johnson-Laird, P. N. (1993). Focusing in reasoning and decision making. *Cognition*, 49(1), 37-66.
- Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, 17(1), 11–17.
- Marsh, H. W., & Richards, G. E. (1986). The Rotter locus of control scale: The comparison of alternative response formats and implications for reliability, validity, and dimensionality. *Journal of Research in Personality*, 20(4), 509–528.
- Moskowitz, G. B. (1993). Individual differences in social categorization: The influence of personal need for structure on spontaneous trait inferences. *Journal of Personality and Social Psychology*, 65(1), 132.
- Nowicki Jr, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of Nonverbal Behavior*, 18(1), 9–35.
- Piaget, J. (2002). *Judgment and reasoning in the child*. Routledge.
- Pinker, S. (2003). *The blank slate: The modern denial of human nature*. Penguin.
- Plomin, R. (1999). Genetics and general cognitive ability. *Nature*, 402, C25-C29.
- Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world*. MIT Press.
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313–345.
- Rose, R. M., Jenkins, C. D., Hurst, M., Herd, J. A., & Hall, R. P. (1982). Endocrine activity in air traffic controllers at work. II. Biological, psychological and work correlates. *Psychoneuroendocrinology*, 7(2), 113–123.

- Saraiya, P., North, C., Lam, V., & Duca, K. A. (2006). An insight-based longitudinal study of visual analytics. *Visualization and Computer Graphics, IEEE Transactions on*, 12(6), 1511–1522.
- Schroeder, D. J., Broach, D., & Young, W. C. (1993). Contribution of personality to the prediction of success in initial air traffic control specialist training (No. DOT/FAA/AM-93/4). Federal Aviation Administration Washington Dc Office of Aviation Medicine.
- Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium* (pp. 336–343). IEEE.
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293.
- Ware, C., Neufeld, E., & Bartram, L. (1999, October). Visualizing causal relations. In *Proceedings of IEEE Information Visualization* (Vol. 99).
- Warbah, L., Sathiyaseelan, M., Vijayakumar, C., Vasantharaj, B., Russell, S., & Jacob, K. S. (2007). Psychological distress, personality, and adjustment among nursing students. *Nurse Education Today*, 27(6), 597–601.
- Yamauchi, T., & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 776.

\

Chapter 2. Using personality factors to predict interface learning performance.

2.1. Overview

This chapter was originally published as *Green, T. M., Jeong, D. H., & Fisher, B. (2010, January). Using personality factors to predict interface learning performance. In System Sciences (HICSS), 2010 43rd Hawaii International Conference on (pp. 1-10). IEEE.* At HICSS, this paper won Best Paper in Track. This paper was the first published on the Personal Equation of Interaction and as such outlines some of the rudimentary assumptions and discusses the differences between interface and reasoning tasks. Because many of the psychometric whole measures did not predict the primary outcome of interface learning efficiency, the individual items from the psychometric measures, including Locus of Control, Beck's Anxiety Inventory, and Budner's Tolerance of Ambiguity were used to build a 9-item psychometric assessment that could be fine-tuned to predict interface learning efficiency (i.e. completion times). The null hypothesis was that there would be no association between the 9-item measure and interface completion times. The primary alternate hypothesis was that the 9-item measure would reasonably predict interface learning efficiency. The alternate hypothesis was supported, as the 9-item measure moderately predicted interface learning efficiency.

2.2 Abstract

This current study explored the impact of individual differences in personality factors on interface interaction and learning performance in both an interactive visualization and a menu-driven web application. Participants were administered 6 psychometric measures designed to assess trait anxiety, locus of control, and other personality traits. Participants were then asked to complete 3 procedural tasks and 3

inferential tasks in each interface. Results demonstrated that participants with an external locus of control completed inferential tasks more quickly than those with an internal locus. Factor analysis of items from the 6 psychometric scales isolated a 9-item short measure, which showed trending with procedural scores. Additionally, data demonstrated that the visualization interface was more effective and efficient for the completion of the inferential tasks. Participants also preferred the visualization to the web interface for both types of task. Implications and future directions of this research are also discussed.

2.2. Introduction

The successful visualization of complex information relies fundamentally on its ability to stimulate human cognition. Humans see what is visualized, emphasize information of interest through focused attention and elimination heuristics, and interact with representations of relational knowledge to reach a goal or complete a task that the human has chosen but that the visualization must facilitate. Each facet of human cognition engaged while using a visualization needs consideration; cognitive processes are typically not linear, and perception, categorization, and problem-solving activities inform and motivate each other throughout the interaction. The loss or impediment of one cognitive process hampers or stymies not only the other thought processes dependent upon it, but the path taken by cognition as a whole (Green, Ribarsky, and Fisher, 2008).

Numerous studies have been undertaken in recent years to evaluate visualizations. Plaisant has outlined four current categories of evaluation: controlled experiments comparing design elements, usability evaluations of tools, controlled between-visualization comparisons, and tool case studies (Plaisant, 2004). Each type of evaluation serves a purpose, but is self-limiting in multiple ways. As is often the case with experimentation generally, these evaluations involve small, simple, sometimes normative or “non-real-world” tasks; interaction in the real world tends to more complex, harder to predict, and thus harder to measure. Additionally, these evaluations focus on the more binary of cognitive processes. In each case, the cognitive variables measured are facets of vision, given attention, and perhaps tactile manipulation. These variables

are indeed important. However, these evaluations are not designed to answer one pivotal question: The human has possession of the target information. Now what?

Perhaps because humans learn so readily, we tend to take the supporting cast of cognition for granted. The visualization literature discusses this in a roundabout way, inferring reasoning and problem-solving from perceptual behaviors. However, there is a general belief in the centrality of learning to visualization, as is evidenced by the continuing discussion about what insight is and how to engender it during interaction (North, 2006, Plaisant, Fekete, and Grinstein, 2008). The question is not whether learning occurs, but how and when. These questions have yet to be tackled by visualization evaluation. One reason these questions remain unanswered is the complexity of the cognition involved.

Reasoning and problem-solving are not uniformly sequential, but rather utilize a variety of heuristics, which can be worked singly or in congress with others (Green, Ribarsky, and Fisher, 2008). Which heuristics are used first or most often depends largely on the task and the individual characteristics of the person undertaking the task. (See also (Gigerenzer and Goldstein, 2008 and Gigerenzer, 1991)) As evaluators, we can control what specific task is undertaken, but what about the variability within the person? Current evaluative methods in visualization are limited in that they ignore this variability. Evaluation focuses on the differences in and between visualizations but tends to treat the persons interacting with the visualization as somehow standardized. This may be an acceptable assumption with regard to basic sensory transduction (however even here there are significant individual differences in color and stereo perception), but it fails altogether as we move to higher-order cognitive processes. One would expect, for example, an expert artist to have chosen their profession based on (perhaps native) abilities in image composition, color perception, innovative use of graphics. One would expect that those abilities would be further developed in the course of their artistic training and practice.

Recognizing the institutional and innate differences, not only between novice and expert, but also between users with varying innate personality factors is key to moving past these evaluative limitations.

In the case of an expert user of visualization environments, we might expect a *priori* individual differences in ability that might be focused and trained through experience in their craft and in the use of a given set of technologies. Any model of the human-computer analytic system must find a way to assess, classify, and incorporate measures of those human characteristics to inform the development and evaluation of visual analytics systems. The comparative uniformity of basic perception (e.g. color space) limits the impact of individual differences to pathological cases (such as dichromacy). Spatial indexing, focused attention, and reasoning processes interact with the user's individual differences in ways that might well obscure or confound analysis of the impact of changes in technological support for those processes in HCI or other evaluative studies. Individual differences in how problems are approached can also affect beliefs and motivation when a user is engaged in goal-oriented behaviors (Heppner and Anderson, 1985). When we evaluate visualizations that people use, we must also understand the built-in learning “pre-sets” of the individual user; it is possible that a visualization that seems intuitive for one subgroup or expert domain could seem a wilderness for another.

Often visualization is assumed to be preferable to traditional types of interface for learning and/or the extension of knowledge. But how is it preferable, and for what types of learning? The science of learning is not generic. This study utilizes two tasks that touch on two broad genres of learning: procedural and inferential. Both genres have broad records in the human behavioral literatures, and represent two very different types of knowledge creation and use. Procedural learning, broadly defined, is the “knowing how” of any sequential task. It is sometimes called skill learning, as it is the learning most common to motor and iterative tasks that require repetition to master (Sun, Merrill, and Peterson (2001)); it is also referred to as script learning, which captures the idea that there is a “recipe” or “roadmap” to be followed. Riding a bike, brushing your teeth, or following a cooking recipe are all very simple examples of procedural learning. Procedural learning is thought to be either top-down (i.e. CLARION) (Sun, Merrill, and Peterson, 2001), or, more commonly, to be bottom up, first assimilating the necessary declarative facts and then the use of that information into the deconstruction of the task procedure (Anderson, 1982). Procedural learning, due in part to repetition, can become “automatic,” requiring little conscious focus. Inference learning, again broadly, is the

ability to draw a conclusion from available data or define a concept in terms of its similarity/dissimilarity to another. Categorization and classification are important building blocks of inference, and inference is used in a variety of reasoning, including induction, deduction, and comparison (Lim, Benbassat, and Todd, 1999). In this study, we study inference by providing an exemplar, and asking participants to find another example that shares/does not share a variety of characteristics. Humans use inferences when we decide whether a four-legged creature is a dog, when we decide whether we will like the new restaurant based on our experiences with others, and when we read body language to understand whether a person is telling a joke or being serious. Inferential learning, unlike procedural learning, does not lend itself to automaticity, and, when complex, involves sustained attention, problem-solving, a variety of reasoning heuristics, and decision-making.

This study was designed to explore 3 research questions. The first question was whether one interface would prove to be more efficient than the other in the performance of procedural tasks. Previous literature comparing motor movement between menu-driven and visualization interfaces (Lim, Benbasat, and Todd, 1996) leads us to hypothesize there will be no significant difference between mean procedural task completion times.

The second question was whether one interface would prove to be more efficient and more effective than the other in the performance of inferential tasks; it was hypothesized that the tasks undertaken in the visualization would have shorter completion times overall and would be more likely to be answered correctly than tasks undertaken in the web application.

The third question was whether the whole scores on one or more of the 6 standard psychometric measures, or some combination of the measures' items would have a significant relationship to the outcome variables in both the procedural and inferential tasks; it was hypothesized that, given the interrelationship between these constructs, one or more constructs would be found to predict completion times and/or error rates in both interfaces.

To isolate individual differences, we chose to use existing measures of human attitudes and abilities that have shown to have explanatory value in a range of applications in cognitive, personality and social psychology. It would be surprising, but not impossible, that one of these scales might be strongly predictive of performance in the quite different situation of problem-solving using a visualization system. It is more likely that some novel combination of these psychologically important measures may interact in an interesting way with performance measures on these tasks. We hope that this work will pave the way for the development of new scales that assess an individual's "personal equation" of interaction (Po, Fisher, and Booth, 2005) with visually-rich information systems. This measure can be used in selecting and training of users and customization of systems as well as a factor to take into account in the development and assessment process.

2.3. Comparative Study

The current study employed 50 participants, all of whom were undergraduate students enrolled in an introductory psychology course; all received course credit for their participation. The participants reported being students in 23 different majors, with the majority reporting a business-related (administration, finance, etc.) academic concentration. Most (47) had taken fewer than 4 laboratory biology courses. Most (44) were right-handed. Only one reported being color-blind (46) reported being comfortable or very comfortable using a computer; 44 reported their computer ability to be "OK" or "Very good." All 50 reported having used a web application previously; 16 claimed to having used a data visualization in the past.

2.3.1. Interfaces

This study asked participants to interact with two genomic interfaces. Both interfaces were fed by the same underlying dataset (GenBank). What differed between the interfaces was the presentation of data and interaction methodology. One interface is the web-based National Center for Biotechnology Information (NCBI) MapViewer for genomic information, which is publically available and can currently be found at <http://www.ncbi.nlm.nih.gov/mapview>. MapViewer utilizes a multiple-row-based

hierarchical representation and is manipulated through primary use of hyperlinks. (See Figure 3.) The use of traditional menus provides access to special genomic features, such as graphical representation of physical structure and mapped genomes.

The other interface is a genomic data visualization (GVis) developed in the Charlotte Visualization Center (Hong et al., 2005) and is not available commercially. (See Figure 4.) GVis was developed to support the visual analysis of large-scale phylogeny hierarchies by visualizing hierarchical relationships between organisms in addition to pictorially representing other essential information, such as the presence of mapped genomes or the phylogenic organization between two related subcategories. It allows quick browsing of the hierarchy from the highest level down to the level of an individual genome for the desired organism of interest via *direct interaction*, a method of data manipulation that minimizes the use of menus, allowing users to “drill down” directly by pressing and holding down a mouseclick near the information of interest.

NCBI Home GenBank BLAST

Map Viewer Home > Help

The Map Viewer provides a wide variety of genome mapping and sequencing data. [More...](#)

Search

Search: Select Group or Organism

for:

Go

Tools Legend

- Search or Browse the Genome
- BLAST
- Clone Finder
- Go to region on a chromosome
- Genome Resources page

News

Human build 37 released Aug 4

An update to the human genome assembly and annotation is now... [more](#)

[Show all](#)

Related Resources

- NCBI Home
- NCBI Web Search
- NCBI Site map
- Genome Browser agreement
- Genome Biology
- Taxonomy
- Entrez (Global Query)
- BLAST
- Map Viewer FTP

Small Genomes

- Bacteria
- Organelles
- Viruses

Vertebrates (17)

Mammals (14)

Primates (3)

Scientific name	Common name	Build	Tools
<i>Homo sapiens</i>	human	Build 37.1 Build 36.3	
<i>Macaca mulatta</i>	rhesus macaque	Build 1.1	
<i>Pan troglodytes</i>	chimpanzee	Build 2.1	

Rodents (2)

Scientific name	Common name	Build	Tools
<i>Mus musculus</i>	laboratory mouse	Build 37.1 Build 36.1	
<i>Rattus norvegicus</i>	rat	RGSC v3.4	

Monotremes (1)

Marsupials (1)

Other Mammals (7)

Other Vertebrates (3)

Invertebrates (12)

Protozoa (18)

Plants (46)

Fungi (17)

Scientific name	Common name	Build	Tools
<i>Aspergillus clavatus</i>		Build 1.1	
<i>Aspergillus fumigatus</i>		Build 2.1	
<i>Aspergillus niger</i>		Build 1.1	
<i>Candida glabrata</i>		Build 1.1	
<i>Cryptococcus neoformans</i>		Build 2.1	
<i>Debaryomyces hansenii</i>		Build 1.1	
<i>Encephalitozoon cuniculi</i>		Build 1.1	
<i>Eremothecium gossypii</i>		Build 3.1	
<i>Gibberella zeae</i>		Build 1.2	
<i>Kluyveromyces lactis</i>		Build 1.1	
<i>Magnaporthe grisea</i>	rice blast fungus	Build 3.1	
<i>Neurospora crassa</i>		Build 2.1	
<i>Pichia stipitis</i>		Build 1.1	
<i>Saccharomyces cerevisiae</i>	baker's yeast	Build 2.1	
<i>Schizosaccharomyces pombe</i>	fission yeast	Build 1.1	
<i>Ustilago maydis</i>		Build 1.1	
<i>Yarrowia lipolytica</i>		Build 1.1	

Figure 3. The NCBI Interface

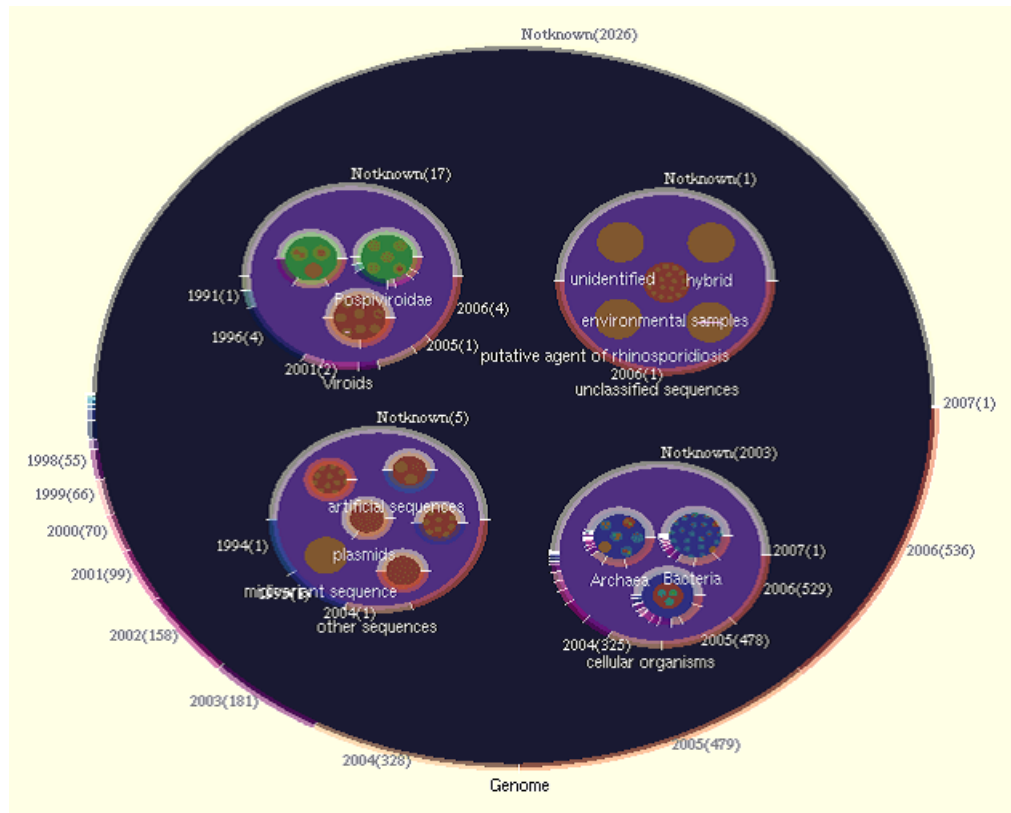


Figure 4. The main view of GVIS

2.3.2. Psychometric Measures

Six psychometric measures were administered: the Locus of Control Inventory, the Beck Inventory, the IPIP 20-item Big Five Neuroticism Scale, the IPIP 20-item Big Five Extroversion Scale, the Self-Regulation Scale, and the Scale of Intolerance-Tolerance of Ambiguity.

The Internal-External Locus of Control Inventory (LOC) (Rotter, 1966) is a 39-item forced choice measure designed to evaluate the degree to which participants attribute life events to some action of their own, or to some uncontrollable action outside of themselves. Lower LOC scores are associated with an “internal locus” of control, an inherent belief that events and outcomes are under a person’s control, and thus, success or failure depends largely on personal behavior and attitudes. Higher scores

indicate an “external locus,” an inherent belief that events and outcomes are *not* under a person’s control, but are largely influenced by other people, unforeseen circumstances, a higher power, or other factors such as “good luck.” Rotter postulated that these loci were individual traits and remained stable over a person’s lifetime (Rotter, 1966).

The Beck Anxiety Inventory (BAI) (Beck et al., 1988) is a 21-item Likert-like scale which asks the participant to evaluate how often common anxiety symptoms were experienced over the previous month, from 0(not at all) to 3(severely-bothered me a lot). The BAI was designed to diagnosis “trait” anxiety, a tendency to be prone to anxiety generally, even absent a generating trigger.

The IPIP 20-item scales for the Big Five Neuroticism and Extraversion (Donnellan, 2006) are both 5-point Likert scales that ask participants the degree to which each listed characteristic applies to themselves. Both scales can be found on the IPIP website: <http://ipip.ori.org/>. Briefly, Extraversion defines the degree to which a person is action-oriented and seeks the society of others. Neuroticism can be viewed as the opposite of Extraversion; it is distinguished by negativity and a propensity to be emotionally sensitive.

The Self-Regulation Scale (SRS) (Schwarzer, Diehl, and Schmitz, 1999) is a 10-item, Likert-like measure which evaluates “postintentional” regulation of focused attention and emotional maintenance throughout the completion of a goal-oriented task, or, in other words, the ability to maintain sustained focus despite distractions, uncertainty, and/or emotional events.

The Scale of Tolerance-Intolerance of Ambiguity (TOA) is a 16-item Likert measure designed to appraise the degree to which the participant self-evaluates novel, complicated, or apparently unsolvable situations as threatening (Budner, 1962). Tolerance of ambiguity, as measured by the TOA, is not, like the SRS, a measure of coping ability *per se*, but an appraisal of self-beliefs, similar to the LOC.

2.3.3. Protocol

After signing up to participate in the study and giving informed consent, participants were invited to complete an online survey that included the 6 psychometric measures as well as demographics questions about self-perceived ability, personal experience, and comfort with computers and computer interfaces. This survey could be completed before arriving in the laboratory, or it could be completed the day of the study. Most participants elected to complete it before the study session. All data were collected for post-hoc analysis with task performance data.

After completion of the self-report measures, participants began the first of the four series of learning tasks. All four series were administered through a web application written in the laboratory for this study, which led the participant through each task. The order of interface was counterbalanced for order effects; half of the participants used GVis first, half used NCBI first. The GVis tasks started with a brief demonstration of GVis 3 basic modes of mouse manipulation: zooming in, zooming out, and panning (moving the visualization within the view for better visibility).

After the demonstration, a short tutorial was administered to introduce participants to essential tools and concepts in the interface, and allow participants to experiment with what was being learned. In some cases, step-by-step instructions were given. A researcher was on hand throughout the study to answer any questions. Following the tutorial was a series of 3 tasks designed to test procedural performance: the participant was asked to identify a piece of information located somewhere within the presented informational hierarchy. The question provided what base categorization or subclass the information was located within, but did not provide step-by-step instructions. Participants were also told to find the item as quickly as possible, as the task was being timed. As soon as the information was located on screen, the participant pushed a “Found It” button on the screen. The time taken from the presentation of the question on-screen to the moment the button was pushed was recorded as *completion time*.

Following the procedural tasks, the participant was administered a series of tasks designed to test inferential performance. A series of 3 questions were asked; each

question asked the participant to evaluate characteristics of one organism and find others that were like or unlike the example; the first question was open response, the other 2 were forced choice. Participants were told not to worry about the time taken to complete the task; this instruction was given to mitigate unmeasured performance anxiety. If the participant answered the question incorrectly, that was recorded post-study as an error; participants were not informed if their answers were incorrect. *Errors* and *completion time* were recorded as outcome variables.

The NCBI tasks series followed the same protocol as the GVis tasks: introduction and demonstration, tutorial, procedural tasks, and inferential task. Special care was taken to create tasks that were very similar in construction and difficulty to the tasks in the GVis. After each participant had completed the tasks in both interfaces, a post-study questionnaire was administered. Participants were asked to specify which interface they liked better and in which interface they were more comfortable working. They were also asked to freely respond to what they liked and disliked about each interface. Finally, they were asked to give each interface a letter grade (“A” (superior) through “F” (failing)). The completion of the post study questionnaire was followed by a short debriefing, which included an opportunity for questions. This ended the study session, and there was no follow-up.

2.3.4. Results

This section is divided into 3 subsections: descriptive results, including participant feedback about each interface, inferential statistical analyses of performance, and generation of new predictors of performance derived from psychometric measures.

Descriptive Results

Mean procedural task completion times (in seconds) in the MapViewer ($M = 136$, $SD = 84$) were faster than in GVis ($M = 162$, $SD = 111$). A paired-sampled t-test between the total completion times in the procedural tasks across the interfaces achieved only borderline significance ($p = .057$). This nonsignificant trend is not entirely congruent with our expectation that we would not find a significant difference between the procedural

completion times across interfaces. The MapViewer times were shorter, but the trend was not strong enough to rule out random chance as a factor.

However, for the inferential task, GVis times ($M = 451$, $SD = 169$) were faster than those in the MapViewer ($M = 922$, $SD = 521$). During the completion of the inferential tasks, participants answered more questions correctly while using the GVis ($M = 1.31$, $SD = .56$) interface than while using the Mapviewer ($M = .77$, $SD = .59$). A paired samples t-test between total inferential completion times for each interface was significant ($t(49) = -7.59$, $p < .01$), as were the total completion times (all 6 tasks) in both interfaces ($t(49) = 6.99$, $p < .01$).

Of the inferential questions answered in the visualization, 40 (78%) answered the first correctly compared with 2 (4%) in the web application. 4 (12%) correctly answered the second visualization question, compared with 8(16%) in the web application. And 5(10%) answered the third GVis question correctly compared with 2(4%) in the web application. Overall, these data support the second hypothesis that tasks completed in the visualization would be done more efficiently and effectively than those completed in the web application.

Correlations between total completion times were all significant: total procedural completion scores in both interfaces ($r(50) = .49$, $p < .01$), total inferential completion times in both interfaces ($r(50) = .61$, $p < .01$), and total completion times (all 6 tasks) in both interfaces ($r(50) = .63$, $p < .01$). These data demonstrate that participants who tended to take longer completing tasks in one interface also tended to take longer completing tasks in the other.

Overall, participants preferred interacting with the visualization to interacting with the web application. This preference was indicated in a variety of ways. For example, when asked to give each interface a letter grade, from A (superior) to F (failing), 36 (71%) gave the GVis an A or B; 18 (25%) gave an A or B to the MapViewer. Additionally, when asked, 33(65%) reported that they both preferred and were more comfortable in the visualization; 15 (29%) preferred and were more comfortable in the web application.

Participants were also asked to freely respond to prompts about likes and dislikes in each interface. Common likes in GVis included the ease of use and use of color and graphical groupings; dislikes included not always knowing where to look for information. Common likes in MapViewer included its alphabetical and hierarchical organization; dislikes included difficult searches and the presentation of too much data.

Inferential task performance

Of the administered psychometric scales, only Rotter's Locus of Control (LOC) whole score predicted inferential learning performance. LOC scores negatively correlated with total inferential task completion scores in the visualization ($r(50) = -.34, p = .02$) but not the web application ($p = .104$). A repeated-measures Analysis of Variance (ANOVA) demonstrated a significant between participant main effect of hi-lo LOC groups for GVis Inferential completion times ($F(13, 36) = 2.06, p = .04, \eta^2 = .43$). No such effect existed for MapViewer completion times.

Procedural task performance

None of the 6 whole psychometric measure scores showed significant trending with procedural completion times in either interface. Therefore, each psychometric scale was analyzed for its principal components, and an analysis of the psychometric items was done to evaluate which constructs might predict performance separately. Intercorrelations between individual items as well as analysis with a metric alternating least squares scaling (ALSCAL) multidimensional scaling (MDS) analysis of each of the 6 measures were used to identify clusters of similar items within each measure. This analysis was followed up with factor analysis with principal component analysis (PCA), which identified the structural "components" of each measure: items whose scores trended together, as well as accounted for substantial proportions of variance in the total score. The results of this analysis were then narrowed by choosing the most influential measure items, or the items that explained the most variance, from components that contained at least one item that significantly correlated with the mean procedural task completion times in either interface. For example, the item "Start conversations" in the IPIP Extraversion scale significantly correlated with the mean completion times from the 3rd procedural task in GVis. This item, "Start conversations" had been identified as part of the 2nd component of the Extraversion measure. That component's top item, or item

that was most influential, had been identified as “Feel comfortable around people,” which was then considered as a potential item for the new measure. In addition to the most influential item in each component, items that correlated with multiple mean task completion scores or had strong intercorrelations, were also considered, whether or not they were the most influential item in a component.

Each item was then analyzed for its contribution to the new measure, and items that did not add to either its measure-task correlation or to its internal consistency of the new measure were eliminated from inclusion. In this way, 126 items from the original 6 administered measures were narrowed to 9 items. (See Table 1.) These 9 items were evaluated together as a separate short measure. A metric ALSCAL MDS with a 2-dimensional solution was conducted to define the underlying structure. The Young’s S-stress value (max of 1, lower numbers indicate less stress) of the solution was .02, and the squared correlation (RSQ) was .99 (higher numbers indicate better intercorrelation), denoting that the data are a good statistical fit to the scale.

Table 1. Items in the short measure.

Item (Originating Measure)	How scored	PCA Component
Unable to Relax (Beck's Anxiety Inventory)	0(never) to 3(severely)	1
Fear the Worst (Beck's Anxiety Inventory)	0(never) to 3(severely)	
Heart Pounding (Beck's Anxiety Inventory)	0(never) to 3(severely)	
What we are used to is always preferable to what is unfamiliar. (Tolerance of Ambiguity)	1(strongly disagree to 7(strongly agree)	2
Talk to a lot of people. (Extraversion)	1(low) to 5(high)	
Hands Trembling (Beck's Anxiety Scale)	0(never) to 3(severely)	3
Numbness (Beck's Anxiety Inventory)	0(never) to 3(severely)	
Don't talk a lot. (Extraversion)	1(low) to 5(high)	
Am easily disturbed. (Neuroticism)	1(low) to 5(high)	

Table 2. PCA components of the short measure

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.734	30.37	30.37
2	1.79	19.99	50.36
3	1.04	11.58	61.94

MDS analysis identified two main clusters: Hands Trembling/Numbness and Unable to Relax/Fearing the Worst. All four of these items in these two clusters originated from the Beck Anxiety Inventory. (Please see Figure 5.)

A factor analysis of the 9-item short measure was conducted using principal component analysis (PCA). Multiple criteria for the correlational factorability were utilized; 8 of the 9 items correlated at a minimum .30 with at least one other item; this suggests a reasonable level of factorability. Also, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .68, which is above the accepted standard of .60. Bartlett's Test of Sphericity was significant ($\chi^2(36) = 97.22, p < .01$). All anti-image correlational diagonals, with the exception of Hands Trembling, were above .50. Internal consistency was measured using Cronbach's alpha; consistency was moderate, $\alpha = .63$.

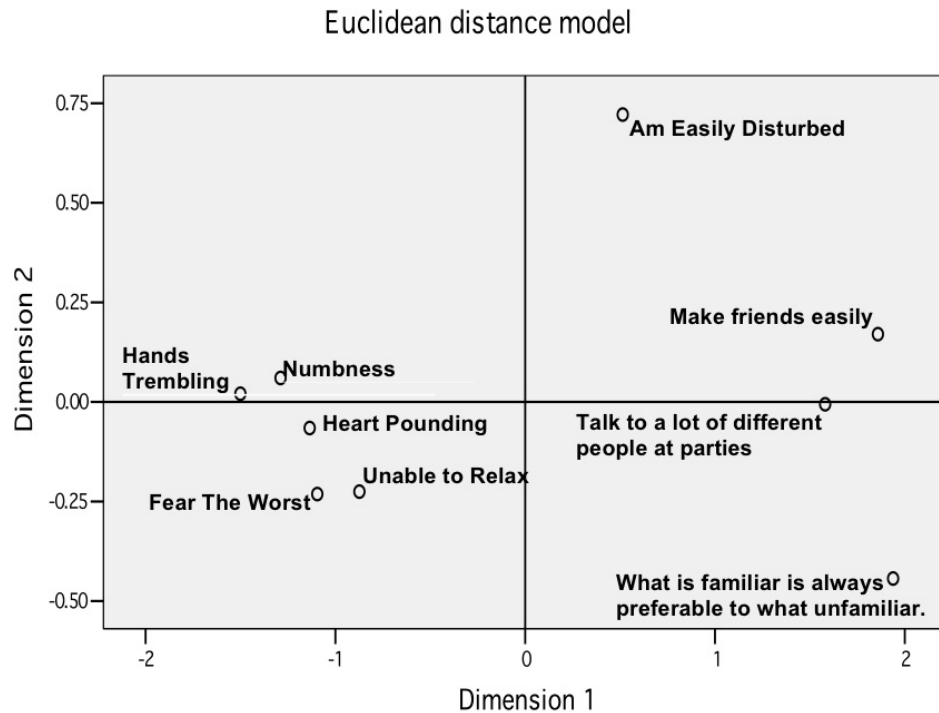


Figure 5. MDS clusters in the 9-item measure

Factor analysis extracted 3 principal components with initial Eigen values > 1.0; these components together accounted for 62% of score variance. (Please see Table 2 for more.) The 9-item short measure moderately correlated with total procedural completion times in each interface (GVis: $r(50) = -.46, p < .01$, MapViewer: $r(50) = -.453, p < .01$); participants with higher scores on the measure tended to take less time completing the procedural tasks in both interfaces. A 2 x 3 (interface x trial) repeated-measures ANOVA was conducted to determine whether the 9-item short measure demonstrated a within-participants main effect of interface ($F(1, 33) = 7.51, p = .01, \eta^2 = .19$). There was also a within-participants main effect of trial ($F(2, 66) = 50.71, p < .01, \eta^2 = .61$.) The interaction of trial x 9-item short measure was also significant ($F(32, 66) = 6.76, p < .01, \eta^2 = .76$). Lastly, a between-participants main effect of 9-item short measure was found ($F(16, 33) = 4.13, p < .01, \eta^2 = .67$).

2.4. Discussion

This study demonstrates that not all interactive learning tasks are created equal. Identify-and-find behaviors like the ones utilized during the procedural tasks use a different combination of tools in the cognitive toolbox than do more complex, iterative reasoning behaviors engaged during the inferential tasks. This was demonstrated by completion times; the simple procedural tasks were not done more efficiently in one interface over the other. But when the tasks became more difficult, requiring the user to categorize, compare, and evaluate multiple choices at once, participants worked more quickly and made fewer errors while using the visualization.

Then too, is the difference in what genre of psychometrics predicted the task behavior. In the procedural tasks, most of the trending items were anxiety-based, or described some fear of the unknown. For the inferential tasks, the biggest indicator of performance was locus of control, a user's self-belief about personal control over circumstances and environment.

Although not considered by interface evaluators, how much control an individual feels over his or her life circumstances has long been regarded as a demonstrative predictor in the human behavioral literature; an internal locus has been associated with such outcomes as better use of problem-solving skills (Krause, 1986), a greater resolve during task difficulty (Krause, 1986), and development of intrinsic motivation (Weiss and Sherman, 1973). This study's finding that an external locus is a predictor of efficient inferential task completion is not explained by much of the extant LOC literature originating from the psychological, learning, health, and HCI disciplines.

It is possible, however, that participants with an external locus of control were more ready to accept the constraints of an unfamiliar environment, and so were more able to quickly work the tasks; similar results were reported in a study in which participants with an external locus who knew that they could not escape a loud, uncomfortable noise adapted more quickly to the environment than those with an internal locus, who tried to escape the environment more quickly or altogether (Hiroto, 1974). Until we can replicate and further explore this trending, we will accept the plausibility that an external locus improves the ability to work within a novel environment

or with novel, complex information by allowing the user to adapt to the environment despite the discomfort of the unknown.

That the external-locus = shorter-completion-time trend is exhibited only in the visualization interaction is easier to explain. GVis was developed in response to a request for a better way to locate and analyze the spatial and semantic relationships between ontological biological structures (Hong et al., 2005); inferential tasks that depend on compare-and-contrast behaviors should, theoretically, be easier to see and solve in GVis. Additionally, the performance outcomes, non-significant trending between MapViewer outcomes and the sporadic psychometric scores, as well as the varying nature of the participant feedback suggest that combination of variables influenced MapViewer complex performance behaviors, perhaps due to the difference in required interaction. Often, tasks that required one or two mouseclicks in a single view in GVis were much more complicated in the MapViewer, requiring multiple mouseclicks and changes of view. For example, unlike the straightforward presentation of mapped genes in GVis through direct interaction (holding down a single mouseclick on the organism of interest), determining the existence of a mapped gene for an organism in MapViewer required the user to hunt for the organism name in the list of organisms, possibly reorganizing the list through primary and secondary sorts, locating and clicking on the small single letter "G" on the far right of the application view, which served as a hyperlink to a separate page. If a gene existed, information about its mapping was presented. If the gene did not exist, the hyperlink led to a page presenting a frustrated-looking male icon and the explanation, "No information found for given taxid." Locus of control played a role in the MapViewer inferential task outcomes, but not one strong enough to show any predictive strength.

In the procedural tasks, the 9-item short measure is moderately negatively correlated with completion times. This suggests that more trait-anxious (i.e. persons that tend to be anxious all the time as compared to anxious only when presented when threatening stimuli), uncommunicative, and/or prone to emotional instability a person is, the less time they tend to take finding requested items while interacting with novel information. This might seem counterintuitive at first glance. However, according to Spence-Taylor Drive Theory (Spence, Farber, and Schmitz, 1999), persons with higher

trait anxiety tend to identify target information more quickly than the non-anxious when the task does not require either iterative or complex reasoning processes. Other studies have found that persons with higher trait anxiety are more attentive to presented information and can identify target threats more quickly than those less anxious (Ionnou, Mogg, and Bradley, 2004). While the causes for this “exception” are still subjects of debate, it has been proposed that trait anxious persons have developed adaptive heuristics than can make advantageous use of their anxiety (Spence et al., 1966). The results of the current study would suggest that certain aspects of trait anxiety tend to make users more attentive and better able to identify target information until the task becomes complex, requiring more complicated reasoning heuristics and lessening the effectiveness of the adaption. Additionally, the 9-item short measure scores positively correlated with LOC scores ($r = .37, p < .01$), suggesting that persons who were more anxious/uncommunicative were also more likely to attribute consequences of life events to forces outside their control, such as luck or divine intervention; scores on the whole Beck’s Anxiety Inventory, however did not correlate. Given that LOC scores were a predictor of more efficient completion times in the more complex, inferential task, it seems reasonable that a relationship exists between the aspects of anxiety captured by the 9-item short measure and locus of control.

The items in the 9-item short measure were culled from 6 measures designed to measure anxiety, personality traits, self-efficacy, and self-beliefs about control over personal circumstances. Subsequent analysis of the 9-item short measure found it to have moderate internal consistency and to meet the requirements for a reliable psychological assessment. However, it is unreasonable to expect that any new measure would be fully validated after one evaluative trial. While we are fairly confident the 9-item short measure has captured trending in this study, we recognize that further trials are required before the 9-item short measure could be considered predictive or reliable in a generalizable way. It is our desire that these results would be subject to replication by ourselves and others; we are currently designing protocols to replicate and extend these findings.

A final note is the use of testing an expert system with non-experts; most of the participants had only a rudimentary knowledge of biology. Even, so, participants were

able to reason through the inferential tasks; overall, participants preferred the visualization to the web application. Also, given the participant pool's overwhelming familiarity with the web application as an interface prototype, it is also telling that participants also found the visualization more comfortable to work within. Our intent was not to test the efficacy of GVis as an expert system; such an evaluation has been reported in other literature (Hong et al., 2005). The aim of the current study was to evaluate which interface proved better at facilitating procedural and/or inferential learning, and to explore whether individual differences in personality factors and self-beliefs could have a large enough influence on outcomes to recommend their consideration in interface design.

For this reason, we sought non-experts who were relatively unfamiliar with the subject matter. Any expertise would have biased the use of the interface; the user would have known exactly where to look for certain information and thus would have been a poor test of how well the each interface design promoted learning.

Overall, the results of the current study suggest that the data visualization is a superior interface for complex, spatial, inferential learning. However, it is still not clear that the same is true for hunt-and-find, simpler tasks.

2.5. Conclusion

The current study has demonstrated that believing in the power of luck (as an external locus does) makes inferencing easier during visualization interaction. It has also shown that certain trait anxiety markers improve seek-and-find iterative behaviors. More generally, this study has established that individual differences between users a) do impact the efficacy of visualization and web application interfaces and so b) should be considered as a part of a maturing theory of visualization and complex interface design.

Domain-specific interface users often share certain common problem-solving tendencies, whether institutional or innate. For example, the tendency of intelligence analysts to adopt specific biases has been a point of study (Heuer, 1999). By studying the group-specific inherent traits or behaviors of an expert cohorts, we may be better

able to create visualizations that are discernibly more intuitively interactive in the environmental set for which they were designed. Because users' engagement with the system is at least somewhat unique due to innate personality factors and to how in-control they feel in their environment, a system that supports visual analysis or other complex cognitive behaviors must be designed to handle this individuality. Yet designing around human "pre-sets" need not be onerous. As this study demonstrates, a better understanding doesn't mean real-time new user screening through use of hundreds of items in multiple psychometric measures; much of this information could be gathered during the requirements phase of interface design through quasi-experimental studies similar to this one. Once highly predictive items are isolated, much in the same way as the 9-item measure, fine-tuning an interface for each user's "personal equation" could theoretically become quicker and more painless.

The best way to design for a user's cognitive individuality is still an emerging field of research. The conclusions of the current study will undoubtedly be edited and expanded as we and others use its findings to design more rigorous future research. But what cannot be disputed is that effective interface design for complex interaction depends on an understanding of the human cognition involved in such interaction. This study was a step toward that understanding.

2.6. References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, 89(4), pp. 369 – 406.
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of consulting and clinical psychology*, 56(6), 893.
- Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of personality*.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological review*, 98(2), 254.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.
- Green, T.M., Ribarsky, W. and Fisher, B., Building and applying a human cognition model for visual analytics, *Information Visualization* 8(1), 2009, pp. 1-13.

- Heppner, P. P., & Anderson, W. P. (1985). The relationship between problem-solving self-appraisal and psychological adjustment. *Cognitive therapy and research*, 9(4), 415-427.
- Heuer, R. J. (1999). *Psychology of intelligence analysis*. Lulu. com.
- Hiroto, D. S. (1974). Locus of control and learned helplessness. *Journal of experimental psychology*, 102(2), 187.
- Hong, J., Jeong, D. H., Shaw, C. D., Ribarsky, W., Borodovsky, M., & Song, C. G. (2005, June). GVis: A Scalable Visualization Framework for Genomic Data. In *EuroVis* (Vol. 5, pp. 191-198).
- Ioannou, M. C., Mogg, K., & Bradley, B. P. (2004). Vigilance for threat: effects of anxiety and defensiveness. *Personality and Individual Differences*, 36(8), 1879-1891.
- Krause, N. (1986). Stress and coping: Reconceptualizing the role of locus of control beliefs. *Journal of Gerontology*, 41(5), 617-622.
- Lim, K. H., Benbasat, I., & Todd, P. A. (1996). An experimental investigation of the interactive effects of interface style, instructions, and task familiarity on user performance. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 3(1), 1-37.
- North, C. (2006). Toward measuring visualization insight. *Computer Graphics and Applications*, IEEE, 26(3), 6-9.
- Plaisant, C. (2004, May). The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 109-116). ACM.
- Plaisant, C., Fekete, J. D., & Grinstein, G. (2008). Promoting insight-based evaluation of visualizations: From contest to benchmark repository. *Visualization and Computer Graphics, IEEE Transactions on*, 14(1), 120-134.
- Po, B. A., Fisher, B. D., & Booth, K. S. (2005). A two visual systems approach to understanding voice and gestural interaction. *Virtual Reality*, 8(4), 231-241.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, 80(1), 1.
- Schwarzer, R., Diehl, M., and Schmitz, G. S. (1999). The Self-Regulation Scale, Berlin: Freie Universität, *Gesundheitspsychologie*.
- Spence, K. W., Farber, I. E., & McFann, H. H. (1956). The relation of anxiety (drive) level to performance in competitive and non-competitive paired-associates learning. *Journal of Experimental Psychology*, 52(5), 296.
- Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive science*, 25(2), 203-244.
- Weiss, H., & Sherman, J. (1973). Internal-external control as a predictor of task effort and satisfaction subsequent to failure. *Journal of Applied Psychology*, 57(2), 132.

Chapter 3. Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction.

3.1. Overview

This chapter was originally published as *Green, T. M., & Fisher, B. (2010, October). Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. In Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on (pp. 203-210). IEEE.* This study builds on chapter two by evaluating the relationships between the psychometric items and additional outcomes. The whole measures (i.e. all items in the measure were used) were Locus of Control, Big Five Extraversion and Big Five Neuroticism. The 3 outcome measures were efficiency (completion times), task errors, and self-reported insights. The null hypothesis was that there would be no significant association between these whole psychometric items and outcomes. The alternate hypothesis was that the whole measures would predict all 3 outcomes. The alternate hypothesis was largely supported, as all 3 whole measures predicted interface learning efficiency and self-reported insights but not task errors.

3.2. Abstract

These current studies explored the impact of individual differences in personality factors on interface interaction and learning performance behaviors in both an interactive visualization and a menu-driven web table in two studies. Participants were administered 3 psychometric measures designed to assess Locus of Control, Extraversion, and Neuroticism. Participants were then asked to complete multiple procedural learning tasks in each interface. Results demonstrated that all three measures predicted

completion times. Additionally, results analyses demonstrated personality factors also predicted the number of insights participants reported while completing the tasks in each interface. We discuss how these findings advance our ongoing research in the Personal Equation of Interaction.

3.3. Introduction

The primary purpose of visual analytics is commonly defined as the facilitation of analytical reasoning through use of interactive visual interfaces (Wong & Thomas, 2004). Facilitating analytical reasoning, however, requires a comprehensive and operational understanding of the cognitive processes that make up analytical reasoning. Complex cognition includes a plethora of smaller processes that work together, including perceptual cognition, categorization, problem-solving, decision-making, judgment, and reasoning. These processes feed and inform each other throughout each stage of the analytical task; simply supporting each process individually is not enough. Visual analytics must also support the temporal and cognitive flow of reasoning. And yet, an operational understanding of analytical cognition has, to date, proven elusive. For example, as is often the case with behavioral experimentation generally, studies of cognition tend to involve small, simple, normative or “toy world” tasks, while interaction in the real world tends to be more complex, harder to predict, and thus harder to measure. Additionally, these evaluations focus on the more binary of cognitive processes. Especially in visualization studies, the cognitive variables measured are usually facets of vision, given attention, and tactile manipulation. While visual and motor effectiveness are important to interface interaction, they are only part of the story.

Complex cognition is not binary nor necessarily sequential. Reasoning, in particular, uses a variety of heuristics, from quick elimination heuristics like Gigerenzer’s Take-the-Best (Gigerenzer and Goldstein, 1996) or satisficing (Simon, 2006) to much more complicated processes such as iterative reasoning, deductive analyses, or rule inferencing. Which heuristics are used and in what order depend on the task, the environment, and the user. These heuristics are often used combinatorially, feeding and informing the analysis until a solution or hypothesis has been satisfactorily reached.

Unfortunately, at this time, analytical reasoning behaviors can be described in part and in whole, but not necessarily predicted. There are no unifying theories of reasoning. And this difficulty of prediction is compounded by three types of user individual differences: institutional, environmental, and inherent. How humans work through reasoning tasks is impacted by institutional differences. Cognition is a social activity (Kaptelinin and Nardi, 2006), and domain-specific knowledge, jargon, learned methodologies, and other cultural factors can influence how analysis tasks are approached and what heuristics are used in solving them. In addition, these domain or expert cultures tend to have similar inherent differences; members of an expert cohort may share personality or learned proclivities (Boyatzis and Kolb, 1995, Heuer, 1999). Environmental differences – such as differences in the interface or tool used during visually enabled interaction – frame the task and can help or hinder the reasoning process. These differences are naturally of particular interest to visual analytics design, as effective interfaces can facilitate analytical reasoning. In this paper, we will highlight the impact of inherent individual differences. Individual differences of whatever variety are obviously not the only factors which demonstrably impact user interactive performance. But as we will show, individual differences – and inherent differences in particular – can predict certain types of performance. Further, these differences seem to influence performance differently, depending on the cognitive task being undertaken. Another reason to study inherent differences is that they, unlike environmental and to some degree institutional differences, are variables over which interface designers have no control.

In our research toward the Personal Equation of Interaction, our goal is to know and understand the impact of these variables, as well as to develop a battery of predictive measures to aid in the development of interfaces which cater to the individuality of the user or user domain. The creation of the Personal Equation of Interaction at this current time is focused on inherent individual differences. Inherent differences are those of learning style, personality factors, self-beliefs, and other cognitive “pre-sets” which the user brings to the interface. We will demonstrate that these inherent differences can and do demonstrably impact interaction outcomes. Further, we can show that, if the inherent differences are known, interaction performance can be predicted, and so could, if part of a robust user profile, be used to develop design

requirements for expert systems design as well as real-time interface individuation. Inherent individual differences in problem-solving approaches can affect task orientation and motivation when a user is engaged in goal-oriented behaviors (Heppner and Anderson, 1984). In particular, personality factors similar to the ones evaluated in the studies reported here have been shown to impact cognition and cognitive performance in other learning environments. For example, personality factors predicted preferences in visual perception of landscapes (Macia, 1979). In an HCI study, Palmer found that interactive behaviors in information search can be categorized by personality factors (1991). In reasoning research, individual differences have been found to impact rationality and metareasoning (Palmer, 1991). These are just a few examples in a broad literature of how personality factors and other individual differences demonstrably affect complex cognition.

The findings we currently report are part of this body of work. The question is not whether individual differences impact cognition, but how and when. We hope, in the creation of the Personal Equation, to answer several of these questions. Furthermore, we can use individual differences to improve our understanding of visually enabled analysis across knowledge domains. Research has demonstrated that users in a particular domain can share personality characteristics and learning preferences, both inherent and institutional. This implies that traits common to the user group can be aggregated into specific user profiles, informing superior design requirements and aiding in evaluation protocols. A personal equation of interaction could both a) provide guidelines for individuated interface designs which could broadly accommodate differences in learning style, reasoning heuristic preferences, and perceptual behaviors and b) develop profiles of expert or non-expert user groups, delineated by either knowledge domain or cognitive task, which would inform the interface design for specific user or task domains.

As we discussed previously, individual differences have been found to have a bearing in traditional learning environments [e.g. 11]. And in an earlier study (Green, Jeong, and Fisher, 2010) we found that certain aspects of trait anxiety had an impact on task efficiency in both inferential and procedural tasks. Also, Rotter's Locus of Control (Rotter, 1966) predicted inferential task efficiency; we will review this finding in Section

3. For user group profiles, characteristics of user domains has been done in a limited fashion (e.g. Hong et al., 2005); this research would further these aims.

Learning is not generic. Learning heuristics and processes vary depending on human individuality, the learning environment, and the learning tasks. In other work, we discussed the impact of locus of control on inference learning in the form of category reasoning (Green, Jeong, and Fisher, 2010). The tasks used in these current studies are procedural. Procedural learning, broadly defined, is the “knowing how” of any sequential task. It is sometimes called skill learning, as it is the learning most common to motor and iterative tasks that require repetition to master (Sun, Merrill and Peterson, 2001); it is also referred to as script learning, which captures the idea that there is a “recipe” or “roadmap” to be followed. Procedural learning is thought to be either top-down (i.e. CLARION) (Sun, Merrill and Peterson, 2001), or, more commonly, to be bottom up, first assimilating the necessary declarative facts and then the use of that information into the deconstruction of the task procedure (Anderson, 1982). Procedural learning, due in part to repetition, can become “automatic,” requiring little conscious focus. *For the purposes of these current studies, procedural learning is the ability to learn to manipulate an interface well enough to find and identify target information, or to answer straightforward questions about the target information.* Procedural or script learning is integral to interface interaction at every level. Some research has been conducted with an eye toward procedural or target-finding tasks. But, as Plaisant has outlined (2004), many of these studies are tool evaluations of specific interfaces, and are designed to designate one interface as “better” than another, or done without an understanding of the learning which underlies task performance.

Individual differences in reasoning ability have been found to impact procedural learning in non-interface task environments (e.g. Hall et al., 1988). These current studies evaluate inherent differences in computer-mediated procedural tasks. In another vein, visualizations are generally considered preferable to other interfaces in generating “insight” (Chang et al., 2009). But this claim to date has been poorly supported by empirical research. Further, research has focused on the visualization and insight generation, but not necessarily on the tasks that support insight generation, or the degree to which user individuality impacts the frequency of insight. *In this study, we*

evaluate the insight generation by comparing the number of reported insights in the two interfaces while completing two types of procedural task: script learning, which involves the use of sequential instructions and interface learnability, and target identification, which can involve hunting for information through several layers of hierarchical organization. In addition, we explore the impact that individual differences have on the number of insights generated in both interfaces across task.

The current studies were designed to explore 2 broad research questions.

The first question was whether and to what degree Locus of Control, Big Five Neuroticism, and Big Five Extraversion would have a significant relationship with the outcome variables in task performance. It was hypothesized that some whole measures or highly-predictive clusters of items would trend with the outcomes. Based on previous work (Green, Jeong, and Fisher, 2010), we expected that the Locus of Control whole score would be one predictor, and that more extraverted and neurotic participants would be quicker in task completion. And based on behavioral literature[e.g. Messer, 1972], we hypothesize that participants with an external locus would be quicker in identifying target information.

The second question was whether and to what degree Locus of Control, Big Five Neuroticism, and Big Five Extraversion would have a significant relationship with the number of insights reported; it was hypothesized that, given the interrelationship between these constructs, whole score or individual items, would be found to predict insight generation in both interfaces. Based on previous locus of control literature [e.g. Weiss and Sherman,1973, Messer,1972], we predicted that participants with an internal locus might be more apt to self-report more insights.

The answers to these questions will aid in the creation of the Personal Equation on Interaction, by identifying influential psychometric items for interactive behaviors and reported insights, which, in the long term will aid in the creation of predictive measures depending on the type of analytical task being undertaken.

3.4. Comparative Studies

Two studies were conducted. Each study employed a within-participants design, and compared procedural learning behaviors in an information visualization and a web table. Study 1 tested procedural learning performance with a series of 5 questions in each interface. Study 2 tested procedural learning performance, with a total 6 questions in each interface (3 training and 3 task).

The procedural task completion times in both studies were combined for the purpose of analysis. The design and findings of Study 2 have also been reported and discussed in Green, Jeong, and Fisher (2010).

3.4.1. Interfaces

Both studies asked participants to interact with two interfaces built to display genomic information. These interfaces were chosen as artifacts because both interfaces were fed by the same underlying dataset (GenBank), both interfaces supported the types of tasks we wanted to study, and the presentation and organization of data and interaction methodology was demonstrable different. One interface is the web-based National Center for Biotechnology Information (NCBI) MapViewer for genomic information, which is publically available and can currently be found at <http://www.ncbi.nlm.nih.gov/mapview>. MapViewer is a multiple-row-based hierarchical representation, and uses standard GUI manipulation, such as menus and hyperlinks. (See Figure 6.)

The other interface is an interactive data visualization (GVis) of genomic relationships (Hong et al., 2005) which is not available publically. (See Figure 7.) GVis primary purpose is to represent relevant relationships (such as mapped genomes or the phylogenic organization) between two organisms. Users manipulate the interface through direct interaction, “drilling down” through each hierarchy of subcategory directly by pressing and holding down a mouseclick near the information of interest.

3.4.2. Psychometric measures

These psychometric measures we have chosen have been shown to capture the impact of these inherent constructs on human cognitive performance and motivation as discussed in the behavioral literatures (as discussed briefly in Section 1). Our purpose was to explore what impact they might have on analytical performance enabled by a visual interface.

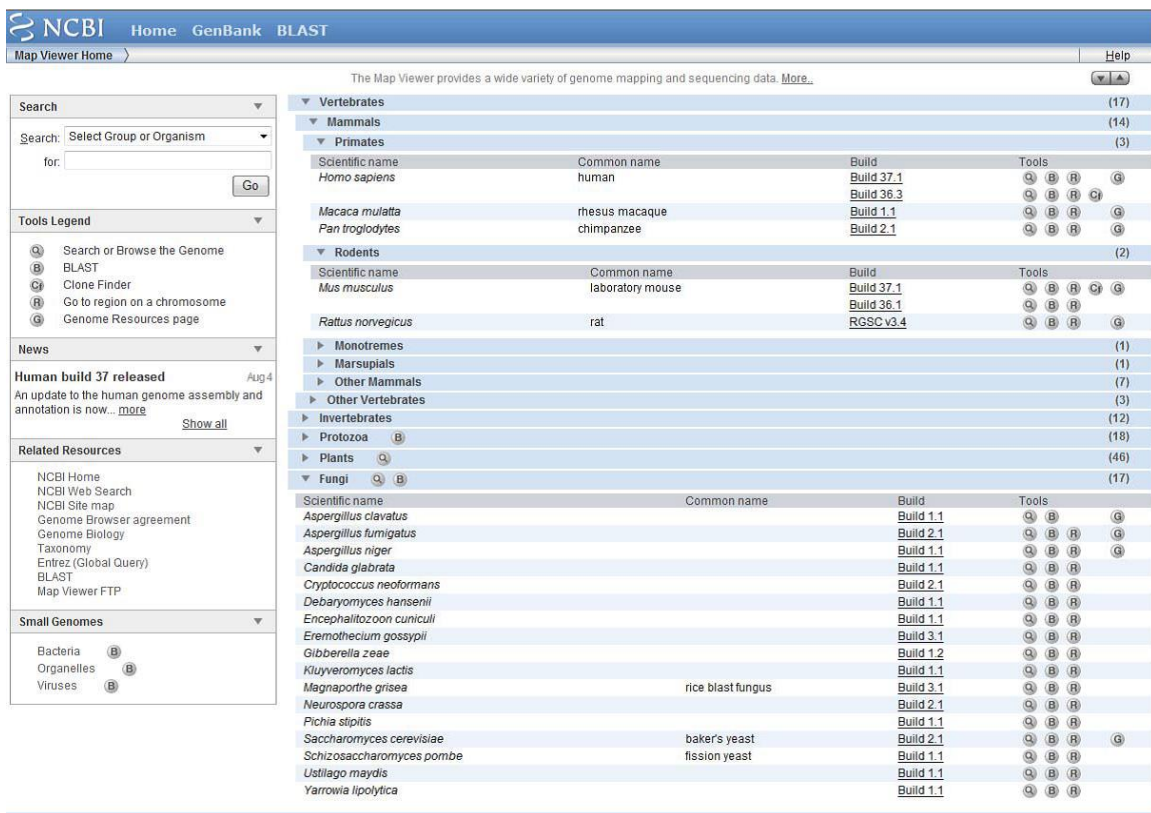


Figure 6. The NCBI MapViewer.

Three psychometric measures were administered: the Locus of Control Inventory, as well as the Neuroticism and Extraversion subscales of the IPIP Mini Big Five Personality Inventory. The Internal-External Locus of Control Inventory (LOC) (Rotter, 1966) is a 39-item forced choice measure designed to evaluate the degree to which participants attribute life events to some action of their own, or to some uncontrollable action outside of themselves.

Lower LOC scores are associated with an “internal locus” of control, an inherent belief that events and outcomes are under a person’s control, and thus, success or failure depends largely on personal behavior and attitudes. Higher scores indicate an “external locus,” an inherent belief that events and outcomes are influenced by external factors such as, unforeseen circumstances, a higher power, or “good luck.” Rotter postulated that these loci were traits remaining stable over a person’s lifetime (1966). Research demonstrates that locus of control has an impact on a wide variety of human outcomes, including academic and workplace performance (.

The Neuroticism and Extraversion subscales of the IPIP 20-item Mini Big Five Personality Inventory (Donnellan et al., 2006) ask participants the degree to which each listed characteristic applies to them. The Big Five factors have a long history in psychology and decades of literature on their scope and impact. Briefly, Extraversion defines the degree to which a person is open-minded, action-oriented and seeks the society of others. Neuroticism is distinguished by negativity and a propensity to be moody. In previous work (Green, Jeong, and Fisher, 2010), these traits have a demonstrated relationship to each other, and, in the case of Neuroticism, to locus of control.

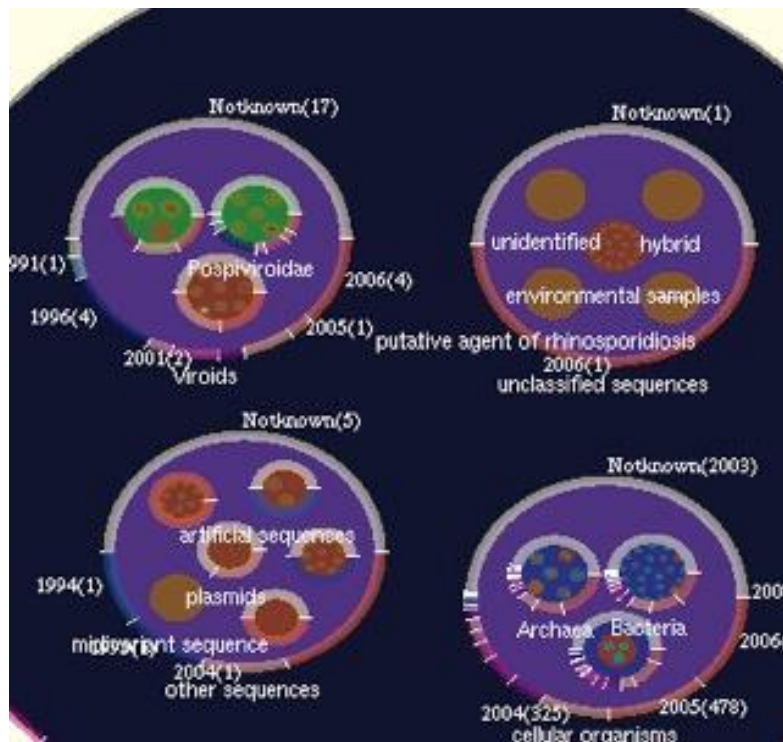


Figure 7. The MainView of GVis

3.4.3. Participants

In total, 106 participants agreed to complete the study: 50 in the first study, 56 in the second study. 94 participants reported being right-handed; 11 were left-handed. Most (101) were undergraduates and received course credit for participation. Students reported having 22 different majors or academic concentrations, including Business, Nursing, Computer Science, and Psychology. The vast majority of all participants 101 (96%) had taken fewer than 4 biology or biology-related classes. Novices were recruited specifically to better evaluate procedural learning with novel information; experts would have had a more advanced understanding of the knowledge ontology, which would have weakened the comparison between interface metaphors. All participants were asked to rate their ability and comfort level with a computer and mouse on a 5-item Likert-like scale. They were also asked to identify whether they had previous experience with the computer interfaces being investigated. 97 reported being comfortable or very comfortable with a computer; 79 reported having “very good” or “expert” computer ability. No one reported a computer comfort or ability level less than a 3 or “OK.” Almost all

(104) participants had used a web-based application before. 35 participants reported having used data visualization previously. None of the participants reported having a medical condition that might interfere with their use of a computer or mouse. 2 participants reported being color-blind.

3.4.4. Study Protocol

After signing the informed consent, participants were asked to fill out an online self-report questionnaire that included the 3 psychometric measures and basic demographic information, with particular emphasis on self-perceived ability, experience and comfort with computers and computer interfaces. Participants in the first study were allowed to complete the questionnaire online before their session in the lab. All data were collected for post-hoc analysis with task performance data.

In both studies, after completion of the self-report measures, participants began the procedural learning tasks in one of the two interfaces. The order of interface was counterbalanced for order effects; half of the participant used GVis first, and half used MapViewer first.

In the first study, the tasks started with a brief demonstration of interface and interaction techniques, such as the use of hyperlinks or how to zoom into the visualization. After the demonstration, a short tutorial was administered to introduce participants to essential tools and concepts in the interface, and to allow participants to experiment with what was being learned. In some cases, step-by-step instructions were given. A researcher was on hand throughout the study to answer any questions.

Following the tutorial was a series of 3 tasks designed to test procedural performance in finding target information: the participant was asked to identify a target located somewhere within the presented informational hierarchy. The question provided what base categorization or subclass the information was located within, but did not provide step-by-step instructions.

Participants were also told to find the item as quickly as possible, as the task was being timed. As soon as the target was located on screen, the participant pushed a

“Found It” button on the screen. The time taken from the presentation of the question on-screen to the moment the button was pushed was recorded as *completion time*.

In the second study, participants were asked to demonstrate script learning or tool skill by answering 5 hunt-and-find questions. All tasks were open response. Each question included step-by-step “cues” to assist in finding the answer to each question. A cue was the next step or concept on the current page or in the current view to look for. Participants were given little or no help from the researchers while working through the question, but were allowed or encouraged to experiment with different interaction paths within the interface in order to find the answer.

If the answer given was incorrect, the error was recorded and the researcher asked the participant to try again, until the correct answer was given. The total time from the initial reading of the question to the indication of the correct answer was recorded as the completion time. Participants were not told explicitly that they were being timed.

A third recorded outcome variable was insight. Participants were asked after finishing each task in both studies to indicate whether they had “learned anything unexpected while finding the solution.” Insight was defined as “unexpected” to prompt for only new knowledge that the participant considered to be novel or surprising. If the participant reported a new insight, they were asked to describe what they had learned.

After each participant had answered the questions in both interfaces, they were asked to specify which interface they liked better, and to give each interface a letter grade (“A” (superior) through “F” (failing)). A short debriefing ended the study session, and there were no follow-up sessions.

3.5. Results

In Study 1, the mean completion times for the procedural learning tasks in the MapViewer ($M = 684.77$, $SD = 235.46$) were more efficient than the completion times in the GVis ($M = 684.77$, $SD = 288.49$). In Study 2, the MapViewer procedural completion times were also faster ($M = 133.54$, $SD = 84.00$) than those in the GVis ($M = 161.64$, SD

= 111.40). Overall, participants preferred interacting with the visualization to interacting with the web table. This preference was indicated by post-study feedback. For example, when asked to give each interface a letter grade, from A (superior) to F (failing), 75 (73%) gave the GVis an A or B; 57 (56%) gave an A or B to the MapViewer. Additionally, when asked, 64(61%) reported that they both preferred the visualization; 39 (37%) preferred the web table.

3.5.1. Completion times and personality factors

The completion times for each condition for the procedural learning tasks in each study were merged into a single statistic, with $N = 106$. Participants completed tasks more quickly in MapViewer ($M = 383.15$, $SD = 32.38$) than in GVis ($M = 426.86$, $SD = 32.15$). A paired t-test between total completion times in GVis and completion times in MapViewer was significant ($t(100) = 2.11$, $p = .037$, suggesting that the differences in completion times was due to more than random chance. A one way Analysis of Variance (ANOVA) was used to test for the impact of Locus of Control (LOC) across interface completion times. The ANOVA for GVis was significant ($F(14, 88) = 1.89$, $p = .039$) but the comparison for MapViewer was not ($p = .099$). In addition, LOC predicted completion times in both interfaces; a Pearson's correlation between LOC and completion times was significant (GVis: $r(105) = .234$, $p = .02$, MapViewer: $r(105) = .254$, $p = .01$). (See Figures 8 and 9.) These findings suggest that participants with a more internal locus (those who believe they have control over personal life events) take less time finding target information than those with a more external locus.

This correlational finding is the opposite of findings reported in an earlier study (Green, Jeong, and Fisher, 2010). This previous study used inferential tasks, and found that participants with a more external locus (those who did not believe that they were in control) tended to solve a series of inferential tasks more quickly than those with a more internal locus. These tasks were more cognitively complex than the current studies, and asked the participants to compare and contrast multi-dimensional objects and make decisions about similarities and differences. We will discuss this further in the Section 4.

ANOVAs to test for the impact of Neuroticism in both interfaces were significant: GVis: ($F(16, 86) = 3.42, p < .001$), MapViewer: ($F(16, 85) = 5.14, p < .001$). Neuroticism also was negatively correlated with completion times in both interfaces. GVis: ($r(103) = -.47, p < .001$), MapViewer: ($r(102) = -.54, p < .001$). (See Figures 8 and 9.) ANOVAs to test for the impact of Neuroticism in both interfaces were significant: GVis: ($F(16, 86) = 3.42, p < .001$), MapViewer: ($F(16, 85) = 5.14, p < .001$). Differences in interface completion times and Extraversion were significant across both interfaces: GVis: ($F(14, 88) = 5.37, p < .001$). MapViewer: ($F(14, 87) = 4.12, p < .001$). These faster participants also tended to be more emotional and sociable. A summary of these findings can be found in Figure 13.

3.5.2. Task Errors and Personality Factors

The two studies measured tasks errors differently, and so must be analyzed separately. In Study 1, procedural tasks asked participants only to indicate when they had located the target (in seconds) across procedural tasks and the Locus of Control, Extraversion, and Neuroticism scores.

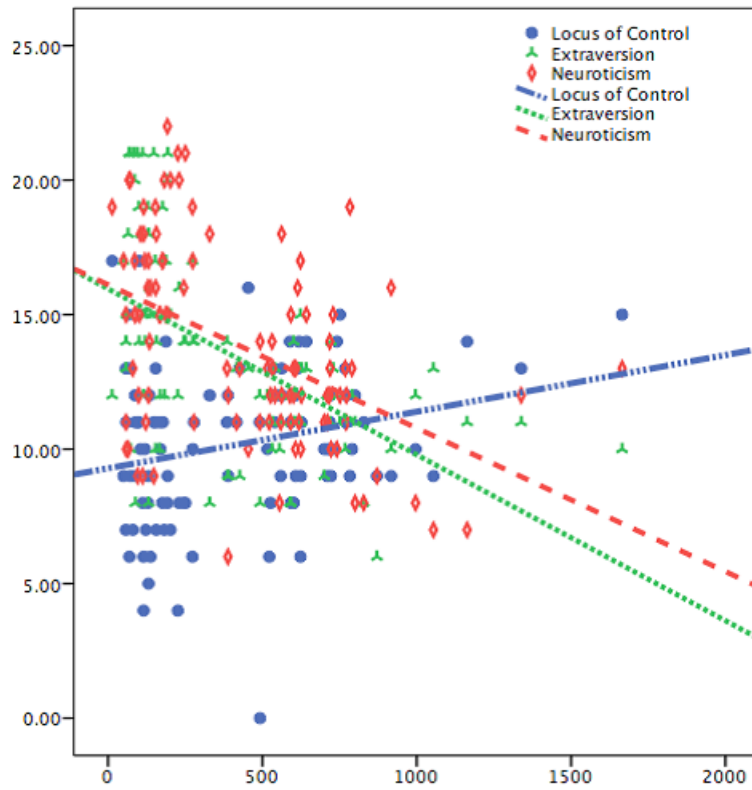


Figure 8. Correlations of GVis total completion times (in seconds) across procedural tasks and Locus of Control, Extraversion and Neuroticism scores.

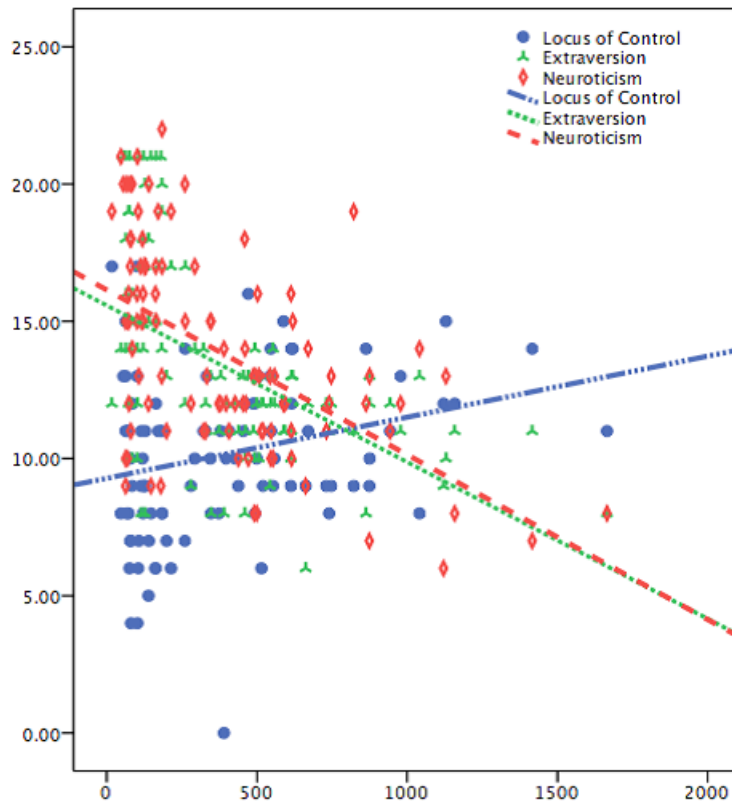


Figure 9. Correlations of Map Views Total completion times (in seconds) across procedural tasks and locus of control, extraversion, and neuroticism scores

Total Completion Times (in seconds) across procedural tasks and the Locus of Control, Extraversion, and Neuroticism scores. In Study 2, error was defined as giving the wrong answer to a question. Upon making an error, participants were asked to continue to try until they correctly solved the task. Each incorrect solution was recorded as an error. Kolomogorov-Smirnov Z was significant in both interfaces (GVis: $p < .001$, MapViewer: $p < .001$). Levene's test of homogeneity was significant in for GVis ($p = .004$), but not MapViewer ($p = .30$), suggesting that sample distributions were not uniformly normal. Due to these two findings, we opted to conduct non-parametric tests for the purposes of the following analyses.

Participants made more errors in GVis ($M = 1.21$, $SD = 1.07$), than they did in MapViewer ($M = .69$, $SD = 1.07$). Friedman's chi square was significant ($X^2(1) = 5.45$, $p = .02$) Kendall's tau was conducted between errors in each interface and psychometric

scores; no significant associations were found. Generally speaking, only the difference in interface had a significant impact on how many errors were made; participants were more effective in the MapViewer interface. A summary can be found in Figure 13.

3.5.3. Insight generation and personality factors

Participants reported having more “unexpected” insights in the GVis (N = 73) than in the web-based MapViewer (N = 70). The distribution of the combined insights reported across both interfaces was not normal according to the Kolomogorov-Smirnov (GVis: $p < .001$, MapViewer: $p < .001$). Levene’s test of homogeneity was significant for GVis ($p < .001$), but not MapViewer ($p = .373$). As the distribution was not normal, a Friedman’s chi square was run between the mean number of insights generated in both interfaces, and was not significant: Friedman’s $X^2 (1) = 1.59$, $p = .208$. Kendall’s Coefficient of Concordance = $.015$. This suggests that interface type did not have a significant impact on the number of insights generated.

In an investigation of the impact of Locus of Control (LOC) on insight generation, a Friedman’s chi square was run between LOC scores and the mean number of insights generated in both interfaces and was significant. GVis: Friedman’s $X^2 (2) = 174.36$, $p < .001$. Kendall’s Coefficient of Concordance = $.83$. MapViewer: Friedman’s $X^2 (1) = 101.04$, $p < .001$. Kendall’s Coefficient of Concordance = $.96$.

As the sample was large ($n > 50$), Spearman’s rho was conducted to evaluate correlations between the psychometric scores and completion times. Locus of Control predicted the number of generated insights (GVis: $R (103) = .20$, $p < .04$; MapViewer: $R (101) = .239$, $p = .016$). Because both studies had a within participants design, a Kendall’s tau-b was conducted. LOC was not associated with the number of generated insights in both interfaces (GVis: $p = .59$, MapViewer: $p = .46$).

These findings demonstrate that LOC had some impact on the number of insights the participants reported; persons with a more external locus tended to report a greater number of insights (Figure 10).

We also explored the impact of Big Five personality traits Extraversion and

Neuroticism on insight generation in both interfaces. A Friedman's chi-square between mean Extraversion scores across interfaces was significant. (GVis: Friedman's $X^2(1) = 105.0, p < .001$. Kendall's Coefficient of Concordance = 1.0. MapViewer: Friedman's $X^2(1) = 105.0, p < .001$. Kendall's Coefficient of Concordance = 1.0). Extraversion was associated with insight generation (GVis: $\tau = -.15, p = .051$, MapViewer: $\tau = -.18, p = .027$), and predicted the number of insights in both interfaces (GVis: $R(103) = -.554, p < .001$; MapViewer: $R(101) = -.543, p < .001$). These findings suggest the more insights were reported by participants that were less extraverted (Figure 4 11).

A Friedman's chi-square between mean Neuroticism scores across interfaces was significant: (GVis: Friedman's $X^2(1) = 105.0, p < .001$. Kendall's Coefficient of Concordance = 1.0. MapViewer: Friedman's $X^2(1) = 105.0, p < .001$. Kendall's Coefficient of Concordance = 1.0). Neuroticism was not significantly associated with insight generation (GVis: $p = .716$, MapViewer: $p = .37$), but did predict the number of generated insights in both interfaces (GVis: $R(103) = -.415, p < .001$; MapViewer: $R(101) = -.509, p < .001$). These findings suggest that more neurotic participants did not report as many insights as those who had lower Neuroticism scores (Figure 12). A summary is in Figure 13.

3.6. Discussion

The findings of these studies demonstrate that, even when the procedural tasks are somewhat different, inherent personality differences can predict interaction and behavioral outcomes across the interfaces. Aside from generally evaluating interface learnability, which we did in both studies, we studied procedural learning tasks in two slightly different ways. The first study focused on target identification; participants were asked to find an organism label on the screen: for GVis, this label was attached to a spherical glyph, for MapViewer, very often the label was also a textual hyperlink. Once the label had been obtained, the participant pushed the "Submit" button and the task was done.

In the second study, we asked participants trivia questions whose answers had to be hunted through the interface. If they gave the wrong answer, we requested that

they keep looking. Like the first study, nothing other than an ability to use the interface and identify target labels was required. In both of these tasks, participants found the targeted information more quickly in the web table MapViewer; in Study 2, they also made fewer errors in MapViewer. Given the wide commercial use of web tables, it seems reasonable that most participants brought some prior knowledge of the interaction metaphor to the MapViewer tasks that they did not have for the data visualization. However, participants still strongly preferred GVis to MapViewer, even if they were not as effective in task performance. This may have been due to the novelty of GVis; most participants had never seen anything like it before. It also may have been due to data organization; many participants, in post-study open response, indicated a clear preference for GVis" organization and interaction.

Locus of Control proved to be an influential personality trait no matter what the interface or task. The faster participants in both interfaces were persons who had a more internal locus of control, which is typified by a belief in personal control over life events. This finding is in close agreement with much of the available literature on locus of control. Persons with a more internal locus have been found to have better problem-solving skills (Krause, 1986), to be more resolved to solve a task when it became difficult, and to be more likely to develop an intrinsic (internal) motivation to finish a difficult task (Weiss and Sherman, 1973). Thanks in part to positive behaviors like these, internal locus has also been found to lead to superior outcomes in academics, hospital recovery, and organizational environments.

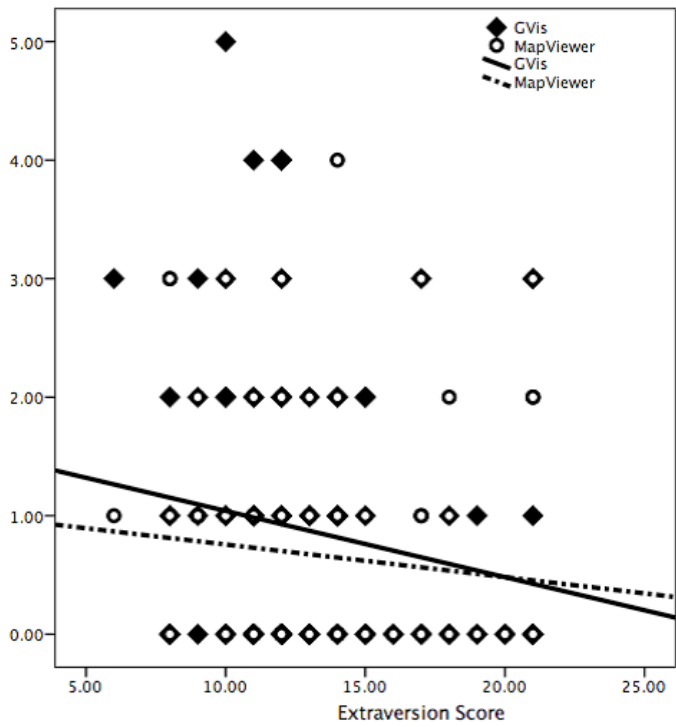


Figure 11. Insights and Extraversion

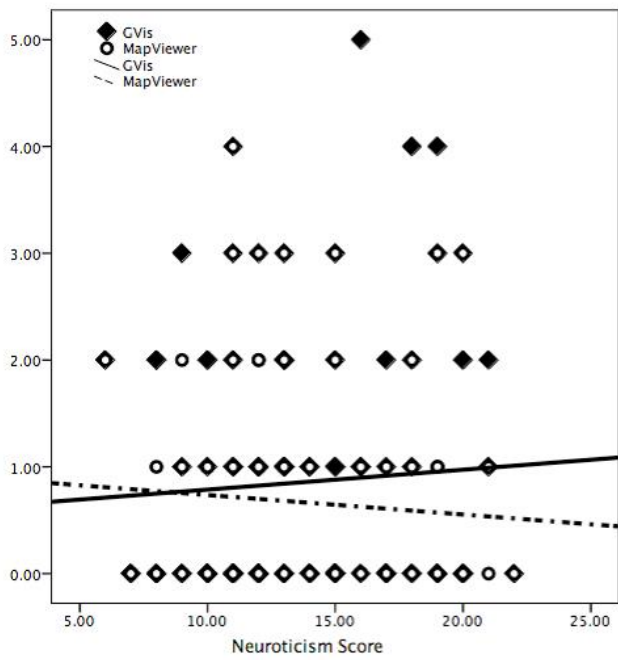


Figure 12. Insights and Neuroticism score

What is intriguing is that, while an internal locus led to faster procedural task outcomes, this is not necessarily the case when the task becomes more cognitively difficult. In a previous paper (Green, Jeong, and Fisher, 2010), we studied inferential learning. The tasks required participants to evaluate a multi-dimensional exemplar, and draw a conclusion about other organisms based on similarities or differences. We reported that participants who had a more external locus – those who believe that they are not in control, and who tend to believe in luck as a cause of events – solved inferential tasks in GVis more quickly than those with an internal locus. For a discussion of these results, please see (Green, Jeong, and Fisher, 2010). The results do not contradict our current findings, but rather expand on them. In these studies, we used a larger N, which likely made our analyses more sensitive to changes in participant scores. Further, we focused on only 3 constructs that seemed more highly predictive, unlike (2010) which used 6 psychometric measures.

For one type of learning task performance to be predicted by the degree of internal locus and another type to be predicted by the degree of external locus lends credence to our introductory statement that, depending on task, inherent individual differences can predict interface performance. Yet while locus of control has been shown to be influential in a wide variety of human performance, as previously discussed, to date, it has not been considered by interface designers and evaluators. Based on our research, as well as a broad locus of control literature, we consider locus of control to be one construct in the Personal Equation of Interaction. In addition to Locus of Control, the Big Five personality factors of Neuroticism and Extraversion also predicted procedural task performance. The more extraverted or neurotic the participant, the more quickly he or she was able to identify target information.

This is interesting, but little in the behavioral literature explains these correlations; for us, it is a subject of our ongoing research. Further, Neuroticism in these studies was found to be negatively correlated with Locus of Control ($r(105) = -.284, p = .003$). This does have some precedent in the literature. For example, Judge et al. (2006) evaluated several personality factors, including Locus of Control and Neuroticism, and found that they were interrelated and could be shown to be a part of the same construct. This means =that items from these measures trended together and were statistically

predictive of the same personality factor(s). Research like this affirms psychometric constructs can and do work together, Further, it lends credence to an approach that seeks to find items or clusters of items which could work together in the prediction of interaction efficacy.

Insights were also predicted by personality as in factor scores. This is compelling because it suggests that the impact of a predictive Personal Equation may go further than efficacy or efficiency; it may extend to being able to predict some learning or problem-solving outcomes as well. Much depends on how the word “insight” is defined. In the visualization and visual analytics literature, insight is often undefined. When defined, it is often broadly defined, as in (North, 2006). This makes “insight” difficult to use as an evaluative interaction outcome, and thus, as briefly discussed earlier, leaves certain claims about the superiority of visual analytics interfaces unproven. Recently, “insight” has been defined within two categories: knowledge-based insight, and spontaneous insight (Chang et al., 2009). Spontaneous insight is a sudden solution to an unsolvable problem, and has often, in the psychological literature, been referred to as an “aha!” moment. Spontaneous insight was not evaluated in these studies.

In these studies, we evaluated the number of knowledge-based insights reported across task and interface, which are generally defined as items or concepts learned or added to the user’s knowledge base. In evaluating the knowledge-based insights reported, we categorized insights on the basis of content: insights about how to use the interface itself were separated from insights about the informational content presented and manipulated.

In both interfaces, roughly twice as many knowledge-based insights were reported about interface learnability (GVis: N = 51, MapViewer: N = 47) as were reported about the informational content (GVis: N = 22, MapViewer: N = 23). In both interfaces, the greatest number of interface learning insights was reported in the first question, which suggests that learnability started early.

As the task set proceeded, the reported count of each insight type tended to even out somewhat, which is not unexpected; users started paying attention to content once manipulating the interface was less of an issue or became more automatic.

Overall, whether learning about the interface or the interface content, personality factors predicted reported learning as well as other interaction outcomes. These findings have immediate implications. For example, these studies have demonstrated that users who tend to be more extraverted and neurotic are also more likely to believe that they are in control of the task situation (internal locus). By extension, this also means highly neurotic or extraverted users tend to be better at interface manipulation and target identification. If the personality factors of the user were known beforehand, we could reasonably predict how quickly he or she would be able to learn a novel interface and find pertinent information. For even when the interaction metaphor was completely unfamiliar, as it was in the GVis visualization, neurotic/extraverted participants were able to learn to manipulate the data more quickly.

	Completion times	Errors	Insights
Interface	faster times in MapViewer	fewer errors in MapViewer	more insights in Gvis
Locus of Control	internal locus faster times	none	external locus more insights
Extraversion	more extraverted faster times	none	less extraverted more insights
Neuroticism	more neurotic faster times	none	less neurotic more insights

Figure 13. Summary of findings.

However, what these findings do not do is demonstrably differentiate between interface and interactive techniques. The three evaluated personality factors impacted both interfaces similarly. Given the cognitive simplicity of the tasks, this is perhaps unsurprising. Ongoing research has been designed to evaluate learning styles which tend to guide focused attention and information organization during task, and where behavior research suggests more delineating personality factors for visualization technique might be found.

A last note is on the use of novices in evaluations using an expert system; most

of the participants had little or no knowledge of biological concepts. However, the participants were still capable of ably find target information in both interfaces. Yet even with the more familiar archetype of the web interface, participants preferred the visualization. The intent of these studies was never to evaluate the efficacy of GVis *per se*; a formal evaluation of GVis as an expert system is reported in other literature [19]. The aim of these studies was to evaluate human cognition during learning interaction using both interfaces as working artifacts of a kind. In addition, we explored whether individual differences in personality factors and self-beliefs could have a large enough impact on interaction outcomes to warrant their inclusion in the Personal Equation of Interaction.

For these reasons, we recruited non-experts who were unfamiliar with the knowledge domain. Expertise would have biased the user's interaction; they would have had an expert knowledge of the genomic hierarchies, and thus known where to look for the requested information. This would have proven a poor evaluation of how each interface promoted learning.

3.7. Conclusion

The Personal Equation of Interaction is still very much a work in progress. In the short-term, it serves as an open discovery and proof of concept. We have shown that inherent differences impact interaction. Our ongoing research seeks to better define what differences impact what type of analytical task (for it seems reasonable to assume that one inherent set of differences will only generalize to one type or set of task constraints). For example, we are currently narrowing our task sets to study multiple decision points in specific types of category or inference reasoning. And further, we hope to explore whether that impact is temporally static or dynamic throughout the analytical process.

In the longer term, we intend to isolate predictive matrices and validate a battery of measures that will successfully inform interface design based on the types of cognitive task undertaken. Ultimately, this is the Personal Equation of Interaction. These measures will likely involve more than personality factor matrices; other areas of

exploration include perceptual logics and use of decision-making heuristics. In addition to informing design, the Personal Equation could be used to provide real-time interface adaptation to accommodate user needs and preferences, and provide a basis for robust group profiles of users who share common differences, such as experts or users of a particular visualization technique. Visual analytics seeks to facilitate analytical reasoning through the use of interactive visual interfaces. In the Personal Equation of Interaction, we will provide a new tool in that pursuit.

3.8. References

- Chang, R., Ziemkiewicz, C., Green, T. M., & Ribarsky, W. (2009). Defining insight for visual analytics. *Computer Graphics and Applications*, IEEE, 29(2), 14-17.
- Green, T. M., Jeong, D. H., & Fisher, B. (2010, January). Using personality factors to predict interface learning performance. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1-10). IEEE.
- Heppner, P. P., & Anderson, W. P. (1985). The relationship between problem-solving self-appraisal and psychological adjustment. *Cognitive therapy and research*, 9(4), 415-427.
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct?. *Journal of personality and social psychology*, 83(3), 693.
- Kaptelinin, V. and Nardi, B.A. (2006). *Acting with technology: Activity theory and interaction design*, MIT Press, Cambridge, Mass.
- Macia, A. (1979, April). Visual perception of landscape: sex and personality differences. In *Proceedings of Our National Landscape Conference*. USDA Forest Service, General Technical Report PSW-35 (pp. 279-295).
- Simon, H. A. (1991). Cognitive architectures and rational analysis: Comment. In *Architectures for intelligence: The 22nd Carnegie Mellon symposium on cognition* (pp. 25-39).
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.
- Weiss, H., & Sherman, J. (1973). Internal-external control as a predictor of task effort and satisfaction subsequent to failure. *Journal of Applied Psychology*, 57(2), 132.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, 89(4), 369.
- Boyatzis, R. E., & Kolb, D. A. (1995). From learning styles to learning skills: the

- executive skills profile. *Journal of Managerial Psychology*, 10(5), 3-17.
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment*, 18(2), 192.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.
- Hall, R. H., Rocklin, T. R., Dansereau, D. F., Skaggs, L. P., O'Donnell, A. M., Lambiotte, J. G., & Young, M. D. (1988). The role of individual differences in the cooperative learning of technical material. *Journal of Educational Psychology*, 80(2), 172.
- Heuer, R. J. (1999). *Psychology of intelligence analysis*. Lulu. com.
- Krause, N. (1986). Stress and coping: Reconceptualizing the role of locus of control beliefs. *Journal of Gerontology*, 41(5), 617-622.
- Messer, S. B. (1972). The relation of internal-external control to academic performance. *Child Development*, 1456-1462.
- North, C. (2006). Toward measuring visualization insight. *Computer Graphics and Applications*, IEEE, 26(3), 6-9.
- Palmer, J. (1991). Scientists and information: II. Personal factors in information behaviour. *Journal of documentation*, 47(3), 254-275.
- Pintrich, P. R., Roeser, R. W., & De Groot, E. A. (1994). Classroom and individual differences in early adolescents' motivation and self-regulated learning. *The Journal of Early Adolescence*, 14(2), 139-161.
- Plaisant, C. (2004, May). The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 109-116). ACM.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1), 1.
- Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, 25(2), 203-244.
- Wong, P. C., & Thomas, J. (2004). Visual analytics. *IEEE Computer Graphics and Applications*, (5), 20-21.

Chapter 4. The personal equation of complex individual cognition during visual interface interaction

4.1. Overview

This chapter was originally published as an invited book chapter: *Green, T. M., & Fisher, B. (2011). The personal equation of complex individual cognition during visual interface interaction. In Human Aspects of Visualization (pp. 38-57). Springer Berlin Heidelberg.* The chapter does not report any new research but identifies and discusses key reasoning categories and elucidates how these categories might be studied as part of building a knowledge domain to improve the intuitive ease of visualization interaction.

4.2. Abstract

This chapter considers the need for a better understanding of complex human cognition in the design of interactive visual interfaces by surveying the availability of pertinent cognitive models and applicable research in the behavioral sciences, and finds that there are no operational models or useful precedent to effectively guide the design of visually enabled interfaces. Further, this chapter explores the impact of individual differences, and in particular, inherent differences such as personality factors, on complex cognition. Lastly, it outlines how knowledge of human individuality, coupled with what is known about complex cognition, is being used to develop predictive measures for interface interaction design and evaluation, a research program known as the Personal Equation of Interaction.

4.3. Introduction

Generally speaking, interactive visualizations are considered to have a number of advantages over more conventional visual interfaces for learning, analysis, and knowledge creation. Much of the support for these claims comes from a variety of sources, such as user evaluations, comparative studies of error rates, time to completion, etc., as well as and designer/developer intuition. One common claim concerns the development of insight. From early on, visualization has been proposed as a preferable interface approach for generating insight (e.g. Card, Mackinlay and Shneiderman, 1999, Saraiya, North and Duca, 2004, Spence, 1956).

As a concept, however, insight is a construct that is often either loosely defined as some variety of meaningful knowledge or is left undefined (e.g. Saraiya, North and Duca, 2005, Springmeyer, Blattner and Marx, 1992). More recently there have been efforts to define insight, although not in ways that might enable it to be quantified. For example, North described insight as a broad construct, which is complex, deep, qualitative, unexpected, and/or relevant (2005), without characterizing the cognitive processes that give rise to it, or the outcomes that are generated by it. Chang et al. defined insight as comprising two categories: knowledge-building insight, which is a form of learning and/or knowledge generation, and spontaneous insight, which is method of problem-solving for previously intractable problems, commonly described as an a-ha! moment (2009).

This dual definition has advantages over a unitary definition in that it supports focused analysis of the component aspects of the overall construct. Spontaneous insight, however, has been an elusive notion for researchers in several disciplines; neuroscientists and psychologists have studied the phenomenon, but as yet do not know how insight is triggered.

By any definition, there is little empirical evidence that supports claims of visualization superiority in insight generation, though user evaluations are often conducted to demonstrate visualization efficacy over other types of interface. Plaisant et al. (2004) identified four current themes in the evaluative literature: controlled experiments comparing design elements, usability evaluation, controlled experiments

comparing two or more tools, and in situ case studies. In all four groups, evaluations and comparative studies have largely focused on perception, motor learning, focal attention, target recognition and/or target acquisition. For example, musing behaviors were used as a predictor of user focus in a geospatial visualization (Wong & Thomas, 2004). Jeong et al. compared two visualization tools to determine in which interface users were more efficient in finding outliers and identifying highly correlated items in a matrix (2010). Nodes were the subject of an evaluation of target identification in large tree tools (Plaisant, Grosjean and Bederson, 2002). And Wang et al. evaluated whether users could focus on the count of visualized objects (in this case, paper proposals) over a period of time (Beth and Piaget, 1966). In these evaluations as well as in cognition as a whole, perceptual, cognitive, and motor processes are important to the overall interaction. However, each of these identified cognitive systems is a feeder processes. That is to say, they support and inform the more complex processes, such as reasoning, problem-solving, and knowledge generation, which form the backbone of systematic analysis or task solution. These complex processes are impacted demonstrably, as we will see, by the individuality of the interface environment, the knowledge domain, and the inherent differences within the user, over which visualization design has no control. To date, visualization evaluation has insufficiently considered the complexity of human cognition. This, in turn, has hampered the design of intuitive interfaces capable of mixed-initiative collaboration.

In this chapter, we will explore a variety of challenges to the consideration of cognitive complexity in visual analytics design, from the current lack of operational models and applicable research to a consideration of individual differences. We will then explore how an consideration of how these complex processes impact the understanding of common visual analytics tasks, and discuss a continuing exploration of how human individuality can be measured and charted, leading to a differentiating set of predictive measures that can not only predict interface performance, but guide visualization design. We call this the Personal Equation of Interaction.

4.3.1. The Challenge of Complex Cognition

Very little research examines the use of what is commonly known as higher

cognition during interaction, which includes processes such as reasoning, problem-solving, and decision-making. Frequently, when a visualization design or evaluation argues that a specific technique or tool improves insight (which is learning and/or problem-solving) or analysis (which involves every major cognitive process), the evidence is actually task completion times for single-step tasks, improved target identification, or other simple outcomes. One reason for this, perhaps, is the practice of inferring the success of complex behaviors from measurements of simpler ones. A common example is the generalization of findings from simple, semantically-unrelated target acquisition tasks to human problem-solving as a whole, without a discussion of which of the many problem-solving theories or heuristics the finding might speak to (e.g. Dou et al., 2009, Robinson, 2008). This practice over-simplifies the complexity of cognition, but is understandable, given that our best complex cognitive models are black box or descriptive. We will now consider the best known of these descriptive models, the sensemaking model.

4.3.2. The Sensemaking Loop

The most familiar approach to descriptively outline task-oriented processes is Pirolli and Cards sensemaking loop (Pirolli and Card, 2005, Russel and Card, 1993). See Figure 14. Russell et al. defined sensemaking as *the process of searching for a representation and encoding data in that representation to answer task-specific questions* (1993). In the rest of this section, we will summarily explore the complexity of analytical cognition through a brief discussion of the sensemaking loop in the broader context of human reasoning. This seems necessary, for, as valuable as the sensemaking loop is to describing certain analytical tasks, its use tends to be overgeneralized in the visualization literature. Indeed, *sensemaking* is often the term given to most or all of the cognitive processes analysts employ during visual analytics tasks (e.g. Stasko et al., 2007, Pirolli and Card, 2005, Heer and Agrawala, 2007).

Sensemaking, as defined in the previous paragraph, creates a mental or physical representation (i.e. a “mental model” or “story”). This singular representation may well be necessary for problem solving, but may not be in itself sufficient for generating valid implications. Analyses may create multiple alternative mental representations of a

situation in order to compare them in a variety of ways, using a variety of evaluative heuristics in order to draw their conclusions.

At this larger scale of problem solving strategy, analytical cognition exhibits a great deal of variability, and is informed by both human and task individuality. For example, the sensemaking loop makes an effort to delineate top-down and bottom-up task descriptions. However, as some of Pirolli and Card's participants indicated, the cognition involved in the early tasks of the loop can rarely be so cleanly categorized. Further, though seemingly simplistic, even the first steps of the sensemaking loop (the "lower-effort" tasks of searching and filter) requires complex cognition in the form of various reasoning heuristics to categorize, evaluate, and assemble pertinent information. These heuristics could be elimination heuristics like elimination-by-aspects (Wang et al., 2008), or satisficing (Kozielecki, 1971) or they could be more complicated, such as the comparing possible shoebox members (concepts the analyst has gathered during sensemaking and think may be related to each other) to an ideal before addition. According to the Loop, pertinent shoebox members become part of an Evidence File which is used as part of the formal structure of the Schema, which is a structured narrative of how the evidence collected thus far fits together (Pirolli and Card, 2005). Thus, sensemaking effectively describes one key component of the complex cognitive processes involved in the generation of insight, it does not claim to capture higher levels of abstraction, i.e. generation of problem-solving strategy, nor does it attempt to capture lower-level pattern recognition processes that support its operation.

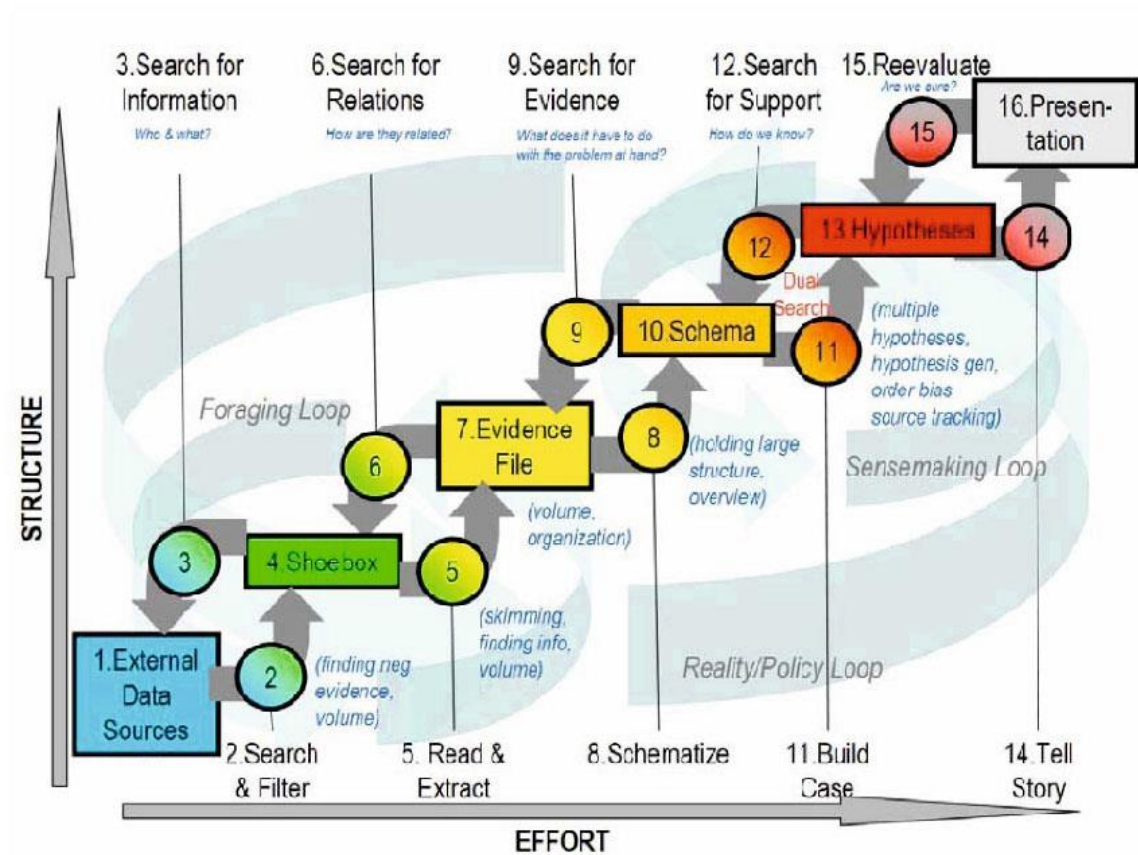


Figure 14. The sensemaking loop. From Russell and Card, 1993.

The latter set of processes are arguably the most critical for visualization to support, since they are precisely those aspects that are most closely tied to the visual systems own information processing capabilities (i.e. Irving Rock’s logic of perception (1983)). Similarly, Cherubini’s “models to rules mechanization” (2006) suggests that the formation of schemata and hypothesis generation are not necessarily higher effort tasks. According to Cherubini, after human reasoning uses the generated model (which does require some cognitive effort to create in novel instantiations), the human infers a rule from the knowledge structure in the mental model. Or in his own words:

After grasping the common structure of the problems, most people should be able to devise a simple rule to solve all problems with the same structure (i.e., a domain-specific rule), disregarding the number of possibilities underlying them. (Cherubini and Mazzocco, 2004)

This rule may be created after only one or two uses of a newly created model. Thus, depending on the information under consideration, hypothesis generation could actually require less cognitive bandwidth than the initial information search.

At a higher level, the sensemaking loop articulately describes one subset of reasoning, that of reasoning types which generate inferred hypotheses or generalized rules, or abduction and induction. Abduction (or abductive reasoning) is process of approaching seemingly unrelated information with the assumption that the data points or concepts are indeed interconnected; abduction creates and infers relations between two previously unrelated data points; the end product of abductions series of inferences is the creation of a hypothesis which explains these relational inferences [41], and usually about a specific item or concept. This fits well with the structure of the sensemaking loop, which situates the search and compartmentalization of small, unassociated details in the early stages of the loop and builds from there to an identifiable hypothesis or story.

The sensemaking loop also illustrates induction. Inductive reasoning, as is described in the behavioural literature, is generally referred to the process of making acceptable generalizations from the similarities of facts or properties (see for example Rips, 1975, Tversky, 1972). The validity or strength of the generalization depends in large degree upon the strength of these similarities. One of the characteristics of induction which makes induction different from abduction is that relationships are not as important to the successful generalization. Once a fact(s) or similarity has been accepted as valid, it or they become the basis of generalization; additional information may or may not be considered. Induction is a powerful form of reasoning that allows us to quickly categorize and infer rules; it is often referred to as a form of bottom-up reasoning, but can utilize top-down cognition as is needed. To some degree, the sensemaking loop describes induction, as this type of reasoning tends to build from smaller facts to a broader concept. However, unlike with hypothesis generation and analysis, induced generalizations do not necessarily require vetting, e.g. careful consideration of available evidence for the generalization. Induced hypotheses can jump past multiple steps, such as the shoebox or schema creation, straight to a generality.

While abduction or induction may accurately describe behavior during discovery

or during exploration of novel problems for which the analyst does not already have a script or mental model to guide her as to what to do next, most problem-solving or decision-making tasks are driven by an articulated goal, theory or hypothesis from the start. These reasoning heuristics are neither abductive, which ends with a hypothesis, nor inductive, ending in an inferred rule, but rather deductive.

One of the more actively studied reasoning types, deductive reasoning falls outside of the sensemaking loop. By deduction (or, for that matter, induction), we are referring to reasoning as the subject of decades of empirically conducted research, which is usually evaluated through use of normative tasks like those that we will briefly itemize in the next section. A discussion of the philosophy of deduction, such as the traditional deduction of Aristotle, propositional (or first order) deduction, or natural deduction (Jaskowski, 1967, Gentzen, 1967), and the similarities or differences between these definitions, is neither intended nor implied.

There are several theories of deduction in the behavioral literature, but we will focus on two of the broader categories of deduction, which assume that human deductive reasoning is either rule based or model based. This dichotomy of rules vs. models raises interesting issues for visual analytics interfaces. Rule-based theories assume that humans use a script or a formal sequential logic while working through a deductive task (e.g. Johnson-Laird, 1991, 1999). From this perspective, the content or information manipulated during the task doesn't demonstrably influence the reasoning, because the inferences drawn during induction are part of this formal process. From the rules perspective, for the development of more intuitive deductive visual analytics tools, it would only be important to uncover the pertinent rules the analyst would use; theoretically, these deductive rules would generalize to all similar visual analytics tasks.

Model-based theories, however, are quite different. Models can be either concrete or abstract, complete or incomplete, pictorial or conceptual (e.g. Johnson-Laird, 1991 and 1999). Models are also flexible; they can change as pertinent information changes, and can quantify degree (such as few or often), as well as causal conditions. Models depend heavily on semantic relationships, and so can be heavily influenced by the content of the information at hand. This, too, influences visualization design, for what

data is presented when, and in what context, can influence the development of the model, as well as its completeness and validity. From the model-based perspective, discovering quickly inferred rules is not nearly as helpful as assuring that the human has all pertinent information readily at hand.

With pertinent information, the human can generate a mental model, which can be manipulated as needed to reasoning through the task to a valid conclusion. Schniederman's Mantra (1996): Overview first, zoom and filter, details-on-demand assumes deductive reasoning. The big picture, or hypothesis, drives the interactive behavior. It is not surprising, then, as powerful as models would seem to be in the successful use of visual analytics interfaces, that they have a place in the visual analytics literature, even if the theory and implications of models are rarely discussed.

This has been only a brief, general discussion of the sensemaking loop and how it fits into the broader context of common reasoning types. Human reasoning is about reaching a usable or verifiable conclusion, but the ways in which we reach these conclusions, as we have seen, can vary widely. For this reason, it is easy to see why analytical reasoning processes have yet to be operationalized in a manner that meaningfully informs research and design. For while descriptive models like the sensemaking loop do much to frame the big picture, intuitive interfaces will require a more detailed working-order understanding of what lies inside the frame.

4.3.3. The Absence of Precedent

As we saw in the last section, there is, as yet, no unifying theory of reasoning (if such a thing is even possible). What does exist is a complication of decades of research into specific laboratory tasks, usually characterized by small-scale problems, which are intended to uncover reasoning heuristics and biases. These are of limited use for real-world applications, and in particular map poorly onto visually enabled human reasoning (e.g. interactive visualization for cognitive tasks). Further, the theories that motivate these studies are often bound to a particular task and environment. Thus the field of behavioural research as a whole is characterized by contradictory, often esoteric theories that fail to explain the narrative of reasoning from beginning of task to end.

For example, deductive reasoning is almost entirely studied in a laboratory trials. Both rule based and model-based deduction has traditionally studied by presenting participants with syllogisms and evaluating the conclusions that are drawn. Phillip Johnson-Laird often uses syllogisms to study aspects of reasoning, which can take forms such as this inference about object properties: *Only one of the following statements is true:*

*At least some of the plastic beads are not red, or
None of the plastic beads is red.*

Is it possible that none of the red beads is plastic? (Johnson-Laird,2008, pg. 150).

Other common uses of syllogisms involve mental reasoning and inferences about spatial relationships, such as:

*The cup is on the right of the plate.
The spoon is on the left of the plate.
The knife is in front of the spoon.
The saucer is in front of the cup.*

What is the relation between the knife and the saucer? (Johnson-Laird,2008, pg. 130)

Cherubini and Johnson-Laird (2004) studied qualified inferences in iterative reasoning through word problems like the following:

*Everybody loves anyone who loves someone.
Anne loves Beth.
Does it follow that everyone loves Anne?*

.. . .

Does it follow that Carol loves Diane?

Cherubini and Mazzocco also evaluated the mental models to rules mechanization through use of a computer program loaded with a series of virtual card problems (2004) as illustrated in Figure 15. The participant was asked whether, based on the presented cards, a proposed sentence was *certainly true*.

Gigerenzer, in his evaluation of “fast and frugal” reasoning heuristics, used what

he considered to be common knowledge about cities in questions about which he asked participants to make quick reasoning decisions. The questions were simple, such as *Is this [city name] the capital of the country?* (Gigerenzer, G., Goldstein, 1996). Gigerenzer postulated that humans could make quick decisions based on very simple elimination heuristics which depended on accumulated general knowledge. These decisions were found to be more accurate than more sophisticated human and computer reasoning simulations. The behavioral literature contains decades of research similar to the research we have discussed, with each study having its own novel, usually non-real world, problem formulation. Study problems are often designed to study some small subcategory of reasoning (iterative inferred, probabilistic, etc.) and very few or no studies are published which are designed to explain how humans solve a complex problem from start to finish.

Perhaps it is not surprising then that, with all of this research, there is still a lack of precedent on how to conduct research into visually enabled reasoning. It is not at all clear how one might evaluate interfaces with respect to their ability to scaffold higher-order cognitive tasks. Further, unlike many of the simpler cognitive tasks, higher cognition is almost never binary, sequential, or singly threaded. It is, in practice, dynamic, combinatorial, and capable (at least to some degree) of parallel processing. Which heuristics are used during complex cognition and when will depend on the task, the environmental framing, and, as we will now discuss, differences in how an individual assimilates and manipulates new information.

4.4. Individual Differences

Complex cognition, for all of its variety, is also influenced by human individuality. There is no standardized unit of human cognition. It is influenced, sometimes profoundly, by the users' distinctive abilities and bottlenecks, beliefs about the world, preferred methods of categorizing and prioritizing information, and other individual differences. This is one reason that the modeling of reasoning has traditionally been difficult. Human behavioral research has demonstrated the impact of individual differences on learning and analysis in traditional environments.

There is also a plethora of examples in the behavioral literature of how individual differences impact cognition; for the sake of brevity, we will focus on the impact of personality factors, which also have a broad literature of their own. For example, personality factors predicted preferences and visual perception of landscapes (Macia, 1979). Visual impairment in children is heavily influenced by individual personality differences (Corn, 1983). Individual differences also affect how humans categorize, including the categorizing of stereotyping and prejudice (Heaven, 2002). Palmer found that interactive behaviors in information search can be categorized by personality factors (Palmer, 1991). Another study found that problem-solving behaviors could be predicted by responses to the Thematic Apperception Test (Ronan, 1996). In reasoning behaviors, individual differences impact rationality and reasoning as well (Stanovich and West, 2000, Stanovich, 1999). These are just a handful of studies in a deep literature of individuality and the impact of these differences on every major cognitive process, as well as behavioural outcomes, such as academic or organizational performance.

The question is not whether individual differences impact cognition, but how we can use individual differences to improve our understanding of visually enabled analysis. In addition, users in a particular domain can share personality characteristics and learning preferences, both inherent and institutional, which implies that some common traits can be aggregated into specific user profiles which can inform superior design requirements and aid in evaluation protocols. These differences will be discussed as part of the Personal Equation of Interaction in a following self-titled section.

4.4.1. The Human Cognition Model

In earlier work (Green, T.M., Ribarsky, W., Fisher, 2008. 2009, Green and Ribarsky, 2008) we outlined an operational framework, the Human Cognition Model (HCM), whose objective was to inform customization of human-computer cognitive collaboration in mixed-initiative interactive systems. Today's information visualization applications tend to be passive; primary interface processes sit and wait for user initiation. This is not a problem if the user knows exactly where to go and what to do. But

for the large semantically-rich datasets which visualizations are increasingly called upon to capture, and the complex analytical reasoning the visualization must scaffold and support, a truly intuitive interface must be capable of initiating a variety of processes on its own.

The HCM identifies many of these tasks and the varieties of cognition the tasks. The central process identified by the HCM is Knowledge Discovery. (See Figure 16.) This was envisioned as a human and computer paired process: the interface presents information and the human user indicates interest in a specific area or point, which the computer in turn presents in a related context. If Knowledge Discovery is goal oriented, the human will, among other processes, use temporally moderated perception, semantic categorization, and elimination reasoning heuristics to search and filter through the information space. If the discovery is not goal-oriented, the user may browse, stopping the explore data items that stand out or that are associated to items of interest.

Other processes in the HCM include information search by pattern and example. Search is an interesting cognitive task, as it is deceptively simple. Some types of search are simply target identification, utilizing perceptual logic, manipulating the interface and information space through a procedural script, and using the simplest of elimination heuristics (a binary filter that asks a question: Is this the specific item I'm looking for?). Other types of search can be much more

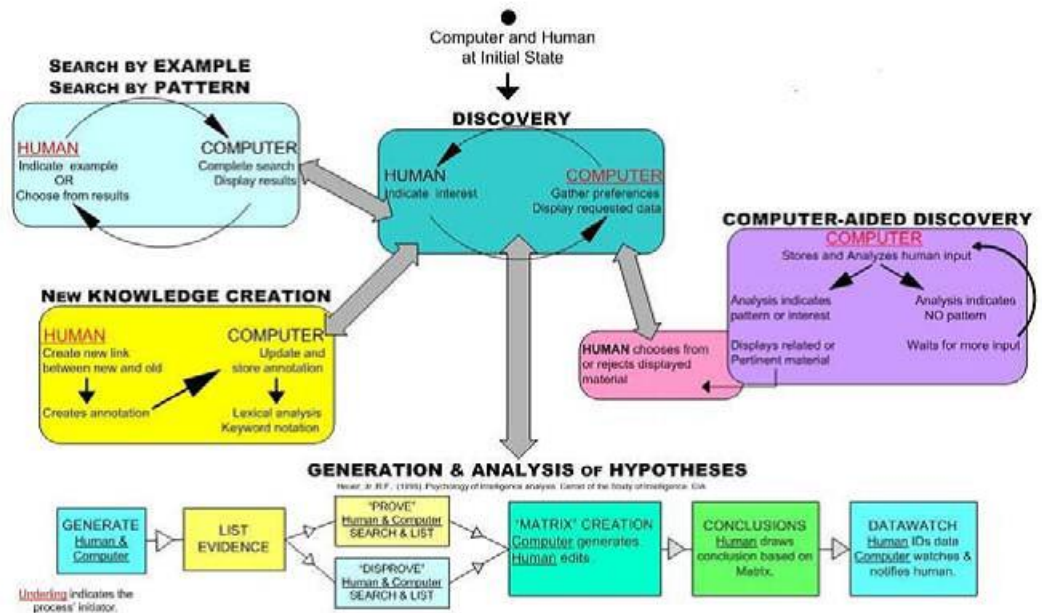


Figure 15. The human cognition model (Green & Ribarsky, 2008).

complex. When the task is to find items similar to an exemplar, for example, all the cognitive processes from the simpler search tasks serve to feed more complex processes, such as inferential and deductive reasoning, which utilize more complicated models or rules for comparison and contrast. Thus, even in the more routine of interface tasks, complex cognition cannot be ignored.

The HCM also outlines the creation and analysis of hypotheses. As discussed in previous sections, hypotheses can be created in a variety of ways, from the loose associations in abduction to the use of more demanding heuristics. Sometimes hypothesis are brought to the task by the user, and drive the interaction from the start. Wherever hypotheses are generated, they serve an important function. They drive

discovery and search behaviors, they determine how information is viewed and filtered, and they promote some data-derived conclusions over others. For these reasons, interfaces which promote the generation of valid hypotheses, either through framing, adaptive search, or more intuitive interaction, might be considered more valuable than others.

Other pertinent processes discussed in the HCM literature (e.g. Green and Ribarsky, 2008) include an interface being aware of and predicting user intent in order to keep pertinent information visible, supporting human working memory, caching data subsets of interest to introduce them in a sequence and timing that will support the flow of reasoning and ongoing discovery, and conducting analyses and providing their findings in a contextual framework, which supports a variety of hypotheses generation. In short, the HCM sketches out an interface that collaborates on cognitive processes per se, informed by a growing understanding of human preferences, abilities and limitations.

4.4.2. The Personal Equation of Interaction (PEI)

Humans are cognitive individuals. As we have seen, a human's individuality influences cognitive performance. These differences, as discussed in Section 3, shape the way we approach and perform cognitive tasks. We have discussed personality and self-beliefs in this chapter for sake of brevity, but we are also aware that humans also exhibit differences in psychophysical characteristics, such as perceptual categorization, focused attention, and haptic preferences. These individual variations interact with each other and the task to which they are applied in manners not yet understood.

Further, as we have seen, there is great variety and complexity to analytical tasks, and so it makes sense that not all cognitive tasks would be equally impacted by the same inherent differences. For example, in our research, we have found that persons who tend to believe in that good things that happen to them are due to "luck" (an external locus of control) are predictably slower in target identification (Green, Jeong, Fisher, 2010). But the same cannot be said for more cognitively complex problems, such as comparing and contrasting multi-dimensional glyphs; for these tasks, believing in luck seems to give users a decided advantage. (See the section on Research in Personal

Equation of Interaction.) It makes sense, then, that not all cognitive tasks would be equally impacted by the same inherent differences; some reasoning tasks may be better predicted by one collection of inherent traits over others.

The Personal Equation of Interaction Defined

Our goal of parameterizing a general model in order to predict performance of a particular individual builds upon foundational work in human perception conducted in the early 19th century by Friedrich Bessel (1979). Bessel recognized that the variability of reports of astronomical observations could be separated into the differences between the average ratings of made by each individual observer (in statistical terms, the between-subject variance around the global mean) and variation within observations made by a given observer (the within-subject variance around that individual's mean for a given task and situation).

While within-subject variability could not be easily factored, the deviation from the overall mean judgment for a given individual was relatively consistent over a range of similar situations. The error for a given individual could be measured to generate a "personal equation" for that individual. This could be used to factor out their characteristic error to bring their data into agreement, more or less, with the global mean. This meant that data from fewer observers were needed in order to achieve the same level of precision. In addition, one could predict any given observer's raw measurement to a reasonable degree of accuracy given the global mean and their personal equation.

Much of the research (Po, Fisher and Booth, 2003, Fisher, 2009, Green, Fisher, Jeong, 2010) in our laboratory has been devoted to defining a modern of the personal equation, the "personal equation of interaction". The PEI uses quantifiable of human perceptual, motor, and cognitive limitations during tasks and perceptual stimuli that are generated by modern visual information systems. In Po et al. al., we demonstrated that system variables such as cursor visibility and display lag interact with individual differences to produce characteristic patterns of behavior for subpopulations of individuals (Po, Fisher and Booth, 2003). These effects are unlikely to be observed by perceptual testing that does not focus on the particular of active visual information

displays and how they differ from those we experience in the physical world that have informed both human evolution and visual experience of individuals.

There are three goals in current efforts in the PEI: first, to predict how a given individual will perform on a given task and information display; second, to build a framework for interface design that to support customization of a visual information system along dimensions that are psychologically valid (i.e. that would track aspects of individual differences in such a way that their accurate fit to an individual's capabilities would measurably improve their performance with the system); and lastly, to build increasingly accurate and comprehensive estimates of personal equations and methods for assessing them. This includes both persistent differences between individuals (e.g. color blindness) and short-term shifts in capabilities (e.g. their performance stress). The latter could potentially be generated "on the fly" by an attentive system and applied as conditions changed to maintain optimal performance over changes in the capabilities of the human operator.

Our approach builds on existing psychometrics methods and materials. However the goal of this particular line of inquiry is to build towards a natural science of human-information interaction. By focusing on the specific kinds of changes in perceptual and interactive experience that are generated by modern visual information systems, we can address how changes in the and statistical regularities of information displays interact with the human visual system in general, and that of an individual observer in particular. For example, many studies in perception (e.g. Marr, 1982) show how our ability to parse complex visual scenes given limited perceptual information (the so-called "poverty of the stimulus") is supported by our internal expectations, which in turn are built through a lifetime of sampling the statistical regularities of our physical environment.

Phenomena such as change blindness (Grimes, 1966), Rensink, O'Regan and Clark, 1997) demonstrate the adaptation of human vision to an environment where abrupt changes are rare. Our visual system is not optimized to detect changes, but pays a small price for this in the physical world. Active updating of visual displays increases the frequency of abrupt display events, and the probability increases that one will coincide with the observer's saccadic eye movement and so escape detection. Thus we

find that the study of human performance with information systems cannot simply rely on applying perceptual and cognitive science research taken from the literature.

It is necessary to actively seek answers to questions that arise from the use of information systems in cognitive task performance, and to do so using methods from the natural sciences. It is an open question to what extent aspects of the PEI are inherent and what aspects are acquired through experience with complex visualizations. Some factors such as low vision, color and stereo blindness etc. clearly fall into the former category. Estimation of these factors might require psychophysical testing of a given individual, but may also be estimated by covariates of the given conditions, whether or not we can establish a causal link between them. To the extent that these covariates are predictive, they can be used to support the design and customization of information systems now, as well as contributing to ongoing research about human.

The first clustering factor for individual differences is through cognitive experience, e.g. in a given institution or profession. Members of a professional or skilled cohort tend to share jargon and conceptual understanding. Additionally, they are often practiced in methodologies and task heuristics that are specific to the group. These methodologies become a part of the way in the user searches for and uses information, and can impact the conclusions drawn. They can also introduce group-specific biases that might not be found in the general user population. For these reasons among others, understanding an institutional user profile is important to the design of an expert system interface. A second factor might be perceptual and perceptuomotor through interaction with some specific environment.

We define the Personal Equation of Interaction as a compilation of predictive measures based upon inherent individual differences, including but not limited to personality; each measure will be validated to predict performance for one type of cognitive task integral to analysis. The PEI has three current and future end goals: the prediction of analytical performance based on differences (Green, Jeong and Fisher, 2010, Green, Fisher and Jeong, 2010, the ability to inform real-time interface individuation, and the creation of fuller-bodied user profiles, broken down by the reasoning tasks performed. The first goal – that of performance prediction – is being

undertaken through a series of human research studies. Participants are asked to complete a series of tasks similar to common interface tasks, such as interface learnability, target identification, categorization, etc. The measured outcomes vary by task, and include completion times, errors, self-reported insights and free response. In addition to the performance and qualitative feedback, participants are asked to complete hundreds of items from a battery of psychometric measures we have chosen for their inter-relatedness and their relationships to learning outcomes in the behavioral literature. Post-study analysis includes evaluating the trending of the psychometric items with measured outcomes. In addition, follow-up testing such as factor analysis is done to isolate highly predictive items or groups of items, such as in (Green, Jeong, Fisher, 2010) for a particular type of reasoning task.

This allows us to delineate our findings by cognitive process and compare multiple interfaces or problem-solving environments in hopefully a more even fashion. Currently, we can predict on simple outcomes like completion times in a variety of common interface tasks; these findings are being replicated and expanded. Not surprisingly, this research is currently quite fluid, and continues to inform the second goal of what matrices will be needed to support real-time interface adaptation. In addition, having hundreds of participants complete these studies has allowed us to sketch out initial user profiles, or describe inherent characteristics of a user based on how the user performs on an analytical task (Green, Fisher and Jeong, 2010).

The goal of the PEI is not, at least in the short term, to tell designers and developers specifically which visualization techniques to use and which to avoid generically, but rather to give interface creators a robust understanding of what the individual analyst needs in order to optimally perform the analytical tasks that must be performed. The Personal Equation of Interaction does not replace interface design; it augments design by making designers aware of user strengths and weaknesses. It cannot replace user studies for a particular interface, but it provides new metrics with which to evaluate study outcomes. And as a research program, it amplifies visual analytics as the study of analytical reasoning supported by interactive visual interfaces by adding to the body of understanding on analytical reasoning and analytical reasoners.

These are aggregated to build an ever more comprehensive and accurate personal equation of interaction that could be used by an application to parametrically modify its display of in such a way as to optimize the cognitive performance of an individual decision-maker. As research we hope to find a way to integrate system models with models of human interaction to better predict the course of fluent human-information interaction.

Research in the Personal Equation of Interaction

Our research has demonstrated that inherent differences can and do influence learning and analytical performance during interface interaction. In combination with the environmental and institutional variations, there is evidence that the impact of inherent differences could be used to derive a personal equation of interaction.

Our recent research has demonstrated that personality factors can predict efficiency during varying task types (Green, Jeong, and Fisher, 2010). We designed a series of tasks we asked participants to complete in two visual interfaces using the same dataset: menu driven web application, and an information visualization using direct interaction on hierarchical graphs. These tasks were designed to test two very different types of learning: procedural and inferential. Procedural learning, as defined for this study, was the ability to use the interface to find target information. This type of learning tends to be inductive: a rule is inferred which generalizes to other similar target identification tasks in the interface. Other the other hand, the inferential learning tasks were highly deductive. Participants were asked to evaluate a multi-dimensional exemplar and find another conceptual object in the hierarchy that was similar to (or different from) the exemplar for the specified dimensions. This type of reasoning involves the creation of a mental model, which is then used to evaluate complex concepts to reach a valid conclusion. For each task, we tracked errors, completions, as well as qualitative feedback.

In addition, we administered several longstanding and well-documented psychometric measures to participants (Green, Jeong and Fisher, 2010). These measures were created to measure personality traits such as (a tendency toward emotional instability), extraversion (a tendency toward sociability or seeking the

company of others), and trait anxiety, which is a tendency to be more anxious generally, regardless of the environment. Trait anxiety differs from state anxiety, which is the tendency to be anxious when in a situation that triggers anxiety.

Item (Originating Measure)	How scored	PCA Component
Unable to Relax (Beck's Anxiety Inventory)	0(never) to 3(severely)	1
Fear the Worst (Beck's Anxiety Inventory)	0(never) to 3(severely)	
Heart Pounding (Beck's Anxiety Inventory)	0(never) to 3(severely)	
What we are used to is always preferable to what is unfamiliar. (Tolerance of Ambiguity)	1(strongly disagree to 7(strongly agree)	2
Talk to a lot of people. (Extraversion)	1(low) to 5(high)	
Hands Trembling (Beck's Anxiety Scale)	0(never) to 3(severely)	3
Numbness (Beck's Anxiety Inventory)	0(never) to 3(severely)	
Don't talk a lot. (Extraversion)	1(low) to 5(high)	
Am easily disturbed. (Neuroticism)	1(low) to 5(high)	

Figure 16. Items in the 9-item short measure. Adapted from Green, Jeong and Fisher, 2010.

Another personality trait that proved to have a demonstrable impact was locus of control, which is a measure of how in control a person feels he or she is over the events in life. Persons with an external locus tend to believe strongly that they are not in control, and attribute events to factors outside themselves, such as luck, other people, or circumstances outside of their control. On the other hand, persons with an internal locus tend to believe that they are responsible for both positive and negative events in their lives. They are more likely to attribute events to some behavior or attitude of their own than to outside influences, and tend to give very little credence to luck.

Other measures designed to test other personality factors, such as a discomfort with problem-solving situations where important factors are unknown (an intolerance of ambiguity) or self-regulation, which is the ability to hold it together emotionally when the

problem or situation difficult, were also evaluated but were not found to be particularly predictive in performance of the tasks under study.

Results of Study 1. Results demonstrated (Green, Jeong and Fisher, 2010) that whole score locus of control predicted inferential task efficiency in the data visualization; participants with a more external locus were able to complete the deductive inferential tasks more quickly.

For the inductive procedural tasks, no singular full measure could predict behavior, which was not unexpected, given that none of these psychometrics were designed to evaluate these traits in this interface environments. But factor analysis uncovered 9 items, largely from the trait measure, that predicted target identification efficiency across both interfaces (Figure 17). Participants who were more trait anxious found target items more quickly, even when the target was buried several layers down in the hierarchy.

Results indicated that no single personality factor measure could predict errors in either interface.

Results of Study 2. Results of a similar study using procedural tasks (Green, Fisher and Jeong, 2010), currently under submission) expanded these findings somewhat. This study used the same interfaces and similar procedural tasks. The scores and outcomes of participants in both studies were combined for greater statistical power (N = 105). Results demonstrated that both neuroticism and extraversion predicted efficiency; the more neurotic/extraverted participants found items more quickly.

Additionally, analysis of the combined set found that locus of control predicted procedural performance, in directly the opposite way to that of the inferential tasks. Participants with an internal locus (a belief that they were in control of life events) found targets more quickly than those with an external locus. This evidence alone demonstrates that not only that personality factors affect interface interaction performance, but that different tasks are impacted differently by inherent individual differences. See Figure 18.

	Completion times	Errors	Insights
Interface	faster times in MapViewer	fewer errors in MapViewer	more insights in Gvis
Locus of Control	internal locus faster times	none	external locus more insights
Extraversion	more extraverted faster times	none	less extraverted more insights
Neuroticism	more neurotic faster times	none	less neurotic more insights

Figure 17. Summary of Findings from Study 2.

Discussion of Results. The existence of significant trending between personality factors and interface interaction outcomes is interesting for a variety of reasons. First, it demonstrates that even complex cognition can, at least to some degree, be predicted. Secondly, it demonstrates that inherent individual differences, over which we as designers have no control, could inform design if we knew the psychometric makeup of our target user group. This holds potential for experts systems, which are designed for users whose differences are likely to trend in similar ways. Thirdly, these studies open a promising doorway; if these few personality factors can predict performance, what else about complex cognition might we be able to predict if we knew more about our users, as well as about the expert cohorts for whom we design visually enabled interfaces?

4.5. Conclusion

The reasoning used during task analysis is complex. In this chapter, we have discussed this complexity by highlighting a handful of reasoning heuristics. We have underscored this complexity with a discussion of Pirolli and Card's sensemaking loop. And we have explored how this complexity complicates the current state of design and evaluation thanks to the absence of applicable reasoning research and pertinent precedent in the behavioural literature (.

We have also broadly discussed the impact of human individuality on every

primary cognitive process, and surveyed our current research in pursuit the generation of new system development models that optimize the cognitive performance of human decision-makers. Optimization in this context must include complex criteria such as insight, innovation, creativity and awareness in uncommon, unique and novel problems and situations. Research has shown that inherent individual differences between users impacts the task and learning performance in visually embedded interfaces. Our previous work in the development of the Human Cognition Model continues to inform our research direction. Our ongoing research in the Personal Equation has highlighted the need to study not only inherent differences in personality factors, but also other user differences, including those in which affect other inherent individualities as well as differences in institutional cohort and environment. These individual differences in human capabilities are great enough that any unitary system will be at best a compromise between the needs of the various sub-populations of users. Our ongoing research seeks to take advantage of human individuality, rather than to ignore it. While still in the early stages of this research, we have already highlighted several inherent differences which predict performance, depending on reasoning task. We intend to explore further, with the expectation of isolating and understanding influential individual differences and how they impact interface interaction, which could benefit visual analytics interface designers by informing design requirements and opening up new areas for innovation.

4.6. References

- Beth, E. W., & Piaget, J. (2013). *Mathematical epistemology and psychology* (Vol. 12). Springer Science & Business Media.
- Brooks, G. P., & Brooks, R. C. (1979). The improbable progenitor. *Journal of the Royal Astronomical Society of Canada*, 73, 9.
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- Cherubini, P., & Johnson-Laird, P. N. (2004). Does everyone love everyone? The psychology of iterative reasoning. *Thinking & reasoning*, 10(1), 31-53.
- Cherubini, P., & Mazzocco, A. (2004). From models to rules: Mechanization of reasoning as a way to cope with cognitive overloading in combinatorial problems. *Acta Psychologica*, 116(3), 223-243.

- Corn, A. L. (1983). Visual function: A theoretical model for individuals with low vision. *Journal of Visual Impairment & Blindness*.
- Dou, W., Jeong, D. H., Stukes, F., Ribarsky, W., Lipford, H. R., & Chang, R. (2009). Recovering reasoning processes from user interactions. *IEEE Computer Graphics and Applications*, (3), 52-61.
- Fisher, B. (2009). Science and smart graphics. *Information Technology*, 3, 142-148.
- Gentzen, G. (1969). Untersuchungen über das logische Schließen. *Mathematische Zeitschrift* 39, 176–210, 405–431 (1935); English Translation: Investigations into logical deduction. The collected Papers of Gerhard Gentzen.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.
- Green, T. M., & Fisher, B. (2010, October). Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on* (pp. 203-210). IEEE.
- Green, T. M., & Fisher, B. (2010, October). Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on* (pp. 203-210). IEEE.
- Green, T. M., & Ribarsky, W. (2008, April). Using a human cognition model in the creation of collaborative knowledge visualizations. In *SPIE Defense and Security Symposium* (pp. 69830C-69830C). International Society for Optics and Photonics.
- Green, T. M., Jeong, D. H., & Fisher, B. (2010, January). Using personality factors to predict interface learning performance. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1-10). IEEE.
- Green, T. M., Ribarsky, W., & Fisher, B. (2008, October). Visual analytics for complex concepts using a human cognition model. In *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on* (pp. 91-98). IEEE.
- Green, T. M., Ribarsky, W., & Fisher, B. (2009). Building and applying a human cognition model for visual analytics. *Information visualization*, 8(1), 1-13.
- Grimes, J. (1996). On the failure to detect changes in scenes across saccades.
- Heaven, P. C., & Quintin, D. S. (2003). Personality factors predict racial prejudice. *Personality and Individual Differences*, 34(4), 625-634.
- Heer, J., & Agrawala, M. (2008). Design considerations for collaborative visual analytics. *Information visualization*, 7(1), 49-62.
- Jaśkowski, S. (1934). On the rules of suppositions in formal logic. Nakładem Seminarjum Filozoficznego Wydziału Matematyczno-Przyrodniczego Uniwersytetu Warszawskiego.
- Jeong, D. H., Green, T. M., Ribarsky, W., & Chang, R. (2010). Comparative Evaluation on Two Interface Tools in Performing Visual Analytics Tasks. In *Proceedings of BELIV Workshop, CHI*.

- Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual review of psychology*, 50(1), 109-135.
- Johnson-Laird, P. N. (2006). *How we reason*. Oxford University Press.
- Johnson-Laird, P. N., & Byrne, R. M. (1991). *Deduction*. Lawrence Erlbaum Associates, Inc.
- Kozielecki, J. (1971). Elements of a psychological decision theory. *Studia Psychologica*, 13(1), 53-60.
- Mac Aoidh, E., Bertolotto, M., & Wilson, D. C. (2008). Understanding geospatial interests by visualizing map interaction behavior. *Information Visualization*, 7(3-4), 275-286.
- Macia, A. (1979, April). Visual perception of landscape: sex and personality differences. In *Proceedings of Our National Landscape Conference*. USDA Forest Service, General Technical Report PSW-35 (pp. 279-295).
- Marr, D. (1982). *Vision: A computational approach*.
- North, C. (2006). Toward measuring visualization insight. *Computer Graphics and Applications*, IEEE, 26(3), 6-9.
- Palmer, J. (1991). Scientists and information: II. Personal factors in information behaviour. *Journal of documentation*, 47(3), 254-275.
- Pirolli, P., & Card, S. (2005, May). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis* (Vol. 5, pp. 2-4).
- Plaisant, C., Grosjean, J., & Bederson, B. B. (2002). Spacetree: Supporting exploration in large node link tree, design evolution and empirical evaluation. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on* (pp. 57-64). IEEE.
- Po, B. A., Fisher, B. D., & Booth, K. S. (2003, January). Pointing and visual feedback for spatial interaction in large-screen display environments. In *Smart Graphics* (pp. 22-38). Springer Berlin Heidelberg.
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological science*, 8(5), 368-373.
- Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of verbal learning and verbal behavior*, 14(6), 665-681.
- Robinson, A. C. (2008, October). Collaborative synthesis of visual analytic results. In *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on* (pp. 67-74). IEEE.
- Rock, I. (1983). *The logic of perception*.
- Ronan, G. F., Senn, J., Date, A., Maurer, L., House, K., Carroll, J., & VanHorn, R. (1996). Personal problem-solving scoring of TAT responses: Known-groups validation. *Journal of Personality Assessment*, 67(3), 641-653.

- Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993, May). The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 269-276). ACM.
- Saraiya, P., North, C., & Duca, K. (2004, October). An evaluation of microarray visualization tools for biological insight. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on* (pp. 1-8). IEEE.
- Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 11(4), 443-456.
- Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 11(4), 443-456.
- Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on* (pp. 336-343). IEEE.
- Spence, R. (2001). *Information visualization* (Vol. 1). New York: Addison-Wesley.
- Springmeyer, R. R., Blattner, M. M., & Max, N. L. (1992, October). A characterization of the scientific data analysis process. In *Proceedings of the 3rd conference on Visualization'92* (pp. 235-242). IEEE Computer Society Press.
- Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and brain sciences*, 23(5), 645-665.
- Stasko, J., Görg, C., & Liu, Z. (2008). Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization*, 7(2), 118-132.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Lawrence Erlbaum.
- Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112(1), 80.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.
- Wang, T. D., Plaisant, C., Quinn, A. J., Stanchak, R., Murphy, S., & Shneiderman, B. (2008, April). Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 457-466). ACM.

Chapter 5. Impact of personality factors on interface interaction and the development of user profiles: Next steps in the personal equation of interaction

5.1. Overview

This chapter was originally published as Green, T. M., & Fisher, B. (2012). Impact of personality factors on interface interaction and the development of user profiles: Next steps in the personal equation of interaction. *Information Visualization*, 11(3), 205-221. This chapter builds on previous chapters by adding a learning profile to what is already known about the Personal Equation of Interaction. We used the participants from Chapter 3, and used the correlations between participants' whole measure scores. In other words, we administered every item from previously normed psychometric assessments, and looked at the statistical relationships between the assessment scores to build an learning profile of highly efficient procedural learners. The profile described these learners as being more verbal than visual, moody, reflective, and uncomfortable with new environments or tasks. As this reported research built on previous work and the profile was built to be descriptive, not test a hypothesis, there is no H_0 or H_a for this chapter.

5.2. Abstract

These current comparative studies explore the impact of individual differences personality factors on interface interaction and learning performance behaviors in both an interactive visualization and a menu-driven web table in two studies. Participants were administered three psychometric measures designed to assess Locus of Control, Big Five Extraversion, and Big Five Neuroticism. Participants were then asked to complete procedural learning tasks in each interface. Results demonstrated that all three measures predicted completion times. Additionally, analyses demonstrated that personality factors also predicted the number of insights participants reported while completing the tasks in each interface. Furthermore, we used the psychometric findings in conjunction with a follow-up psychometric survey with a further 50 participants to build

initial user profiles based on the cognitive task being undertaken. We discuss how these findings advance our ongoing research in the Personal Equation of Interaction.

5.3. Introduction to comparative studies

The primary purpose of visual analytics is commonly defined as the facilitation of analytical reasoning through the use of interactive visual interfaces (Wong and Thomas, 2004). Facilitating analytical reasoning, however, requires a comprehensive and operational understanding of the cognitive processes that make up analytical reasoning. Complex cognition includes a plethora of smaller processes that work together, including perceptual cognition, categorization, problem-solving, decision-making, judgment, and reasoning. These processes feed and inform each other throughout each stage (target identification, mental modeling, rule inferencing, etc.) of an analytical task. Thus, simply supporting each process individually is not enough. Visual analytics must also support the temporal and cognitive flow of reasoning. And yet, an operational understanding of analytical cognition has, to date, proven elusive. For example, as is often the case with behavioral experimentation generally, studies of cognition tend to involve small, simple, normative or “toy world” tasks, whereas interaction in the real world tends to be more complex, harder to predict, and thus harder to measure.

Additionally, these evaluations focus on what are often considered less complex cognitive processes. Especially in visualization studies, the cognitive variables measured are usually facets of vision, given attention, and tactile manipulation. While visual and motor effectiveness are important to interface interaction, they are only part of the story.

Another point of interest is the generally accepted superiority of visualizations. Research to date has largely focused on the visualization and insight generation, but not necessarily on the tasks that support insight generation, or the degree to which user individuality impacts the frequency of insight. Insight, as defined for the purposes of this study, is knowledge gained from the use of the interface. This knowledge can be about the content, i.e. about genomes and their ontological relationships. Or it can be about

the interface, and how best to navigate it. We asked participants after each task to tell us if they learned anything novel or unexpected while completing the task. We chose to use words such as “novel” in the directions to encourage the participant to focus on knowledge they had gained uniquely from the performance of the task.

This insight differs from spontaneous insight, or the so-called “a-ha!” moment, which is a seemingly spontaneous solution to what is considered to be an unsolvable problem. In this study, we evaluate the insight generation by comparing the number of reported insights in the two interfaces while completing two types of procedural task: script learning, which involves the use of sequential instructions and interface learnability; and target identification, which can involve hunting for information through several layers of hierarchical organization. In addition, we explore the impact that individual differences have on the number of insights generated in both interfaces across the task.

Complex cognition is not necessarily sequential. That is to say, human thinking is not something that is easily turned on or off, nor is it appropriate to define the hard thinking that humans can do as simply a series of step-by-step processes. Reasoning, in particular, can be complicated. It uses a variety of heuristics, from quick elimination heuristics like Gigerenzer's Take-the-Best (Gigerenzer and Goldstein, 1996) or satisficing (Simon, 1991) to much more complicated processes such as iterative reasoning, deductive analysis, or rule inferencing. Which heuristics are used and in what order depend on the task, the environment, and the user. These heuristics are often used combinatorially, feeding and informing the analysis until a solution or hypothesis has been satisfactorily reached (Stanovich, 1999). Unfortunately, at this time, analytical reasoning behaviors can be described in part and in whole, but cannot necessarily be predicted. There are no unifying theories of reasoning. And three types of user individual differences compound this difficulty of prediction: institutional, environmental, and intrapersonal.

How humans work through reasoning tasks is impacted by institutional differences. Cognition is a social activity (Kaptelinin and Nardi, 2006), and domain-specific knowledge, jargon, learned methodologies, and other cultural factors can

influence how analysis tasks are approached and what heuristics are used in solving them. In addition, these domain or expert cultures tend to have similar intrapersonal differences; members of an expert cohort may share personality leanings or learned proclivities (Boyatzis and Kolb, 1995, Heuer, 1999). Environmental differences – such as differences in the interface or tool used during visually enabled interaction – frame the task and can help or hinder the reasoning process. These differences are naturally of particular interest to visual analytics design, as effective interfaces can facilitate analytical reasoning.

In this paper, we will highlight the impact of intrapersonal individual differences. Individual differences of whatever variety are obviously not the only factors which demonstrably impact user interactive performance. But, as we will show, individual differences – and intrapersonal differences in particular – can predict certain types of performance. Intrapersonal differences interact with environmental and institutional differences to influence human behavior. A complete Personal Equation of Interaction will take all three types of difference into consideration. But, for now, we will start with intrapersonal differences, which, in their relationship to interaction and analytical behaviors, are the least well understood. Further, these differences seem to influence performance differently, depending on the cognitive task being undertaken.

Another reason to study intrapersonal differences is that they, unlike environmental and to some degree institutional differences, are variables over which interface designers have no control. Learning is not generic. Learning heuristics and processes vary depending on human individuality, the learning environment, and the learning tasks. In other work, we discussed the impact of locus of control (LOC) on inference learning in the form of category reasoning (Green, Jeong and Fisher, 2010). The tasks used in these current studies are procedural. Procedural learning, broadly defined, is the “knowing how” of any sequential task. It is sometimes called skill learning, as it is the learning most common to motor and iterative tasks that require repetition to master (Sun, Merrill and Peterson, 2001); it is also referred to as script learning, which captures the idea that there is a “recipe” or “roadmap” to be followed. Procedural learning is thought to be either top down (i.e. CLARION (Sun, Merrill and Peterson, 2001)) or, more commonly, to be bottom up, first assimilating the necessary declarative

facts and then the use of that information into the deconstruction of the task procedure (Anderson, 1982). Procedural learning, due in part to repetition, can become “automatic,” requiring little conscious focus. For the purposes of these current studies, procedural learning is the ability to learn to manipulate an interface well enough to find and identify target information, or to answer straightforward questions about the target information.

Tasks such as these are not as complex as real-world tasks in and of themselves. But search-and-find behaviors interface learning, and other procedural tasks are the backbone of interaction in any interface. These simple but powerful behaviors are used over and over at every level of the visually-enabled analytical process.

The task protocols used in these studies are not meant to be indicative of real-world tasks so much as they are intended to be an early test area from which to explore the degree to which personality factors impact human reasoning behaviors. Future research will build on these tasks in protocols that will attempt to more closely replicate real-world tasks.

Procedural or script learning is integral to interface interaction at every level. Some research has been conducted with an eye toward procedural or target-finding tasks. But, as Plaisant has outlined (2004) many of these studies are tool evaluations of specific interfaces, and are designed to designate one interface as ““better”” than another, or done without an understanding of the learning which underlies task performance.

Individual differences in reasoning ability have been found to impact procedural learning in non-interface task environments (Hall et al.,1988) These current studies evaluate intrapersonal differences in computer-mediated procedural tasks.

In our research toward the Personal Equation of Interaction, our goal is to know and understand the impact of these variables, as well as to develop a battery of predictive measures to aid in the development of interfaces which cater to the individuality of the user or user domain. The creation of the Personal Equation of

Interaction at this current time is focused on intrapersonal individual differences. Intrapersonal differences are those of learning style, personality factors, self-beliefs, and other cognitive “pre-sets” which the user brings to the interface. We will demonstrate that these intrapersonal differences can and do demonstrably impact interaction outcomes. Further, we can show that, if the intrapersonal differences are known, interaction performance can be predicted, and so could, if part of a robust user profile, be used to develop design requirements for expert systems design as well as real-time interface individuation.

Intrapersonal individual differences in problem-solving approaches can affect task orientation and motivation when a user is engaged in goal-oriented behaviors. In particular, personality factors similar to the ones evaluated in the studies reported here have been shown to impact cognition and cognitive performance in other learning environments. For example, personality factors predicted preferences in visual perception of landscapes (Macia, 1979).

NCBI Home GenBank BLAST

Map Viewer Home Help

The Map Viewer provides a wide variety of genome mapping and sequencing data. [More...](#)

Search

Search:

for:

Tools Legend

- Search or Browse the Genome
- BLAST
- Clone Finder
- Go to region on a chromosome
- Genome Resources page

News

Human build 37 released Aug 4

An update to the human genome assembly and annotation is now... [more](#)

[Show all](#)

Related Resources

- [NCBI Home](#)
- [NCBI Web Search](#)
- [NCBI Site map](#)
- [Genome Browser agreement](#)
- [Genome Biology](#)
- [Taxonomy](#)
- [Entrez \(Global Query\)](#)
- [BLAST](#)
- [Map Viewer FTP](#)

Small Genomes

- Bacteria
- Organelles
- Viruses

Vertebrates (17)

Mammals (14)

Primates (3)

Scientific name	Common name	Build	Tools
<i>Homo sapiens</i>	human	Build 37.1 Build 36.3	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Macaca mulatta</i>	rhesus macaque	Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Pan troglodytes</i>	chimpanzee	Build 2.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Rodents (2)

Scientific name	Common name	Build	Tools
<i>Mus musculus</i>	laboratory mouse	Build 37.1 Build 36.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Rattus norvegicus</i>	rat	RGSC v3.4	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Monotremes (1)

Marsupials (1)

Other Mammals (7)

Other Vertebrates (3)

Invertebrates (12)

Protozoa (18)

Plants (46)

Fungi (17)

Scientific name	Common name	Build	Tools
<i>Aspergillus clavatus</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Aspergillus fumigatus</i>		Build 2.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Aspergillus niger</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Candida glabrata</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Cryptococcus neoformans</i>		Build 2.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Debaryomyces hansenii</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Encephalitozoon cuniculi</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Eremothecium gossypii</i>		Build 3.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Gibberella zeae</i>		Build 1.2	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Kluyveromyces lactis</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Magnaporthe oryzae</i>	rice blast fungus	Build 3.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Neurospora crassa</i>		Build 2.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Pichia stipitis</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Saccharomyces cerevisiae</i>	baker's yeast	Build 2.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Schizosaccharomyces pombe</i>	fission yeast	Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Ustilago maydis</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<i>Yarrowia lipolytica</i>		Build 1.1	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 18: The NCBI Viewer

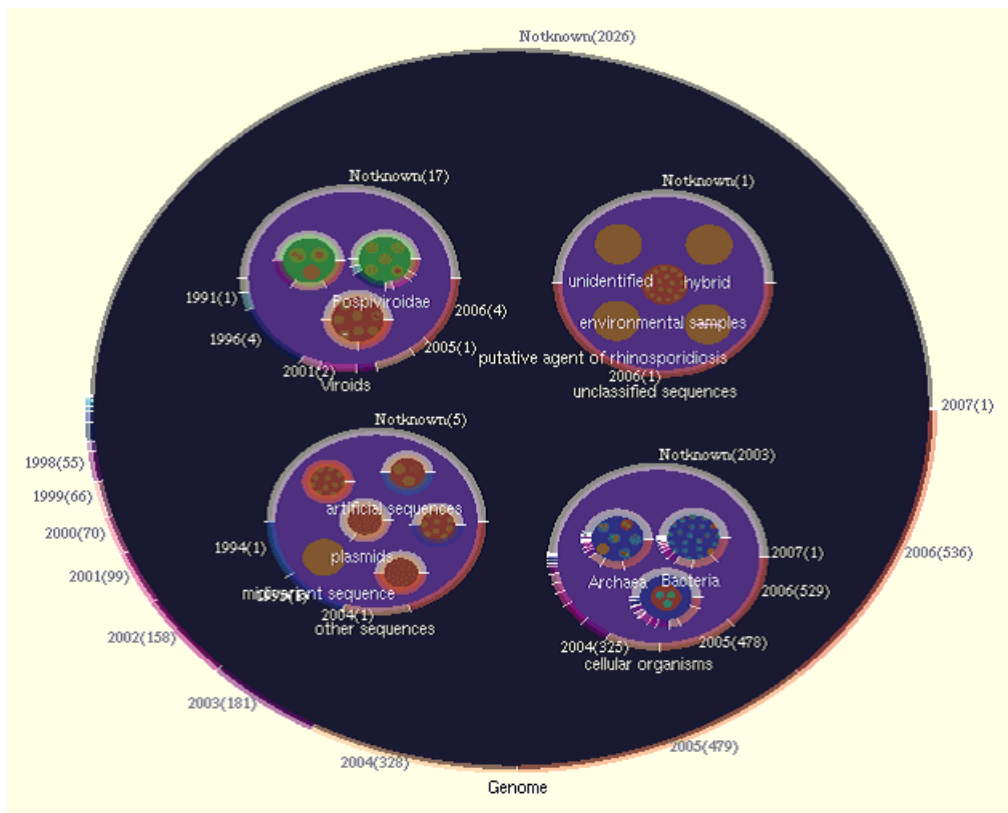


Figure 19: The main view of GVIs.

There are not many examples, but procedural learning similar to the tasks in this study have also been found to be impacted by personality factors. For example, in a human–computer interaction study, Palmer found that interactive behaviors in an information search could be categorized by personality factors (1991). In reasoning research, individual differences have been found to impact rationality and metareasoning (Stanovich, 1999). These are just a few examples in a broad literature of how personality factors and other individual differences demonstrably affect complex cognition. The findings we currently report are part of this body of work.

The current studies were designed to explore two broad research questions. The first question was whether and to what degree LOC, Big Five Neuroticism, and Big Five Extraversion would have a significant relationship with the outcome variables in task

performance. It was hypothesized that some whole measures would trend with the outcomes.

Based on previous work (Green, Jeong, and Fisher, 2010) we expected that the LOC whole score would be one predictor, and that more extraverted and neurotic participants would be quicker in task completion. And based on behavioral literature, 16 we hypothesize that participants with an external locus would be quicker in identifying target information.

The second question was whether and to what degree LOC, Big Five Neuroticism, and Big Five Extraversion would have a significant relationship with the number of insights reported; it was hypothesized that, given the interrelationship between these constructs, it would be found to predict insight generation in both interfaces. Based on previous LOC literature (e.g. Weiss and Sherman, 1973), we predicted that participants with an internal locus might be more apt to self-report more insights.

5.4. Comparative studies

Two studies were conducted. Each study employed a within-participants design, and compared procedural learning behaviors in an information visualization and a web table. Study 1 tested procedural learning performance with a series of five questions in each interface.

Study 2 tested procedural learning performance, with a total of six questions in each interface (three training and three task). The procedural task completion times in both studies were combined for the purpose of analysis. The design and findings of Study 2 have also been reported and discussed (Green, Jeong and Fisher, 2010)

5.4.1. Interfaces

Both studies asked participants to interact with two interfaces built to display genomic information. These interfaces were chosen as artifacts because both interfaces

were fed by the same underlying dataset (GenBank), both interfaces supported the types of tasks we wanted to study, and the presentation and organization of data and interaction methodology was demonstrably different. One interface is the web-based

National Center for Biotechnology Information (NCBI) MapViewer for genomic information, which is publically available and can currently be found at <http://www.ncbi.nlm.nih.gov/mapview>. MapViewer is a multiple-row-based hierarchical representation, and uses standard graphical user interface (GUI) manipulation, such as menus and hyperlinks. (See Figure 19).

The other interface is an interactive data visualization (GVIs) of genomic relationships (which is not available publically (see Figure 20)). The primary purpose of GVIs is to represent relevant relationships (Hong, Jeong and Shaw, 2005) such as mapped genomes or the phylogenetic organization) between two organisms. Users manipulate the interface through direct interaction, “drilling down” through each hierarchy of subcategory directly by pressing and holding down a mouseclick near the information of interest.

5.4.2. Psychometric measures

The psychometric measures we have chosen have been shown to capture the impact of these intrapersonal constructs on human cognitive performance and motivation, as discussed in the behavioral literature (as discussed briefly in the “Introduction to comparative studies” and “Comparative studies” sections). Our purpose was to explore what impact they might have on analytical performance enabled by a visual interface. Three psychometric measures were administered that were common to both studies: the LOC Inventory, as well as the Big Five neuroticism and Big Five Extraversion subscales of the International Personality Item Pool (IPIP) Mini Big Five Personality Inventory.

The Internal–External LOC Inventory¹⁹ is a 39- item forced-choice measure designed to evaluate the degree to which participants attribute life events to some action

of their own, or to some uncontrollable action outside of themselves. Lower LOC scores are associated with an “internal locus” of control, an intrapersonal belief that events and outcomes are under a person’s control, and thus, success or failure depends largely on personal behavior and attitudes. Higher scores indicate an “external locus,” an intrapersonal belief that events and outcomes are influenced by external factors, such as unforeseen circumstances, a higher power, or “good luck.” Rotter postulated that

these loci were traits remaining stable over a person’s lifetime.¹⁹ Research demonstrates that LOC has an impact on a wide variety of human outcomes, including academic and workplace performance (Hong, Jeong and Shaw, 1973, Cacioppo, Petty and Feinstein, 1996).

The Big Five Neuroticism and Big Five Extraversion subscales of the IPIP 20-item Mini Big Five Personality Inventory (Donnellan, Oswald and Baird, 2006) ask participants the degree to which each listed characteristic applies to them. The Big Five factors have a long history in psychology and decades of literature on their scope and impact. Briefly, Big Five Extraversion defines the degree to which a person is open minded, action oriented and seeks the society of others. Big Five Neuroticism is distinguished by negativity and a propensity to be moody. In previous work, as well in other literature (Judge, Erez and Bono 2002) these traits have a demonstrated relationship to each other, and, in the case of Big Five Neuroticism, to LOC.

As previously stated, we hypothesized that participants with an external locus would be quicker in identifying target information, that persons with a higher internal locus would self-report more insights, and that Big Five Extraversion and Big Five Neuroticism would also be associated with superior performance.

5.4.3. Participants

In total, 106 participants agreed to complete the study: 56 in the first study, 50 in the second study. Ninety-four participants reported being right handed; 11 were left handed. Most (101) were undergraduates and received course credit for participation. Students reported having 22 different majors or academic concentrations, including business, nursing, computer science, and psychology.

The vast majority of all participants (101, 96%) had taken fewer than four biology or biology-related classes. Novices were recruited specifically to better evaluate procedural learning with novel information; experts would have had a more advanced understanding of the knowledge ontology, which would have weakened the comparison between interface metaphors. Experts would have needed to learn the interfaces like any other user, but their institutionalized methods of exploration – the way they were taught to organize the domain-specific information – would have made this a study more about interface suitability than about procedural learning and target identification.

All participants were asked to rate their ability and comfort level with a computer and mouse on a five-item Likert-type scale. They were also asked to identify whether they had previous experience with the computer interfaces being investigated. Ninety-seven reported being comfortable or very comfortable with a computer; 79 reported having ““very good”” or ““expert”” computer ability. No one reported a computer comfort or ability level less than a three or ““OK.”” Almost all (104) participants had used a web-based application before. Thirty-five participants reported having used a data visualization previously. None of the participants reported having a medical condition that might interfere with their use of a computer or mouse. Two participants reported being color blind.

5.4.4. Study Protocols

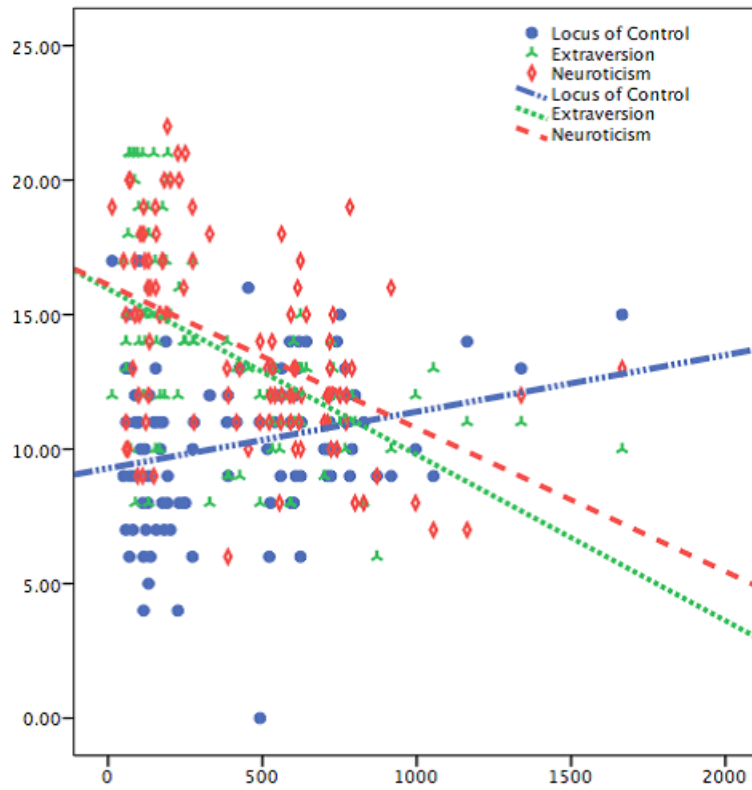
After signing the consent form, participants were asked to fill out an online self-report questionnaire that included the three psychometric measures and basic demographic information, with particular emphasis on self-perceived ability, experience and comfort with computers and computer interfaces. Participants in the first study were allowed to complete the questionnaire online before their session in the lab. All data were collected for post-hoc analysis with task performance data.

In both studies, after completion of the self-report measures, participants began the procedural learning tasks in one of the two interfaces. The order of interface was counterbalanced for order effects; half of the participant used GVis first, and half used MapViewer first.

In the first study, the tasks started with a brief demonstration of interface and interaction techniques, such as the use of hyperlinks or how to zoom into the visualization. After the demonstration, a short tutorial was administered to introduce participants to essential tools and concepts in the interface, and to allow participants to experiment with what was being learned. In some cases, step-by-step instructions were given when the user requested them. A researcher was on hand throughout the study to answer any questions.

Following the tutorial was a series of three tasks designed to test procedural performance in finding target information: the participant was asked to identify a target located somewhere within the presented informational hierarchy as quickly as possible. The question provided what base categorization or subclass the information was located within, but did not provide step-by-step instructions. As soon as the target was located on screen, the participant pushed a “Found It” button on the screen. The time in seconds taken from the presentation of the question on screen to the moment the button was pushed was recorded as completion time.

In the second study, participants were asked to demonstrate script learning or tool skill by answering five hunt-and-find questions. All tasks were open response. Each question included step-by-step “cues” to assist in finding the answer to each question. A cue was the next step or concept on the current page or in the current view to look for. Participants were given little or no help from the researchers while working through the questions, but were allowed or encouraged to experiment with different interactions



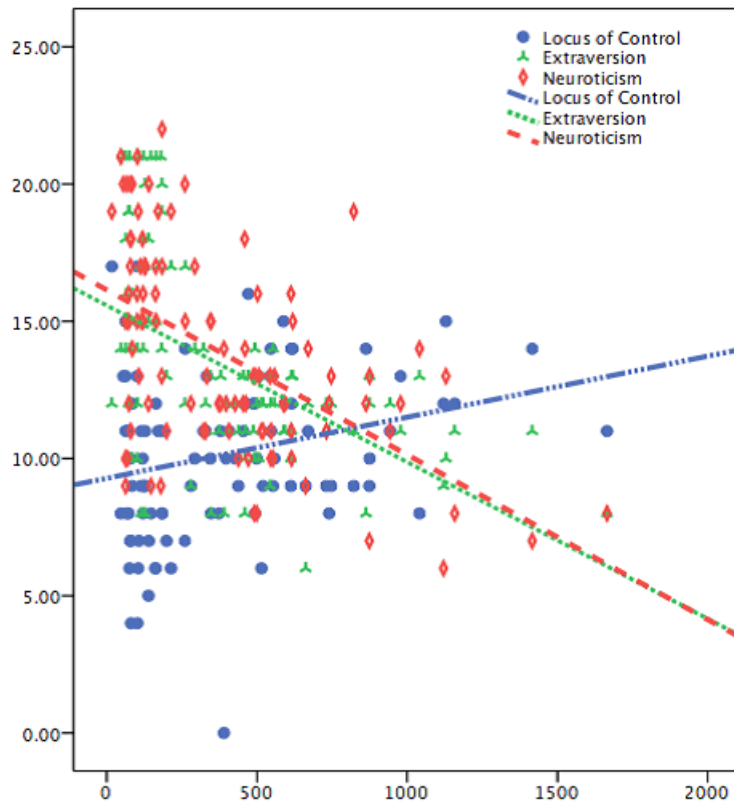


Figure 21. Correlation across tasks in MapViewer between completion times and personality traits.

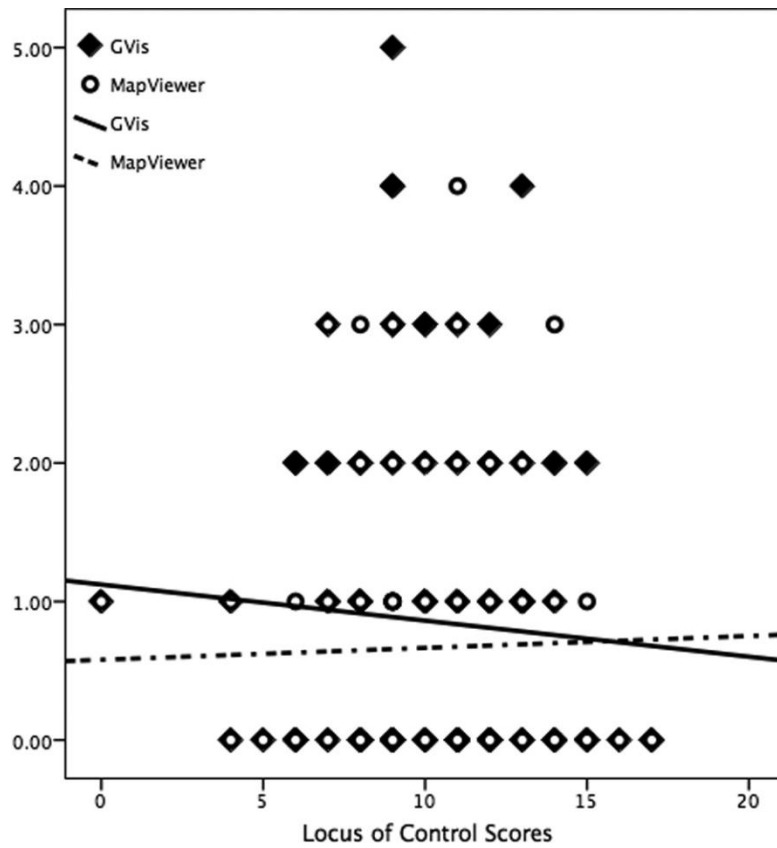


Figure 22. Scatter Plot overlay of Correlation of Generated Insights and Locus of Control

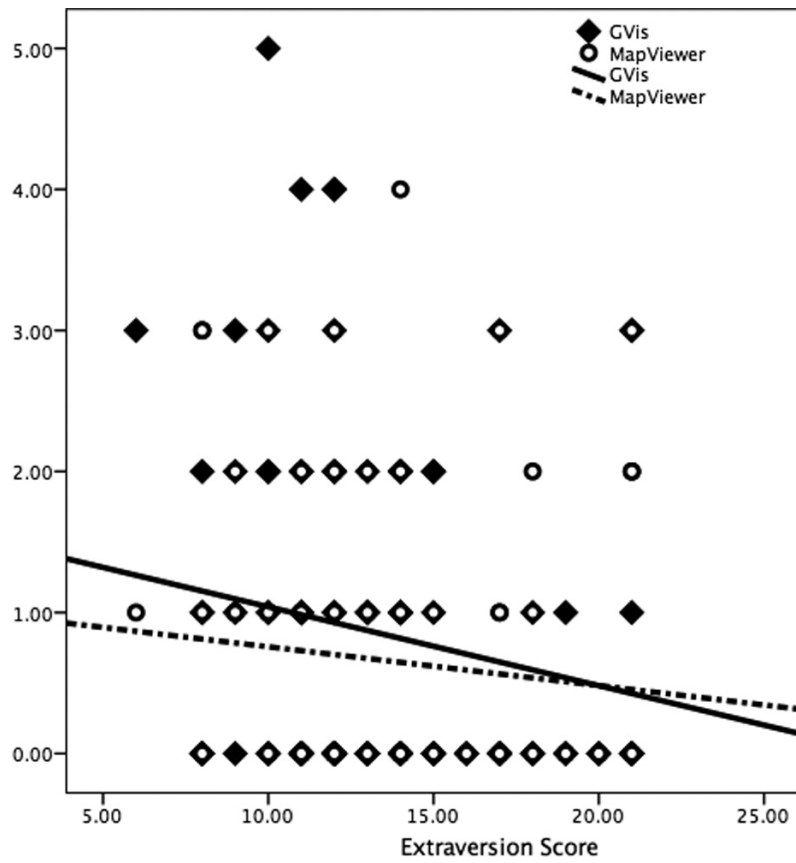


Figure 23. Scatterplot overlay of generated insights and Big Five Extraversion.

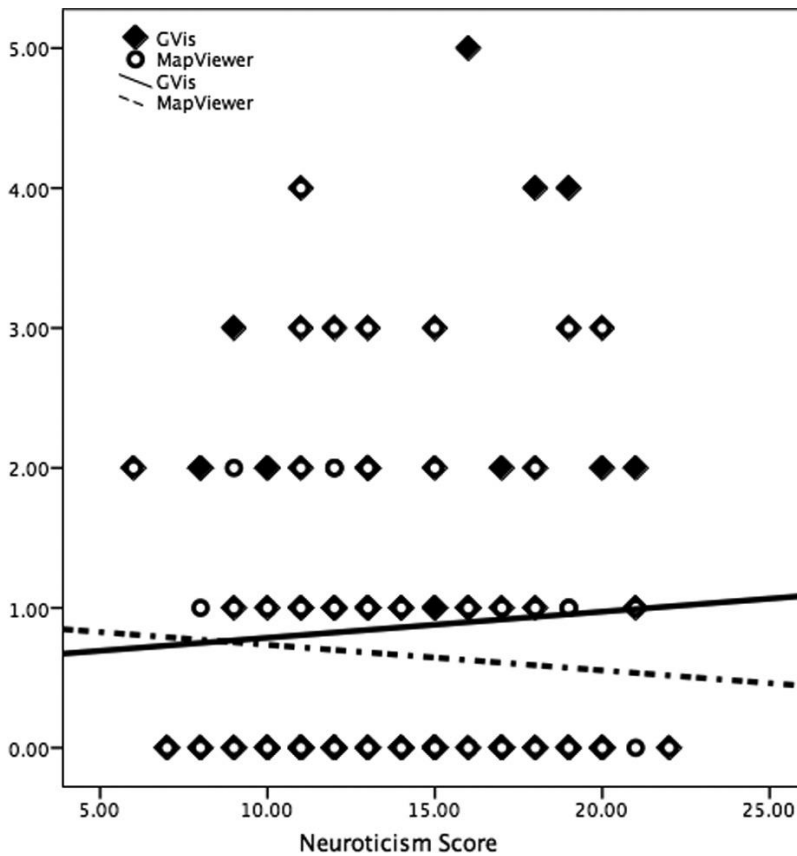


Figure 24. Scatterplot overlay of correlation between generated insights and Big Five Neuroticism.

within the interface in order to find the answer. If the answer given was incorrect, the error was recorded and the researcher asked the participant to try again until the correct answer was given. These errors were the only type of error recorded; this study was not designed to evaluate interaction logs or perceived “errors” or deviations from researcher-defined normative interaction paths. The total time in seconds from the initial reading of the question to the indication of the correct answer was recorded as the completion time.

Participants were not asked to move as quickly as possible in the second study, largely because while the first study was designed to more closely approximate a speed test, the second was somewhat more complicated and it was felt that an emphasis on speed would falsely increase the rate of reported errors. As the results for both studies in completion time were similar, we felt both study protocols provided an interesting dichotomy for comparison.

After participants had answered the questions in both interfaces, they were asked to specify which interface they liked better, and to give each interface a letter grade (“A” (superior) through “F” (failing)). Task sessions in the laboratory for both studies lasted approximately one hour. A short debriefing ended the study session, and there were no follow-up sessions.

5.5. Results

In Study 1, the mean completion times for the procedural learning tasks in the MapViewer ($M = 635.81$, $SD = 288.49$) were slightly faster than the completion times in the GVis ($M = 684.77$, $SD = 235.46$). In Study 2, the MapViewer procedural completion times were also faster ($M = 133.54$, $SD = 84.00$) than those in the GVis ($M = 161.64$, $SD = 111.40$).

Overall, participants preferred interacting with the visualization to interacting with the web table. This preference was indicated by post-study feedback. For example, when asked to give each interface a letter grade, from A (superior) to F (failing), 75 (73%) gave the GVis an A or B; 57 (56%) gave an A or B to the MapViewer. Additionally, when asked, 64 (61%) reported that they preferred the visualization; 39 (37%) preferred the web table.

5.5.1. Completion times and personality factors

The completion times in seconds for each condition for the procedural learning tasks in each study were merged into a single dataset, with $n = 106$, by running one set of analyses on both sets together. Participants completed tasks more quickly in MapViewer ($M = 383.15$, $SD = 32.38$) than in GVis ($M = 426.86$, $SD = 32.15$). A paired t -test between total completion times in GVis and completion times in MapViewer was significant ($t(100) = 2.11$, $p = 0.037$, suggesting that the differences in completion times was due to more than random chance.

A one-way analysis of variance (ANOVA) was used to test for the impact of LOC across interface completion times. The ANOVA for GVis was significant ($F(14,88) =$

1.89, $p = 0.039$) but the comparison for MapViewer was not ($p = 0.099$). In addition, LOC predicted completion times in both interfaces; a Pearson's correlation between LOC and completion times was significant (GVis: $r(105) = 0.234$, $p = 0.02$, MapViewer: $r(105) = 0.254$, $p = 0.01$) (see Figure 21).

These findings suggest that participants with a more internal locus (those who believe they have control over personal life events) take less time to find target (Figure 21). Correlations of GVis Total Completion Times (in seconds) across procedural tasks and the Locus of Control, Big Five Extraversion, and Big Five Neuroticism scores. This correlational finding is the opposite of findings reported in an earlier study (Green, Jeong, and Fisher, 2010). This previous study used inferential tasks, and found that participants with a more external locus (those who did not believe that they were in control) tended to solve a series of inferential tasks more quickly than those with a more internal locus. These tasks were more cognitively complex than the current studies, and asked the participants to compare and contrast multidimensional objects and make decisions about similarities and differences. We will discuss this further in the "Personality factors and predictors" section.

We used ANOVAs to test for the impact of Big Five Neuroticism in both interfaces were significant: GVis: ($F(16,86) = 3.42$, $p < 0.001$), MapViewer: ($F(16,85) = 5.14$, $p < 0.001$). Big Five Neuroticism also was negatively correlated with completion times in both interfaces. GVis: ($r(103) = -0.2047$, $p < 0.001$, MapViewer: $r(102) = -0.2054$, $p < 0.001$) (see Figure 22). ANOVAs to test for the impact of Big Five Neuroticism in both interfaces were significant: GVis: ($F(16,86) = 3.42$, $p < 0.001$), MapViewer: ($F(16,85) = 5.14$, $p < 0.001$). Big Five Neuroticism also was negatively correlated with completion times in both interfaces (Figure 22): GVis: ($r(103) = -0.2047$, $p < 0.001$) MapViewer: ($r(102) = -0.2054$, $p < 0.001$).

Differences in interface completion times and Big Five Extraversion were significant across both interfaces : GVis: ($F(14,88) = 5.37$, $p < 0.001$); MapViewer: ($F(14,87) = 4.12$, $p < 0.001$).

In summary, these faster participants tended to be more emotional (high Big Five Neuroticism) and more sociable (high Big Five Extraversion). A summary of these findings can be found in Figure 26.

5.5.2. Task errors and personality factors

The two studies measured tasks errors differently, and so must be analyzed separately. In Study 1, procedural tasks asked participants only to indicate when they had located the target information, so no errors were made or recorded. In Study 2, error was defined as giving the wrong answer to a question. Upon making an error, participants were asked to continue to try until they correctly solved the task. Each incorrect solution was recorded as an error.

Kolmogorov–Smirnov Z was significant in both interfaces (GVis: $p < 0.001$, MapViewer: $p < 0.001$). Levene’s test of homogeneity was significant in for GVis ($p = 0.004$), but not MapViewer ($p = 0.30$), suggesting that sample distributions were not uniformly normal. Owing to these two findings, we opted to conduct non-parametric tests for the purposes of the following analyses.

Participants made more errors in GVis ($M = 1.21$, $SD = 1.07$) than in MapViewer ($M = 0.69$, $SD = 1.07$). Friedman’s chi-squared was significant ($\chi^2(1) = 5.45$, $p = 0.02$). Kendall’s tau was conducted between errors in each interface and psychometric scores; no significant associations were found.

Generally speaking, only the difference in interface had a significant impact on how many errors were made; participants were more effective in the MapViewer interface. A summary can be found in Figure 26.

Insight generation and personality factors Participants reported having more “unexpected” insights in the GVis ($n = 73$) than in the web-based MapViewer ($n = 70$). The distribution of the combined – knowledge and interface – self-reported insights reported across both interfaces was not normal according to the Kolmogorov–Smirnov test (GVis: $p < 0.001$, MapViewer: $p < 0.001$). Levene’s test of homogeneity was significant for GVis ($p < 0.001$), but not MapViewer ($p = 0.373$). As the distribution was

not normal, a Friedman's chi-squared was run between the mean number of insights generated in both interfaces, and was not significant: Friedman's $\chi^2(1) = 1.59$, $p = 0.208$. Kendall's coefficient of concordance was found to be 0.015. This suggests that interface type did not have a significant impact on the number of insights generated.

In an investigation of the impact of LOC on insight generation, a Friedman's chi-squared was run between LOC scores and the mean number of insights generated in both interfaces and was significant. GVis: Friedman's $\chi^2(2) = 174.36$, $p < 0.001$, Kendall's coefficient of concordance = 0.83; MapViewer: Friedman's $\chi^2(1) = 101.04$, $p < 0.001$, Kendall's coefficient of concordance = 0.96.

Because both studies had a within-participants design, a Kendall's tau-b was conducted. LOC was not associated with the number of generated insights in both interfaces (GVis: $p = 0.59$; MapViewer: $p = 0.46$).

We also explored the impact of Big Five personality traits Big Five Extraversion and Big Five Neuroticism on insight generation in both interfaces. A Friedman's chi-squared between mean Big Five Extraversion scores across interfaces was significant (GVis: Friedman's $\chi^2(1) = 105.0$, $p < 0.001$, Kendall's coefficient of concordance = 1.0; MapViewer: Friedman's $\chi^2(1) = 105.0$, $p < 0.001$, Kendall's coefficient of concordance = 1.0). Big Five Extraversion was associated with insight generation (GVis: $t = 20.15$, $p = 0.051$; MapViewer: $t = 20.18$, $p = 0.027$), and predicted the number of insights in both interfaces (GVis: $R(103) = 20.554$, $p < 0.001$; MapViewer: $R(101) = 20.543$, $p < 0.001$). These findings suggest the more insights were reported by participants who were less extraverted (Figure 24).

A Friedman's chi-squared between mean Big Five Neuroticism scores across interfaces was significant: (GVis: Friedman's $\chi^2(1) = 105.0$, $p < 0.001$, Kendall's coefficient of concordance = 1.0; MapViewer: Friedman's $\chi^2(1) = 105.0$, $p < 0.001$, Kendall's coefficient of concordance = 1.0). Big Five Neuroticism predicted the number of generated insights in both interfaces (GVis: $R(103) = 20.415$, $p < 0.001$; MapViewer: $R(101) = 20.509$, $p < 0.001$). These findings suggest that more neurotic participants did not report as many insights as those who had lower Big Five Neuroticism scores (Figure 25). A summary of these findings can be found in Figure 26.

5.6. Introduction to an early profile of efficient users

The Personal Equation of Interaction has several end goals. The first is to be able to predict how well, if certain cognitive factors about the user/user group are known, those users will perform in a given environment.

Our report in “Comparative studies” and “Results” sections has been framed in such a way as to apply to this objective. Another end-goal, as previously discussed, is to adapt interfaces to allow the user to change certain aspects of the interface to best suit their known cognitive, perceptual or learning styles. This real-time individuation will become easier as we learn more about user individual differences across tasks and interface metaphors.

A third end-goal is to develop methodologies for the creation of robust user profiles: profiles which consider individual differences in the user group within the construct of a knowledge or expert domain. This goal in and of itself has several possible offshoots. One is an in-depth understanding of the knowledge domain with its instructional differences. This has been done elsewhere. ⁷ Another is to be able to describe what types of tasks users will perform best based not only on an understanding of their intrapersonal differences, such as personality factors described here, but on a knowledge of what types of learning or reasoning the user will employ in order to reach analytical goals. This is a paradigm shift from the traditional first-wave human–computer interaction perspective of seeing the human as an information-processing machine with a standard taxonomy of interactions. It requires looking at what users do with an interface, and specifically what types of cognitive task will be utilized to reach those goals.

	Completion times	Errors	Insights
Interface	faster times in MapViewer	fewer errors in MapViewer	more insights in Gvis
Locus of Control	internal locus faster times	none	external locus more insights
Extraversion	more extraverted faster times	none	less extraverted more insights
Neuroticism	more neurotic faster times	none	less neurotic more insights

Figure 25. A summary of chapter findings.

In this section, we will sketch out early user profiles based on the work reported here as well as newly completed research. These profiles will not be based on the user's membership in an expert cohort or the user's use of a specific interface, but rather on what type of cognitive task the user was employing during interaction and completion of assigned task.

	Locus of Control	Beck's Anxiety Inventory	9-item short measure	Tolerance of Ambiguity	Extraversion	Neuroticism
Locus of Control		.394 .000	.366 .009			
Beck's Anxiety	.394 .000			.543 .000	-.333 .001	-.448 .000
9-item short (n=50)	.230 .022	.747 .000		.493 .000	-.229 .022	-.442 .000
Tolerance of Ambiguity		.543 .000			-.276 .005	-.418 .000
Extraversion		-.229 .022	.585 .000	-.276 .005		
Neuroticism		-.448 .000		-.418 .000		

Figure 26. A summary of the correlations between all aspects of the Personal Equation.

We can use individual differences to improve our understanding of visually enabled analysis across knowledge domains. Research has demonstrated that users in a particular domain can share personality characteristics and learning preferences, both intrapersonal and institutional (Heuer, 1999). This implies that traits common to the user group can be aggregated into specific user profiles, informing superior design requirements and aiding in evaluation protocols. A personal equation of interaction could both (a) provide guidelines for individuated interface designs which could broadly accommodate differences in learning style, reasoning heuristic preferences, and perceptual behaviors and (b) develop profiles of expert or non-expert user groups, delineated by either knowledge domain or cognitive tasks that would inform the interface design for specific user or task domains.

In other words, the Personal Equation's goal is to predict interface performance by measuring personality factors and other user cognitive proclivities. Based on that, and because we know that experts of a certain kind tend to be similar in certain ways, if we can describe a superior analyst – his or her personality, learning styles, and world-view – we can use that information to predict the expert cohort's strengths and weaknesses, and, when developed to a predictive degree, use that information to inform interface design.

As previously discussed, we hope to describe groups of users based on their membership in an expert group. There are two ways to do this: describe the users and observe which tasks the users perform in a superior manner, or observe the performance and evaluate user personality factors after the task. In this paper we report early user profiles based on both methods. We use the findings of empirical studies to delineate users into groups of superior performers/inferior performers.

We then use the personality traits matrices of these performers to start a profile. However, only a few personality factors are reported in these current empirical studies, and we would like to explore a more robust profile, even at these early stages.

Because profiling is a form of descriptive modeling and requires large number of participants, we use additional participants who completed only a battery of personality measures in conjunction with the empirical study results. Using the user performance

groups delineated by the empirical study, we will evaluate correlations between persons who share personality characteristics and explore other learning proclivities which might be associated.

These profiles are not meant to be predictive models; they are intended as a chalk outline of what characteristics a superior performer in a given task might have. Describing an analyst and predicting behavior based on that description is a primary goal of the Personal Equation, and these profiles are a quick peek at what that description might look like.

5.7. Personality factors as predictors

The field of personality psychology has a long history in defining and assessing personality. Most of the research done to define personality theories can roughly fall into one of two categories: structure-oriented theories and process-oriented theories.

Structure-oriented theories of personality are those that focus on how personality is structured and on traits which are stable, life-long, and can be reliably discriminated from others through psychometric testing. ²³ Process-oriented theories, on the other hand, tend to focus on how a person organizes his or her thoughts and actions, such as how well a person copes with negative life events or whether or not a person is antisocial. ²³ In this paper, we primarily discuss structure-oriented theories.

Further, structure-oriented personality theories can be grouped in multiple ways. Single-trait theories, such as LOC, trait anxiety, self-regulation, need for cognition, and tolerance of ambiguity, which we will describe in greater detail in the following sections, tend to be invariant between persons, play a demonstrative role in individual behavior, and can be reliably measured on a continuum, where some people have the trait to a higher or larger degree than others.²³ On the other hand, multi-trait theories, such as the Big Five²¹ personality inventory, which we also use, tend to describe the individual at a higher level, and are driven by the assumption that there are a limited number of broad personality traits which more holistically describe human behavior. This is one reason we include scores from all five of the Big Five personality index in our descriptive

expert profile; even though every subscale may not seem to be an obvious predictor of behavior, these traits are broad, and do interact.

Lastly, learning styles, such as the Index of Learning Styles (Felder and Soloman, 2001) measures the manner in which an individual acquires knowledge and approaches changes in information. Single-trait theories have a deep literature on their impact on cognitive and learning performance. For the purposes of this short review, we will emphasize examples in technology-related research. For example, Kagan and Douthat demonstrated that extraversion and neuroticism predicted FORTRAN exam scores (Kagan and Douthat, 1985) As we discussed previously, individual differences have been found to have a bearing in traditional learning environments (Pintrich, Roeser and Wand De ReGroot, 1994) And in an earlier study (Green, Jeong, and Fisher, 2010), we found that certain aspects of trait anxiety had an impact on task efficiency in both inferential and procedural tasks.

Also, Rotter's Locus of Control(1966) predicted inferential task efficiency in our previous work; we will review this finding in the "Introduction to an early profile of efficient users" section. For user group profiles, personality characteristics of user domains have been done in a limited fashion⁷; this research would further these aims.

In addition, one or all of the Big Five personality traits have been shown to influence behavior. Each of the traits was found to predict academic performance in college-level students in Australia (Poropat, 2009). The Big Five trait conscientiousness was found to predict the programming performance of students in a pair programming class (Salleh et al., 2010) Further, our previous work has demonstrated that extraversion and neuroticism are both predictors of procedural tasks.

Research in learning styles is somewhat more scattered, but there is evidence that the way in which a human organizes information can impact outcomes. For example, studies done with 170 college students in introductory Java programming courses found that students who were more reflective, intuitive, verbal, and global thinkers tended to have higher performance scores on coursework, but especially on exams.²⁹ This was a primary motivator in the early expert profiles we discuss in the "Personality factors as predictors" section. We use these intercorrelations to describe

user groups based on how they organize information and the type of task they are undertaking.

Lastly, these personality constructs can and do trend together, either as multiple contributors to a single outcome or as corresponding factors. This can demonstrate some overlap in the tested constructs, but it may also demonstrate a synergy of traits that work together to impact performance. For example, Judge et al. found that LOC and neuroticism were intercorrelated and tended to trend together in their study of trait assessment (2002). Early user profile For the purposes of this early profile, we will use psychometric measures previously reported (LOC, Big Five Extraversion, and Big Five Neuroticism) in addition to several new measures described below. One hundred participants took part; 50 of whom were those from Study 2. In addition, 50 participants also completed the psychometrics but did not the study tasks, and were used to explore correlations between psychometric whole scores. Participants from Study 1 were not used for this profile.

These profiles are not predictive models. This work is in its early stages, and what is presented here is a first stab at a description of an analyst profile and what it might look like. What follows will be used to inform future work.

Our exploration of what personality factors impact cognitive task outcomes and interface interaction is ongoing. As part of that continuing exploration, we asked 50 undergraduate students in interdisciplinary studies classes to complete a new battery of psychometric measure online. These participants in Study 3 did not perform interface tasks and were included in this analysis and in these profiles only to support and extend the descriptions of efficient users, as previously discussed in the Introduction. In other words, the performance of participant data in Study 2 was used to define what an “efficient” learner is. We isolated participants whose completion times were the fastest in both tasks and both interfaces, or in the bottom 25 percentile of all participant completion times (GVis: 4.88 seconds, MapViewer: 4.76 seconds) as efficient learners. Six participants were more efficient for all procedural tasks in both interfaces. We evaluated the psychometric scores of these participants, specifically how these more efficient learners trended on each of the original three measures: Big Five Extraversion and

Neuroticism, as well as LOC. From there, we extrapolated these early informal profiles based on how the participants in Study 3 performed on the original three measures. We used bivariate correlations to search for trending and associations between all psychometric measures.

The findings from the analyses of these measures will be used in this paper to sketch a fuller-bodied profile of more efficient users segregated only by the types of learning/reasoning they are employing during interface interaction, not by expert domain or interface type. This profile is a work in progress, and will very certainly adapt to new findings in ongoing research as they accrue.

Several psychometric measures were added to those original three measures already reported. These additional measures were as follows: the other three Big Five personality factors of Agreeableness, Intellect/ Imagination, and Conscientiousness (Donnellan et al., 2006), the Need for Cognition scale (Cacioppo et al., 2006) and the Index of Learning Styles, which has four subscales: (Active/Reflective, Sensing/Intuitive, Sequential/Global, and Visual/Verbal) (Felder and Soloman, 2006).

For the other Big Five personality inventory, we used a condensed version of a 20-item Mini-IPIP, and used only the psychometric items of Big Five Extraversion and Big Five Neuroticism which were common to both studies. The other three factors in the Big Five have been included in this profile here as all five are submeasures of the one holistic Big Five personality model or inventory.

The Big Five personality factor of Agreeableness (Donnellan et al., 2006) is generally defined as a person's tendency to get along or go along with others. Agreeable persons are considered considerate of others and desirous of working cooperatively (Schwarzer, Diehl and Schmitz, 1999)

The Big Five personality factor of Intellect/Imagination is also called Openness (Donnellan et al., 2006). Persons with high Intellect/Imagination scores tend to be intellectually curious, and appreciative of novel ideas and experiences.

The Big Five Factor of Conscientiousness²¹ is used to describe the degree to which a person is organized, self-motivated to achieve, and effective at defining goals and creating plans to carry out those goals, as compared with persons who are more spontaneous and perhaps more unstructured.

Cacioppo and Petty's Need for Cognition scale (1996) evaluates the degree to which a person seeks out cognitively challenging experiences. Persons with a high need for cognition enjoy challenging problems or puzzles, and find fulfillment in tackling difficult problems. We used the short form of the measure, which has 18 items.

The Beck Anxiety Inventory (BAI) (Beck et al., 1988) is a 21-item Likert-type scale which asks the participant to evaluate how often common anxiety symptoms were experienced over the previous month, from 0 (not at all) to 3 (severely – bothered me a lot). The BAI was designed to diagnosis “trait” anxiety, a tendency to be prone to anxiety generally, even when stressors are not present. Persons with high trait anxiety have been shown to be alert and more responsive in other laboratory performance environments, and trait anxiety has been shown to be an intercorellary of LOC (Archer, 1979). The Self-Regulation Scale (SRS) (Schwarzer, Diehl and Schmitz, 1999) is a 10-item Likert-type measure which evaluates “post-intentional” regulation of focused attention and emotional maintenance throughout the completion of a goal oriented task, or, in other words, the ability to maintain sustained focus despite distractions, uncertainty, and/or emotional events.

The Scale of Tolerance–Intolerance of Ambiguity (TOA) is a 16-item Likert measure designed to appraise the degree to which the participant self-evaluates novel, complicated, or apparently unsolvable situations as threatening (Budner, 1962). In other words, it measures how comfortable the user is with uncertainty. Tolerance of ambiguity, as measured by the TOA, is not, like the SRS, a measure of coping ability per se, but an appraisal of self-beliefs, similar to the LOC.

Both self-regulation and tolerance of ambiguity as constructs are, at least superficially, related to the previously mentioned traits; we chose these measures to explore other persistent personality characteristics that might have an impact on procedural or inference learning.

The Index of Learning Styles (ILS)²⁴ evaluates four dimensions or continuums of cognitive or learning styles; learning styles are often defined as trait-like proclivities to perceive and process information in distinctive ways. These measures were added to our battery after observing that many participants offered evocative feedback over which interface was better organized, suggesting that their observations were due to more than just opinion or preference for familiarity, but might be due to the way the participant themselves preferred to organize information, and this influences their choices.

The first of the ILS subscales, Sensing/Intuitive, defines the two ends of ILS's first continuum; scores closer to Sensing indicate a practical learning style oriented to facts and empirical knowledge, whereas Intuitive scores indicate a more theoretical style with an emphasis on underlying ideas.

The ILS Visual/Verbal continuum puts participants who prefer information presented as pictures and graphs (Visual) at one end and participants who prefer words, including written or oral explanations, at the other (Verbal). The ILS Active/Reflective scale defines learners who prefer to think and learn by doing, perhaps through collaboration with others, as compared with learners who prefer to learn or think by reflecting on the problem, usually alone or with a few others.

Finally, the ILS Sequential/Global dimension reflects two ends of another spectrum, with learners who prefer to see a problem one step at a time, tackling problems in a bottom-up fashion at one end and learners who are top-down, big-picture holistic learners at the other.

These additional psychometrics were found to be closely related to many of the previously discussed personality factors, as we will now discuss. Relationships between factors For the evaluation of personality factors evaluated both in the study reported here and in the current work, we aggregated the datasets ($n = 100$). Generally speaking,

we found that there were relationships and/or associations between the constructs of LOC, trait anxiety, the tolerance of ambiguity, extraversion and neuroticism.

For the relationships between the personality factors specific to the ongoing work, we constrained our analyses of all factors to just that dataset ($n = 50$). We will itemize those findings here, and discuss them further in the following subsections. See the previous subsection for descriptions of the psychometric measures.

Self-regulation was negatively correlated with Beck's Anxiety Inventory ($r(49) = 0.520, p < 0.001$) and positively correlated with Tolerance of Ambiguity ($r(49) = 0.310, p = 0.029$). This suggests that persons with higher self-regulation tend to have lower trait anxiety and are more tolerant of ambiguity.

The Big Five factor of Agreeableness was positively correlated with Big Five Extraversion ($r(49) = 0.281, p = 0.048$), Tolerance of Ambiguity ($r(49) = 0.347, p = 0.013$), the Big Five factor Intellect/Imagination ($r(49) = 0.434, p = 0.002$), and the Need for Cognition ($r(49) = 0.314, p = 0.028$). This would indicate that persons who are more agreeable are also likely to be more extraverted, tolerant of ambiguity, have higher intellect (i.e. open to new experiences), and be more open to challenging and novel stimuli.

The Big Five factor of Intellect/Imagination was positively correlated with Tolerance of Ambiguity ($r(49) = 0.434, p = 0.002$), Agreeableness ($r(49) = 0.434, p = 0.002$), Need for Cognition ($r(49) = 0.385, p = 0.006$), and the ILS Sensing/Intuitive ($r(49) = 0.424, p = 0.002$). These findings suggest that persons with a higher intellect score are also more tolerant of ambiguity, are more agreeable, have a higher need for cognition, and tend to be intuitive.

The Big Five factor of Conscientiousness was positively correlated with the Index of Learning Styles (ILS) Sequential/Global ($r(49) = 0.331, p = 0.019$). In other words, more conscientious persons tend also to be global thinkers. The ILS subscale Active/Reflective was positively correlated with Big Five Neuroticism ($r(49) = 0.352, p = 0.012$). These statistics indicate that persons who are more reflective are more neurotic, and persons who are more active are less so.

The ILS subscale Sensing/Intuitive was positively correlated with Need for Cognition ($r(49) = 0.391, p = 0.002$) and Intellect/Imagination ($r(49) = 0.424, p = 0.002$).

This would suggest that participants who are more intuitive are also more likely to be more open to novel and challenging experiences.

The ILS subscale Visual/Verbal was negatively correlated with Tolerance of Ambiguity ($r(49) = -0.292, p = 0.040$). In other words, participants who tend to be more visual also tend to be intolerant of ambiguity. The ILS subscale Sequential/Global was positively correlated with Conscientiousness ($r(49) = 0.331, p = 0.002$). This suggests that participants who are more global thinkers may also be more conscientious.

In the next subsections, we will use these statistical findings to sketch out profiles of users segregated by superior cognitive task performance in one or both of the interfaces, which was defined in “Introduction to an early profile of efficient users” section as faster completion times on the interface tasks. We will use Pearson’s r correlations to establish trending and relationships between personality factors, as these are often used as the basis of establishing relationships between psychometric constructs.²² Future work will utilize follow-up tests in addition to Pearson’s r , such as Cronbach’s alpha and factor analysis, to evaluate the strength and subtleties of some of these relationships, but the use and implications of these additional analyses are beyond the scope of our current discussion.

5.8. Early profile of the efficient procedural learner

In this subsection we will discuss participants who were more efficient in the procedural learning tasks or those tasks that involved target identification and interface learnability and were cognitively simple. These participants tended to have an internal locus; in other words, these participants believe that they have some control over events that happen to them.

More efficient procedural learners are more likely to be trait anxious, i.e. anxious all the time, whether or not there is a trigger for stress present. Further, research has demonstrated that certain aspects of trait anxiety seem more predictive than others, especially those constructs that involve fear or foreboding. Not surprisingly, then, more efficient procedural learners in both interfaces also tend to be less open to new

experiences (Donnellan, 2006) and are relatively intolerant of environments, situations, and tasks that are unfamiliar or contain the unknown (Budner, 1966) In a similar vein, these participants were more distractible, and less likely to stay on task when distracted. They are more moody or emotional, but also more social than their counterparts.

More efficient procedural learners have a lower need for cognition; they do not seek challenges and like the predictable. Put another way, they generally do not need to be intellectually stimulated in order to feel satisfied with their lives. They tend to think about problems by reflecting, i.e. they prefer not to make a move until they have thought through the steps of a problem's solution. They like well-established approaches to (preferably) simple problems. They are better at memorizing facts than discovering new ways to solve a problem, and tend to be practical in their choice of methods. Further, these more efficient procedural learners are more comfortable exploring new concepts through words, whether spoken or written, over pictures, graphs, or images. For a summary of this profile, please see Figure 22.

Thus, in conclusion, based on completion times, users who were able to complete target identification tasks more quickly tended to be low information users who disliked uncertainty, social situations, and new challenges, and were more likely to allow their emotions to dictate their behavior.

5.9. Discussion

Aside from generally evaluating interface learnability, which we did in both studies, we studied procedural learning tasks in two slightly different ways. The first study focused on target identification. Participants were asked to find an organism label on the screen: for GVis, this label was attached to a spherical glyph; for MapViewer, very often the label was also a textual hyperlink. Once the label had been obtained, the participant pushed the "Submit" button and the task was done.

In the second study, we asked participants trivia questions, the answers to which had to be hunted for through the interface. If they gave the wrong answer, we requested that they keep looking. Like the first study, nothing other than an ability to use the interface and identify target labels was required. In both of these tasks, participants found the targeted information more quickly in the web table MapViewer; in Study 2, they also made fewer errors in MapViewer. Given the wide commercial use of web tables, it seems reasonable. See Figure 28.

	Completion times	Errors	Insights
Interface	faster times in MapViewer	fewer errors in MapViewer	more insights in Gvis
Locus of Control	internal locus faster times	none	external locus more insights
Extraversion	more extraverted faster times	none	less extraverted more insights
Neuroticism	more neurotic faster times	none	less neurotic more insights

Figure 27 . A summary of the user profile of a more efficient procedural learning in both interfaces

Most participants brought some prior knowledge of the interaction metaphor to the MapViewer tasks that they did not have for the data visualization. However, participants still strongly preferred GVis to MapViewer, even if they were not as effective in task performance. This may have been due to the novelty of GVis; most participants had never seen anything like it before. It also may have been due to data organization; many participants, in post-study open response, indicated a clear preference for GVis” organization and interaction.

LOC proved to be an influential personality trait no matter what the interface or task. The faster participants in both interfaces were persons who had a more internal

LOC, which is typified by a belief in personal control over life events. This finding is in close agreement with much of the available literature on LOC.

Persons with a more internal locus have been found to have better problem-solving skills,²⁰ to be more resolved to solve a task when it became difficult, and to be more likely to develop an intrinsic (internal) motivation to finish a difficult task (Weiss and Sherman, 1973). Thanks in part to positive behaviors like these, internal locus has also been found to lead to superior outcomes in academics, hospital recovery, and organizational environments.

What is intriguing is that, while an internal locus led to faster procedural task outcomes, this is not necessarily the case when the task becomes more cognitively difficult. In a previous paper (Green, Jeong, and Fisher, 2010) we studied inferential learning. The tasks required participants to evaluate a multidimensional exemplar, and draw a conclusion about other organisms based on similarities or differences. We reported that participants who had a more external locus – those who believe that they are not in control, and who tend to believe in luck as a cause of events – solved inferential tasks in GVis more quickly than those with an internal locus. The results, which we discuss, do not contradict our current findings, but rather expand on them. In these studies, we used a larger sample size, which likely made our analyses more sensitive to changes in participant scores. Further, we focused on only three constructs that seemed more highly predictive, unlike previous work, which used six psychometric measures.

For one type of learning task performance to be predicted by the degree of internal locus and another type to be predicted by the degree of external locus lends credence to our introductory statement that, depending on task, intrapersonal individual differences can predict interface performance. Yet while LOC has been shown to be influential in a wide variety of human performance, as previously discussed, to date, it has not been considered by interface designers and evaluators.

Based on our research, as well as a broad LOC literature, we consider LOC to be one construct in the Personal Equation of Interaction. In addition to LOC, the Big Five personality factors of Big Five Neuroticism and Big Five Extraversion also predicted

procedural task performance. The more extraverted or neurotic the participant, the more quickly he or she was able to identify target information.

This is interesting, but little in the behavioral literature explains these correlations; for us, it is a subject of our ongoing research. Further, Big Five Neuroticism in these studies was found to be negatively correlated with LOC ($r(105) = -0.284, p = 0.003$). This does have some precedent in the literature.

For example, as previously discussed, Judge et al. (2002) evaluated several personality factors, including LOC and Big Five Neuroticism, and found that they were inter-related and could be shown to be a part of the same construct. This means that items from these measures trended together and were statistically predictive of the same personality factor(s). Research like this affirms that psychometric constructs can and do work together. Further, it lends credence to an approach that seeks to find items or clusters of items which could work together in the prediction of certain interaction behaviors.

This was also the case as we sought to use whole psychometric scores to describe users who target identified information more quickly than their counterparts. This early user profile is one of the first to describe user personality by cognitive task rather than by membership in a particular user group (e.g. computer science students or intelligence analysts). Given future work, user profiles based on cognitive task could be used to aid interface design by painting a picture of the “ideal” user group for an interface, depending on which types of task the interface was designed to support. For example, in our profile, ideal users are described as being intolerant of uncertainty; this informs interaction design by suggesting that interface affordances and interaction metaphors involved in target identification should be painfully clear and more learnable that might be required in other areas of the interface more tailored to other types of cognitive task.

And in a similar vein, this profile encourages interface designers to see the interface as a tool to support cognition, to consider that types of cognitive task and process that interface is being designed to support, and to consider the needs of those processes during design and evaluation.

Insights were also predicted by personality as in factor scores. This is compelling because it suggests that the impact of a predictive Personal Equation may go further than efficacy or efficiency; it may extend to being able to predict some learning or problem-solving outcomes as well. In these studies, we asked participants to self-report what they viewed to be novel, newly learned knowledge. This is a simple procedure often used in the learning and behavior sciences, and, although imperfect, served its purpose as a boundary-finding measure. Further studies will be designed to more carefully evaluate what learning occurs during task. Much depends, too, on how the word “insight” is defined. In the visualization and visual analytics literature, insight is often undefined, and when defined, it is often broadly defined (Chang et al., 2009). This makes “insight” difficult to use as an evaluative interaction outcome, and thus, as briefly discussed earlier, leaves certain claims about the superiority of visual analytics interfaces unproven. Recently, “insight” has been defined within two categories: knowledge-based insight, and spontaneous insight. Spontaneous insight is a sudden solution to an unsolvable problem, and has often, in the psychological literature, been referred to as an “aha!” moment. Spontaneous insight was not evaluated in these studies.

In these studies, we evaluated the number of knowledge-based insights reported across task and interface, which are generally defined as items or concepts learned or added to the user’s knowledge base.

In evaluating the knowledge-based insights reported, we categorized insights on the basis of content: insights about how to use the interface itself were separated from insights about the informational content presented and manipulated. In our studies, we asked participants to report knowledge-based insights.

In both interfaces, roughly twice as many knowledge-based insights were reported about interface learnability (GVis: $n = 51$, MapViewer: $n = 47$) as were reported about the informational content (GVis: $n = 22$, MapViewer: $n = 23$). In both interfaces, the greatest number of interface learning insights was reported in the first question, which suggests that learnability started early. As the task set proceeded, the reported count of each insight type tended to even out somewhat, which is not unexpected; users

started paying attention to content once manipulating the interface was less of an issue or became more automatic.

Overall, whether learning about the interface or the interface content, personality factors predicted reported learning as well as other interaction outcomes. These findings have immediate implications. For example, these studies have demonstrated that users who tend to be more extraverted and neurotic are also more likely to believe that they are in control of the task situation (internal locus). By extension, this also means highly neurotic or extraverted users tend to be better at interface manipulation and target identification.

If the personality factors of the user were known beforehand, we could reasonably predict how quickly he or she would be able to learn a novel interface and find pertinent information. For even when the interaction metaphor was completely unfamiliar, as it was in the GVis visualization, neurotic/extraverted participants were able to learn to manipulate the data more quickly.

However, what these findings do not do is demonstrably differentiate between interface and interactive techniques. The three evaluated personality factors impacted both interfaces similarly. Given the cognitive simplicity of the tasks, this is perhaps unsurprising.

Ongoing research has been designed to evaluate learning styles that tend to guide focused attention and information organization during tasks, and where behavior research suggests more delineating personality factors for visualization technique might be found. A last note is on the use of novices in evaluations using an expert system; most of the participants had little or no knowledge of biological concepts. However, the participants were still capable of ably finding target information in both interfaces. Yet even with the more familiar archetype of the web interface, participants preferred the visualization.

The intent of these studies was never to evaluate the efficacy of GVis per se; a formal evaluation of GVis as an expert system is reported in other literature (Hong et al., 2005). The aim of these studies was to evaluate human cognition during learning

interaction using both interfaces as working artifacts of a kind. In addition, we explored whether individual differences in personality factors and self-beliefs could have a large enough impact on interaction outcomes to warrant their inclusion in the Personal Equation of Interaction.

For these reasons, we recruited non-experts who were unfamiliar with the knowledge domain. Expertise would have biased the user's interaction; they would have had an expert knowledge of the genomic hierarchies, and thus known where to look for the requested information. This would have proven a poor evaluation of how each interface promoted learning.

5.10. Conclusion

The Personal Equation of Interaction is still very much a work in progress. In the short term, it serves as an open discovery and proof of concept. We have shown that intrapersonal differences impact interaction. Our ongoing research seeks to better define what differences impact what type of analytical task (for it seems reasonable to assume that one intrapersonal set of differences will only generalize to one type or set of task constraints).

For example, we are currently narrowing our task sets to study multiple decision points in specific types of category or inference reasoning. And, further, we hope to explore whether that impact is temporally static or dynamic throughout the analytical process.

In the longer term, we intend to isolate predictive matrices and validate a battery of measures that will successfully inform interface design based on the types of cognitive task undertaken. Ultimately, this is the Personal Equation of Interaction. These measures will likely involve more than personality factor matrices; other areas of

exploration include perceptual logics and use of decision-making heuristics. In addition to informing design through use of user profiles like the one described here, the Personal Equation could be used to provide real-time interface adaptation to accommodate user needs and preferences, and provide a basis for robust group profiles of users who share common differences, such as experts or users of a particular visualization technique. Visual analytics seeks to facilitate analytical reasoning through the use of interactive visual interfaces. In the Personal Equation of Interaction, we will provide a new tool in that pursuit.

5.11. References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, 89(4), 369.

Archer, R. P. (1979). Relationships between locus of control and anxiety. *Journal of Personality Assessment*, 43(6), 617-626.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of consulting and clinical psychology*, 56(6), 893.

Boekaerts, M. (2000). Personality and the psychology of learning. *Psychology of Education: Pupils and learning*, 2(10), 414.

Boyatzis, R. E., & Kolb, D. A. (1995). From learning styles to learning skills: the executive skills profile. *Journal of Managerial Psychology*, 10(5), 3-17.

Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of personality*.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological bulletin*, 119(2), 197.

Chang, R., Ziemkiewicz, C., Green, T. M., & Ribarsky, W. (2009). Defining insight for visual analytics. *Computer Graphics and Applications*, IEEE, 29(2), 14-17.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment*, 18(2), 192.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650.

Green, T. M., Jeong, D. H., & Fisher, B. (2010, January). Using personality factors to predict interface learning performance. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1-10). IEEE.

Hall, R. H., Rocklin, T. R., Dansereau, D. F., Skaggs, L. P., O'Donnell, A. M., Lambiotte, J. G., & Young, M. D. (1988). The role of individual differences in the cooperative learning of technical material. *Journal of Educational psychology*, 80(2), 172.

Hepner, P. P., & Anderson, W. P. (1985). The relationship between problem-solving self-appraisal and psychological adjustment. *Cognitive therapy and research*, 9(4), 415-427.

Heuer, R. J. (1999). Psychology of intelligence analysis. Lulu. com.

Hong, J., Jeong, D. H., Shaw, C. D., Ribarsky, W., Borodovsky, M., & Song, C. G. (2005, June). GVis: A Scalable Visualization Framework for Genomic Data. In *EuroVis* (Vol. 5, pp. 191-198).

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct?. *Journal of personality and social psychology*, 83(3), 693.

Kagan, D. M., & Douthat, J. M. (1985). Personality and learning FORTRAN. *International journal of man-machine studies*, 22(4), 395-402.

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. Mit Press.

Macia, A. (1979, April). Visual perception of landscape: sex and personality differences. In *Proceedings of Our National Landscape Conference*. USDA Forest Service, General Technical Report PSW-35 (pp. 279-295).

Palmer, J. (1991). Scientists and information: II. Personal factors in information behaviour. *Journal of documentation*, 47(3), 254-275.

Pintrich, P. R., Roeser, R. W., & De Groot, E. A. (1994). Classroom and individual differences in early adolescents' motivation and self-regulated learning. *The Journal of Early Adolescence*, 14(2), 139-161.

Plaisant, C. (2004, May). The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 109-116). ACM.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin*, 135(2), 322.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, 80(1), 1.

Salleh, N., Mendes, E., Grundy, J., & Burch, G. S. J. (2010, September). The effects of neuroticism on pair programming: an empirical study in the higher education context. In *Proceedings of the 2010 ACM-IEEE international symposium on empirical software engineering and measurement* (p. 22). ACM.

Schwarzer R, Diehl M and Schmitz GS. (1999). The selfregulation scale. Berlin: Freie Universitat, *Gesundheitspsychologie*.

Simon, H. A. (1991). Cognitive architectures and rational analysis: Comment. In *Architectures for intelligence: The 22nd Carnegie Mellon Symposium on Cognition* (pp. 25-39).

Soloman, B. A., & Felder, R. M. (2005). Index of learning styles questionnaire. NC State University. Available online at: <http://www.engr.ncsu.edu/learningstyles/ilsweb.html> (last visited on 14.05. 2010).

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive science*, 25(2), 203-244.

Thomas, L., Ratcliffe, M., Woodbury, J., & Jarman, E. (2002). Learning styles and performance in the introductory programming sequence. *ACM SIGCSE Bulletin*, 34(1), 33-37.

Weiss, H., & Sherman, J. (1973). Internal-external control as a predictor of task effort and satisfaction subsequent to failure. *Journal of Applied Psychology*, 57(2), 132.

Wong, P. C., & Thomas, J. (2004). Visual analytics. *IEEE Computer Graphics and Applications*, (5), 20-21.

Chapter 6. Discussion

So much depends upon reasoning, but it cannot be seen. Only its outcomes can be measured. Thus, reasoning's journey must be inferred through what actions it takes after a decision is reached. Human reasoning is the process by which interface learning solves novel problems. To a large degree, reasoning is concealed. This might explain why comparatively few cognitive researchers study reasoning *per se* but prefer to study its outcomes through judgment, learning, and decision-making.

However, what if we could do more than just observe its faint trail as it leads us through the cognitive process? What if instead we could predict its outcomes? In some ways, this current research was an attempt by one reasoning researcher to tackle reasoning from an unexpected angle and to predict what reasoning would do before the task was attempted, and through this prediction, understand the reasoner not only by performing the task at hand but also by using the predictors that the reasoner brings to the task.

The understanding of reasoning cognition is the primary driver of this research. Prediction is the means of obtaining that understanding. Being able to predict performance through even simple measures, such as time-to-target and self-reported insights or learnings, is one way to overcome reasoning's invisibility. The mere idea that one could predict visual analysis by first administering a three-minute, nine-item Likert survey before task serves to lift the invisibility curtain for a peek at the mastermind behind it.

From the beginning of this line of research, the idea that the complexity of reasoning and learning could be bounded by the way a user or analyst answers a short series of questions has been a key motivator. Truthfully, it would not have mattered what the questions were, only that they were statistically normed to the degree that differences could be detected between groups and that those groups would systemically

behave in predictable ways. This explains why this research corpus has made little attempt to defend the concept of personality or, indeed, to take a position in a debate on the source of personality. What does it matter? All that matters is that the surveys administered elicit consistent, predictive answers to the question of learning performance.

Although the concept of interface individuation was discussed briefly in Chapter 1, there has been little effort to map personality onto the task of interface design. Careful thought was given to the choice of interface across the studies; the interfaces shared a dataset and a knowledge base but little else. Design could certainly glean much from this line of research, but we do not purport that one interface paradigm is better than another. Indeed, if this research demonstrates anything, it is that no one interface is best for all users.

6.1. Chapter 2

In Chapter 2, the results demonstrated that task and interface were predicted by the psychometrics. Most of the items in the 9-item PEI were anxiety-based, or described some fear of the unknown. For the inferential tasks, the biggest indicator of performance was locus of control. As we discussed in Chapter 2 the degree of control an individual feels over his or her life circumstances has long been regarded as a demonstrative predictor in the human behavioral literature.

In this study, the belief in a lack of control – or an external locus – was only predictive of completion times in the data visualization. GVis was developed in response to a request for a better way to locate and analyze the spatial and semantic relationships between ontological biological structures (Chapter 2.3); compare-and-contrast behaviors should, then, be easier to see and solve in GVis. Additionally, the performance outcomes, non-significant trending between MapViewer outcomes and the sporadic psychometric scores, as well as the varying nature of the participant feedback suggest that combination of variables influenced MapViewer complex performance behaviors, perhaps due to the difference in required interaction. Often, tasks that required one or two mouse clicks in a single view in GVis were much more complicated in the

MapView, requiring multiple mouse clicks and changes of view. For example, unlike the straightforward presentation of mapped genes in GVIs through direct interaction (holding down a single mouse click on the visualized target), determining the existence of a mapped gene for an organism in MapViewer required the user to hunt for the organism name in the list of organisms, possibly reorganizing the list through primary and secondary sorts, locating and clicking on the small single letter “G” on the far right of the application view, which served as a hyperlink to a separate page. If a gene existed, information about its mapping was presented. If the gene did not exist, the hyperlink led to a page presenting a frustrated-looking male icon and the explanation, “No information found for given taxid.” Locus of control played a role in the MapViewer inferential task outcomes, but not one strong enough to show any predictive strength.

In the procedural tasks, the 9-item short measure is moderately negatively correlated with completion times. This suggests that more trait-anxious (i.e. persons that tend to be anxious all the time as compared to anxious only when presented when threatening stimuli), uncommunicative, and/or prone to emotional instability a person is, the less time they tend to take finding requested items while interacting with novel information. This might seem counterintuitive at first glance. However, according to Spence-Taylor Drive Theory (Spence, Farber, and Schmitz, 1999), persons with higher trait anxiety tend to identify target information more quickly than the non-anxious when the task does not require either iterative or complex reasoning processes. Other studies have found that persons with higher trait anxiety are more attentive to presented information and can identify target threats more quickly than those less anxious (Ionnou, Mogg, and Bradley, 2004). While the causes for this “exception” are still subjects of debate, it has been proposed that trait anxious persons have developed adaptive heuristics that can make advantageous use of their anxiety (Spence et al., 1966). The results of the current study would suggest that certain aspects of trait anxiety tend to make users more attentive and better able to identify target information until the task becomes complex, requiring more complicated reasoning heuristics and lessening the effectiveness of the adaptation. In other words, trait anxiety helps analysts develop cognitive tools that are highly sensitive to new stimuli, and which allow rapid identification of items of interest (or potential threat).

Additionally, the 9-item short measure scores positively correlated with LOC scores ($r = .37$, $p < .01$), suggesting that persons who were more anxious/uncommunicative were also more likely to attribute consequences of life events to forces outside their control, such as luck or divine intervention. Scores on the whole Beck's Anxiety Inventory, however did not correlate. Given that LOC scores were a predictor of more efficient completion times in the more complex, inferential task, it seems reasonable that a relationship exists between the aspects of anxiety captured by the 9-item short measure and locus of control.

The items in the 9-item short measure were culled from 6 measures designed to measure anxiety, extraversion, neuroticism (or emotional instability), self-efficacy, and self-beliefs about control over personal circumstances. Subsequent analysis of the 9-item short measure found it to have moderate internal consistency and to meet the requirements for a reliable psychological assessment. However, it is unreasonable to expect that any new measure would be fully validated after one evaluative trial. While we are fairly confident the 9-item short measure has captured trending in this study, we recognize that further trials are required before the 9-item short measure could be considered predictive or reliable in a generalizable way.

This replication would include a variety of types of task developed for the procedural and inferential learning. End-goal target identification is only one type of "how-to" interface learning. Use of specific interface functionalities such as the use search box or the use of help capabilities could also be tasks. Inference tasks could involve a wide variety of visualized content. Expert systems which use knowledge schemas other than genomics would also be utilized. By using a variety of knowledge schemas, we demonstrate that the 9-item measure can predict interface learning and not just genomic interface learning.

Additionally, and perhaps even more importantly, a variety of interfaces would also be used to test the breadth of the learning generalizability. There are numerous styles and types of GUI. And data visualization's presentation and interaction paradigms increase in both quantity and novelty year over year. Because the goal of the PEI is to predict analysis outcomes and not the superiority of one interface over another, testing

its predictive capacity repeatedly with a variety of visual interfaces allows us to demonstrate the degree of its ability to predict reasoning no matter the environment.

In summary, Chapter 2 set the stage for the study of cognitive outcome predictors. We found the boundaries of the degree to which a standard data visualization evaluation could be predicted by inherent individual differences. In addition, having found that this protocol did an acceptable job of testing participant procedural and inferential learning, we used this protocol in the studies that followed.

The research in Chapter 2 is joined by other research in data visualization and interface interaction. Ziemkiewicz, Ottley, and Crouser used locus of control to predict the performance of participants which were presented information in a variety of visualized contexts or metaphors (2013). Psychometrics such as locus of control has been used in the evaluation of data visualization techniques (see for example Yi, 2010).

6.2. Chapter 3

In Chapter 3, we built on Chapter 2 by evaluating the relationships between the PEI and additional reasoning outcomes. We predicted two additional outcome measures: task errors and self-reported insights. Whole measure scores (i.e. all items in the measure were used) from Locus of Control, Big Five Extraversion and Big Five Neuroticism were used as predictors.

Locus of Control proved to be an influential personality trait no matter what the interface or task. The faster participants in both interfaces were persons who had a more internal locus of control, which is typified by a belief in personal control over life events. This finding is in close agreement with much of the available literature on locus of control. Persons with a more internal locus have been found to have better problem-solving skills (Krause, 1986), to be more resolved to solve a task when it became difficult, and to be more likely to develop an intrinsic (internal) motivation to finish a difficult task (Weiss and Sherman, 1973). Thanks in part to positive behaviors like these, internal locus has also been found to lead to superior outcomes in academics, hospital recovery, and organizational environments.

This research has a variety of implications and informs new work. One obvious implication is that it changes the way interfaces are evaluated or ‘graded’. Interaction paradigms, as we saw in Chapters 2, 3 and 5, often employ cognitive-based schemes, such as GUI’s fondness for text, even when supporting target identification. Visualizations, by contrast, lead the user to interact with glyphs or other graphical depictions. Because the user may have a demonstrable proclivity for either text or graphics, but not both, an interface’s evaluation should include a description of the user’s individual differences. If the GUI is seen as effective for users who prefer text, then it may be seen as performing well. However, understanding that users who prefer graphics or spatial displays will not adapt to the interaction paradigm as easily prevents the interface from being reviewed too harshly, provided that the evaluators understand the user’s proclivities beforehand.

Much the same way as we did in Chapter 2, our tasks required participants to evaluate a multi-dimensional exemplar, and draw conclusions about other related concepts based on similarities or differences. We reported that participants who had a more external locus – those who believe that they are not in control, and who tend to believe in luck as a cause of events – solved inferential tasks in GVIs more quickly than those with an internal locus. For a discussion of these results, please see Chapter 2. The results do not contradict Chapter 3’s reported findings, but actually expand them. In Chapter 3, we used a larger participant group, which likely made our analyses more sensitive to changes in participant scores. Further, we focused on only 3 whole constructs that seemed more highly predictive, unlike in Chapter 2 which culled items from 6 whole measures.

In Chapter 3, we demonstrated that one type of learning task performance was predicted by the degree of internal locus and another type was predicted by the degree of external locus. And this lends credence to one underpinning assumption that inherent individual differences can predict multiple types of analysis performance, albeit differentially. Based on our research, as well as a broad locus of control literature, we consider locus of control to be one construct in the Personal Equation of Interaction. In addition to Locus of Control, the Big Five personality factors of Neuroticism and Extraversion also predicted procedural task performance. The more extraverted or

neurotic the participant, the more quickly he or she was able to identify target information.

Further, Neuroticism in these studies was found to be negatively correlated with Locus of Control ($r(105) = -.284, p = .003$). This does have some precedent in the literature. For example, Judge et al. (2006) evaluated several personality factors, including Locus of Control and Neuroticism, and found that they were interrelated and could be shown to be a part of the same construct. This demonstrates that items from these measures trended together and were statistically predictive of the same personality factor(s). Research like this affirms psychometric constructs can and do work together, Further, it lends credence to an approach that seeks to find items or clusters of items which could work together in the prediction of interaction efficacy, especially if the whole measure scores fail to reach an acceptable degree of prediction.

And lastly, the interrelationships between these whole measures or constructs highlights again in both Chapters 2 and 3 why we were able to identify specific trends between psychometrics and outcomes. This is true even though in Chapter 2 we used individual items and in Chapter 3 we used whole measures. The intercorrelations between different but complementary inherent constructs allows us to build not only predictive tools but – as we will do in Chapter 5 – build profiles that allow us to describe and thus identify individuals in any target user population that are likely to use our interfaces the most effectively or perhaps intuitively.

Chapter 3 gave us a more complete picture of how inherent differences predict interface learning cognition. Insights were also predicted by personality as in factor scores. This is compelling because it suggests that the impact of a predictive Personal Equation may go further than efficacy or efficiency; it may extend to being able to predict some learning or problem-solving outcomes as well. Much depends on how the word “insight” is defined. In the visualization and visual analytics literature, insight is often undefined. When defined, it is often broadly defined, as in (North, 2006). This makes “insight” difficult to use as an evaluative interaction outcome, and thus, as briefly discussed earlier, leaves certain claims about the superiority of visual analytics interfaces unproven. Recently, “insight” has been defined within two categories:

knowledge-based insight, and spontaneous insight (Chang et al., 2009).

Spontaneous insight is a sudden solution to an unsolvable problem, and has often been referred to as an “aha!” moment. In problem solving and related neuroscience research, spontaneous insight has been defined as the name of “the process by which a problem solver suddenly moves from a state of not knowing how to solve a problem to a state of knowing how to solve it (Mai et.al , 2004).

Spontaneous insight differs from the self-reported research in a variety of ways. For example, spontaneous insight does not depend on many of the more gradual cognitive heuristics we engage in the tasks in these studies (Kounios, 2004). Secondly, it tends to happen when the analyst is not focused on the problem to be solved, unlike the focus required in these studies in order to complete those tasks, as is typical of interface learning (Mai et.al , 2004). Interface learning depends on iterative learning that builds a scaffold between what is known and what is about to be known. This step-by-step process may use a variety of cognitive tools, but does so in a comparatively straightforward way.

Or in other words, interface learning uses defined heuristics, a clear (if highly repetitive) path to task completion, and conscious choice between exemplars or choices. In spontaneous insight, the reasoner rarely can describe how the problem solution was achieved. The process appears to depend on unconscious problem re-organization and spontaneous understanding (Kounios, 2004).

In these studies, by contrast, we asked for conscious observations of learning. The participant not only recognized that they had learned something, but what they had learned and often, how they learned it. We evaluated the number of knowledge-based insights reported across task and interface, which are generally defined as items or concepts learned or added to the user’s knowledge base. In evaluating the knowledge-based insights reported, we categorized insights on the basis of content: insights about how to use the interface itself were separated from insights about the informational content presented and manipulated.

In both interfaces, roughly twice as many knowledge-based insights were

reported about interface learnability (GVis: N = 51, MapViewer: N = 47) as were reported about the informational content (GVis: N = 22, MapViewer: N = 23). In both interfaces, the greatest number of interface learning insights was reported in the first question, which suggests that learnability started early.

As the task set proceeded, the reported count of each insight type tended to even out somewhat, which is not unexpected. Users started paying more attention to content once manipulating the interface was less of an issue or became more automatic.

These self-reported insights are not unlike those reported in Saraiya, North, and Duca (2005), which grouped their self-reported insights by what the context of those insights and allowed the participants to freely respond, or Plaisant (2004), which described them as basic building blocks for interface evaluation. Pousman, Stasko, and Mateas also suggested that insights should be garnered and separated by the context or substance of insight. (2007 See also Yi, et al, 2008.)

Overall, whether learning about the interface or the interface content, personality factors predicted reported learning performance as well as insights. These findings have immediate implications. For example, these studies have demonstrated that users who tend to be more extraverted and neurotic are also more likely to believe that they are in control of the task situation (internal locus). By extension, this also means highly neurotic or extraverted users tend to be better at interface manipulation and target identification. If the personality factors of the user were known beforehand, we could reasonably predict how quickly he or she would be able to learn a novel interface and find pertinent information. For even when the interaction metaphor was completely unfamiliar, as it was in the GVis visualization, neurotic/extraverted participants were able to learn to manipulate the data more quickly.

A last note is on the use of novices in evaluations using an expert system; most of the participants had little or no knowledge of biological concepts. However, the participants were still capable of ably finding target information in both interfaces. Yet even with the more familiar archetype of the web interface, participants preferred the visualization.

The intent of these studies was never to evaluate the efficacy of GVis *per se*; a formal evaluation of GVis as an expert system is reported in other literature (Hong et al., 2005). The aim of these studies was to evaluate human cognition during learning interaction using both interfaces as working artifacts of a kind. In addition, we explored whether individual differences in personality factors and self-beliefs could have a large enough impact on interaction outcomes to warrant their inclusion in the Personal Equation of Interaction.

For these reasons, we recruited non-experts who were unfamiliar with the knowledge domain. Expertise would have biased the user's interaction; they would have had an expert knowledge of the genomic hierarchies, and thus known where to look for the requested information. This would have proven a poor evaluation of how each interface promoted learning.

6.3. Chapter 4

In Chapter 4, we sought to fit the complexity of the PEI into the complexity of reasoning itself. In particular, we discussed the types of reasoning used during task analysis, what made each unique, and how each fit into the whole of the lifecycle of visual analysis. Chapter 4 did not report new research but it did fit the PEI research within the boundaries of our previous work, such as the Human Cognition Model or Pirolli and Card's Sensemaking Model (2005).

We also broadly discussed the impact of human individuality on every primary cognitive process, and surveyed our current research in pursuit the generation of new system development models that optimize the cognitive performance of human decision-makers. Optimization in this context must include complex criteria such as insight, innovation, creativity and awareness in uncommon, unique and novel problems and situations. Research has shown that inherent individual differences between users impacts the task and learning performance in visually embedded interfaces. Our previous work in the development of the Human Cognition Model continues to inform our research direction. The Personal Equation has highlighted the need to study not only inherent differences in personality factors, but also other user differences, including

those in which affect other inherent individualities as well as differences in institutional cohort and environment. These individual differences in human capabilities are great enough that any unitary system will be at best a compromise between the needs of the various sub-populations of users. Or phrased another way, Chapter 4 demonstrates why no one interface visualization and interaction paradigm could possibly be considered ideal for each and every analyst and each and every analytical goal.

The great white whale of the 'intuitive' interface is not only complicated by the complexity of information, the limited scale of the presentation screen, the varying degrees of confidence in the data and its validity, but by the analyst themselves. As we saw in Chapter 2,3,and 5, there is enough variation within each analyst and each expert cohort to a) distinguish between analyst groups or clusters and b) systemically predict outcomes as an interaction of both the analyst and interaction differences.

6.4. Chapter 5

In Chapter 5, we used what we have learned in the previous chapters to build a psychometric profile of superior procedural learners. Or put another way, we described users that are best at target identification. What we learned was that trends between psychometric whole measures such as those we reported in Chapter 3 were very useful not only in predicting desired learning behaviors but also in describing why these analysts were superior to their less efficient counterparts. By extending on the work in Chapter 3, we were also able to extend the profile beyond the 3 whole measures used in Chapter 3 to multiple complementary inherent constructs such as tolerance of ambiguity, the need for cognition, and the ability to self-regulate. We used these additional whole measures to build a profile which described these superior learners in a more meaningful way.

Profiles previously have been used in the psychological and organizational research to describe and understand an expert profile. The intent for these profiles is to understand not only what makes these cohorts similar but also why they are superior to the general population at their chosen skills or expertise. Sometimes these profiles are built by administering a battery of previously normed measures of personality or ability,

as in Fazel-Zarandi and Fox (2011). Warbah et al (2007) used personality measures other psychological constructs to better understand an expert cohort of nurses. Another way to build an expert profile is to use objective measures of expertise, such as those used in the Expert Seeker as implemented at National Aeronautics and Space Administration (NASA), which culls human resource records for similarities within an industry expert cohort in training, education, and experience (Becerra-Fernandez and Fox, 2000. For similar research outside of NASA, see Balog and De Rijke, 2007.) This type of profile would likely enhance a PEI-based profile, and allow associations to be drawn between the PEI and what types of training and experience profile-members sought and achieved. For example, let us hypothetically assume that all data visualization developers have a global-visual profile. These developers prefer information to be presented in a top-down graphical fashion. More diagrams, fewer paragraphs of text. Big ideas first, then the details. The “why” before the “what.” It seems not only logical but probable that these visual developers would tend to seek and achieve training and expertise in computer graphics, visual design, and/or computational knowledge representations. The PEI sits in the intersection or interplay of inherent differences between what makes us unique and what we self-select as desired knowledge and career.

Of course, this chapter leaves the profile of the analysts best at inferential research yet to be done. In some ways it would likely be similar to the current profile and quite different in others. As we saw in Greensmith, 2016, the overlap between classification and inferential predictors was such that one assessment could be created that predicted the outcomes of both tasks, even if the assessment predicts these outcomes with varying degrees of prediction. Classification in categorization is similar in some ways to the work of target identification in procedural interface learning. Both are deductive reasoning process that use rules to correctly label concepts. Both can involve iterative sessions of comparison between multi-attribute artifacts or complex ideas. Thus, we should not be surprised if the profile of the superior inferential interface learner is market similar to that of the superior procedural learner.

The Personal Equation of Interaction can be used in a variety of ways to predict analysis outcomes and inform interface design. We can tightly predict one type of

cognitive task outcome, using the PEI to create a predictor such as target identification. We can use previously normed whole measures to predict the superiority of one interface over another for more loosely defined group of target users. And, by extension, we can use the relationships between these whole measures to build an analyst persona or profile which can by turn define or inform the design of a more intuitive interface for those analysts. It is this interplay between the interface, the Personal Equation of Interaction and cognition that we will explore further in the next chapter.

6.5. References

- Balog, K., & De Rijke, M. (2007, January). Determining Expert Profiles (With an Application to Expert Finding). In *IJCAI* (Vol. 7, pp. 2657-2662).
- Becerra-Fernandez, I. (2000, October). Facilitating the Online Search of Experts at NASA using Expert Seeker People-Finder. In *PAKM*.
- Chang, R., Ziemkiewicz, C., Green, T. M., & Ribarsky, W. (2009). Defining insight for visual analytics. *Computer Graphics and Applications, IEEE*, 29(2), 14-17.
- Fazel-Zarandi, M., & Fox, M. S. (2011, September). Constructing expert profiles over time for skills management and expert finding. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies* (p. 5). ACM.
- Green, T. M., Ribarsky, W., & Fisher, B. (2009). Building and applying a human cognition model for visual analytics. *Information visualization*, 8(1), 1-13.
- Green, T. M., Wakkary, R., & Arias-Hernández, R. (2011, January). Expanding the scope: Interaction Design perspectives for visual analytics. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (pp. 1-10). IEEE.
- Hong, J., Jeong, D. H., Shaw, C. D., Ribarsky, W., Borodovsky, M., & Song, C. G. (2005, June). GVis: A Scalable Visualization Framework for Genomic Data. In *EuroVis* (Vol. 5, pp. 191-198).
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct?. *Journal of personality and social psychology*, 83(3), 693.
- Kounios, J., Frymiare, J. L., Bowden, E. M., Fleck, J. I., Subramaniam, K., Parrish, T. B., & Jung-Beeman, M. (2006). The prepared mind neural activity prior to problem presentation predicts subsequent solution by sudden insight. *Psychological Science*, 17(10), 882-890.
- Krause, N. (1986). Stress and coping: Reconceptualizing the role of locus of control beliefs. *Journal of Gerontology*, 41(5), 617-622.

- Mai, X. Q., Luo, J., Wu, J. H., & Luo, Y. J. (2004). "Aha!" effects in a guessing riddle task: An event-related potential study. *Human brain mapping*, 22(4), 261-270.
- North, C. (2006). Toward measuring visualization insight. *Computer Graphics and Applications*, IEEE, 26(3), 6-9.
- Pirolli, P., & Card, S. (2005, May). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis* (Vol. 5, pp. 2-4).
- Plaisant, C. (2004, May). The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 109-116). ACM.
- Pousman, Z., Stasko, J. T., & Mateas, M. (2007). Casual information visualization: Depictions of data in everyday life. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6), 1145-1152.
- Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 11(4), 443-456.
- Spence, K. W., Farber, I. E., & McFann, H. H. (1956). The relation of anxiety (drive) level to performance in competition and non-competition paired-associates learning. *Journal of Experimental Psychology*, 52(5), 296.
- Warbah, L., Sathiyaseelan, M., Vijayakumar, C., Vasantharaj, B., Russell, S., & Jacob, K. S. (2007). Psychological distress, personality, and adjustment among nursing students. *Nurse Education Today*, 27(6), 597-601.
- Weiss, H., & Sherman, J. (1973). Internal-external control as a predictor of task effort and satisfaction subsequent to failure. *Journal of Applied Psychology*, 57(2), 132.
- Yi, J. S. (2010, April). Implications of individual differences on evaluating information visualization techniques. In *Proceedings of the BELIV Workshop*(pp. 2010-2012).
- Yi, J. S., Kang, Y. A., Stasko, J. T., & Jacko, J. A. (2008, April). Understanding and characterizing insights: how do people gain insights using information visualization?. In *Proceedings of the 2008 Workshop on Beyond time and errors: novel evaluation methods for Information Visualization* (p. 4). ACM.
- Ziemkiewicz, C., Ottley, A., Crouser, R. J., Yauilla, A. R., Su, S. L., Ribarsky, W., & Chang, R. (2013). How visualization layout relates to locus of control and other personality factors. *Visualization and Computer Graphics, IEEE Transactions on*, 19(7), 1109-1121.

Chapter 7. Conclusion

At the beginning of this dissertation, we stated the following research question:

What is the Personal Equation of Interaction for Interface Learning?

We have explored the idea of using self-report psychometric items that are culled from normed measures as attributes of a mathematical equation that would allow us to predict cognitive performance during interface learning. We have created tasks that, although administered in a laboratory, mimicked the types of tasks an analyst would undertake when exploring a graphical representation of data, including target identification and inferring information about one category through the understanding of another. By using items in the Big Five Personality Factors and Locus of Control measures, we have built the first Personal Equation of Interaction (PEI) for Interface learning, which satisfactorily predicted variance in the accuracy of interface learning.

As a research program, the PEI tackled the question of how to predict the accuracy, efficacy, and learning in a visual interface. In Chapter 2, we introduced the interfaces and tasks that would be used in all studies. We tabulated a dimension-reduced measure to predict task performance, isolating items from trait anxiety, extraversion, neuroticism, and locus of control measures. We explored the domain of the PEI by evaluating analysts' performance in both a traditional GUI interface and a data visualization. By comparing two very different representational paradigms, we were able to control for certain assumptions that tend to be built into software design, such the superiority of data visualization over the GUI interface. We found both novel and familiar effects, but the PEI was able to predict the task performance of participants in both interfaces.

In Chapters 3 and 5, we also explored the influence of the individual. Using the same interfaces and tasks as used in Chapter 2, we again used items from normed

psychometric measures as predictors of task performance. Similar to our findings in Chapter 2, we found that items isolated from the measurement of constructs such as Neuroticism (sometimes called moodiness), Extraversion, and Locus of Control were influential in predicting interface learning outcomes.

In Chapter 5, we went from prediction to profiling, using the self-report data from more than 100 participants to evaluate the personalities of analysts who were the most accurate. The analysts that had an internal locus and were both more extraverted and more neurotic were faster than those who did not. Conversely, the externally focused, less extraverted, and less neurotic participants learned more from the interface than their faster counterparts learned.

In Chapter 4, we sought to place the PEI within the context of other cognitive research done in information visualization and visual analytics. We briefly categorized the types of reasoning that have often been studied, and then we used those category definitions to describe the findings of other work, such as the sensemaking loop and the human cognition model. The sensemaking loop was a solid example of abduction, for example. In the work presented here, we claim to understand reasoning by predicting it and by allowing prediction to group similar cognitive tasks.

Now that we have a PEI, what can we do with it? The Holy Grail of research that uses the differences between analysts as a map to interface design would naturally be interface individuation. Interface individuation is the use of the deep understanding of the analyst's preferences, needs, and goals, which allows the interface to modify itself in real time to best support the analyst. To date, interface individuation has focused on the human computer interaction, as emphasized in GUI design, such as color, token placement, and white space. The PEI extends individuation into the realm of design for interface reasoning. Each visual interface—wittingly or unwittingly—is designed to support a specific list of inter-related cognitive tasks. As we demonstrated, the performance of these cognitive tasks can be predicted through the measurement of the PEI. This requires the understanding of the target user beyond personas and cognitive walkthroughs to a quantifiable formula for optimal design. For example, if a designer knows that the target use of an upcoming project is the quick identification of crime

hotspots during a city-wide emergency (classification), and if the designer understands that the target user is an analyst who is likely to experience some performance anxiety, the designer will understand that the analyst will likely lean towards inherent and institutional proclivities. The PEI gives the designer a way to measure these proclivities and profile the target user in a more predictive way. Because the PEI identifies user weaknesses (e.g., users that tend to score lower also tend to take more time during target identification), the designer now has been informed that the visualization approach needs to be individuated to the target user in order to mitigate this weakness. In the example of emergency services, if we know that the analyst (or the expert cohort to which the analyst belongs) has a tendency to learn best through a global organization style and has very little tolerance for ambiguity, the designer would likely choose an interface that started with a bird's eye overview of the situation and would allow for the analyst to then focus on items of interest. Furthermore, an intolerance for the obtuse would encourage the use of clean design that seeks to disambiguate both the interface and its functionality. Clear choices about the representation of uncertainty in data would also be wise.

On the other hand, if the designer is tasked with an interface to be used by primary school teachers managing the scheduling of multiple classrooms simultaneously, similar knowledge would likely lead to a quite different design. For analysts who tend to think sequentially and from the bottom up, prefer text to pictures, and are likely to have an internal locus and higher trait anxiety, an interface that is global, pictorial, and minimized text would not be the best choice. A better choice would be an interface that allowed the analyst to enter information one step at a time, give verbal feedback about issues or conflicts, and provide textural confirmation. This interface might never really give an overview because one would likely not be necessary. Furthermore, because these analysts have an internal locus and higher trait anxiety, we know that they would likely be very quick to find the functionality they need, but if not, would stick to the task until they have achieved it. (We explored this briefly in Chapter 5, when we contributed this efficacy to Spence Drive Theory, which suggests that anxious persons adapt to new situations and therefore can find targets more quickly.) This is a high-level example, but it illustrates that only a little knowledge about the PEI could contribute to design. Specific examples, such as using the PEI to

determine whether a target audience is more or less visual, can also determine the degree to which visualized semantic representations, such as composite glyphs should be used.

The idea that individual differences predict cognitive performance is not new; it has been used in the behavioral sciences for decades. The PEI is novel because it attempts to quantify specific differences that directionally predict cognitive outcomes during interface interaction. However, the proposed research does not seek a single unified metric of prediction. This would immediately be met with a myriad of exceptions, in the same way that complete cognitive architectures have been challenged by the infinite ingenuity of cognitive adaptation and accommodation. Human cognition is too fraught with potential to rest the dull inevitability of a single explanatory variable. Rather than seek to predict the entirety of a cognitive process using a single PEI, this research leans on the Gestalt wisdom that the whole is greater than the sum of its parts. By integrating the measurement of multiple assessments, the PEI has its own way of emerging in making predictions for a larger reasoning outcome. It must be said, however, that forecasting a future is not the cause of that future. Thus, the PEI is not the outcome but is one of many answers to the question of why people—the user, the analyst, and the designer—do the things they do.

7.1.1. Ever onward and upward

Where do we go from here? Chasing the invisible across the space of complexity tends less to be less linear than it is a Gestalt. Having stumbled upon a seemingly consistent method for predicting learning outcomes, we find the ongoing program of research cluttered with more questions than ever. Breaking down visual learning into its reasoning modularities and studying those modularities through the many iterations of the analysis (e.g. classification, categorical inference, rule mechanization, and so forth) begs this driving question: Is there a personal equation for each type of reasoning cognition? Alternatively, will one equation predict all?

This current research is a study of final outcomes. There was no attempt to study how analysts may have reasoned through each iteration of the study tasks; only the final

answer was captured. This suggests that, at least on some level, the PEI may indeed serve multiple purposes. However, demonstrating such generalizability demands a systemic plod through the reasoning heuristics, one cognitive task at a time.

The PEI for Interface Learning is a short measure that can predict how well an analyst can find and synthesize the information presented through an interface. It would be unlikely to predict a very different cognitive task; the work in the previous chapters demonstrates that what makes one superior at task completion does not make one superior at synthesis. Within this limited scope, it explains *what* interface likely would work the best for the target user group, *why* some analysts would be able to work more effectively or efficiently, and *where* there likely would be opportunities to improve design.

7.1.2. Implications

Based on this research, it is easy to see the interface as the coordination of cognitive tasks. Each overarching analytical goal employs a variety of cognitive tasks or processes. In particular, analytical goals use iterations of decision making. And each iteration involves other reasoning heuristics that enable and impel the decision being made. Consider the comparison of two concepts. Regardless of interface, comparison requires the learning of the conceptual category before it can be compared to anything else. This learning contributes to basic categorizational tasks such as classification – the naming of a concept – as well as inferential categorization. Inferential categorization is the understanding of a category’s definitions well enough to infer what the category will look like from its name or label. For example, we learn to define a ‘dog’ by learning its component shapes or template, its associated colors and textures, as well as its movements, sounds, and smells. Thus, whenever one observes a four-legged animal that wags its tail, pants, barks, and seems social, one might define that animal as a dog. This is classification.

Perception of object → attenuation → breakdown of determining observable characteristics → classification

Inferential categorization, on the other hand, could be called classification in reverse. If I see a dog, I will assume its characteristics. It should have 4 legs and might lick my hand if I am friendly. Classification and inferential categorization employ different

reasoning heuristics. Classification depends heavily on deductive and normative rules. I should be able to define the category in such a way that separates that category from other categories. Conversely, inference assumes details from category labels. That dog should have 4 legs. If it does not, this violation of the category rules will force the reasoner to evaluate the category inductively, using satisficing or heuristics such as elimination-by-aspects until one finds a satisfactory explanation.

Classification of object → identification of observable category attributes → inference of missing attributes

Interfaces, intentionally or not, support the completion of a wide variety of analytical goals, which in turn are made possible by a wide taxonomy of cognitive heuristics. This research demonstrates that analyst proclivities impact the outcomes of those cognitive heuristics to demonstrative degree. For this reason, this research supports the building of a cognitive task taxonomy to inform visual interface design. This taxonomy would include analytical goals that the interface supports. In these studies, for example, the interfaces were built to support the classification of genomes and an evaluation of their characteristics. These analytical goals include the use of reasoning tasks or processes with outcomes such as perception, attenuation, classification, decision-making, judgement, as well as generation and comparison of hypotheses. Each one of these processes is a component that feeds and informs the analytical process outcome. And as our research demonstrates, each of these processes is influenced by individual differences.

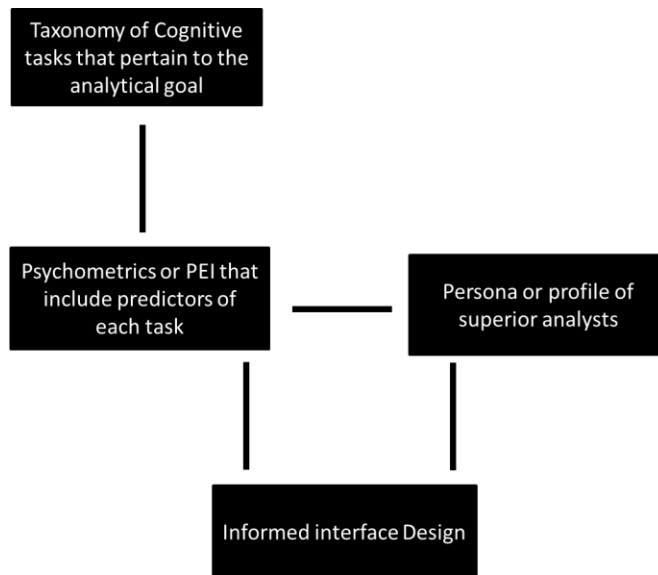


Figure 29. A simple diagram of the integration of the PEI in informed design.

By breaking down the analytical goals that the interfaces were intended to support, a hierarchical and/or conceptual node-link diagram emerges which demonstrates the order, weight, and inter-dependencies between these processes in the accomplishment of the goal. What remains is to define the Personal Equation of Interaction (PEI) that predicts superior performance for each of these cognitive outcomes. These PEIs will likely overlap to some degree, as recent research as demonstrated (Greensmith, 2016). In other words, there will likely be a number of psychometric items that predict most of all of the cognitive tasks involved in the completion of the analytical goal. By defining what characterizes a superior user of the interface, we can define design guidelines for design that better supports analytical goals.

Understanding analyst proclivities depending on the cognitive processes undertaken allows us to design interfaces that best support those proclivities. As we discussed earlier, our more recent research has demonstrated that the best visual classifiers are have an internal locus and tend to see the world from the bottom-up or sequentially. These are analysts who believe they are in charge of events that affect them, and they prefer to learn by accomplishing one step at a time. So we should design an interface that is text-based and makes it very clear through both interaction paradigm and sandbox or workspace layout what steps are necessary to accomplish classification.

Further, this PEI-informed interface may be great for associated types of task that are best accommodated by textual, sequential analysts. A robust PEI which describes and predicts superior performers for each type of cognitive process would allow for a richer design process. If a designer knows how his target audience differs from the world (e.g. more extraverted is more apt to prefer graphics, etc), and he knows what cognitive tasks are required to accomplish the analytic goal (for example, identifying which glyph or node is part of the target concept), then a PEI-based taxonomy provides an evidence-based method for designing an interface which seems immediately familiar. The Personal Equation can aid that design process by identifying how the target user likes to think about the world and organize information.

So, in summary, one PEI is a building block in what would be a larger program of evidence-based, user-centric interface design. The PEI can predict performance of each key cognitive task by associating individual differences with task performance. The individual psychometric constructs can be used as a profile of the best task performers. Or in other words, the PEI tells us why the best users are the best.

In addition, by breaking down a larger analytical goal into a taxonomy of cognitive tasks, the PEI becomes a system of equations that can predict performance aspects of the analytic goal. This is done by describing superior outcomes for each cognitive task that contributes to the goal, and identifying the psychometric persona of analysts who undertake the successful completion of the goal. This persona describes how these analysts like to organize information; implementing the preferred organizational schema not only makes the interface seem more intuitive, it aids in superior visual analytics.