DISTAL TRAIT AND PROXIMAL STRATEGIC PREDICTORS OF TECHNOLOGICAL FLUENCY

A Dissertation
Presented to
The Academic Faculty

by

Sibley F. Lyndgaard

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Psychology

Georgia Institute of Technology
December, 2024

DISTAL TRAIT AND PROXIMAL STRATEGIC PREDICTORS OF TECHNOLOGICAL FLUENCY

Approved by:

Dr. Phillip L. Ackerman, Co-Advisor
School of Psychology
Georgia Institute of Technology

Dr. Ruth Kanfer, Co-Advisor
School of Psychology
Georgia Institute of Technology

Dr. Sashank Varma
School of Psychology
Georgia Institute of Technology

Dr. Margaret E. Beier
Department of Psychological Sciences
Rice University

Dr. Richard Catrambone
School of Psychology
Georgia Institute of Technology

Date Approved: November 20, 2024

ACKNOWLEDGEMENTS

I would first like to thank my co-advisors, Dr. Ruth Kanfer and Dr. Phillip Ackerman, for their mentorship. The guidance they have provided during my graduate education was pivotal to both my philosophy of science and my development as a researcher – I consider myself fortunate to have been able to learn from both of them.

Thanks are also due to my committee members, Dr. Margaret Beier, Dr. Richard Catrambone, and Dr. Sashank Varma, for their insight throughout my dissertation process. This study would not have been possible without the contributions of my research assistants: Katie Cowart, Carolyne Halfon, Annie He, Yusra Khan, Haley Merritt, Utkarsh Nattamai Subramanian Rajkumar, and Andy Waisser. I am particularly grateful for Katie Cowart's efforts in scoring over 3,100 open-ended items across nearly 150 assessments.

My time at Georgia Tech would have been far less enjoyable without the incredibly supportive community in the PARK Lab and the friendship of fellow graduate students throughout the psychology department. Thank you all for the writing sessions and happy hours over the years – even (especially) when the former accidentally turned into the latter.

To my parents, Deb and Dan Lyndgaard, who have been my biggest cheerleaders: thank you for everything. I am incredibly lucky to have your love and encouragement through this and all other milestones in my life. Finally, I would like to thank Corey Tatel, whose support and partnership has been a constant buoy since day one. There is no one with whom I would rather have shared this experience – I cannot wait to write our next chapter together.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Modern workplaces are characterized by the necessity of just-in-time, non-routine problem-solving, often aided by technology. Although the Internet is an increasingly core resource for such problem-solving, the competencies which support leveraging technology for problem-solving are not well understood. Effective use of online resources depends upon a complex of individual differences including ability and non-ability traits, the accumulation of relevant knowledge and skills, and the use of adaptive strategic approaches during interactions with technological resources. I refer to this as technological fluency, a trait complex which describes individuals' propensity (i.e., ability/willingness) to leverage technological resources to solve real-world problems. The present study examined the relative contributions of distal trait and proximal strategic variables to explaining individual differences on an assessment of technological fluency. While item-level floor effects limited the interpretability of aggregate performance data, there were several interesting findings related to individual differences in process variables. Cluster analyses, for example, showed that minimally satisficing participants (who had very low intrapersonal variability in item scores) tended to be more positively oriented toward technology, while the two clusters of participants with greater intrapersonal variability in item scores could be distinguished by patterns of differences in both ability and non-ability variables. Follow-up qualitative analyses identified several key process variables for which ability differences were most striking, including inference quality (including the ability to self-correct faulty inferences), the ability to leverage visual problem cues, and the nature of AI use during problem-solving. Finally, exploratory regression analyses suggest that

proximal strategic variables do improve prediction of available process indicators above and beyond distal trait complexes, providing support for a key proposition of the study. In addition to providing methodological recommendations for the assessment of technological fluency (e.g., more direct assessment of metacognitive processes, inclusion of domain knowledge tests), the present study's findings suggest that the quality of metacognitive processing in technology-supported problem-solving may be correlated with individual differences in ability – this is a critical direction for future research.

CHAPTER 1. INTRODUCTION

Tool use has been broadly ubiquitous throughout human history (e.g., see early development of composite tools such as blades or hammers with attached handles; Ambrose, 2001; Basalla, 1988). In contrast to tools that are primarily functional (i.e., designed to extend humans' physical capabilities), informational tools are a more recent development which support problem-solving by extending a user's understanding of a novel task or informing their development of a potential solution. The late 20th and early 21st century rise in Internet accessibility has somewhat democratized access to informational tools by providing a widely available means of searching for problem-solving support in both occupational and personal settings. The integration of Internet capabilities into decision support systems, for example, made it simpler and more efficient for organizations to improve workers' access to job-relevant information beyond their extant knowledge, thereby improving the efficiency and effectiveness of their decision-making (Shim et al., 2002). More informal searches may also yield productive results. As discussed by Tidler et al. (2022), individuals seeking to repair or assemble a relatively unfamiliar object may have access to a wide variety of online manuals, 'how-to' videos, or social forums which might bridge knowledge/skill gaps.

The latter case (i.e., informal use of the Internet to accommodate skill gaps in novel problem-solving) is the focus of the current study. Increasingly, organizations seek to either select for, or train 'on-the-spot' problem-solving competencies, which are becoming more important, considering trends toward less routine work tasks and more frequent problem-solving in a variety of sectors (Neubert et al., 2015). I posit that because nonroutine work tasks are more likely to require a worker to employ procedures with which they may or may not be familiar, adults' propensity to use informational tools to solve novel problems may be a key means by which to identify high performers. In some cases, these informational tools may be available through intra-organizational resources such as in-house technical manuals. In other cases, such as when no documentation has been created or when a situation arises that cannot be solved based on existing organizational expertise, the Internet may support a much broader search.

The World Wide Web is a widely used, accessible source of problem-solving support, providing just-in-time access to innumerable resources. In a survey of internal medicine residents, for example, 63% of respondents reported that they used Google to support care decision-making at least once a day (Duran-Nelson et al., 2013). Efficient, effective use of online resources for novel problem-solving depends upon a complex of individual difference characteristics including ability and non-ability traits (e.g., Moehring et al., 2016), the accumulation of relevant knowledge and skills (e.g., Desjarlais & Willoughby, 2007; Willoughby et al., 2009), and the use of adaptive strategic approaches (e.g., Elshout, 1987) during interactions with technological resources. Broadly, these characteristics can be thought of in terms of variance in the extent to which an individual *can* and *will* utilize technological resources to accomplish some goal which cannot be accomplished relying solely on their existing knowledge and skills (Ackerman et al., 2023). I refer to this as *technological fluency*.

The major goals of the current study were: (1) to develop criterion measures of technological fluency which can be administered online; (2a) to identify and develop a set of individual-difference measures representing learners' strategic approaches to engaging online resources during problem-solving (i.e., predictors of technological fluency); and (2b) to demonstrate the criterion-related validity of these measures for technology-supported problem-solving outcomes above and beyond more traditional approaches to predicting task performance (i.e., those focused primarily on distal trait measures of cognitive, affective, and conative constructs). The project makes three contributions:

- *Contribution 1:* Delineates the technological fluency construct.
- *Contribution 2:* Provides evidence for the distal trait and proximal strategic correlates of technological fluency, including directions for research into additional promising correlates (i.e., domain knowledge).
- *Contribution 3:* Provides methodological recommendations for assessing technological fluency, including evidence for the importance of assessing process variables.

In the literature review below, I provide an overview of relevant research which justifies and informs the study. In the first section, I define technological fluency and distinguish it from conceptually-related constructs (e.g., digital literacy; Bhatt & MacKenzie, 2019). I then review extant online assessments of technology-supported problem-solving, identify limitations of these measures, and discuss how the study addressed the first expected contribution by developing a criterion measure that avoids these issues.

The remainder of the review both: (1) situates the proposed technological fluency assessment from the first section of the review in the context of traditional selection procedures; and (2) argues for the development of a battery of proximal variables which should outperform a more traditional battery for predicting technological fluency. The first issue is addressed by a brief discussion of distal trait-based approaches. Historically, distal and stable traits have been useful for predicting general task performance and job performance (e.g., Schmidt & Hunter, 2004; Sonnentag et al., 2008), but they may be too broad to explain substantial variance in the relatively narrower construct of technological fluency. The final section of the literature review addresses the second issue by introducing a set of proximal strategic variables which are expected to influence the qualitative approach an individual might take in using technology to solve a novel problem. I draw on the theoretical definition of technological fluency, which is laid out in the first section of the review, to explain why the proximal measures are expected to be more strongly related to performance than distal traits.

CHAPTER 2. REVIEW OF RELATED LITERATURE

2.1 Operationalization and Assessment of Technological Fluency

2.1.1 Defining Technological Fluency

Technological fluency, as described in the previous section, is the general capability and propensity of individuals to leverage technology to solve a novel problem. The strength of this propensity is conjectured to be a function of both distal and proximal psychological factors (Ackerman et al., 2023). Relatively distal influences on technological fluency include stable traits (e.g., cognitive traits such as verbal ability). Relatively proximal influences include acquired knowledge and skills (e.g., related to solving similar problems or utilizing technology in general), motivational factors (e.g., related to the effort-utility function; Kanfer, 1987), and differences in individuals' strategic approaches to problem-solving.

Previous attempts to define online search-related abilities have focused on various aspects of computer use, information search, and comprehension which are related to, but do not fully encompass technological fluency. Early assessments of 'computer literacy,' for example, addressed both basic use (e.g., word processing, file management) as well as the ability to interact with Internet-based search engines (Hoffman & Blake, 2003). As computer use skills became more generally ubiquitous among both students and working-age adults, however, more recent research (e.g., on digital literacy) has focused on learners' ability to navigate and interpret problem-relevant information gleaned from searches (Bawden, 2008).

“Digital literacy” was proposed as an overarching construct related to how people find and evaluate information in online contexts (Bhatt & MacKenzie, 2019; Graesser et al., 2022), and it has some conceptual overlap with technological fluency. However, where the construct of digital literacy tends to have a more targeted evaluative focus (e.g., can learners sufficiently judge the quality or credibility of a source, or avoid misinformation; Bhatt & MacKenzie, 2019), technological fluency represents a broader complex of influences on novel problem-solving. Technological fluency can therefore be distinguished from constructs such as computer and digital literacy in at least two critical ways, as follows:

First, whereas computer literacy and digital literacy focus on narrower aspects of competence in ‘online’ actions as the primary phenomenon of interest (e.g., correctly navigating a file management system; correctly identifying an article which contains misinformation), technological fluency is concerned with individuals’ ability and willingness to leverage technology for addressing a novel problem in real-world settings. Because the propensity to do so effectively and efficiently is likely dependent on multiple sets of psychological variables and their interactions with a given task or environment (Ackerman et al., 2023), technological fluency can best be understood as a “trait complex” (i.e., a constellation of traits which together influence learning; Ackerman, 2003; see also “aptitude complexes”; Snow, 1989). The impact of technological fluency should be observable both in criterion measures of realistic problem-solving *performance* (i.e., an individual with high technological fluency should be more capable of leveraging technological resources to solve a novel problem than an individual with low technological fluency) as well as qualitative accounts of the problem-solving *process* (i.e., an individual with high technological fluency should demonstrate more effective strategies for leveraging technological resources to solve a novel problem than an individual with low technological fluency).

The second critical distinction between technological fluency and related constructs is related to assessment. Assessment of technological fluency requires both: (1) a task which is applicable to novel problems adults might encounter in the real-world, if not one that requires hands-on implementation; and (2) realistic access to technological resources. In contrast, computer and digital literacy can be assessed using non-realistic tasks (or in the case of computer literacy, realistic tasks that are unlikely to be novel to many adults – for example, navigating a file management system) and relatively controlled search environments. This is because the central focus of research on these constructs is individuals' interactions with specific features of technology, rather than their propensity to leverage the affordances of technological resources for real-world problem solving. The development of a criterion measure for technological fluency was the first major goal of the current study. Below, I review tasks used in prior research on online problem-solving and identify shortcomings of these tasks for assessing technological fluency.

2.1.2 Prior Assessments of Technology Use in Problem-Solving

One way that researchers have assessed individuals' technology-supported problem-solving with technology is to provide online assessments with a series of problem scenarios and ask participants to solve those problems using the technological resources provided (e.g., Fraillon et al., 2013). This is not fundamentally different than the approach which was used in the current study to assess participants' technological fluency. However, many existing assessments of this type have considerable limitations for assessing technological fluency as it is defined here (i.e., a propensity to leverage technological resources for solving novel real-world problems). Specifically, they may be limited in terms of the realism of technological resources (e.g., Internet access), the appropriateness of the task, and/or scoring processes. Examples of assessments which display one or more of these limitations are briefly reviewed below.

2.1.2.1 Assessments With Technology and Task Limitations

The first set of exemplar assessments (Greiff et al., 2013; Wüstenberg et al., 2014) were restricted in both access to technological resources and realism of the problem-solving tasks. The task paradigm used in these studies, MicroDYN (Greiff et al., 2012), required learners to deduce causal relationships between a set of variables by exploring a simulated microworld-style environment. Greiff et al. (2013) assigned participants to solve problems within the system related to the effects of advertising content on tourism rates in a fictional location, while Wüstenberg et al. (2014) used the MicroDYN system to create a scenario where participants assessed the relationships between different handball training methods and various team outcomes (e.g., power of throws, player exhaustion). The MicroDYN task paradigm and similar approaches were developed for research on primary and secondary students' complex problem-solving skills, including reasoning about dynamic or multivariate systems (e.g., Funke, 2001; Organization for Economic Co-operation and Development [OECD], 2010). While these tasks may be appropriate for assessing reasoning about complex systems, they are not well-suited for assessing technological fluency. This is because: (1) the task (i.e., manipulating a small number of variables to identify linear changes in an outcome variable) is not likely to be encountered in daily life; and (2) participants are not permitted to explore the Internet or other informational resources to help with generating a solution, which may not reflect problem-solving in real-world scenarios.

2.1.2.2 Assessments With Technology Limitations

A second class of online problem-solving assessments is more concerned with the development of realistic tasks, but again does not permit full use of real-world technological resources (e.g., the Internet). Fraillon et al. (2013), for example, described an assessment where participants were given access to a series of webpages and applications created by the researcher, and asked to use these resources to complete a series of tasks related to planning the design and construction of a garden. The resources provided to participants were intended to simulate real-world communication and informational resources that might support task completion (e.g., gardening-related internet forums, plant species databases), but participants were not permitted to actually conduct their own external searches beyond these resources. As noted in the previous section, digital literacy assessments may not require full Internet access to assess particular behaviors of interest (e.g., searching a database, identifying and avoiding false information) – these could be instead assessed within a controlled/simulated web environment such as the one constructed by Fraillon et al. (2013). This decision may have been made because the task was designed to assess 'computer and information literacy' (i.e., a construct similar to digital literacy). In that context, reduced variability in participants' strategic approaches to the task may actually be desirable to the extent that researchers are interested in participants' processing of some information rather than their ability to locate it in an uncontrolled environment. In contrast, I suggest that because individuals solving real-world problems would typically have access to the Internet, providing Internet access while assessing technological fluency improves the ecological validity of the assessment.

2.1.2.3 Assessments With Task Limitations

A final class of problem-solving assessments relevant to the current study includes those which do permit participants to engage in realistic information search but have task characteristics which make them inappropriate for assessing technological fluency. For example, Moehring et al. (2016) asked participants to use the Internet to answer a series of multiple choice and short answer (i.e., entering one word into a text box) items related to health and medical topics (e.g., using the Internet to identify a surgical procedure based on a single X-ray image). In this case, both the content of the task as well as its response format are unlikely to be useful for assessing an individual's propensity to use technology to solve a novel, real-world problem (Tidler et al., 2022). Other studies have come closer to designing problems which are more applicable to daily life (e.g., reasoning about food safety; Brand-Gruwel et al., 2009; understanding a relative's recent medical diagnosis; Sharit et al., 2015), but are purely informational (i.e., searching the Internet for a single answer to questions of varying complexity) in nature. In contrast, an assessment of technological fluency should include problem scenarios which have (at least in principle) some hands-on component for solution of the problem.

2.1.2.4 Assessments With Scoring Limitations

One study which sought to investigate the 'digital divide' in Internet skills both used relatively realistic problem-solving tasks and gave participants realistic access to technological resources (van Deursen et al., 2011). Even in this case, however, such an assessment may be insufficient for assessing technological fluency because it only scored participants based on the number of tasks successfully completed, and did not assess any process variables (e.g., how long participants took for each task, or qualitative differences in their approach to leveraging technology for support). The exclusion of process measures becomes problematic in the case of assessing technological fluency because this construct is concerned not only with whether a person ends up with a correct or satisfactory answer, but also with the efficiency and strategic quality of their means of problem-solving (Ackerman et al., 2023).

In contrast to van Deursen et al. (2011), Tidler et al. (2022) assessed both performance (i.e., successful completion) and process factors (i.e., time to completion, frequency of references to informational resources) for a hands-on computer upgrade task where participants were given Internet access in addition to other informational resources (e.g., technical manuals, a 'technical assistance' phone line staffed by research assistants). The current study captured similar process data (via screen recordings) which could support investigations of participants' strategic approaches to technology-supported problem-solving. However, although the hands-on task used by Tidler et al. (2022) provided good ecological validity, higher-fidelity procedural tasks are characterized by substantial time and effort demands for administration, making larger-scale assessment (e.g., sampling tasks across multiple domains; group administration) challenging. For this reason, the current study developed online criterion measures of technological fluency which were developed in an attempt to combine the added process focus of Tidler et al. (2022) with the simplified administration procedures of van Deursen et al. (2011).

2.1.3 Summary

At least two important distinctions can be made between technological fluency and similar constructs such as computer and digital literacy. First, technological fluency refers to a general propensity for leveraging technology in *real-world* problem-solving, while computer and digital literacy research focuses more narrowly on users' competence with tasks which can only be completed using technology and may or may not be highly relevant to solving novel problems encountered in daily life. Because technological fluency is concerned with applicability to daily life, the construct encompasses both 'can do' and 'will do' aspects of performance (Ackerman et al., 2023; for a related discussion on maximal versus typical behavior, see Cronbach, 1990). Second, assessments of individuals' technological fluency should be characterized both by: (1) novel real-world tasks and (2) open access to technological resources such as the Internet, so that respondents' behaviors reflect as much as possible how they might proceed with leveraging technology to solve issues encountered in the real world.

In reviewing prior research on online assessments of problem-solving, I found that most assessments fail to meet one or both of these criteria. One study (van Deursen et al., 2011) provided both realistic technology access and tasks reasonably applicable to daily life, but did not account for participants' search processes, meaning that the study did not provide a means by which to examine individuals' strategic approaches to problem-solving. More recently, Tidler et al. (2022) assessed both successful task completion as well as process variables, but the hands-on task used in that study does not easily translate to large-scale assessment. The current study sought to develop technological fluency criterion measures that address each of these issues:

1. Participants were allowed realistic (i.e., not restricted to designed or pre-approved resources) Internet access.
2. Problem-solving tasks were designed to represent scenarios that might realistically be encountered by the average person. Further, these tasks were: (a) complex enough that successful completion requires multiple steps; (b) scored based on demonstration of adequate problem-solving in addition to their identification of an acceptable solution; and (c) administered online.

I also identified sets of distal (e.g., stable traits) and more proximal (i.e., motivational and strategic influences) predictors of technological fluency. The next section reviews literature related to the first set of distal predictors.

2.2 Distal Trait Influences on Technological Fluency

The traits reviewed in this section (i.e., cognitive, affective, and conative traits) each have long histories which are beyond the scope of the current paper (for more in-depth discussion of how these classes of traits relate to adult behavior and skilled task performance, see Ackerman & Kanfer, 2004). In this section, I briefly provide an overview of empirical evidence for the relationship between each class of traits and performance. Where possible, I identified studies that have examined the relationship between a given trait and performance on tasks related to the current study. When studies using a comparable task are not available, I refer to data on broader performance outcomes (e.g., job performance). However, I note that the relationships between distal traits and technological fluency may be expected to be less strong than the relationships between the same traits and a broader outcome measure such as job performance. All else being equal, larger effects are expected when there is a match between the specificity of a given predictor variable and criterion variable (Wittmann & Suß, 1999), as is the case between broad trait assessments and higher-level performance measures.

2.2.1 *Cognitive Traits*

General cognitive ability is among the most well-studied predictors of performance across domains (e.g., job performance and educational achievement; Schmidt & Hunter, 2004). Two studies (Moehring et al., 2016; Sharit et al., 2015) have examined how individual differences in cognitive ability traits are related to adults' Internet search performance. Both asked participants to use the Internet to answer questions related to health and medical scenarios. Moehring et al. (2016) asked participants to use the Internet to identify a series of diseases, surgical procedures, or other medical terminology. They found that fluid intelligence (operationalized as spatial reasoning) accounted for approximately 56% of variance in performance on this assessment, although this dropped to approximately 35% of variance after accounting for computer experience (also see the previous section for limitations of the task– a more realistic task may have resulted in smaller contributions of fluid intelligence).

Sharit et al. (2015) gave participants an assessment consisting of problems which required searching for information online, and investigated relationships between search accuracy (i.e., the number of correct answers found) and several ability traits. Those most relevant to the current study include verbal ability ($r_{\textit{search accuracy}} = 0.39$), spatial ability ($r_{\textit{search accuracy}} = -0.28$), and processing speed ($r_{\textit{search accuracy}} = 0.28$). They also explored how individual differences in these traits related to several process variables: time spent searching, the number of distinct websites visited, and the number of transitions made between websites. Participants were categorized into 'high' or 'low' groups for each trait based on a median split, and *t*-tests were conducted to examine group differences in each search process variable.

Although median splits may be problematic (e.g., by obscuring or restricting variance in the sample; Cronbach & Snow, 1977), the results of this study do suggest some interesting process differences on the basis of individual differences in cognitive ability. Notably, there were (1) moderate effects of verbal ability, spatial ability, and processing speed on search time such that higher trait scores were associated with faster search (d = 0.57 – 0.63), (2) moderate to large effects on distinct website visits such that higher trait scores were associated with visiting more websites (d = 0.64 – 1.02), and (3) moderate to large effects of these three traits on transitions between websites such that higher trait scores were associated with more frequent switching (d = 0.57 – 0.85). Qualitative analysis of search strategies within high and low-ability groups also suggested that higher scores on ability trait measures tended to be associated with 'bottom-up' (e.g., strategic approaches to searching based on specific problem components) rather than 'top-down' (e.g., 'serendipitous' approaches to searching based on browsing links) search strategies. Finally, Sharit et al. (2015) noted that their conclusions about the effect of cognitive ability on internet search performance may have been limited by a lack of inclusion of more proximal variables.

2.2.2 *Affective Traits*

Affective (i.e., personality) traits are commonly used as non-ability predictors of performance (e.g., job performance; for a review see Sonnentag et al., 2008). For example, modern personality assessments in military selection contexts (see Stark et al., 2014) are largely based on the five-factor model of personality (Costa & McCrae, 1992). Because non-ability trait measures tend to assess typical rather than maximal behavior, they are most predictive in weak environmental situations or when behavior is aggregated over time (Ackerman & Kanfer, 2004; Rushton et al., 1983). However, non-ability traits may still be expected to have some utility for predicting more specific problem-solving behaviors.

Two traditionally assessed traits of particular interest to the current study are conscientiousness and openness to experience. I expected conscientiousness to be positively correlated with technological fluency because it incorporates tendencies toward (among other facets) achievement striving and deliberate action (Costa & McCrae, 1998), both of which should increase individuals' propensity to generate and enact well-planned strategies for solving a novel problem (see also Barrick et al., 1993 for a discussion of the relationship between conscientiousness and goal-setting). Likewise, individuals high in openness to experience tend to prefer intellectually stimulating experiences, and this preference may aggregate over time to reflect a greater investment of effort into the types of novel tasks which increase crystallized intelligence and potentially performance (von Stumm et al., 2011).

One study has examined personality traits as an antecedent of more general complex problem-solving capabilities. D'Zurilla and colleagues defined 'social problem-solving' as a set of cognitive, affective, and behavioral processes by which individuals attempt to resolve various problems which might be encountered in daily life (e.g., D'Zurilla & Nezu, 1990). D'Zurilla et al. (2011) investigated the utility of the Big Five personality traits for predicting several possible 'problem-solving styles.' They found that conscientiousness and openness to experience had small to moderate (Cohen, 1988) relationships with: (1) rational problem-solving, defined as deliberate and systematic approaches to identifying a solution ($r_{conscientiousness}$ = 0.34; $r_{openness}$ = 0.27); (2) impulsivity/carelessness ($r_{conscientiousness}$ = -0.35; $r_{openness}$ = -0.17); and (3) avoidance, defined as inaction toward identifying a solution ($r_{conscientiousness}$ = -0.43; $r_{openness}$ = -0.17).

2.2.3 Conative Traits

Conative (i.e., volitional) traits include motivational and interest components, and generally refer to traits informing individuals' willingness to invest effort (e.g., assuming low environmental press, interest in a domain is likely to be associated with greater effort on tasks in that domain; Ackerman & Kanfer, 2004). Holland's (1959, 1963) model of vocational interests is the predominant approach for assessing interest. Six interest themes are proposed in the model: (1) Realistic (i.e., physical, 'hands-on' tasks), (2) Investigative (i.e., tasks that help to "organize and understand" the world), (3) Artistic (i.e., tasks requiring aesthetic expression), (4) Social (i.e., interpersonal, helping tasks), (5) Enterprising (i.e., interpersonal tasks involving persuasion or leadership), and (6) Conventional (i.e., highly ordered tasks).

Although interests have received less attention in the selection literature than other non-ability traits (Van Iddekinge et al., 2011), they have experienced a resurgence in recent decades (e.g., Nye, 2022) and are now considered to be meaningful predictors of job/task performance. A meta-analysis by Nye et al. (2017), for example, found a moderate relationship ($r = 0.32$) between interest congruence (i.e., the extent to which an individual's relative interest scale scores are well-aligned with the interest profiles of their job) and job performance. Although no studies have assessed the relationship between interests and technology-supported problem-solving performance, realistic and investigative interests tend to be well represented by a 'Science/Math' trait complex which is in turn related to academic knowledge acquisition (Ackerman et al., 2011), including technology-related knowledge (Ackerman, 2003). Conventional interests may be negatively correlated with technological fluency because individuals high in this interest theme tend to prefer highly structured tasks (Holland, 1973), while novel problem-solving situations may be less well structured.

2.2.4 Summary

This section reviewed a set of traits comprising traditional approaches to predicting successful task performance. Cognitive, affective, and conative traits represent useful predictors of more general performance outcomes such as overall job performance (e.g., Nye et al., 2017; Schmidt & Hunter, 2004) and were expected to have some relationship with the technological fluency assessment developed in the current study. However, because they are not well-matched to the specificity of the technological fluency criterion, their contribution to its prediction was expected to be smaller than indicated in past research (Ackerman et al., 2023). I expected that a more proximal set of measures related to individuals' strategic approaches to technological problem-solving would explain greater amounts of variance in technological fluency. The development of these measures and the demonstration of their criterion-related validity for technological fluency above and beyond traditional selection batteries is the second major expected contribution of the current study. In the next section, I review literature which informed the development of these additional measures.

2.3 Proximal Influences on Technological Fluency

In contrast to distal cognitive, affective, and conative traits, which were expected to have more indirect effects on technological fluency, the current study also addresses several important proximal predictors of technological fluency. As discussed in the model of technological fluency proposed by Ackerman et al. (2023; see also **Figure 1**), three key individual differences expected to exert proximal influences on individuals' strategic approaches to problem-solving with technology include prior technology attitudes and experiences, self-efficacy for technological problem-solving (including the tendency to overestimate one's capabilities), and self-regulatory capabilities (e.g., the extent to which an individual is able to set and implement appropriate goals for completing the task).

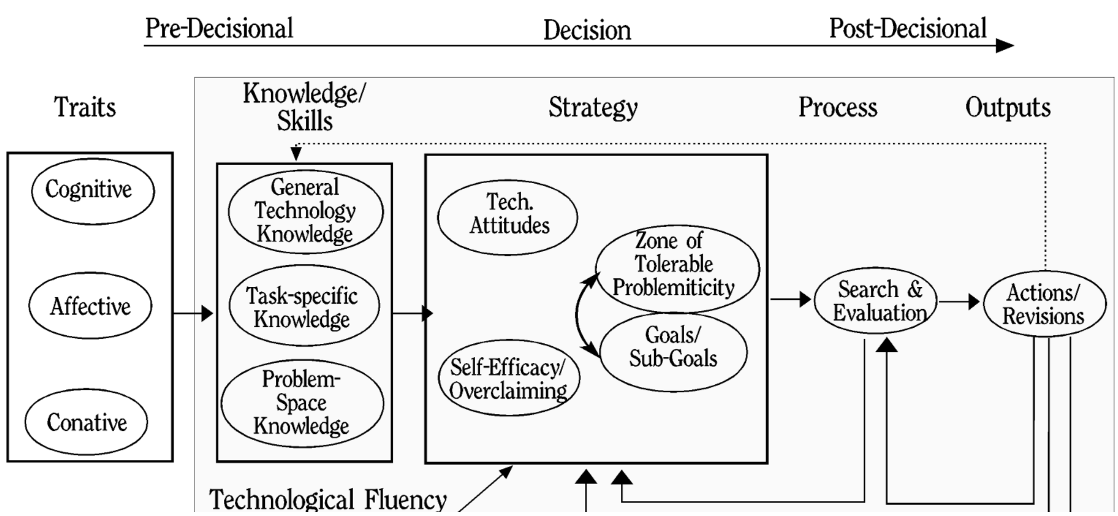


Figure 1 - A Model of Technological Fluency (Reprinted From Ackerman et al., 2023)

These constructs are discussed in the context of how they relate to individuals' willingness to engage with technology to develop solutions to novel problems. Each person has a range of difficulty (i.e., the 'zone of problematicity') within which a task is perceived to be both worthwhile of expending effort (i.e., not so easy that it is perceived as boring) and feasible (i.e., not so difficult that it increases anxiety or surpasses ability) (Elshout, 1987; see also Snow, 1989). Elshout (1987) described the zone of problematicity as a region of difficulty/complexity where ability and extant knowledge/skill is insufficient to complete the task – an individual must apply self-regulatory strategies (e.g., heuristics) in order to develop a novel solution. Below the threshold of this zone (also referred to as the algorithmic-heuristic threshold; Snow, 1989), performance is smooth and relatively automatic because the individual can rely fully on prior experiences to solve the problem. At difficulty levels far beyond the threshold, even the most successful problem-solving strategies begin to fail, and further progress is improbable.

Within an individual's zone of tolerable problematicity, however, there are appreciable differences in the extent to which the individual is capable of devising and enacting strategies which will permit the development of a solution to the problem. Non-ability individual differences such as the ones described in this section are expected to be important predictors of technological fluency because they should inform the likelihood that an individual is able and willing to apply their extant knowledge and skills to generate a solution to a novel problem. The interactions of these variables with features of the task environment may be particularly important for understanding problem-solving processes outcomes in real-world situations (such as the ones assessed in the current study) because people may have accumulated typical strategic responses to similar situations they have previously encountered (Snow, 1989). In the remainder of this section, I review relevant literature related to each of the three key constructs specified above and explain how individual differences in these constructs are expected to affect technological fluency through their influence on individuals' zone of tolerable problematicity.

2.3.1 Technology Attitudes and Experiences

2.3.1.1 Technology Attitudes

Attitudes (i.e., the tendency to respond to some object with favorable or unfavorable evaluations; Ajzen & Cote, 2008) are posited to be a major factor influencing (through increased intentions to perform a behavior) specific actions undertaken by an individual (Ajzen, 1985). Because the current study aimed to predict the likelihood that an individual is able and willing to effectively and efficiently leverage technology to solve problems, it is important to assess participants' attitudes toward the use of technological aids.

One popular model of technology attitude-behavior links is the Technology Acceptance Model (TAM; Davis, 1989). Building on the Theory of Reasoned Action (Fishbein & Ajzen, 1977) and the Theory of Planned Behavior (Ajzen, 1985), which suggest that attitudes are a core driver of intentions to engage in a particular behavior, TAM posits that intentions to adopt/use technology are in part a function of the user's perceived technological utility and ease of use (i.e., the extent to which they believe that the target technology will both provide value and not require undue effort) (Davis, 1989). Around the beginning of the 21st century, this model was applied to study individual differences in Internet adoption and utilization. Spacey et al. (2004), for example, reported a descriptive study of library staff attitudes toward the use of information and communication technology (including the Internet) at work, finding generally positive attitudes which were interpreted in the lens of TAM to suggest a general willingness of the sample to integrate the Internet into their daily work tasks. Shih (2004) found a strong positive relationship ($r = 0.54$) between Internet attitudes and self-reported performance on work tasks requiring the Internet. More recently, van Laar et al. (2019) reported that perceived ease of use was the strongest predictor of Internet-based problem-solving in a set of other personal (e.g., demographics, goal orientation) and social (e.g., social support) variables.

TAM is often applied to focus on prediction of a single behavior point (i.e., the decision to adopt or not adopt a technological tool for a certain use case) but has not been widely used to study the qualitative process differences that occur *during* use of the technology. The current study extends prior research on technology attitudes by examining their influence on both the ability and the willingness to engage in the iterative problem-solving processes which constitute technological fluency. Although positive technology attitudes may not actually shift the zone of problematicity (i.e., they are unlikely on their own to cause an individual to experience a given technology problem-solving task as less complex), they may have a meaningful impact on individuals' strategic approach to the task. For example, the presence of positive attitudes toward technological aids' utility and ease of use may increase the likelihood that an individual chooses to continue rather than terminate search processes upon experiencing difficulty finding a solution to a task.

2.3.1.2 Technology Experiences

It is worth noting that an individual's history of direct experience with the object of an attitude is expected to influence the development of that attitude over time (Ajzen &

Fishbein, 2005; Fazio, 1995). In the case of the current study, the relevant experiences are prior instances of using technological aids to solve real-world problems. Tidler et al. (2022) provided initial support for the beneficial role of technology experiences on the problem-solving tasks of interest to the current study, finding significant relationships between computer experience and performance on a hardware upgrade task ($d_{successful\ vs.\ unsuccessful}$ = -0.82) as well as process variables ($r_{experience,\ completion\ time}$ = -0.45).

Although technology experiences are not explicitly identified as a pre-decisional or decisional factor in Ackerman et al.'s (2023) model of technological fluency, relevant experiences are expected to develop as a function of more distal (i.e., pre-decisional stage) factors and to influence decisions to engage with technological resources in concert with proximal (i.e., decisional stage) factors. That is, experiences are correlated with distal ability and non-ability traits that influence intellectual development through individuals' selection into situations well-matched to their aptitudes (for more discussion see Ackerman, 1996). Experiences may also raise individuals' threshold of problematicity by increasing relevant knowledge (e.g., knowledge of how to navigate technological resources to locate required information; Ackerman et al., 2023). When considering 'can do' and 'will do' influences on technological fluency, in other words, technology attitudes might be considered a 'will do' factor that develops in part out of prior experiences (which contribute to 'can do' factors).

2.3.2 Self-Efficacy and Overclaiming

From a motivational perspective, individuals' efficacy for technological problem-solving is also expected to be an important predictor of technological fluency (Ackerman et al., 2023). Tidler et al. (2022), for example, found a moderate negative correlation (r = -0.40) between internet search self-efficacy and completion time for a task involving several computer hardware upgrades, where participants were permitted to search the Internet for guidance. Self-efficacy judgments focus primarily on whether one has the capability to engage in behaviors related to successful task completion (Bandura, 1977) and are good predictors of performance on relatively specific tasks (as opposed to broader/aggregated outcomes; Bandura, 1997). For example, self-efficacy has been estimated to have a moderate to strong relationship with academic performance (r = 0.38; Multon et al., 1991; r = 0.59; Richardson et al., 2012) and persistence (r = 0.34; Multon et al., 1991). Positive self-efficacy beliefs are posited to increase use of self-regulatory strategies and engage motivational processes related to effort expenditure (Bandura, 1991, 2013; Schunk & DiBenedetto, 2021).

An important assumption of the expected positive relationship between efficacy and performance, however, is that efficacy judgments are relatively accurate (i.e., individuals are assumed to have a reasonable idea of what they can do). A learner who has substantially overestimated their competence, for example, may apply insufficient effort when facing a challenging problem because they do not expect to have to work hard to devise a solution; this reduces their likelihood of succeeding and may reduce their willingness to attempt similar tasks in the future (e.g., Schunk & Pajares, 2009). In sum, the tendency to overclaim (i.e., over-estimate one's competence) was of interest to the current study because inaccurate judgments may cause individuals to adopt ineffective strategic approaches.

Overclaiming measures assess the tendency of an individual to claim knowledge they do not actually have (e.g., Ackerman & Ellingsen, 2014; Paulhus et al., 2003). In one form of this assessment (Paulhus et al., 2003), participants are provided with a checklist of items (20% of which are false but plausible 'foils') and asked to indicate known items. Overclaiming scores are then calculated as a function of the number of foils selected. As noted by Ackerman and Ellingsen (2014), however, this technique has two key methodological limitations: (1) there is no validation of respondents' actual knowledge, and (2) the impact of overclaiming on foil selection may be confounded with foil plausibility – that is, the same individual may be unequally likely to select two foils if they are substantially different in plausibility. The latter authors posit that more appropriate assessments of overclaiming should provide only real items and compare participants' claimed knowledge against an objective test of the same items.

Although most research on overclaiming has assessed knowledge of various academic and popular culture domains (e.g., vocabulary; Ackerman & Ellingsen, 2014; multiple domains ranging from literature to physical sciences; Paulhus et al., 2003), a few studies have explored overclaiming in technology (Gravill et al., 2006; Vonkova et al., 2021) and information literacy (Pennycook & Rand, 2020). However, these studies largely rely on problematic assessments (e.g., operationalizing overclaiming as claiming foils). The current study included a measure of self-efficacy in which participants were asked to rate their confidence in their ability to provide solutions of increasingly high quality for the first six scenarios on the technological fluency assessment. They were not informed that they would later be asked to develop solutions for these problem scenarios. This supported potential exploration of individual differences in the tendency to overclaim technology problem-solving capabilities, using a method which (1) does not require foils and (2) can be validated by an assessment of respondents' actual ability to use technology to solve novel/challenging problems.

2.3.3 Self-Regulatory Capabilities

Self-regulatory capabilities (e.g., in planning or self-monitoring) were the final set of major strategic influences on technological fluency addressed in the current study (Ackerman et al., 2023). Prior research has identified self-regulatory skill as a factor which can separate minimally satisficing problem-solving solutions from more effortful, optimized solutions. Using the example of daily informal reasoning (another real-world task), for example, Perkins et al. (1991) posited that individuals often cut short their exploration of a problem space at the first reasonable solution (i.e., a "makes-sense epistemology"). While such solutions may be adequate for simple or relatively inconsequential problems, they are often suboptimal compared to more carefully or thoroughly considered solutions (i.e., a "critical epistemology"). Perkins and colleagues (1991) suggested that individual differences in self-regulatory knowledge and skills (i.e., "heuristic know-how", p. 100) may explain why people do not reach optimal solutions. Consistent with Elshout (1987), this perspective suggests that for novel, real-world problems which go beyond the reach of an individual's extant knowledge or expertise, the willingness to expend effort on strategic problem-solving may be more important than general cognitive abilities alone.

One self-regulatory behavior relevant to the current study is breaking down a complex or novel task into simpler or more familiar component parts (i.e., subgoals). Goals and subgoals are valuable in problem-solving because they simplify a problem state so that (1) each component part of the solution is more attainable and (2) the individual has some direction in how to navigate the task (e.g., Catrambone, 1998; Simon & Newell, 1971). However, identifying useful subgoals may be a challenging or effortful task in and of itself. Research on self-regulated learning (SRL) in achievement contexts suggests that many learners may lack either sufficient prior knowledge or self-regulatory skills required to engage in early-phase SRL activities such as subgoal formation and strategic planning (Pressley et al., 1997), inhibiting their ability to engage in effective metacognitive strategies during learning (Zimmerman & Kitsantas, 2005).

Accordingly, individual differences in self-regulatory capabilities were expected to predict performance on a measure of technological fluency. Breaking a novel task down into manageable subgoals may bring the larger task within an individual's zone of problematicity if the subgoals represent more familiar tasks which allow learners to draw on existing knowledge or skills (Ackerman et al., 2023). Consider an individual who is attempting to use the Internet to troubleshoot a faulty connection between a monitor and laptop. If they are unsure how to proceed with an initial search, they may be described as being at an 'impasse' (VanLehn, 1988), which can be resolved by identifying smaller, more easily searchable subgoals. They might, for example, first engage with the subgoal of learning how to identify a faulty cable or check for physical obstructions such as dust in the laptop ports. In addition to effects on problem-solving outcomes (i.e., *whether* a solution is reached), individual differences in the ability to generate subgoals may also result in appreciable effects on problem-solving process (i.e., *how* a solution is reached) (e.g., see Elio & Scharf, 1990).

Although the most direct measurement of this capability would be to have an individual generate a series of goals and subgoals for a sample task and score their responses against an 'expert' task analysis, this may introduce challenges. In addition to being labor intensive, this approach may confound knowledge of the sample task domain with more general tendencies toward being willing to engage in effortful self-regulatory behavior. This is because domain expertise may make it easier to identify subgoals for tasks related to that domain (Gick, 1986). Accordingly, another option is to use self-report measures of self-regulation which are focused on goal-setting and planning outside of any individual task. Schraw and Dennison's (1994) Metacognitive Awareness Inventory (MAI), for example, includes a planning scale (an example item is "I set specific goals before I begin a task.").

2.3.4 Summary

In this section, I proposed technology attitudes and experiences, overclaiming of technological fluency, and self-regulatory capabilities as key proximal influences on individuals' strategic approaches to solving novel problems with technology. Collectively, I expected these factors to outperform more distal cognitive, affective, and conative trait measures as predictors of technological fluency. I expected each of these constructs to influence problem-solving processes and outcomes as a function of their interactions with an individual's zone of problematicity (i.e., the range of task difficulty/complexity within which finding a solution is feasible but not automatic; Elshout, 1987).

Technology attitudes, for example, may exert motivational effects on search processes such that individuals with more positive attitudes towards the utility and usability of such tools are more persistent in search efforts, while prior experiences are a potential source of relevant problem space knowledge (i.e., knowledge about how to navigate informational technological resources; Ackerman et al., 2023) that could raise an individual's threshold of problematicity (i.e., reduce the perceived difficulty of a given task). Self-regulatory capabilities were also expected to be positively associated with technological fluency. Although greater self-regulatory capabilities might not alter an individual's zone of problematicity for a task as a whole, individuals who are willing to expend sufficient effort may be able to identify and pursue a series of simpler or more familiar subgoals which each exist below the top of the zone of problematicity. In contrast to these positive effects, overclaiming may be negatively associated with technological fluency to the extent that inaccurate perceptions of capability lead individuals to adopt inappropriate problem-solving strategies or falsely believe low levels of effort will be sufficient to find a solution.

CHAPTER 3. THE CURRENT STUDY

Modern workplaces are characterized by the necessity of just-in-time problem-solving (Neubert et al., 2015), often aided by technology and occurring in domains with which adults may have had widely varying prior experience. Technological fluency is a trait complex describing individuals' propensity (i.e., ability and willingness) to leverage informational technology resources to solve real-world problems relevant to this need for just-in-time learning (Ackerman et al., 2023).

The first major goal of the study concerned the introduction of the technological fluency construct. The delineation of this construct constitutes a theoretical contribution to the literature on real-world problem-solving performance, which is particularly important as informational technology resources continue to be integral to 21st century work and adult life. A criterion measure of technological fluency was designed to avoid task-based, technology-based, or scoring limitations of assessments previously used to assess conceptually related constructs such as digital literacy. The measure includes a series of real-world problem-solving tasks covering several domains (e.g., fitness, home repair) that the average person might encounter and for which they may or may not have prior relevant experience. For example, one item on the measure asks participants to devise a way to repair sizable holes left from wall-mounting a television. Given open access to the Internet, participants were asked to provide a solution to each problem (for which there may be one or more optimal solutions). Performance on this open-ended task was scored as a function of the efficiency and effectiveness of the solution (e.g., is the solution optimal rather than minimally satisficing, how well does it account for situational constraints of the problem, etc.). Individuals with a higher level of technological fluency were expected to be more able to leverage the Internet to devise a high-quality solution.

The second major goal of the current study concerned the development of a set of distal and proximal influences on technological fluency. Four key variables (technology attitudes and experiences, self-efficacy, and self-regulatory capabilities) were expected to be important proximal predictors. Building on theoretical and empirical research on how these variables influence achievement and task performance in other contexts, I developed or adapted measures of each that were designed specifically for the context of predicting technological fluency. Given the potential practical implications of technological fluency as an in-demand work competency, I also aimed to demonstrate the criterion-related validity of the proximal strategic measures above and beyond distal traits typically assessed in selection contexts (e.g., general cognitive ability, personality traits, vocational interests).

3.1 Hypotheses

As described above, one major goal of the current study was to evaluate potential distal and proximal factors, which were expected to have differential relationships with performance on an assessment of technological fluency. Generally, I expected that a pattern would emerge among the proposed predictors of technological fluency such that proximal factors are stronger predictors than distal traits. Hypotheses 1-7 provide specific predictions about these relationships and were based on expectations related to convergent validity for proximal predictors as well as discrimination between proximal and distant predictors[1]. Hypotheses 8-9 provide predictions about the amount of variance in technological fluency accounted for by the sets of distal and proximal predictors. See **Table 1** for a summary of hypotheses.

Table 1- Summary of Hypotheses Related to Convergent, Discriminant, and Criterion-Related Validity

Hypotheses	Aim of Hypotheses
Hypotheses 1-3 Related to discriminant validity	Predict the extent to which each of the distal predictors that is expected to have a weaker relationship with technological fluency than proximal factors are actually less strongly related.
Hypotheses 4-7 Related to convergent validity	Predict the extent to which each of the proximal predictors that is expected to have a stronger relationship with technological fluency than distal traits are actually more strongly related.
Hypotheses 8-9 Related to criterion-related validity	Predict: (1) the proportion of variance in technological fluency accounted for by the set of distal traits, and (2) the proportion of variance in technological fluency accounted for by the set of proximal factors, above and beyond that accounted for by distal traits.

3.1.1 Hypotheses Related to Distal Predictors of Technological Fluency

Consistent with a proposition made by Ackerman et al. (2023) and adopted in the current study (see *Hypothesis 9*), which is that cognitive ability traits are expected to be statistically significant but ultimately weaker predictors of technological fluency compared to more proximal variables, I expected that:

[1] In addition to the traits expected to have modest but significant relationships with technological fluency, there are also several personality and interest traits not addressed in this section which are not expected to be meaningfully related to technological fluency. Because the technological fluency assessment does not require social interaction, for example, I did not expect that extraversion or social interests would have the same criterion-related validity as more theoretically related traits such as openness to experience or investigative interests. However, because the latter are already expected to have relatively modest relationships with the criterion measure, detections of differences between theoretically relevant and irrelevant distal traits would require sample sizes exceeding 1,000. For this reason I did not make predictions regarding discrimination between these traits.

Hypothesis 1: Cognitive ability will be less strongly associated with technological fluency than proximal variables ($r_{cognitive\ ability,\ TF}$ = 0.20; $r_{proximal,\ TF}$ = 0.50; Δr = 0.30).

I expected that conscientiousness and openness to experience would both be positively related to performance on an assessment of technological fluency, due respectively to the former's tendency toward achievement striving (Costa & McCrae, 1998) and the latter's tendency toward participation in novel or intellectually stimulating experiences (e.g., von Stumm et al., 2011).

Hypothesis 2a: Conscientiousness will be less strongly associated with technological fluency than proximal variables ($r_{conscientiousness,\ TF}$ = 0.20; $r_{proximal,\ TF}$ = 0.50; Δr = 0.30).

Hypothesis 2b: Openness to experience will be less strongly associated with technological fluency than proximal variables ($r_{openness\ to\ experience,\ TF}$ = 0.20; $r_{proximal,\ TF}$ = 0.50; Δr = 0.30).

Although no empirical studies have assessed the relationships between interest traits and performance on technology-supported problem-solving tasks, there are theoretical reasons to expect effects associated with several of Holland's (1959, 1963) vocational interest themes. In the absence of empirical evidence to support estimates of effect size, I predicted a similar magnitude of relationships for investigative, realistic, and conventional interests as was hypothesized for other distal traits.

Hypothesis 3a: Investigative interests will be less strongly associated with technological fluency than proximal variables ($r_{investigative,\ TF}$ = 0.20; $r_{proximal,\ TF}$ = 0.50; Δr = 0.30).

Hypothesis 3b: Realistic interests will be less strongly associated with technological fluency than proximal variables ($r_{realistic,\ TF}$ = 0.20; $r_{proximal,\ TF}$ = 0.50; Δr = 0.30).

Hypothesis 3c: Conventional interests will be less strongly associated with technological fluency than proximal variables ($r_{conventional,\ TF}$ = -0.20; $r_{proximal,\ TF}$ = 0.50; $\Delta|r|$ = 0.30).

3.1.2 Hypotheses Related to Proximal Predictors of Technological Fluency

Problems that cannot be solved based on existing knowledge and skills require the individual to use self-regulatory strategies (i.e., heuristics) to devise a solution (Elshout, 1987). Although no studies have assessed the relationship between subgoal generation and technology-supported problem-solving as defined in the current study, I predicted a large effect based on prior research of the effects of strategies such as goal-setting on task performance (e.g., d = 0.52 – 0.82; Locke & Latham, 1990).

Hypothesis 4: Self-regulatory skills will be more strongly associated with technological fluency than distal traits ($r_{self\text{-}reg,\ TF}$ = 0.50; $r_{distal,\ TF}$ = 0.20; Δr = 0.30).

Dominant models of influences on behavior (e.g., Ajzen, 1985; Fishbein & Ajzen, 1977) suggest that attitudes are a critical antecedent of intentions to engage in various actions. This has also been supported empirically in a context relevant to the current study: Shih (2004) found that positive Internet attitudes were strongly (r = 0.54) correlated with performance on job-relevant tasks requiring use of the Internet. I expected that technology attitudes would be strongly related to performance on an assessment of technological fluency.

Hypothesis 5: Technology attitudes will be more strongly associated with technological fluency than distal traits ($r_{attitudes,\ TF} = 0.50$; $r_{distal,\ TF} = 0.20$; $\Delta r = 0.30$).

Because prior experiences relevant to problems on an assessment of technological fluency increase knowledge about the domain or problem space, they were expected to be strong predictors of performance in the current study. It is worth noting that Sharit et al. (2015) did not find a significant relationship between Internet experience and problem-solving performance. However, this may have been a result of the operationalization of experience as 'regular frequency of using the Internet' rather than 'experience using the Internet to solve problems', as well as range restriction in the experience measure (over 70% of the sample reported typically using the Internet at least once a day, and less than 10% reported using it less than once a week). Given these issues with the measurement of experience in Sharit et al. (2015), I expected to replicate Tidler et al.'s (2022) finding that experience is a strong predictor of technological fluency.

Hypothesis 6: Technology experience will be more strongly associated with technological fluency than distal traits ($r_{experience,\ TF} = 0.50$; $r_{distal,\ TF} = 0.20$; $\Delta r = 0.30$).

Finally, I expected that self-efficacy would be positively related to performance on the assessment of technological fluency (see **Figure 1**; conceptual model from Ackerman et al., 2023). Consistent with both evidence on self-efficacy's relationship with more distal measures of academic performance (e.g., grades; Richardson et al., 2012) as well as more specific, novel problem-solving tasks (e.g., Evans et al., 2021), I expected a moderate to large relationship with technological fluency:

Hypothesis 7: Self-efficacy will be more strongly associated with technological fluency than distal traits ($r_{overclaiming,\ TF} = 0.50$; $r_{distal,\ TF} = 0.20$; $\Delta r = 0.30$).

3.1.3 Hypotheses Related to Criterion-Related Validity

Finally, because the proximal predictors included in the current study (e.g., technology attitudes, ability to identify subgoals within a novel problem, etc.) are (1) better matched to the specificity of the criterion (Wittmann & Suß, 1999) and (2) are expected to have more direct effects on individuals' strategic approaches to problem-solving, I expected that the set of proximal variables would increment prediction of performance on the technological fluency assessment above and beyond the variance accounted for by distal traits alone.

Hypothesis 8: Distal predictors alone will account for 10-15% of variance (i.e., $0.11 < f^2 < 0.18$) in the technological fluency assessment.

Hypothesis 9: Proximal predictors will account for a moderate ($0.15 < f^2 < 0.35$; Cohen, 1988) increase in variance accounted for in the technological fluency assessment above and beyond the distal predictors.

It is worth noting that the predictors of interest for the current study were not expected to capture 100% of variance in the assessment of technological fluency. In fact, substantial additional variance was expected to be attributable to other factors, including knowledge structures and motivational factors (**Figure 1**; for additional discussion, see Ackerman et al., 2023). Although investigations of the predictive utility of these additional individual differences are beyond the scope of the current study, they are listed in **Table 2** to provide a more complete (although non-exhaustive) conceptual accounting of potential influences on technological fluency which might be addressed in future research.

Table 2- Additional Proposed Influences on Technological Fluency

Knowledge-Related Predictors	Motivational Predictors
• Traditional g_c Measures	• Task Engagement
• General Technology Knowledge	• Motivational Skills
• Task-Specific Knowledge	• Approach/Avoidance Goals
• Problem-Space Knowledge	• Self-Concept

3.2 Exploring Variation in Problem-Solving Processes

The hypotheses above focus on individuals' performance on an assessment of technological fluency, which reflects the previously stated primary aims of the current study (i.e., developing an assessment of technological fluency, developing a battery of proximal predictors and assessing the criterion-related validity of this battery above and beyond more traditional measure). However, I expect that in addition to variance in performance on the technological fluency assessment, there will be qualitative differences in the processes or strategies used by participants as they engage with problems. I also expect that these strategic approaches may be associated with performance.

Two studies have examined process differences in technology-supported problem-solving. Tidler et al. (2022) found that participants who successfully completed a computer hardware upgrade task referenced task instructions (d = 0.64) and the Internet (d = 0.78) less frequently than their unsuccessful peers. Sharit et al. (2015) qualitatively compared Internet search strategies for the five highest scoring and the five lowest scoring individuals on their Internet search task and found that the low performers: (1) made fewer abstractions from the problem instructions when selecting search keywords, and (2) seemed to be less 'stable' in their problem orientation (i.e., more easily lost, disoriented, or distracted). Because there is limited prior research on this topic, I do not make specific predictions about the process measures that are expected to most clearly differ between high and low scorers on the technological fluency assessment. Instead, I ask the following exploratory research question:

> *Research Question 1:* What qualitative differences in problem-solving processes exist between high and low performers on an assessment of technological fluency?

CHAPTER 4. METHOD

4.1 Participants

Participants were recruited from the population of Georgia Tech undergraduates and received four hours of SONA course credit as compensation for participation. I conducted a series of power analyses using G*Power (Faul et al., 2009) to assess the sample size required to reliably detect each predicted effect. The largest sample was required to detect a conservative estimate of the effect predicted in Hypothesis 8. Specifically, a sample of 131 participants was required to achieve 0.80 power for a regression model explaining 10% of variance ($f^2 = 0.11$) in technological fluency. Accordingly, I sought to recruit at least 140 participants to account for potential data loss (e.g., insufficient effort, incomplete data).

In total, 152 individuals participated in the study. Two procedural changes (revision of instructions to participants, as well as minor revisions to the assessment) were made after the first three participants completed the study. These three responses were treated as additional pilot data and excluded from further analysis. Additionally, data from sixteen participants were excluded prior to analysis for the following reasons: significant outliers on criterion assessment (N = 7), careless/implausible responding and/or insufficient effort (N = 4), failing to return for the second session (N = 2), failure to follow instructions (N = 2), or insufficient English proficiency (N = 1). After exclusions, the final sample consisted of 133 participants. The average age within the sample was 19.14 years (*SD* = 1.21 years; range = 18 – 23), and the sample's distribution of gender and race/ethnicity is provided in **Table 3**. Most participants reported their academic majors as either computer science/computational media (43.8%) or one of several engineering disciplines (35%; disciplines include mechanical, industrial/systems, biomedical, etc.). Additional majors included the natural sciences (3.6%), humanities (4.3%), physical sciences (3%), social sciences (3.6%), business (3%), and mathematics (1%).

Table 3- Sample Gender and Race/Ethnicity

	Female	Male	Non-Binary	Total
Asian	17	29	0	**46 (34.6%)**
African American/Black	8	7	0	**15 (11.3%)**
Caucasian/White	13	36	1	**50 (37.6%)**
Hispanic/Latino	0	2	0	**2 (1.5%)**
Other (self-specified)	1	0	0	**1 (0.8%)**
ME	*1*	*0*	*0*	**-**
2+ Reported	5	14	0	**19 (14.3%)**
AA/Bl.; Asian	*0*	*1*	*0*	**-**
AA/Bl.; His./Lat.	*2*	*0*	*0*	**-**
Asian; Cau./Wh.	*1*	*4*	*0*	**-**
Asian; Cau./Wh.; His./Lat.	*1*	*0*	*0*	**-**
Asian; His./Lat.	*0*	*1*	*0*	**-**
His./Lat; Cau./Wh.	*0*	*8*	*0*	**-**
Other (ME); Cau/Wh.	*1*	*0*	*0*	**-**
Total	**44 (33.1%)**	**88 (66.2%)**	**1 (0.8%)**	**133 (100%)**

Notes. Rows in italics provide additional descriptive detail on the multiracial and self-specified "other" categories. Totals are calculated by the sum of non-italicized frequencies. ME = Middle Eastern.

4.2 Measures and Materials

4.2.1 Technological Fluency Assessment

The assessment of technological fluency required participants to solve open-ended problems that an average adult might encounter in daily life across a variety of domains.

The allowed time to complete the speeded assessment was 90 minutes (1.5 hours). Participants were given access to an Internet-enabled computer and instructed that they may use the Internet to solve the problems, but that they may not enlist social support (e.g., texting a friend) or use their personal devices. The assessment included 21 items, which were pilot tested by research assistants and refined before full data collection began (e.g., to ensure that it was very difficult to provide fully correct, high-quality answers to all items in the allotted time). The items were designed to cover a variety of domains, including: (1) Furniture and home maintenance/repairs, (2) Making simple appliance repairs, (3) Home and personal technology (e.g., cameras, smart home tools), (4) Personal care and aesthetics (e.g., home decoration, mending clothes), (5) Home organization and space optimization, and (6) Planning self-improvement projects (e.g., medium- to long-term education or fitness goals). An example item from the first category is shown in **Figure 2** below, and a full copy of the assessment is included in **Appendix A**.

The order of items on the assessment was determined based on several factors. First, it is important to note that item order was not counterbalanced across participants. The major reason for this was because the first six items were used as problem scenarios for the self-efficacy measure, so I wanted to ensure that all participants were able to respond to these items. The other primary considerations related to item order included varying the distribution of (1) items with content from different domains and (2) easy versus difficult items. The former was intended to reduce confounds with unmeasured variables such as domain knowledge in the self-efficacy measure and aggregate assessment score. For example, if all the first six problems were related to DIY home repair, someone with higher knowledge of this domain would likely report higher self-efficacy regardless of their more specific expectations of their ability to use technology to develop solutions. Likewise, if this domain were clustered at the beginning of the assessment, someone with extensive prior knowledge in this domain may have received a higher aggregate score because they did not need to conduct as much search to develop high-quality answers (therefore leaving them with more time to work on additional problems). Distributing item content is unlikely to have eliminated potential confounds of domain knowledge (see Chapter 9 for additional discussion) but may have reduced its impacts on the results. Likewise, items were intentionally not ordered from easy to difficult. This was intended to increase variability in the self-efficacy measure which, as mentioned above, used the first six items from the assessment (as presumably, self-efficacy judgments would be consistently high for easy items and consistently low for difficult items).

2. When you moved into your current apartment, you caused some damage (pictured below) while mounting electronics and hiding cables. You want to repair this damage before moving out to avoid being charged part of your security deposit.
    a. What tools and supplies will be needed to fix the damage? Provide links to all items you would need for the repair.
    b. What procedures are involved to fix the damage?

Figure 2 - Example Item From Technological Fluency Assessment

Each item was scored by two raters on a scale of 0-5, where 0 corresponded to a missing answer, and attempted problems received a score of 1 (anchors of "completely inadequate solution", letter grade 'F') to 5 (anchors of "excellent to optimal solution", letter grade "A"). In addition to these generic anchors, a rubric was created for each problem that contained more content-specific criteria at each anchor point. For each rubric, an optimal solution that addressed all problem constraints was scored higher than a minimally satisficing solution, which was in turn scored higher than a poor solution or no attempted solution. In the item shown above, for example, a solution that distinguished between materials that would be sufficient to spackle a small nail hole versus those sufficient to patch and repair a large opening would be scored higher than one which does not make that distinction. The final item score was the average of the two independent ratings, and the total score on the assessment was the sum of the average ratings. For additional detail on the calculation of interrater reliability in the final sample, see the portion of the results section which presents preliminary and descriptive analyses.

4.2.2 Distal Predictors

4.2.2.1 <u>Cognitive Traits</u>

Several specific cognitive abilities typically related to performance on complex tasks (i.e., verbal ability, spatial ability, numerical ability) were expected to be related to technological fluency. Participants completed a battery of ability tests sampling from these domains. The content of these tests is summarized in **Table 4**. Although not directly relevant to the formal hypotheses, self-concept was also assessed for these domains using the self-concept measures developed by Goff (1994) and further refined by Ackerman et al. (2001), as well as a new measure of technology-related self-concept (example item: "I am good at navigating unfamiliar websites or other online resources to locate information that I need"). All self-concept scales used a 6-point Likert-type response scale (1-6 scale; Strongly disagree – Strongly agree).

Table 4- Assessments in the Cognitive Ability Battery

Domain	Assessment (Length)
Verbal	• Extended Range Vocabulary (7 minutes) • Multidimensional Aptitude Battery (MAB) - Similarities (8 minutes) • MAB Comprehension (9 minutes)
Spatial	• Verbal Test of Spatial Abilities (12 minutes) • Spatial Analogies (12 minutes) • Paper Folding (8 minutes)
Numerical	• Math Knowledge (13 minutes) • Number Series (6 minutes) • Math Approximation (10.5 minutes)

4.2.2.2 Affective Traits

Personality traits were assessed using the Conscientiousness and Openness to Experience scales from the NEO-FFI (Five-Factor Inventory; Costa & McCrae, 1992). Each scale included 12 items, and participants reported responses on a Likert-type scale (1-6; Very untrue of me – Very true of me).

4.2.2.3 Conative Traits

Vocational interests were addressed using the Realistic, Investigative, and Conventional scales from the O*NET Interest Profiler Short Form (10 items each; Rounds et al., 2010). Participants were asked to indicate their preference for various activities on a Likert-type response scale (1-6; Strongly dislike – Strongly like). In addition, participants completed the short form of the Motivational Trait Questionnaire (MTQ), which included six scales for assessing personal mastery, competitive excellence, and anxiety-related motivation (Kanfer & Ackerman, 2000; items scored on a 1-6 scale; Strongly disagree – Strongly agree). Although the MTQ was not hypothesized to explain substantial additional variance in the criterion measure, its scales may provide evidence for convergent validity for other assessed variables.

4.2.3 Proximal Predictors

4.2.3.1 Technology Attitudes.

To assess participants' attitudes toward technology, I adapted a scale from Ackerman (2021). Participants were asked to indicate whether they thought given technological tools were overall detriments or improvements to the world using a 6-point

Likert-type scale ranging from “All or nearly all bad” to “All or nearly all good.” Extant items that had at least some relevance to the current study (e.g., “YouTube”) were retained, and new items were added that also reflected the focus of the current study (e.g., “Auto-complete search assistance”). See **Appendix B** (**Table B1**) for all items.

4.2.3.2 Prior Technology Experiences

I developed a self-report measure to assess participants’ prior experiences related to using technology to solve problems. This measure was designed to cover experiences related to technological problem-solving across the content domains included in the criterion measure. Consistent with extant methods of assessing biodata-type items (e.g., Owens & Schoenfeldt, 1979; Tesluk & Jacobs, 1998), I included both a qualitative experiences scale (i.e., assessing the extent to which participants agreed that a certain behavioral tendency has generally been true of them in the past) as well as a quantitative experiences scale (i.e., assessing the frequency at which participants reported engaging in particular experiences or behaviors). The former was assessed on a 6-point Likert-type scale ranging from “Very untrue of me” to “Very true of me,” while the latter was assessed on a 6-point Likert-type scale ranging from “Never” to “Very frequently.” See **Appendix B** (**Table B2**) for all items on both scales.

4.2.3.3 Self-Regulatory Skill

To assess participants’ tendency to plan their approach to problem-solving tasks, including generating subgoals, I used a scale consisting of items from the ‘Planning’ scales of the Metacognitive Awareness Inventory (Schraw & Dennison, 1994; example item: “I set specific goals before I begin a task.”), the Self-Report Self-Regulation Scale (Herl et al., 1999; example item: "I try to determine what the task requires."), and the Metacognitive Self-Regulation scale of the Motivated Strategies for Learning Questionnaire (Pintrich et al., 1993; example adapted item: "When I start working on a problem, I set goals for myself in order to direct my activities."). Responses were on a 6-point Likert-type scale ranging from "Very untrue of me" to "Very true of me." See **Appendix B** (**Table B3**) for all items.

4.2.3.4 Self-Efficacy

I developed a measure of self-efficacy for technological fluency, which consisted of the first six problems from the criterion measure of technological fluency. For each problem scenario, participants were asked to report their confidence in their ability to use the Internet to provide solutions of increasing quality to each problem (eight levels of difficulty per scenario), on a 10-point scale ranging from "No confidence" to "Certain I can do it"). See **Appendix B** (**Table B4**) for all items.

4.2.3.5 Domain Interest and Domain Curiosity

To provide potential context on hypothesized relationships, participants completed two brief measures assessing their interest in and curiosity about each of the primary content domains represented on the criterion measure. These measures both used 6-point Likert-type response scales (1-6; "Very boring" to "Very interesting" and "Not at all curious" to "Very curious" respectively). See **Appendix B** (**Table B5**) for all items.

4.2.3.6 AI Tool Use Items

To provide potential context on hypothesized relationships, participants were asked to respond to three items about their use of/beliefs about AI after completing the technological fluency assessment. First, they were asked to indicate the extent to which they used any AI tool during the assessment, using a 5-point scale ranging from exclusive use of non-AI resources (e.g., regular web search) to exclusive use of AI resources (e.g., ChatGPT or similar). Second, they were asked to respond to two items concerning AI's general utility for problem-solving: (1) "For everyday problem-solving, I can find a higher quality solution with AI tools than non-AI tools," and (2) "For everyday problem-solving, I can find a solution faster with AI tools than non-AI tools." Both items were scored on a 6-point Likert-type response scale (1-6 scale; Strongly disagree – Strongly agree).

4.3 Procedure

The study consisted of two in-person, group laboratory sessions, each of which lasted approximately two hours. At maximum capacity, twelve participants could be run simultaneously in each session, although most sessions contained significantly fewer participants. In the first session, participants first completed consent forms (including disclosure that portions of the second session will be screen-recorded but that this recording will not capture audio or video of the participant) and then completed the distal cognitive, conative, and affective trait measures as well as the measures of prior technology experiences and technology attitudes. The battery of ability tests was completed on paper immediately following the consent procedures. Participants were provided with a five-minute break approximately one hour into the ability testing portion of the session, during which time they were not permitted access to cell phones or any personal devices. Following conclusion of the paper-and-pencil ability test battery, the self-report measures were completed online via Qualtrics. After completing all assessments, participants were provided with instructions to return for a second session separated by at least one day but no more than one week.

In the second session, all measures/assessments were completed online via Qualtrics. Measures of proximal predictors (i.e., self-regulatory skill and self-efficacy) and the domain interest/curiosity measures were completed first, followed by the assessment of technological fluency. Upon completing the self-report measures, the online protocol provided instructions for completing the assessment of technological fluency. Although answers to items on the assessment were entered on the Qualtrics form, participants were provided with a paper copy of the assessment, with one item per page. This decision was made after piloting, in order to reduce the ease of direct pasting of item content into web searches (e.g., introducing variability in the representation of re-typed content, requiring participants to interpret visual cues rather than using image search).

Participants were advised that they are permitted to search the Web for solutions to each problem and that they may not ask experimenters for assistance (except for problems with the Qualtrics platform, Internet connection, or other technical issues). They were also (1) advised that their answers would be scored on quality, including whether the level of detail provided was sufficient for another person to replicate the solution using their response as instructions, (2) instructed to take their time and move carefully through each item, rather than rushing through the assessment at the expense of the quality of their answers, and (3) given a warning that the Qualtrics form would not allow them to click back, so they must complete all items one at a time, in the order shown on the paper handout. Before participants began the assessment, the experimenter initiated the screen recording. Participants were instructed not to pause, terminate, or otherwise interfere with the screen recording application. After the participant finished the assessment (or time expired) and submitted the Qualtrics survey, the experimenter turned off the screen recording. The participant then completed the final three items related to AI tools before being debriefed and dismissed. After all participants in a session were dismissed, the experimenter saved all recording files and cleared the browsing history of each computer used during the session.

CHAPTER 5. PRELIMINARY ANALYSIS

Analyses and results are reported in the next four chapters. In this chapter, I report a series of preliminary aggregations and analyses including: (1) construction of ability test composites, (2) construction of self-report scale composites, (3) calculations of basic descriptive statistics (*M*, *SD*, range) and internal consistency reliability (Cronbach's α) for all scales/composites, (4) interrater reliability (Krippendorff's α) for the criterion measure, (5) construction of a correlation matrix including all variables in the study, (6) summary results of the originally proposed hypothesis tests, and (7) exploration of psychometric features of the criterion measure. Examination of the criterion measure (e.g., inter-item correlation matrix, item-level performance) revealed several concerning psychometric features. The final section of this chapter describes analyses that attempt to diagnose the underlying cause of these issues. These findings and their implications for changes to the analysis plan are described in more detail at the end of the chapter.

Ultimately, I argue that the data presented in this chapter support a shift in analytic approach to understanding individual differences in problem-solving *processes/strategies*, rather than determinants of problem-solving *performance*. In Chapter 6, I report the results of person-centered analyses, which identified clusters of participants who displayed similar patterns of behavior during the assessment, and mean differences between clusters in key predictors of interest. In Chapter 7, I report analyses addressing the exploratory research question related to qualitative differences in the problem-solving strategies or processes displayed among individuals. Finally, in Chapter 8 I build on the previous two chapters' findings by assessing the utility of distal versus proximal predictors for predicting strategic components of technological fluency. Data cleaning and analyses were conducted in R Studio, version 2024.04.2, using packages boot (Canty & Ripley, 2024; Davison & Hinkley, 1997), cluster (Maechler et al., 2023), cocor (Diedenhofen & Musch, 2015), esquisse (Meyer & Perrier, 2024), GPArotation (Bernaards & Jennrich, 2005), irr (Gamer & Lemon, 2019), janitor (Firke, 2023), lmtest (Zeileis & Hothorn, 2002), MASS (Venables & Ripley, 2002), openxlsx (Schauberger & Walker, 2023), plotly (Wickham, 2007), pscl (Jackman, 2024), psych (Revelle, 2023), RColorBrewer (Neuwirth, 2022), readxl (Wickham & Bryan, 2022), reshape2 (Sievert, 2020), and tidyverse (Wickham et al., 2019).

5.1 Factor Structure of Ability Battery

The nine tests selected to assess verbal, spatial, and numerical abilities were aggregated into three composite scores based on their theorized content domain. To provide empirical justification for this aggregation, I performed an exploratory factor analysis on the ability test data. Parallel analysis (Horn, 1973; Humphreys & Ilgen, 1969) suggested extracting three factors, as expected. Because I expected positive correlations among the three ability domains (i.e., evidence of positive manifold among ability tests; Carroll, 1993; Spearman, 1904), I used an oblique rotation (oblimin) which allowed factors to correlate. Factor loadings and correlations are shown in **Table 5** and **Table 6** respectively.

Table 5- Rotated Factor Solution for Ability Tests

Test	Factor 1 (Numerical)	Factor 2 (Verbal)	Factor 3 (Spatial)
Math Knowledge	**0.69**	-0.06	0.14
Number Series	**0.83**	0.10	-0.10
Math Approximation	**0.85**	-0.05	0.02
Extended Range Vocabulary	-0.14	**0.70**	0.01
MAB Similarities	0.11	**0.68**	-0.02
MAB Comprehension	0.04	**0.67**	0.05
Verbal Test of Spatial Abilities	**0.39**	0.09	**0.34**
Spatial Analogies	**0.38**	0.18	**0.36**
Paper Folding	-0.02	0.01	**0.87**

Note: "MAB" refers to "Multidimensional Aptitude Battery". Loadings greater than 0.30 are bolded.

Table 6- Ability Factor Correlation Matrix

	Numerical	Verbal	Spatial
Numerical	1		
Verbal	0.47	1	
Spatial	0.47	0.37	1

Note: Correlations significant at the $p < 0.05$ level are bolded.

After examining the factor structure of the battery of ability tests and making decisions about aggregation, I calculated verbal, spatial, and numerical ability composite scores as shown in **Table 4**, using unit-weighted *z*-scores. Except for two instances of cross-loading (two of the tests selected to represent the spatial domain also loaded on the same factor as the numerical ability tests), the expected factor structure was generally present in the data.

5.2 Aggregation of Self-Report Measures

With the exception of the self-efficacy scales and domain interest/curiosity measures, composite values for all self-report measures (i.e., all conative and affective traits, self-concept, technology attitudes/experiences, and self-regulatory skill) were constructed by taking an average of every item in the measure. The six items on each of the domain interest/curiosity measures were not aggregated. Because the self-efficacy measure consisted of six eight-item scales (i.e., six problem-solving scenarios in which participants reported their self-efficacy for developing problem solutions at eight increasingly difficult levels of quality), composites were constructed in two steps. All responses were first summed within the six scenario-level scales, and then averaged across scenarios to create a composite score for self-efficacy.

5.3 Descriptive Overview of Predictor Variables

For a descriptive overview of scores on all predictors related to formal hypotheses as well as supplemental variables, see **Table 7** and **Table 8** respectively.

Table 7- Summary of Descriptive Statistics for Predictors Related to Formal Hypotheses

	M	*SD*	Possible Range	Actual Range	Cronbach's α
Verbal Ability Composite					
Extended Range Vocabulary	10.41	3.37	0 – 24	2.00 – 17.00	0.54
MAB Similarities	27.75	3.28	0 – 34	15.00 – 34.00	0.59
MAB Comprehension	21.39	3.45	0 – 28	6.00 – 27.00	0.43
Spatial Ability Composite					
Verbal Test of Spatial Abilities	15.56	3.78	0 – 24	5.00 – 23.00	0.72
Spatial Analogies	21.23	5.42	0 – 30	6.00 – 30.00	0.83
Paper Folding	7.71	2.50	0 – 12	0.00 – 12.00	0.67
Numerical Ability Composite					
Math Knowledge	26.74	6.31	0 – 35	6.00 – 35.00	0.88
Number Series	11.78	3.15	0 – 20	4.00 – 19.00	0.89
Math Approximation	17.44	5.83	0 – 35	5.00 – 33.00	0.86
Conventional Interests	3.10	0.83	1 – 6	1.30 – 5.80	0.84
Realistic Interests	3.40	0.88	1 – 6	1.10 – 5.50	0.84
Investigative Interests	3.91	0.96	1 – 6	1.30 – 6.00	0.87
Conscientiousness	4.38	0.77	1 – 6	1.42 – 5.75	0.87
Openness to Experience	4.18	0.65	1 – 6	2.50 – 5.42	0.70
Technology Attitudes	4.34	0.38	1 – 6	3.41 – 5.44	0.79
Prior Experiences	4.10	0.61	1 – 6	2.40 – 5.55	0.82
Self-Regulatory Skill	4.45	0.64	1 – 6	2.59 – 5.94	0.89
Self-Efficacy	59.54	13.71	0 – 80	8 – 80	0.86

Note. "MAB" refers to "Multidimensional Aptitude Battery". Descriptives for ability tests were calculated prior to standardization.

Table 8- Summary of Descriptive Statistics for Supplemental Predictors

	M	*SD*	Possible Range	Actual Range	Cronbach's α
Verbal Self-Concept	4.62	0.91	1 – 6	1.50 – 6.00	0.83
Spatial Self-Concept	4.49	0.94	1 – 6	1.50 – 6.00	0.84
Numerical Self-Concept	4.82	0.87	1 – 6	1.83 – 6.00	0.88
Technology Self-Concept	5.05	0.69	1 – 6	2.17 – 6.00	0.82
MTQ – Competitiveness	3.67	0.99	1 – 6	1.67 – 5.83	0.85
MTQ – Other-Oriented Goals	4.26	0.83	1 – 6	2.00 – 6.00	0.80
MTQ – Mastery	4.75	0.70	1 – 6	1.38 – 6.00	0.79
MTQ – Desire to Learn	4.67	0.72	1 – 6	2.75 – 6.00	0.81
MTQ – Emotionality	3.40	1.01	1 – 6	1.44 – 5.89	0.86
MTQ – Worry	4.32	0.90	1 – 6	1.80 – 6.00	0.85
Domain Interest					
Furniture/home repairs	3.52	1.44	1 – 6	1.00 – 6.00	N/A
Appliance repairs	3.67	1.35	1 – 6	1.00 – 6.00	N/A
Home/personal technology	4.53	1.26	1 – 6	1.00 – 6.00	N/A
Personal care/aesthetics	4.09	1.40	1 – 6	1.00 – 6.00	N/A
Organization	4.62	1.25	1 – 6	1.00 – 6.00	N/A
Self-improvement projects	4.83	1.22	1 – 6	1.00 – 6.00	N/A
Domain Curiosity					
Furniture/home repairs	3.03	1.62	1 – 6	1.00 – 6.00	N/A
Appliance repairs	3.18	1.57	1 – 6	1.00 – 6.00	N/A
Home/personal technology	4.07	1.60	1 – 6	1.00 – 6.00	N/A
Personal care/aesthetics	3.67	1.62	1 – 6	1.00 – 6.00	N/A
Organization	4.03	1.39	1 – 6	1.00 – 6.00	N/A
Self-improvement projects	4.39	1.53	1 – 6	1.00 – 6.00	N/A
AI Solution Quality	3.68	1.34	1 – 6	1.00 – 6.00	N/A
AI Solution Speed	4.70	1.26	1 – 6	1.00 – 6.00	N/A

Note: "MTQ" refers to "Motivational Trait Questionnaire" (Kanfer & Ackerman, 2004). AI Solution Quality and AI Solution speed refer, respectively, to participants' agreement that using AI tools can help them to generate a better solution or to find a solution more quickly.

5.4 Interrater Reliability for Criterion Scoring

As described in the Measures section, two raters (myself and a research assistant) independently scored each assessment. Krippendorff's α (Krippendorff, 1970, 2004) was used as a measure of interrater reliability. Linear distance was assumed between ratings. Unlike simpler indices such as percent agreement, Krippendorff's α accounts for agreement that may have occurred by chance. Although Cohen's κ (Cohen, 1960) is a popular interrater reliability metric which can also account for chance in rating, I chose to use Krippendorff's α because it is generally a more flexible metric and better handles key characteristics of the technological fluency data. Most critically, Cohen's κ was designed to handle nominal data, while Krippendorff's α was designed to handle a wider/more complex range of data types (Krippendorff, 2004). This is important in the context of the technological fluency assessment because the scoring categories have inherent ordered values (i.e., a '1' is considered 'worse' than a '2', which is 'worse' than a '3', and so on). Cohen's κ does not account for theoretical distance between scoring categories and treats agreement/disagreement as a binary issue. Accordingly, it would fail to distinguish between an instance where an item receives the scores '1' and '2' and another where the item receives the scores '1' and '4', despite these scenarios having theoretically distinct implications for the raters' interpretation of the data.

In accordance with recommendations from the literature, I set a minimum threshold for sufficient interrater reliability (i.e., the point at which the rater training process would be considered complete) at 0.67 (on a possible scale of 0 – 1; Krippendorff, 2004). Although differences in study goals and the content of data to be rated mean that "there are no magical numbers" in evaluating interrater reliability (Krippendorff, 2004, p. 429), this threshold was chosen as a reasonable minimum given the overall context of the current study as a first attempt to measure the construct of technological fluency. After the rubric for each item on the assessment was drafted, I met with the research assistant to discuss the scoring rubric and procedures in detail, but without reference to any actual examples from the data in order to avoid biasing independent judgments. We then each scored a single assessment, which I selected based on the fact that all items were attempted (so that an estimate of reliability could be made that incorporated data from all items). Because

Krippendorff's α for this assessment (0.62) was below the pre-selected threshold, we repeated this process for a second round with a new assessment. In between the first and second rounds, we met to discuss points of confusion on the rubric (particularly for items where we disagreed on scores), and I made several revisions to the rubric (e.g., providing additional item-level example quotes pulled from the pilot data, as well as generalized guidelines for what constitutes a score of '1', '2', '3', and so on). In the second round, Krippendorff's α was 0.73. After the second round, we met for a final discussion on best practices, and ratings on all remaining assessments were then completed independently. Krippendorff's α for the entire set of assessments was 0.81.

5.5 Correlations Between Study Variables

For a correlation matrix showing relationships between all study variables, including both those related to formal hypotheses as well as those included for the purpose of construct validity assessment and potential exploratory analyses, see **Table 9**.

Table 9- Correlations Among Distal/Proximal Predictors, Supplementary Variables, and Criterion Measure

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Verbal Ability	1													
2. Spatial Ability	0.433	1												
3. Numerical Ability	0.373	0.613	1											
4. Conventional Interests	-0.137	-0.030	0.038	1										
5. Investigative Interests	0.120	0.087	0.132	-0.037	1									
6. Realistic Interests	0.068	0.207	0.230	0.190	0.147	1								
7. Conscientiousness	-0.022	0.038	-0.035	0.209	0.185	-0.078	1							
8. Openness to Experience	0.353	0.153	0.104	-0.194	0.301	-0.016	-0.158	1						
9. Verbal Self-Concept	0.383	0.049	-0.006	0.079	0.247	-0.027	0.357	0.289	1					
10. Spatial Self-Concept	0.045	0.353	0.210	0.046	0.046	0.253	0.195	0.240	0.180	1				
11. Math Self-Concept	0.053	0.426	0.450	0.197	0.083	0.259	0.286	0.075	0.064	0.486	1			
12. Technology Self-Concept	0.163	0.293	0.145	0.048	0.002	0.161	0.071	0.243	0.134	0.543	0.443	1		
13. MTQ – Competitiveness	0.071	0.250	0.183	0.075	0.014	0.003	0.194	-0.016	0.079	0.268	0.204	0.157	1	
14. MTQ – Other-Oriented Goals	0.090	0.147	0.029	0.165	0.102	-0.101	0.139	0.052	0.201	0.076	0.105	0.128	0.538	1
15. MTQ – Mastery	0.044	0.069	-0.054	-0.002	0.240	0.061	0.558	0.065	0.352	0.235	0.269	0.186	0.280	0.337
16. MTQ – Desire to Learn	0.104	0.148	0.097	0.059	0.382	0.115	0.382	0.201	0.333	0.403	0.528	0.333	0.088	0.156
17. MTQ – Emotionality	-0.189	-0.335	-0.282	0.054	-0.130	-0.299	-0.271	-0.047	-0.117	-0.390	-0.383	-0.283	-0.370	0.006
18. MTQ – Worry	-0.129	-0.204	-0.210	0.075	0.012	-0.265	-0.156	-0.023	-0.002	-0.284	-0.324	-0.166	-0.362	0.113
19. Technology Attitudes	-0.151	-0.065	-0.061	0.213	-0.226	0.012	0.083	-0.209	0.077	0.180	0.170	0.257	-0.050	-0.049
20. Prior Experiences	-0.029	0.091	-0.053	0.093	-0.049	0.167	-0.026	0.059	-0.045	0.309	0.187	0.530	-0.007	-0.067
21. Self-Regulatory Skill	-0.212	-0.046	-0.131	0.139	0.172	0.001	0.567	-0.094	0.244	0.377	0.252	0.231	0.139	0.096
22. Self-Efficacy	0.147	0.203	0.183	-0.038	0.341	0.207	0.224	0.142	0.151	0.283	0.280	0.244	0.106	0.096
23. Domain Interest	-0.098	0.007	-0.181	0.232	0.103	0.343	0.223	0.017	0.058	0.207	0.140	0.102	-0.079	-0.010
24. Domain Curiosity	-0.151	-0.149	-0.167	0.175	-0.002	0.182	0.106	-0.002	0.003	0.103	0.089	0.125	-0.056	-0.041
25. AI Quality Beliefs	-0.223	-0.149	-0.057	0.142	-0.163	0.055	0.009	-0.253	-0.136	-0.035	-0.041	-0.070	0.067	0.056
26. AI Speed Beliefs	-0.152	-0.076	0.087	0.206	-0.166	0.133	0.067	-0.120	0.081	0.029	0.135	0.052	0.216	0.191
27. Technological Fluency	0.103	0.162	0.135	0.039	-0.080	-0.042	-0.010	0.013	0.167	0.028	0.070	0.155	0.059	0.060

Note: MTQ = Motivational Trait Questionnaire. AI Quality/Speed Beliefs = agreement that AI helps generate higher-quality solutions/reach solutions faster.

Table 9 - Correlations Among Distal/Proximal Predictors, Supplementary Variables, and Criterion Measure (cont.)

	15	16	17	18	19	20	21	22	23	24	25	26	27
1. Verbal Ability													
2. Spatial Ability													
3 Numerical Ability													
4. Conventional Interests													
5. Investigative Interests													
6. Realistic Interests													
7. Conscientiousness													
8. Openness to Experience													
9. Verbal Self-Concept													
10. Spatial Self-Concept													
11. Math Self-Concept													
12. Technology Self-Concept													
13. MTQ – Competitiveness													
14. MTQ – Other-Oriented Goals													
15. MTQ – Mastery	1												
16. MTQ – Desire to Learn	0.635	1											
17. MTQ – Emotionality	-0.179	-0.196	1										
18. MTQ – Worry	-0.088	-0.158	0.777	1									
19. Technology Attitudes	-0.009	0.087	-0.108	0.036	1								
20. Prior Experiences	0.103	0.297	-0.159	-0.039	0.357	1							
21. Self-Regulatory Skill	0.534	0.465	-0.120	-0.041	0.113	0.182	1						
22. Self-Efficacy	0.111	0.230	-0.173	-0.082	-0.105	0.083	0.124	1					
23. Domain Interest	0.152	0.150	-0.074	0.040	0.032	0.195	0.358	0.310	1				
24. Domain Curiosity	0.101	0.076	-0.025	0.095	0.065	0.224	0.328	0.158	0.716	1			
25. AI Quality Beliefs	-0.102	-0.180	-0.010	0.022	0.206	-0.032	0.087	-0.059	-0.056	0.056	1		
26. AI Speed Beliefs	0.133	-0.005	0.007	-0.011	0.228	-0.060	0.020	-0.135	-0.096	0.049	0.350	1	
27. Technological Fluency	-0.006	-0.027	-0.075	0.006	0.228	0.164	-0.096	-0.055	-0.134	0.011	0.222	0.182	1

Note: MTQ = Motivational Trait Questionnaire. AI Quality/Speed Beliefs = agreement that AI helps generate higher-quality solutions/reach solutions faster.

5.6 Performance Data and Hypothesis Testing

As described in Chapter 4, scores on the assessment of technological fluency were calculated as the sum of all averaged ratings. Total possible scores ranged from 0 (no items attempted) to 105 (all 21 items receiving the top score). The average total score in the sample was 49.5 (SD = 10.4) with a range of 18.5 to 69. Prior to conducting the proposed formal hypothesis tests, I examined predictor-criterion relationships in **Table 9**, and identified several patterns that raised concerns for the interpretability of the hypothesis testing results, given that the proposed analyses are derived from the correlation matrix.

Namely, most correlations with technological fluency were either in the expected direction but weaker than expected (e.g., $r_{\text{ability}} = 0.103\text{-}0.162$, $r_{\text{attitudes}} = 0.228$, $r_{\text{experience}} = 0.164$), in the opposite direction as would be theoretically expected (e.g., $r_{\text{investigative interests}} = -0.080$, $r_{\text{self-regulation}} = -0.096$), or near zero when more substantial relationships were expected (e.g., $r_{\text{conscientiousness}} = -0.010$, $r_{\text{self-efficacy}} = -0.055$). Across all core and supplementary predictors, 19 of 27 total correlations with technological fluency were less than 0.15, and all were less than 0.25. As a result of these concerns, and to improve conciseness in the primary results chapters, details of the originally proposed hypothesis testing (including tables with full results for non-significant tests) are provided in **Appendix C** rather than in the main text. An overview of key takeaways from the hypothesis testing is provided here.

I conducted two primary sets of analyses. First, I conducted a series of Steiger's z-tests to explore whether there were significant differences in the magnitude of correlations between distal traits and technological fluency on the one hand and proximal variables and technological fluency on the other hand. A total of 32 tests were performed (comparing differences between combinations of eight distal traits against four proximal variables, see **Table C1** in **Appendix C**), of which five were statistically significant prior to correction for Type I error. None were significant post-correction, suggesting a lack of support for formal Hypotheses 1-7. Of the five tests significant prior to correction, only three were in the expected direction: the correlation between technology attitudes and technological fluency was significantly larger than the assessment's relationship with conscientiousness (z = -2.03, p = 0.04), investigative interests (z = -2.23, p = 0.02), and realistic interests (z = -2.21, p = 0.03). Second, I tested Hypotheses 8 and 9 with a set of hierarchical regression equations to assess whether the proximal predictors significantly improved the prediction of the criterion above and beyond the variance predicted by distal traits alone. Results for these hypotheses were mixed: although adding the proximal variables significantly improved the prediction of technological fluency scores (F = 2.64, p = 0.04; see **Table C2** in **Appendix C** for full regression results), the magnitude of the effect was less than predicted with an increase in variance explained of only 8% compared to the predicted increase of approximately 15-25%.

5.7 Exploration of Criterion Measure

To investigate possible causes of the unexpected features of the performance data (i.e., failure to correlate meaningfully with theoretically relevant predictors), I first calculated internal consistency reliability for the assessment, in case low reliability was placing an upper bound on possible relationships between constructs. However, Cronbach's α was 0.77, suggesting that internal consistency reliability of the criterion measure was not likely to be significantly constraining the magnitude of correlations. I next conducted several exploratory item-level analyses to investigate potential explanations for the theoretically inconsistent predictor-criterion relationships, including examination of: (1) an inter-item correlation matrix, (2) potential floor/ceiling effects, and (3) within-person variability in performance.

5.7.1 Item-Level Correlations and Floor/Ceiling Effects

The original goal of these exploratory analyses was to determine whether there might be potentially problematic items constraining predictor-correlation relationships, and if so, to identify the problematic items for removal. Examination of the inter-item correlation matrix suggested that the issues were unlikely to be isolated to any one or several problematic items. As seen in **Table 10**, the inter-item correlation matrix can be characterized by a simplex-like pattern, wherein a series of measurements are most strongly correlated with those immediately preceding or following them in a sequence (i.e., a correlation matrix characterized by order effects; Guttman, 1955). This pattern is routinely found in situations where repeated measurements are taken over time (in this case, repeated problem-solving trials), and is not necessarily indicative of a psychological cause – rather, a simplex-like pattern may imply that the relative influence of performance determinants is changing over the course of the assessment (for discussion, see Ackerman, 1989).

Importantly, however, for an assessment where all items are intended to represent the same construct, there should still be evidence of a more general positive manifold across the entire item set even if a simplex-like pattern is present (i.e., generally positive inter-item correlations which are larger along the diagonal and smaller for non-adjacent items). In contrast, **Table 10** includes a large number of near-zero and negative inter-item correlations, particularly for correlations between early and late items.

Table 10 - Item-Level Correlation Matrix for Technological Fluency Assessment

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Q1	1													
Q2	0.372	1												
Q3	0.046	0.102	1											
Q4	0.160	0.271	0.244	1										
Q5	0.046	0.162	0.082	0.070	1									
Q6	0.140	0.181	0.121	0.240	0.061	1								
Q7	0.024	0.074	0.277	-0.015	0.057	0.108	1							
Q8	0.010	0.083	0.080	0.163	0.107	0.382	0.205	1						
Q9	-0.061	0.100	-0.147	-0.083	0.042	0.110	0.145	0.274	1					
Q10	-0.038	-0.018	0.051	0.081	-0.026	0.128	0.042	0.180	0.423	1				
Q11	0.004	0.000	-0.108	-0.036	-0.215	0.103	-0.127	0.095	0.339	0.515	1			
Q12	-0.065	-0.104	-0.017	-0.077	-0.122	0.055	-0.058	0.093	0.409	0.392	0.528	1		
Q13	-0.051	-0.100	-0.074	-0.143	-0.126	0.109	-0.014	0.109	0.238	0.360	0.520	0.664	1	
Q14	-0.077	-0.190	-0.141	-0.247	-0.140	-0.055	-0.077	0.016	0.074	0.202	0.312	0.486	0.681	1
Q15	-0.126	-0.282	-0.160	-0.339	-0.175	-0.071	-0.047	-0.126	0.159	0.211	0.312	0.454	0.585	0.658
Q16	-0.005	-0.224	-0.131	-0.283	-0.186	-0.178	-0.035	-0.169	0.148	0.164	0.269	0.392	0.486	0.578
Q17	-0.075	-0.233	-0.148	-0.319	-0.272	-0.185	-0.011	-0.196	0.137	0.095	0.173	0.298	0.398	0.447
Q18	-0.038	-0.294	-0.236	-0.322	-0.192	-0.169	-0.143	-0.213	0.046	0.082	0.162	0.256	0.346	0.458
Q19	-0.060	-0.287	-0.205	-0.183	-0.238	-0.214	-0.192	-0.276	0.025	0.092	0.154	0.238	0.276	0.364
Q20	0.010	-0.247	-0.129	-0.056	-0.173	-0.152	-0.133	-0.175	0.020	0.028	0.156	0.273	0.218	0.289
Q21	-0.048	-0.259	-0.161	-0.135	-0.168	-0.224	-0.163	-0.217	-0.008	0.037	0.156	0.208	0.202	0.292

Table 10 - Item-Level Correlation Matrix for Technological Fluency Assessment (cont.)

	Q15	Q16	Q17	Q18	Q19	Q20	Q21
Q1							
Q2							
Q3							
Q4							
Q5							
Q6							
Q7							
Q8							
Q9							
Q10							
Q11							
Q12							
Q13							
Q14							
Q15	1						
Q16	0.750	1					
Q17	0.670	0.737	1				
Q18	0.600	0.692	0.763	1			
Q19	0.510	0.613	0.706	0.767	1		
Q20	0.371	0.609	0.579	0.588	0.801	1	
Q21	0.439	0.513	0.604	0.624	0.742	0.793	1

Because examining the correlation matrix did not suggest obvious candidate 'problem items', I next looked at score distributions of the items on the assessment. Whereas the correlation matrix provides insight into the similarity of participant rank order for early versus late items, item-level score distributions provide more context on average performance differences between early and late items – the latter could help to identify instances of ceiling or floor effects. **Figure 3** shows boxplots of scores for each item on the assessment, with a smoothed line connecting item means. As shown in the figure, average performance was relatively stable until it began to decline about halfway through the assessment.

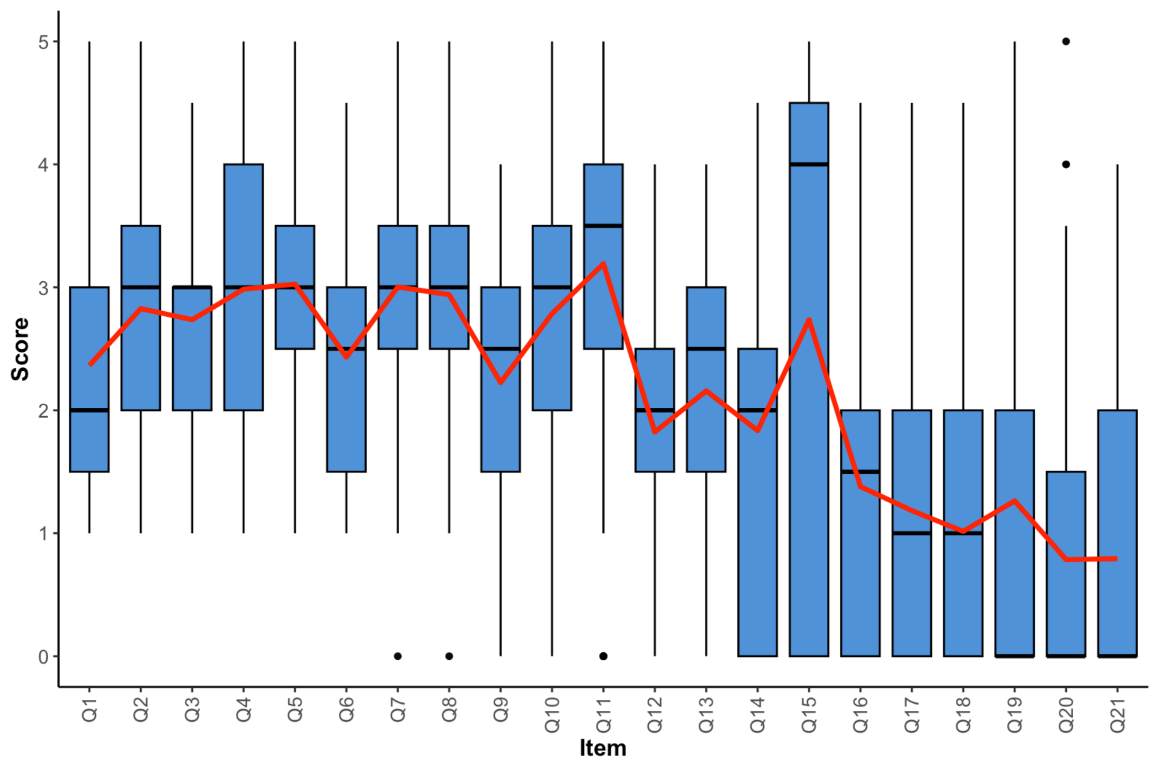


Figure 3 - Boxplots of Technological Fluency Item Scores

Scores in **Figure 3** included both attempts and non-attempts, the latter of which were scored as zero. To account for the possibility that non-answers could have been depressing the means (especially in later items), I also plotted the proportion of correct attempted answers for each item. ‘Correct’ answers were operationalized as achieving an average (across coders) score of at least 3.5, which means that at least one of the two coders rated the score a 4 or 5. **Figure 4** shows that this approach did not differ substantially from the performance patterns shown in **Figure 3**, and in fact provides additional evidence of floor effects in most items.

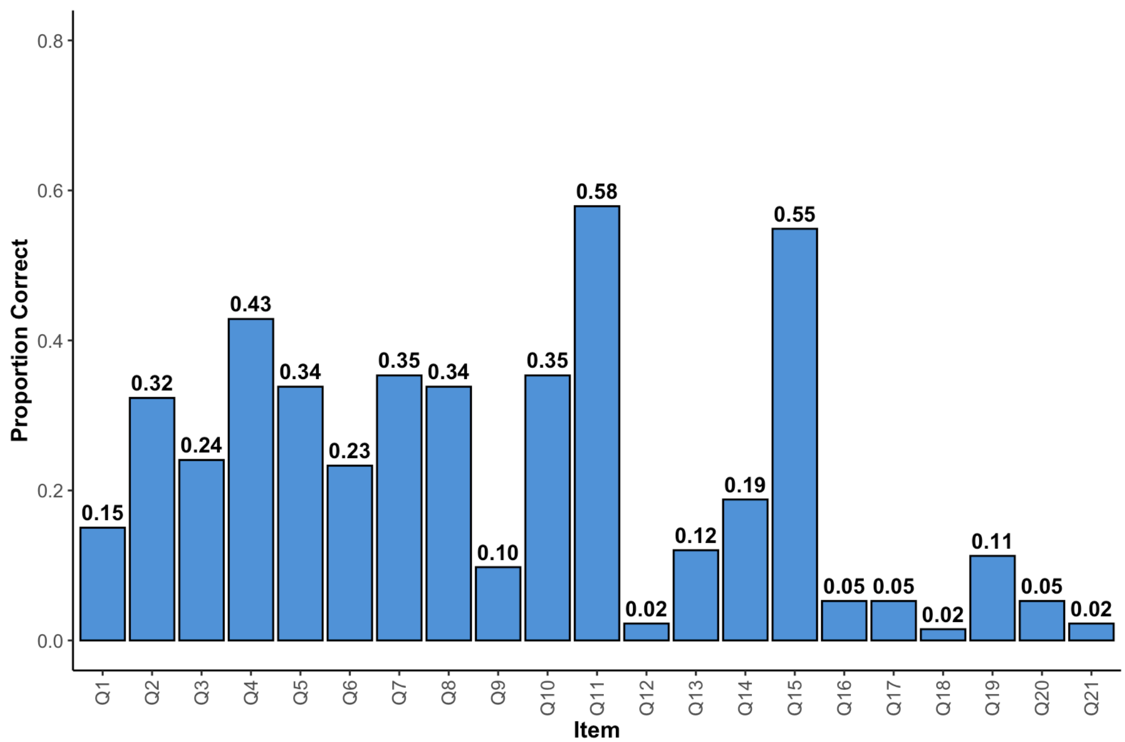


Figure 4 - Proportions of Attempted Technological Fluency Items Marked Correct

Based on the available data, I cannot conclude the exact reason that the proportion of correct answers were so low. Because the severity of floor effects varies across items, it is likely that multiple issues were involved which may have affected items to varying extents. Potential sources of the effect could, for example, have varied both in terms of their specificity/generality as well as whether they were associated with some feature of the assessment/procedure or were related to participant characteristics (e.g., varying quality of strategies or resources used during the assessment; within-person variation of strategy at different points in the assessment). Specific effects related to the assessment might include item difficulty (in that most items were too difficult). A more general possible effect related to assessment procedures concerns the instructions that were given to participants. Although participants were instructed that (1) solution quality was the primary criterion and (2) they should not move on to the next item until they were satisfied that their answer was high-quality, these instructions may not have been clear enough. Alternatively, the time limit placed on the assessment may have placed pressure on participants to move more quickly through items than they would in a real-life situation. Regardless of the specific cause of the item-level floor effects, however, these findings raise concerns about my ability to meaningfully interpret differences in performance on the assessment as a whole.

5.7.2 Intrapersonal Variability in Item Scores

In addition to identifying item-level floor effects as a primary issue impacting predictor-criterion correlations, I sought to identify additional relevant individual differences in assessment behavior/performance that were not apparent in previous exploration of the criterion data. For example, different within-person strategic approaches to the assessment (minimally satisficing behavior versus attempting to provide maximal detail on each item) may have contributed to the number of items attempted and therefore contributed to the total score which was a sum of all 21 items. I calculated within-person

SD's across the 21 items and found a strong negative correlation with aggregate performance ($r = -0.52$) such that lower intrapersonal variability in item scores (i.e., more consistent item scores) was associated with a higher total score. As shown in **Table 11**, however, the correlation between within-person variability and item-level score varied widely across the course of the assessment. For early (Items 1-7), middle (Items 8-14), and late (Items 15-21) item sets, the correlations between within-person SD and total score were $r = 0.68$, -0.22, and -0.84 respectively. That is, greater variability in item-level scores was strongly positively correlated with scores on the first third of items on the assessment, and strongly negatively correlated with scores on the last third of items on the assessment.

Table 11 - Item-Level Score Correlations with Within-Person Standard Deviation

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
r	0.19	0.44	0.32	0.45	0.32	0.37	0.29	0.3	-0.07	0.05	-0.13
	Q12	**Q13**	**Q14**	**Q15**	**Q16**	**Q17**	**Q18**	**Q19**	**Q20**	**Q21**	
r	-0.28	-0.32	-0.46	-0.61	-0.67	-0.73	-0.76	-0.75	-0.66	-0.72	

5.7.3 Summary

My original research questions concerned determinants of performance on an assessment of technological fluency. However, a number of predictor-criterion relationships were inconsistent with theoretical expectations. Exploratory analyses suggested that the issue was unlikely to be primarily a function of assessment reliability or of a few poor-quality items. Additional examination of item-level data revealed evidence of relatively pervasive floor effects. Although it is not possible to determine the exact cause(s) of the item-level floor effects, it is likely that one or more issues related to item difficulty, assessment procedures, or participants' heterogenous approaches to the task may have restricted predictor-criterion correlations by restricting variance at the item-level.

Collectively, these preliminary analyses suggest that my ability to interpret differences in total assessment performance may be limited. However, I did find that individual differences in strategic behavior (e.g., item-level variability within persons) were strongly related to scores on the assessment of technological fluency, and that the nature of this relationship changed throughout the assessment (i.e., variability had different relationships with performance early versus late in the assessment). The implication is that there may be meaningful individual differences in process which can tease out important aspects of technological fluency (i.e., a focus on determinants of performance-related processes rather than performance itself). Therefore, the next two chapters took a more person-centered approach to tease out prototypical *patterns* of performance across items (Chapter 6) and individual differences in problem-solving processes which might be related to those patterns (Chapter 7).

CHAPTER 6. CLUSTER ANALYSIS AND DETERMINANTS OF CLUSTER MEMBERSHIP

In this chapter, I took a person-centered approach to (a) identify groups of participants who approached the assessment similarly (cluster solution itself) and (b) look at standardized differences between clusters (i.e., exploring determinants of behavior during assessment rather than determinants of performance per se). Below, I report procedures for conducting a cluster analysis of item-level assessment data and for selecting the most appropriate cluster solution. I then explore potential determinants of cluster membership by examining mean differences in distal and proximal predictor means within clusters. Finally, I provide raw/standardized item mean plots by cluster and conclude with a summary and interpretation of cluster profiles.

6.1 Cluster Analysis Solution

I began by conducting a sparse k-means cluster analysis on the 21 technological fluency items. This is a non-hierarchical clustering approach, which accounts for high dimensionality in clustering data (i.e., including 21 input variables) by up-weighting variables that contribute the most to the solution (for discussion, see Witten & Tibshirani, 2010). All items were standardized before being entered in the cluster analysis. I compared solutions ranging from two to five clusters, and evaluated each on (1) silhouette scores, (2) theoretical interpretability, and (3) parsimony. The silhouette score metric compares intra-cluster difference between participants (cluster tightness) to inter-cluster difference between participants (cluster separation) (Rousseeuw, 1987). Scores range from -1 to +1, where scores closer to +1 indicate a good match of participants to their assigned cluster and a poor match to alternate clusters.

Of the various solutions tested, the two-cluster solution was the best fitting (silhouette score = 0.15) and most interpretable/parsimonious. However, it did not seem to be a good-fitting model overall, for two reasons. First, although the silhouette score was highest of those tested (the three – five cluster solutions had scores ranging from 0.09 to 0.06), it was still relatively low. Second, as a result of the weighting process used to account for high dimensionality, nearly 100% of cluster assignment was attributable to only one input variable (item 17). Together, these factors suggest that the sparse 2-cluster solution was not an appropriate model for my data. One possibility is that the sample size is too small to produce reliable solutions for a large (>10) number of input variables.

Therefore, I next conducted a regular k-means cluster analysis (better suited for data with <10 input variables), with three variables consisting of item sets: (1) sum of items 1-7, (2) sum of items 8-14, and (3) sum of items 15-21. I aggregated variables in this way because the final set of preliminary analyses in Chapter 5 suggested that there are meaningful differences in strategic assessment behavior between early, middle, and late items. As in the previous clustering approach, scores on the three item sets were standardized prior to being entered in the cluster analysis.

Because there were only three input variables, I only compared a two-cluster (silhouette score = 0.34) and three-cluster solution (silhouette score = 0.35) to preserve parsimony. Since the fit metrics were comparable between these two models, the choice of solution was based on theoretical interpretability. As shown in **Table 12**, the three-cluster solution seems to distinguish two performance patterns within the first group in the two-cluster solution: in both patterns, performance is above average at the beginning of the assessment and below average at the end of the assessment, but there are appreciable differences in *when* this drop-off occurs (e.g., after the early vs. middle item sets). Because this difference provides potentially theoretically interesting information about process differences in the sample, I moved forward with the three-cluster solution. Differences in average time allocation across the assessment by cluster are shown in **Figure 5**. As described in the following sections, Cluster 2 can be considered ‘minimally satisficing’, while Cluster 1 and Cluster 3 can be described as demonstrating ‘early maximum effort’ and ‘sustained maximum effort’ respectively.

Table 12 - Standardized Item Scores By Cluster

	Early Items (1-7)	Middle Items (8-14)	Late Items (15-21)
2-Cluster Solution			
Cluster 1 (N = 63)	0.57	-0.30	-0.91
Cluster 2 (N = 70)	-0.51	0.27	0.82
3-Cluster Solution			
Cluster 1 (N = 20)	0.67	-1.84	-1.06
Cluster 2 (N = 65)	-0.55	0.24	0.89
Cluster 3 (N = 48)	0.47	0.44	-0.76

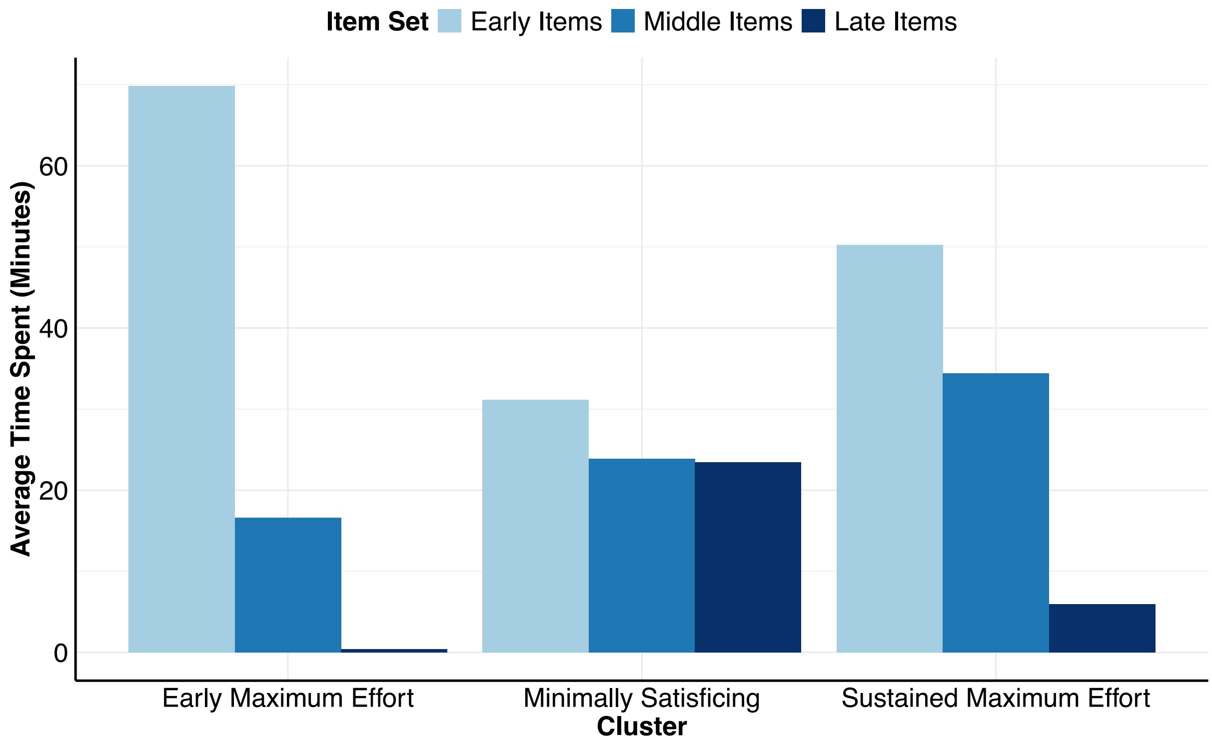


Figure 5 - Time Allocation By Cluster

6.2 Determinants of Cluster Membership

Table 13 shows the average and standard deviation, by cluster, for all distal, proximal, and supplementary predictor variables included in the study, as well as within-person SD. All predictors were standardized *before* calculating cluster averages, so that each value in the table refers to a cluster's mean (standardized) value of that predictor against the entire sample.

Table 13 - Means and Standard Deviations of Standardized Predictors By Cluster

	Cluster 1 Mean (SD)	**Cluster 2** Mean (SD)	**Cluster 3** Mean (SD)
Distal			
Verbal Ability	**-0.35 (1.13)**	**-0.13 (1.00)**	**0.33 (0.86)**
Spatial Ability	**-0.24 (1.01)**	-0.06 (0.97)	**0.18 (1.03)**
Numerical Ability	-0.09 (1.03)	-0.05 (1.04)	0.10 (0.94)
Realistic Interests	0.14 (0.92)	-0.08 (1.04)	0.05 (0.99)
Investigative Interests	**0.25 (1.04)**	**-0.16 (1.09)**	0.11 (0.82)
Conventional Interests	0.11 (1.17)	-0.08 (1.06)	0.06 (0.84)
Conscientiousness	0.09 (0.75)	-0.12 (0.97)	0.12 (1.12)
Openness to Experience	0.01 (0.79)	-0.10 (1.00)	0.13 (1.07)
Proximal			
Attitudes	0.00 (0.98)	**0.20 (0.99)**	**-0.27 (0.97)**
Prior Experience	-0.15 (0.84)	0.14 (1.12)	-0.13 (0.87)
Self-Regulation Skill	**0.42 (0.90)**	**-0.17 (1.02)**	**0.06 (0.98)**
Self-Efficacy	0.04 (1.07)	**-0.16 (0.95)**	**0.20 (1.02)**
Supplementary			
Verbal Self-Concept	**-0.42 (0.98)**	**-0.06 (1.04)**	**0.26 (0.89)**
Spatial Self-Concept	**0.33 (0.79)**	**-0.07 (1.01)**	**-0.04 (1.06)**
Math Self-Concept	**0.27 (0.93)**	**-0.07 (1.03)**	**-0.02 (0.99)**
Technology Self-Concept	0.14 (0.89)	-0.08 (0.98)	0.05 (1.08)
MTQ – Competitiveness	-0.06 (0.82)	-0.01 (1.08)	0.04 (0.97)
MTQ – Other-Oriented Goals	-0.13 (1.10)	-0.06 (1.03)	0.13 (0.93)
MTQ – Mastery	**0.24 (0.84)**	**-0.03 (1.11)**	**-0.06 (0.91)**
MTQ – Desire to Learn	0.10 (1.09)	-0.11 (1.04)	0.11 (0.90)
MTQ – Emotionality	0.04 (0.95)	0.03 (0.97)	-0.06 (1.07)
MTQ – Worry	0.13 (0.99)	0.06 (0.96)	-0.13 (1.06)
Beliefs about AI Quality	**-0.24 (1.06)**	**0.30 (1.00)**	**-0.30 (0.87)**
Beliefs about AI Speed	0.00 (0.89)	**0.13 (1.02)**	**-0.17 (1.01)**
Within-Person SD	**0.99 (0.58)**	**-0.89 (0.43)**	**0.79 (0.51)**

Note: Variables which show differences of at least 0.3 between at least one pair of clusters were bolded.

6.3 Item Means By Cluster

Figures 6-7, **8-9**, and **10-11** show raw and standardized item means respectively for Cluster 1, Cluster 2, and Cluster 3. For the latter, data were standardized prior to clustering.

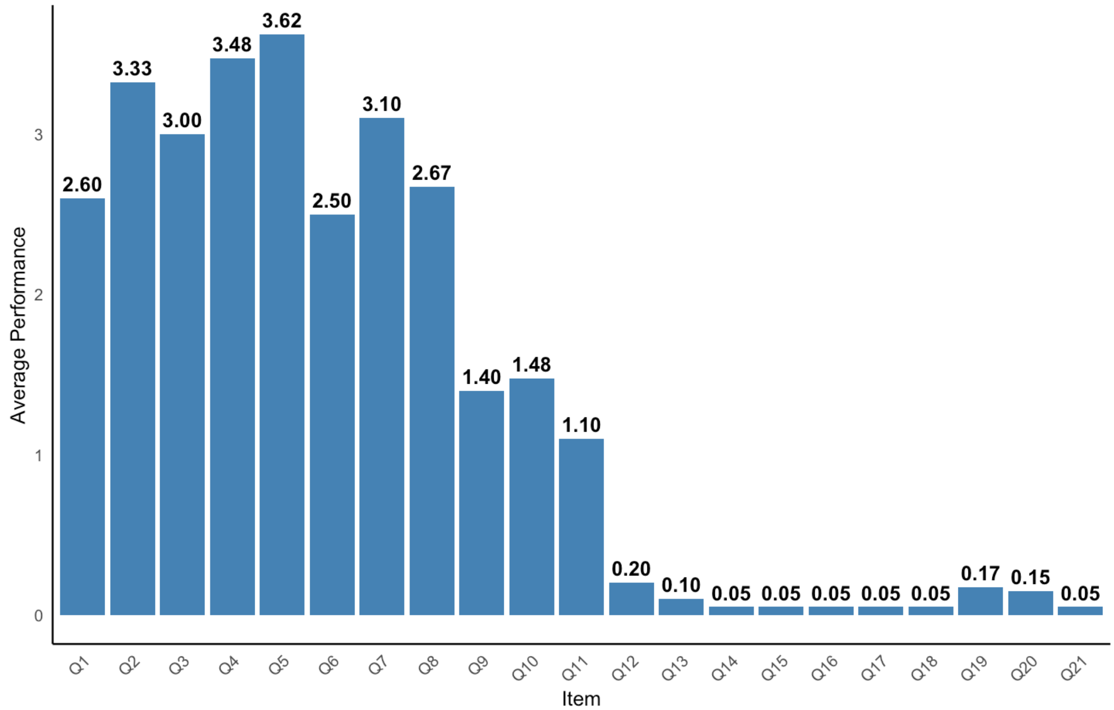


Figure 6 - Item-Level Scores on Technological Fluency; Cluster 1

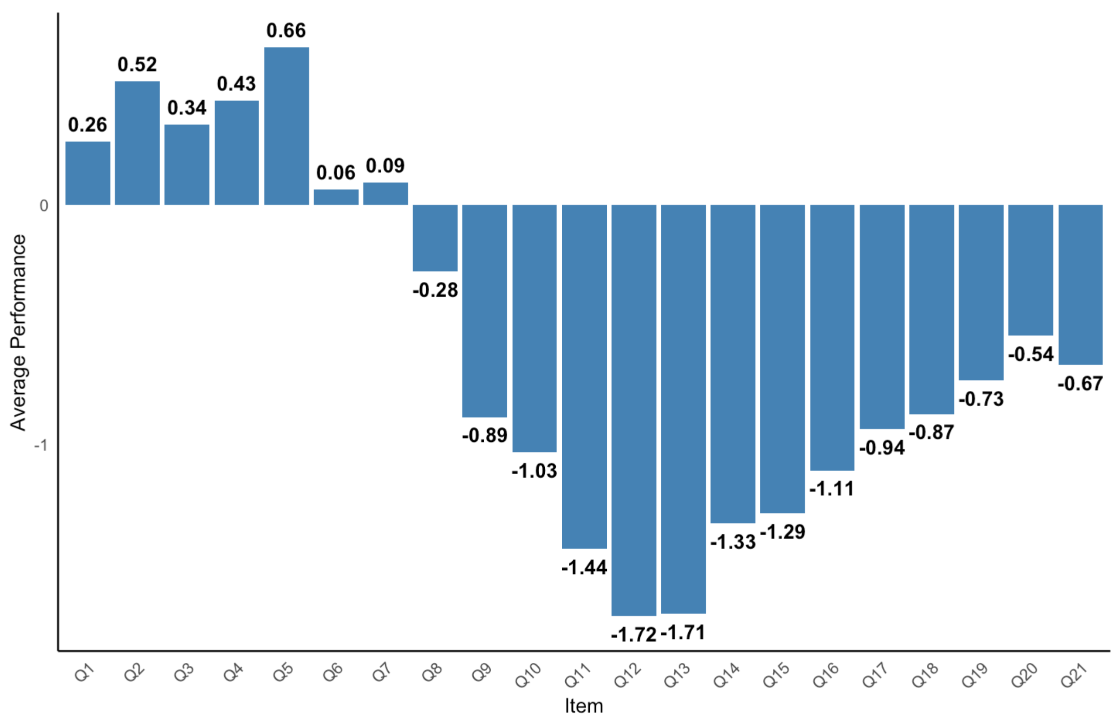


Figure 7 - Standardized Item-Level Scores on Technological Fluency; Cluster 1

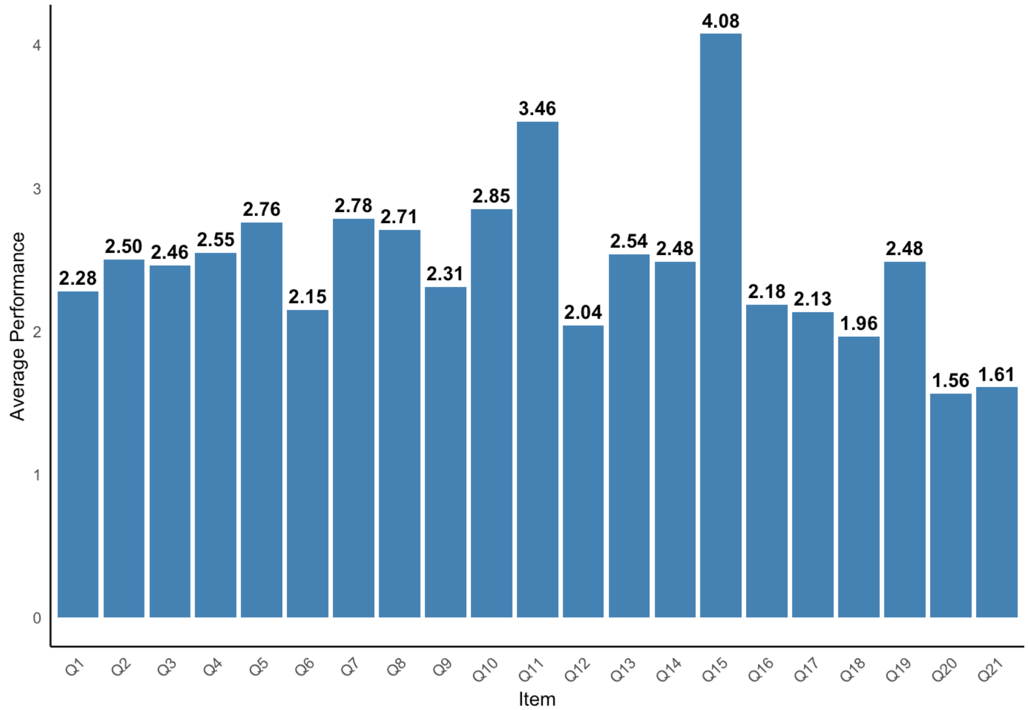


Figure 8 - Item-Level Scores on Technological Fluency; Cluster 2

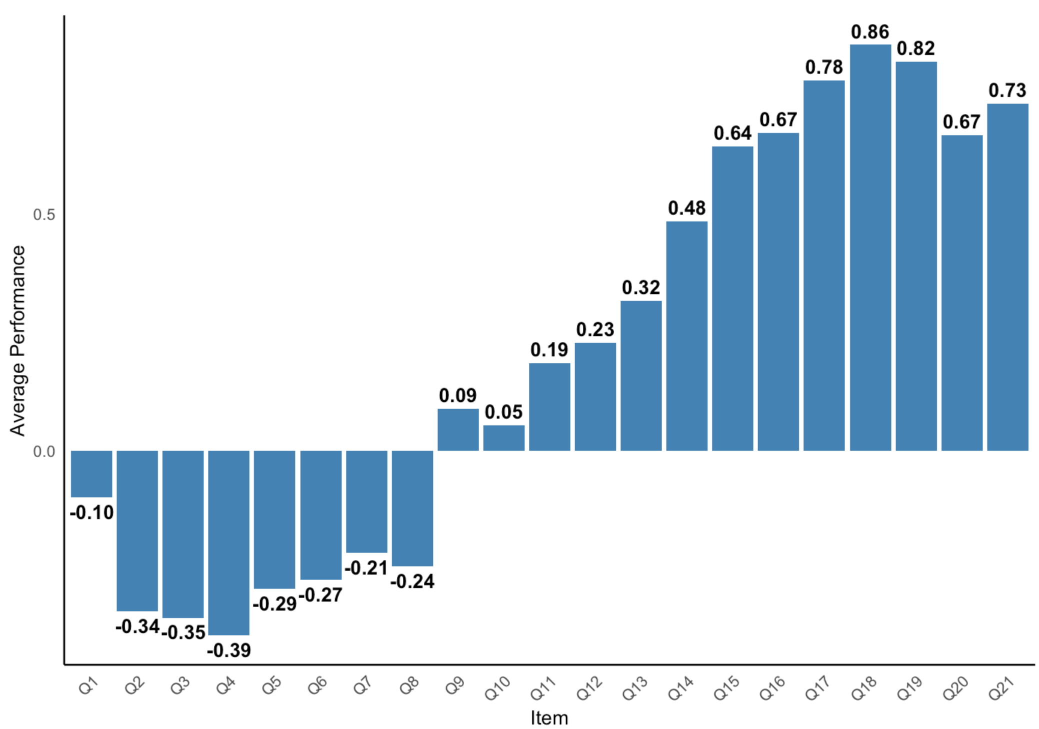


Figure 9 - Standardized Item-Level Scores on Technological Fluency; Cluster 2

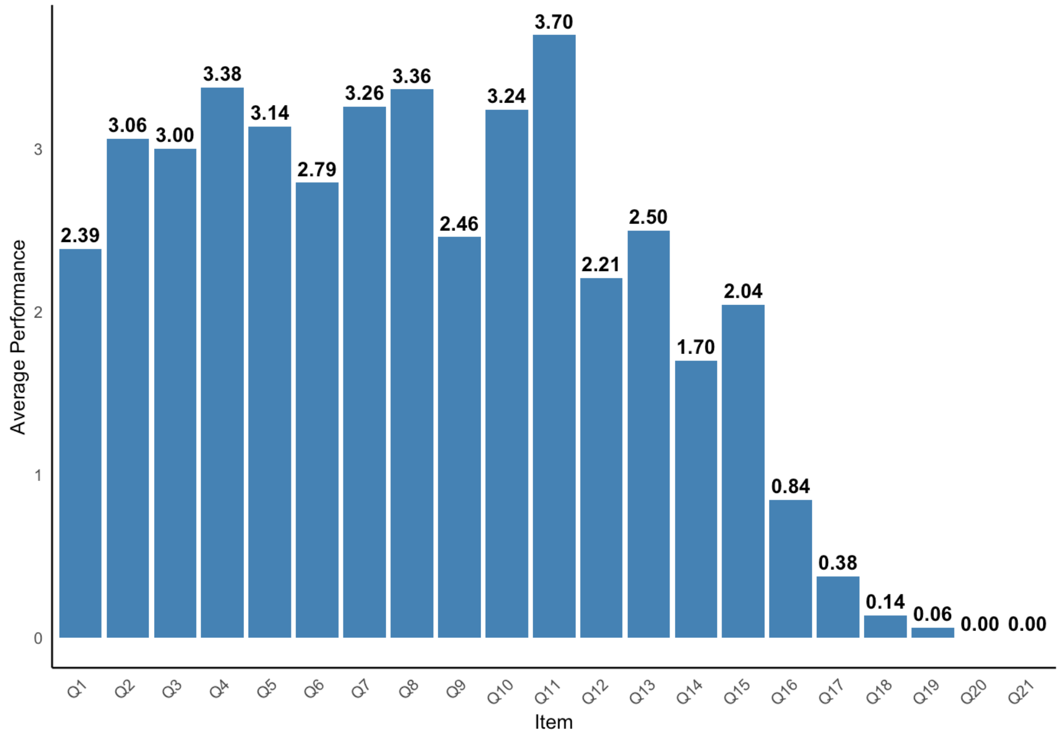


Figure 10 - Item-Level Scores on Technological Fluency; Cluster 3

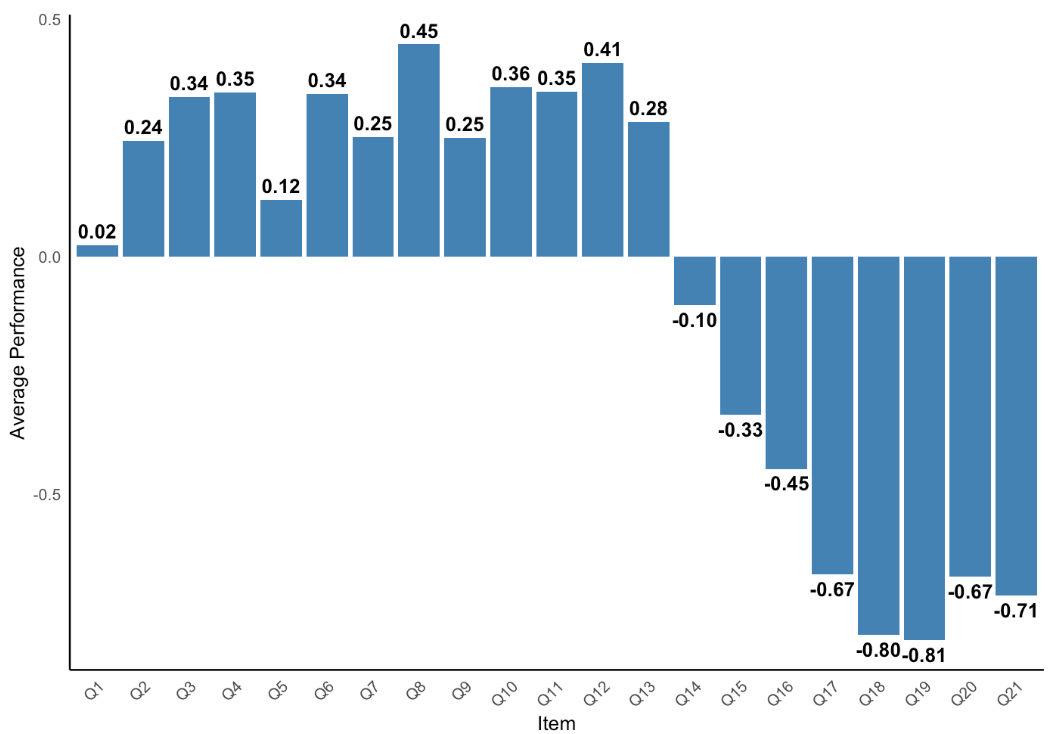


Figure 11 - Standardized Item-Level Scores on Technological Fluency; Cluster 3

6.4 Summary of Cluster Characteristics

6.4.1 Cluster 1: Early Maximum Effort

Cluster 1 was the smallest group, comprising about 15% of the sample (N = 20). Individuals in this cluster completed the fewest number of items on average (*M* = 9.9) although there was more variability than in the other clusters (*SD* = 3.11, range = 6-21). This cluster had the highest average within-person variability (*M* = 0.99). Item plots suggest that participants in this cluster tended to perform well only on the first third of the assessment, followed by a very sharp drop in relative performance in the middle set of items which appears to be partially offset by similar performance declines for the late items in Cluster 3. Compared to other groups, Cluster 1 had lower average scores on verbal (*M* = -0.35) and spatial (*M* = -0.24) ability, as well as verbal self-concept (*M* = -0.42). They also had the highest average scores on investigative interests (*M* = 0.25) and self-reported self-regulation skill (*M* = 0.42). Smaller effects (but still potentially interesting given item completion patterns) included that Cluster 1 had the highest scores of any group on both the Mastery (*M* = 0.24) and Worry (*M* = 0.13) scales of the MTQ.

6.4.2 Cluster 2: Minimally Satisficing

Cluster 2 was the largest group, comprising about 50% of the sample (N = 65). Individuals in this cluster completed the most items on average (*M* = 20.52), with low variability compared to other clusters (*SD* = 0.90, range = 18-21). Specifically, this cluster had an average within-person SD of item scores of -0.89. Examining this value in combination with the raw item mean plots shows that performance was highly stable throughout the entire assessment, with scores between 2 and 3 for most items (on a scale of 0 – 5). While Cluster 2 had a moderate ability profile compared to other groups (i.e., average verbal and spatial ability *z*-scores for this cluster were approximately halfway between those of Cluster 1 and Cluster 3), its more distinguishable characteristics include non-ability traits, attitudes, and beliefs. For example, this cluster had the lowest average scores on investigative interests (M = -0.16), self-reported self-regulation skills (M = -0.17), and self-efficacy (M = -0.16). In contrast, they also reported the most positive technology attitudes (M = 0.20) as well as the most positive beliefs about the potential of AI to help develop high quality solutions (M = 0.30).

6.4.3 Cluster 3: Sustained Maximum Effort

Cluster 3 was moderate in size, comprising about 35% of the sample (N = 48). Participants completed the second-highest number of items on average (M = 14.71) and had the second-highest variability in scores (SD = 1.97, range = 11-19). Within-person SD was above average compared to the full sample (M = 0.79), and while this variability manifested itself in a generally similar manner to Cluster 1 (i.e., means of around 3 on early test items), participants in Cluster 3 tended to maintain that performance for longer before item-level performance began to drop, and scores dropped more gradually once declines did begin. Compared to other groups, Cluster 3 is distinguishable by a pattern of having the highest scores on several predictors related to ability and motivational self-beliefs, including both verbal (M = 0.33) and spatial (M = 0.18) ability as well as verbal self-concept (M = 0.26) and self-efficacy for technological fluency (M = 0.20). Although the effect is small, they had lower average scores on the Worry scale of the MTQ (M = -0.13) compared to Cluster 1. Notably, they are also distinguishable from Cluster 2 in particular by having the most negative technology attitudes (*M* = -0.27) as well as the most negative beliefs about the potential of AI to help develop high quality solutions (*M* = -0.30).

6.5 Summary of Differences Between Clusters

The results of the cluster analysis provide support for two primary conclusions. The first key takeaway is that there appear to be theoretically meaningful process differences in the technological fluency assessment data (i.e., different strategic approaches to the task). In Chapter 5, preliminary analysis results suggested that item-level performance data was strongly correlated with within-person variability and that the direction of this effect depended on item order. Results from this chapter provide more insight into this relationship. Specifically, based on item mean plots and the number of items completed per cluster, the prototypical member of Cluster 2 tended to engage in minimally satisficing behavior throughout the entire assessment, while those in Clusters 1 and 3 were more likely to attempt to maximize item scores early on (with Cluster 3 succeeding at maximizing performance for more items on average compared to Cluster 1). While at first glance **Table 12** appears to suggest that Cluster 2 is learning something about technological fluency during the assessment (because performance increases through the assessment relative to other groups), the raw item mean plots suggest this is not the case – rather, Cluster 2's performance stays stable and therefore they perform *relatively* better on the late item set compared to Clusters 1 and 3 who either did not finish the late items, or did but performed poorly.

The second takeaway from the results in the current chapter is that the distinct strategic approaches identified by the cluster analysis appear to be correlated with individual differences in a number of both distal and proximal predictors. In addition to the behavioral differences described above, the clusters can be distinguished by three overarching patterns within the predictor variables included in the study (i.e., cluster differences in ability, motivation, and orientation toward technology use).

The pattern of ability differences and technology use orientations across clusters were the most straightforward. For the former, Cluster 3 and Cluster 1 had the highest and lowest average ability scores, respectively, across all three domains (verbal, spatial, numerical), although the effect was small for the numerical domain. For the latter, Cluster 3 also had the most negative technology attitudes and the least belief in AI's propensity to support high-quality problem-solving, with Cluster 2 endorsing the strongest/most positive orientation toward technology use across the same two variables. In addition to answering two post-assessment questions about whether they believed AI could help (1) produce high-quality solutions and (2) produce faster solutions, participants also answered a question about the extent to which they used any AI resource (e.g., ChatGPT) during the assessment. I checked to see if the technology orientation differences between clusters corresponded with patterns of self-reported AI use and found that it did: approximately 26.5% of participants in Cluster 2 reported using mostly or exclusively AI resources, compared to approximately 10% and 6.3% of Cluster 1 and Cluster 3 respectively.

Motivational differences across clusters were more mixed. Despite having the highest self-efficacy and verbal self-concept scores of any group, for example, Cluster 3 scored lower on spatial/math self-concept and the self-regulation scale compared to Cluster 1. Although the effects were relatively small, several interesting differences also emerged among Clusters 1 and 3 on the MTQ scales. Although their average scores on the Desire to

Learn scale were comparable, for example, Cluster 1 scored higher on both the Mastery and the Worry scales. One possible interpretation of these results (taken together with ability differences and item mean plots) is that while Cluster 1 and Cluster 3 both tended to be more thoughtful and conscientiousness about their problem solutions than Cluster 2, Cluster 1 may have overcompensated for performance concerns by allocating too much time and effort on the early items, at the expense of performance on a large majority of the assessment. Although self-regulation skills are typically positively associated with task performance, it may be that Cluster 1 exhibits some metacognitive skill deficits which negatively affected their ability to judge when to move on from a problem. That is, participants might actively attempt to engage in self-regulation (e.g., set goals/subgoals, self-monitor goal progress, or self-evaluate performance), but fail to do so accurately.

Overall, the cluster analysis suggests that behavioral patterns are a critical differentiator of strategic approaches to the assessment, but also that they do not tell the whole story. Notably, for example, Clusters 1 and 3 had relatively similar patterns of performance when considering raw item means, but they appear to have very different ability and motivational profiles. Taken together, the results of the last two chapters raise the following question: If the original focus on ‘good’ technological fluency performance/outcomes resulted in a lack of interpretable predictor-criterion relationships, how could I assess ‘good’ technological fluency differently? It is clear from the results of the cluster analysis that there are meaningful individual differences in strategic engagement with the assessment and that these differences are associated with multiple psychological characteristics. It may that process differences are more informative for understanding individual differences in technological fluency than raw sum score performance differences. For example, perhaps ‘good’ technological fluency requires individuals to be able to identify and use optimal search strategies. Are individual differences in ability or other variables associated with different search procedures? Where/how do people ‘go off course’? The next chapter explores these process questions using qualitative methods.

CHAPTER 7. QUALITATIVE ANALYSES OF PROBLEM-SOLVING PROCESSES

This chapter reports on a set of analyses related to the exploratory research question posed in this study, which focused on qualitative differences in participants' strategic approaches to the problem solving task. Whereas the originally posed question focused on differences in process between low and high performers on the assessment, however, these analyses examined differences in process between participants with low and high scores on the ability tests. This decision was made based on results from the preceding two chapters, which respectively (1) identified problematic aspects of the scoring methodology and (2) identified ability as an individual-differences variable that had a meaningful relationship with strategic approaches to the assessment.

I created a general ability composite by summing *z*-scores of the three domain-level ability test aggregate variables and then identified participants with the five lowest and highest scores on the composite (ten total). The qualitative analyses focused on the first six items on the assessment. I first reviewed the screen recordings and created a spreadsheet for each participant that consisted of a log of timestamped actions. Each new search, opened link, selected image, switched tab, or other action related to 'switching resources' was counted as a new line. Scrolling up and down within a page was not counted as a new action.

After recording all actions for a single item on a given participant's log, I reviewed the series of problem-solving procedures and took notes on several different aspects of participants' problem-solving processes. Some process components were selected to reflect the specific context of the current study, including the range of specific internet search strategies or common resources, as well as the frequency/nature of participants' AI use for problem-solving). Others were selected based on prior research on expert problem-solving. In a series of studies on novice versus expert approaches to problem-solving in physics, Chi et al. (1982) found that expert problem-solving processes were distinguishable in large part by the quality of inferences made during problem-solving, rather than by individuals' ability to identify key components of the problem or by the discrete steps taken (e.g., use of particular equations). Although technological fluency is inherently a less well-defined domain than physics problem-solving, I included columns for (1) participants' ability to extract key problem components and (2) inference quality and the ability to self-correct, in order to explore whether those findings would be replicated in the context of the current study.

7.1 Process Similarities Across Ability Profiles

7.1.1 Breadth of Search Behaviors/Resources

There was minimal variation in the actual search behaviors engaged by low versus high-ability participants. In most cases, participants began by navigating to Google to search some components of the problem, while a minority (approximately 23% of total items reviewed – mostly driven by one participant) went directly to ChatGPT. Those who began at Google primarily did so with an initial text search, although some started with an image search (most common in Items 1 and 2, which showed key problem components only visually). In all web search cases, participants did not navigate past the first page of results for the searches they made. After selecting some resource from the search results (e.g., a DIY forum; a blog post; a YouTube video), in almost all cases the participant tended to stay on the first 'level' of the resource. That is, although many of the websites and videos viewed by participants provided highly visible links to related resources, most people would only review their first selection. Instead, if they decided the original link was not useful, they either returned to the main results page to choose a different resource or initiated a new search altogether. Only four of the ten participants reviewed clicked a second-level link at least once, and even for these participants this behavior was not typically repeated across items.

7.1.2 Quality of Problem Component Extraction

Most problems either had multiple parts, or contained multiple criteria for what would constitute a ‘correct’ answer (e.g., in Question 5, proposed solutions must involve common household goods only and be able to be enacted quickly). Across the two ability groups, all participants demonstrated reasonable proficiency at identifying the key components of each problem, as reflected both in the fact that their search terms and eventual answers had coverage of the whole problem. This is consistent with Chi et al.’s (1982) findings that novice problem-solvers’ ability to identify key terms in a physics problem is not substantially less than that of expert problem-solvers. As will be discussed below, the meaningful distinctions between groups came from the quality of searches rather than their nominal coverage of relevant problem components.

7.2 Process Differences Across Ability Profiles

7.2.1 Inference Quality and Use of Visual Cues

Consistent with findings from Chi et al. (1982), a key differentiator between the problem-solving processes of participants in the low and high-ability groups was the quality of the inferences they made during problem-solving. Often, participants from each group had initial searches of comparable quality, but their paths diverged based on assumptions and inferences made early in the problem-solving process. This pattern occurred across items included in the review. In this section, I illustrate the primary pattern of results using responses to the first problem, which required participants to identify from a provided image the source of a persistent refrigerator smell, propose a temporary odor reduction method, and then list procedures that would resolve the issue.

High-ability participants were more capable of using information from the problem to infer which of several possible causes of the smell to investigate and seemed to have a more intuitive understanding of why a given potential solution would or would not be appropriate. Because phrases most often used by participants in search engines tended to lead to similar shortlists of common potential odor causes (e.g., dirty air filter, drain pan, coils, etc.) the high-ability group tended to encounter the same potentially misleading information as the low-ability group. Whereas low-ability participants tended to simply converge on potential causes that were described as the most common, high-ability participants were more likely to select a cause that was corroborated by information explicitly provided in the problem scenario.

Specifically, the high-ability group's primary means of locating the most likely source of the odor appeared to be the use of the visual cues provided in the problem (see

Item 1, **Appendix A**). In several cases, they began their search by proceeding directly or almost directly to image searches rather than regular text-based web search. Once engaged in these searches, they seemed better able to either select resources that looked similar to the prompt, or to revise the search to intentionally bring up labeled appliance diagrams for more information. One participant in the high-ability group, for example began their search by simply searching for “why does fridge smell,” and then conducted a series of revised image searches (see **Figure 12**), which eventually allowed them to recognize that the problem in the assessment handout depicted dirty condenser coils. After identifying the image in the assessment handout as dirty condenser coils, the participant was able to search for and successfully identify a resource with step-by-step procedures for cleaning this appliance component (although their final solution had credit deducted for failing to include major components of this procedure).

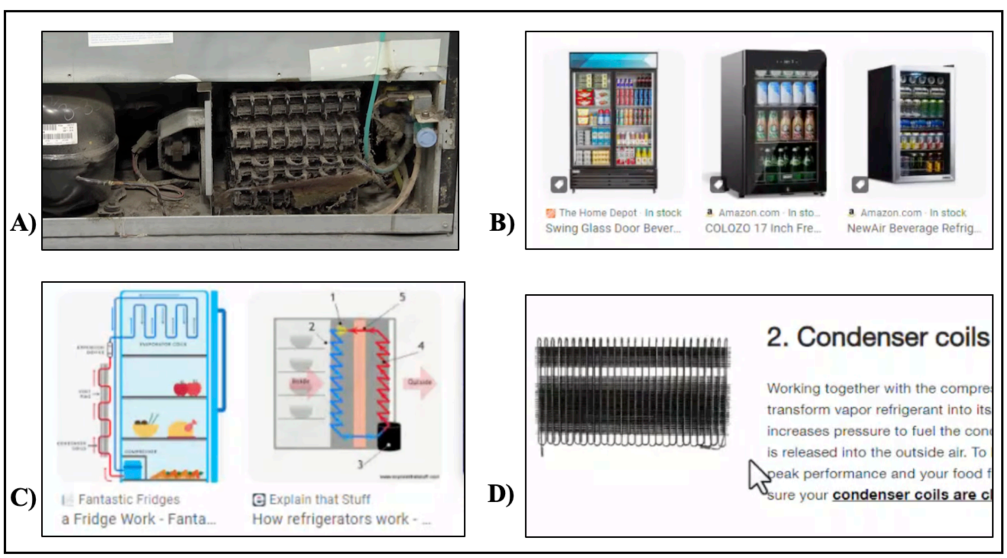


Figure 12 - Example of Good Quality Inference Process Based on Visual Problem Cues

Note: Panel A depicts the visual cue provided in the assessment handout. Panels B-D depict the iterative search procedures of a sample participant in the high-ability group, which get successively closer to and ultimately serve to identify the correct solution. Search activities accompanying panels B-D are, in alphabetical order, an image search for “fridge cooler”, an image search for “parts of a fridge explained”, and an excerpt from an article obtained from the home results page of the prior search.

In contrast, the low-ability group did not show the same capability to make effective inferences about the appropriateness of various solutions (i.e., evaluating the match between information explicitly given in the problem scenario and plausible-sounding related information found in online resources). **Figure 13** depicts the process by which a participant in the low-ability group quickly identified stagnant water in the drain pan as the most likely cause of the problem and begins to report this answer. Interestingly, this participant's answer claimed to be basing their response off of the visual cues given in the assessment handout, but did not validate their answer prior to deciding on it (e.g., by looking up images of drain pains to confirm it matched the image shown in the assessment handout). Their prompt to ChatGPT also provides a good illustration of the difference between identifying key problem components and abstracting from that information to make high-quality inferences: that is, they successfully picked up on cues that eliminate the possibility of rotted food or spills as a potential cause, and included that information as a qualifier in their search procedures.

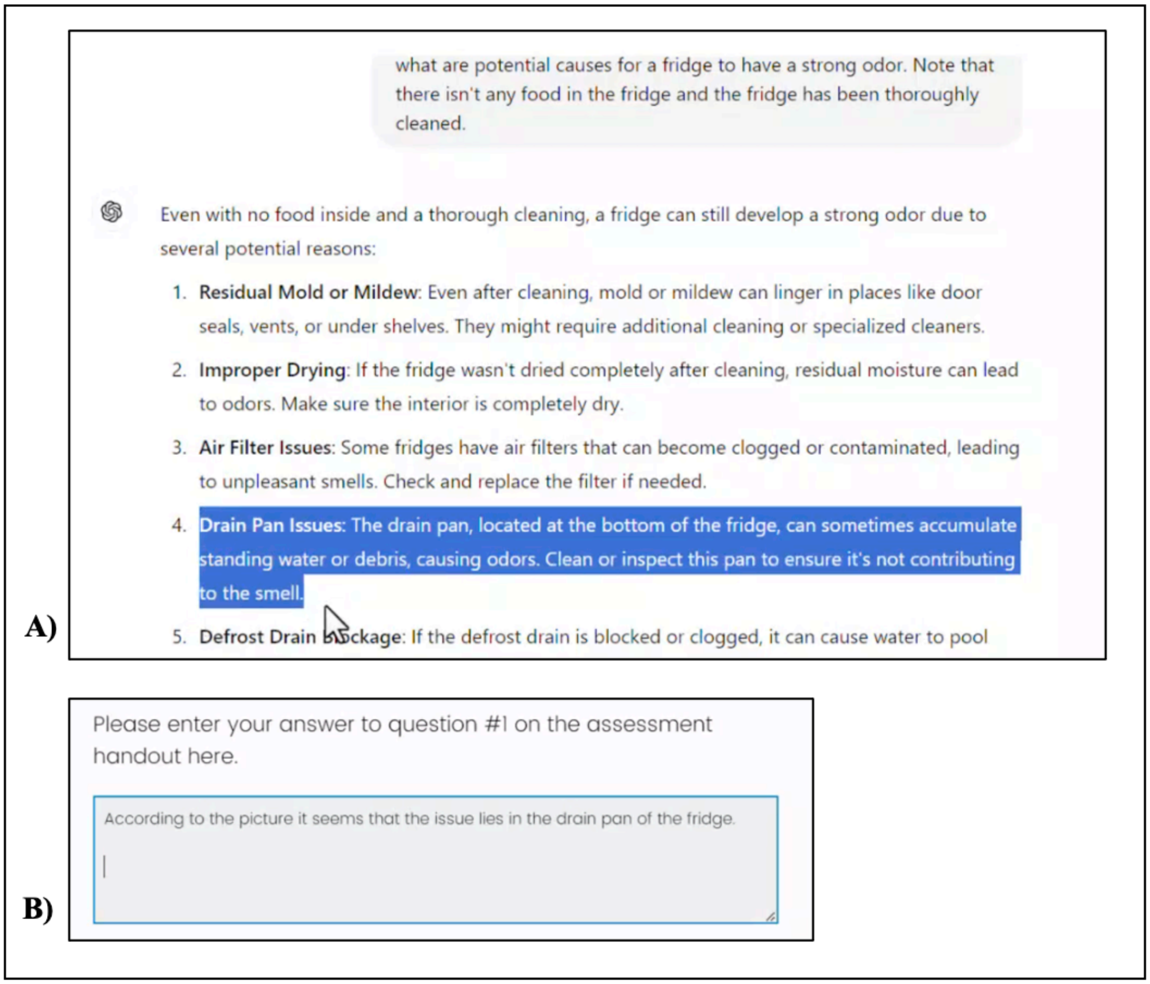


Figure 13 - Example of Poor Quality Inference Process
Note: Panel A depicts the ChatGPT-generated response that was used by the participant as the sole resource for identifying the cause of the odor. Panel B depicts the participant moving directly back to Qualtrics and beginning to enter their answer.

7.2.2 Ability to Self-Correct Faulty Influences

Although participants in both the low and high-ability groups may have reviewed potentially misleading information, the latter were more capable of correcting faulty assumptions, which may have led them down an incorrect path. For Item 1, four of the five high-ability group participants initially selected a cause of the odor (e.g., electrical outages, drain pan, mold) which did not correctly identify the issue depicted in the prompt. After continuing their search process, all four successfully self-corrected to identify condenser coils as the depicted issue, although their final item scores varied from 2.5 – 4 based on other components of their answers. In contrast, all five participants in the low-ability group initially selected an incorrect cause, but only one was able to self-correct during further search. This suggests that a key process difference between low and high-ability individuals may not be whether they get the right answer on the first try, but whether they are able and willing to recognize their potential errors and get back on track.

7.2.3 Nature of AI Use and Trust

A final notable difference related to participants' inference quality and capacity for self-correction concerns differences in low versus high-ability scorers' use of AI resources during the problem-solving process. There was some evidence that the low-ability group used AI more frequently: of 30 items reviewed per group (first six items for each of five participants), the low-ability group used an AI tool at least once during 14 items versus only 5 items total for the high-ability group. This was complicated, however, by the fact that Google launched a new beta tool that provides automatic AI-based summaries of search results while data collection was ongoing. This tool constituted the majority of AI use within the qualitative sample, and the low-ability group had more participants from the end of the data collection period, which occurred after the tool's launch. Given the small sample size associated with qualitative methods, it is difficult to say from this analysis whether ability-based differences in AI use frequency would replicate in the larger sample.

Clearer differences emerged when examining the qualitative nature of AI tool use within each group. The high-ability group more often used AI resources as more of an additional source of information or to generate search ideas, while participants in the lower-ability group were more likely to take the AI tool's response at face value without doublechecking or conducting additional searches. For example, consider the differences, shown in **Table 14**, between two participants' (one from each group) responses to Item 3, which asked them to (a) diagnose the cause of, and (b) provide procedures to fix the issue of a clothes steamer leaving stains on fabric. Both participants engaged with the Google AI summary of search results. Notably, the participant from the low-ability group did not further investigate likely causes provided by the results summary, moving directly from reviewing that resource to entering in a response for the first part of the problem. In contrast, the participant from the high-ability group read through both (1) resources linked in the Google AI compilation and (2) additional blog posts from further down the results page, and additionally revised their search terms to confirm their understanding of the probable cause/solution before beginning to enter their solution in Qualtrics.

Table 14 - Examples of Qualitative Differences in AI Tool Use

	Participant A - Low-ability group	Participant B - High-ability group
Action 1	• New search: ' why does my handheld steamer keep leaving faint stains on my clothes'	• New search: 'handheld steamer staining clothes'
Action 2	• Reviewed Google AI summary results which discussed mineral deposits; switched back to Qualtrics tab to answer part A	• Reviewed Google AI summary results which discussed mineral deposits; opened new tab with AI-provided link to article on cleaning procedures
Action 3	• Returned to results list; clicked first link which is to a yahoo.com article on cleaning a steamer	• Skimmed article (ctrl-F 'stains'); returned to main search results
Action 4	• Switched back and forth between yahoo.com article and Qualtrics tab to copy in procedures for part B	• Scrolled through search results p1; opened Philips article linking to several 2nd level sources for causes/fixes.
Action 5		• Tried to follow second level link to specific procedures for descaling (showing understanding of terminology) but link was broken.
Action 6		• New search: 'how to descale clothes steamer'
Action 7		• Scrolled through search results p1; opened a Conair article on using/decalcifying steamers
Action 8		• Read through Conair article; switched to Qualtrics tab and began answering question with text copied from multiple sources

7.3 Summary of Qualitative Process Analysis

The qualitative analyses build on the cluster analysis results by providing initial evidence about the nature of process differences related to individual differences in ability (i.e., a predictor with one of the larger differences between clusters). Notably, there are key areas that do *not* appear to differ based on ability – consistent with prior research on expert problem-solving (Chi et al., 1982), the separating factor seems to be how well individuals make inferences about information explicitly provided in a given problem, rather than how proficient they are at recognizing/extracting that information at the beginning of the problem-solving process. Relatedly, the majority of participants went off track at some point, and those with high composite ability scores were better able to self-correct upon encountering information which suggested their original conclusions were incorrect. Although mid-data collection changes in participants' exposure to AI makes it difficult to extrapolate, there also appear to be some differences in the frequency and nature of AI resource use during problem-solving favoring more restricted use by high-ability participants. Given the preliminary nature of conclusions drawn from a small sample in this chapter, the final chapter of the results section returns to the full sample to re-assess distal and proximal predictors of process indicators.

CHAPTER 8. EXPLORATORY QUANTITATIVE ANALYSIS OF KEY PROCESS INDICATORS

Collectively, the results of the preceding chapters suggest that individual differences in problem-solving processes may be more informative of the quality of technological fluency than aggregate performance on the assessment. In Chapter 6, I identified three clusters of participants across whom the largest mean difference was the process variable of within-person standard deviation on item-level performance. In Chapter 7, I observed qualitative patterns of key differences between high- and low-scorers on the ability battery, including in participants' use of AI resources during the search process. In this chapter, I leverage those findings to re-assess one of the original goals of the study, which was to develop a battery of proximal measures to improve prediction of technological fluency above and beyond the utility of distal trait measures. Specifically, I conduct two sets of regression analyses assessing the criterion-related validity of distal vs. proximal measures for the two process variables identified as potentially fruitful in Chapter 6 and Chapter 7 that had corresponding quantitative measures in the full sample: respectively, (1) within-person standard deviation of item scores and (2) self-reported AI use frequency.

8.1 Distal Trait Complexes

Before conducting the regression analyses, I created trait complex composite variables comprising distal ability, personality, and interest variables, as well as the self-concept and motivational trait scales. I took this initial step because while there appeared to be motivational differences among clusters in Chapter 6 that could plausibly be correlated with process indicators, adding another 10 predictors (four self-concept scales and six MTQ scales) to the regression analysis would likely introduce substantial issues related to multicollinearity and interpretability. The trait complex approach improves parsimony in the subsequent analyses while still allowing for the inclusion of additional motivational predictors.

I constructed the trait complex composite variables in two steps. First, I conducted an exploratory factor analysis (with oblimin rotation) to identify distal predictors which shared meaningful common variance. Parallel analysis suggested six factors; however, I also examined a five-factor solution for two reasons: (1) with 18 indicators (three ability variables, two personality, three interest, four self-concept, and six motivational trait), a six-factor solution was stretching the minimum preferable limit on indicators per factor, and (2) in the six-factor solution, realistic interests did not load saliently on any factor – this was surprising given its pattern of correlations with other distal variables (see **Table 9**) and its theoretical relevance to many of the mechanical or 'hands-on' items included on the assessment of technological fluency.

Because the five-factor solution was more parsimonious and showed a more theoretically consistent loading for realistic interests, I used this solution to construct the trait complex composites. Factor solutions and factor correlations for the five-factor solution are shown in **Table 15** and **Table 16** respectively. For variables with cross-loadings (i.e., multiple factor loadings above 0.30), indicators were assigned to the factor with the higher loading. Conventional interests did not load on any factor and were excluded from further analysis. The five extracted trait complexes were named as follows (factors 3 and 5 included only scales for higher-level traits in the MTQ and were named for those traits): I. Mastery/Intellectual; II. Math/Science/Technology; III. Motivation-Related Anxiety; IV. Verbal Ability/Openness to Experience; and V. Competitive Excellence. In the second step after the exploratory factor analysis, I calculated composite variables which consisted of sums of unit-weighted *z*-scores for indicated traits within each complex. A correlation matrix including the five trait complex composite variables plus the proximal predictors is provided in **Table 17**.

Table 15 - Distal Trait Rotated Factor Solution

	Mastery/ Intellectual	Math/Science/ Technology	Motivation-Related Anxiety	Verbal Ability/ Openness	Competitive Excellence
Verbal Ability	-0.010	0.010	-0.080	**0.730**	0.050
Spatial Ability	-0.190	**0.560**	-0.050	**0.340**	0.180
Numerical Ability	-0.250	**0.550**	-0.060	0.270	0.080
Realistic Interests	-0.030	**0.330**	-0.190	-0.030	-0.140
Investigative Interests	**0.360**	-0.020	0.000	0.270	-0.070
Conventional Interests	0.050	0.240	0.150	-0.280	0.150
Conscientiousness	**0.560**	-0.040	-0.160	-0.170	0.160
Openness to Experience	0.180	0.020	0.040	**0.560**	-0.140
Verbal Self-Concept	**0.510**	-0.170	0.000	**0.390**	0.070
Spatial Self-Concept	0.230	**0.470**	-0.170	0.020	0.000
Math Self-Concept	0.210	**0.780**	-0.070	-0.160	0.000
Technology Self-Concept	0.170	**0.440**	-0.050	0.090	0.000
MTQ – Competitiveness	-0.030	0.030	-0.270	-0.050	**0.740**
MTQ – Other-Oriented Goals	0.140	0.030	0.240	0.050	**0.740**
MTQ – Mastery	**0.730**	-0.030	-0.050	-0.020	0.220
MTQ – Desire to Learn	**0.750**	0.290	0.010	0.060	-0.110
MTQ – Emotionality	-0.070	-0.070	**0.850**	-0.070	-0.030
MTQ – Worry	0.020	0.020	**0.900**	0.020	0.030

Note: "MTQ" refers to "Motivational Trait Questionnaire". Loadings greater than 0.30 are bolded.

Table 16- Distal Trait Rotated Factor Correlation Matrix

	Mastery/ Intellectual	Math/Science/ Technology	Motivation-Related Anxiety	Verbal Ability/ Openness	Competitive Excellence
Mastery/ Intellectual	1				
Math/Science/ Technology	0.20	1			
Motivation-Related Anxiety	-0.11	-0.36	1		
Verbal Ability/ Openness	0.02	0.20	-0.08	1	
Competitive Excellence	0.18	0.15	-0.16	0.10	1

Table 17 - Correlation Matrix of Distal Trait Complex Composites and Proximal Predictors

	1	2	3	4	5	6	7	8	9	10	11
1. Mastery/ Intellectual	1										
2. Math/Science/ Technology	0.266	1									
3. Motivation-Related Anxiety	-0.194	-0.452	1								
4. Verbal Ability/ Openness	0.230	0.293	-0.125	1							
5. Competitive Excellence	0.258	0.206	-0.185	0.068	1						
6. AI Quality	-0.163	-0.074	0.006	-0.289	0.07	1					
7. AI Speed	0.031	0.090	-0.002	-0.165	0.232	0.35	1				
8. Technology Attitudes	0.003	0.123	-0.038	-0.219	-0.056	0.206	0.228	1			
9. Prior Experiences	0.080	0.307	-0.105	0.019	-0.042	-0.032	-0.060	0.357	1		
10. Self-Regulation	0.564	0.170	-0.085	-0.186	0.134	0.087	0.020	0.113	0.182	1	
11. Self-Efficacy	0.301	0.349	-0.135	0.175	0.115	-0.059	-0.135	-0.105	0.083	0.124	1

8.2 Distal and Proximal Predictors of Within-Person Variability in Technological Fluency

I first modeled a regression equation in which the five distal trait complexes were entered simultaneously to explain variance in within-person standard deviation for technological fluency. Effect sizes of both R^2 and Cohen's f^2 are reported. Next, I added in the proximal predictors (technology experiences, technology attitudes, self-regulatory skill, self-efficacy, as well as the supplemental items which asked about participants' beliefs that AI use would be helpful for solution quality and/or speed) and assessed whether this second step improved prediction of within-person variability. **Table 18** shows the summary results of both models and model comparison.

Table 18 - Summary of Hierarchical Multiple Regression Results for Within-Person Variability in Technological Fluency

	Step 1				Step 2			
	B	***SE***	***t***	***p***	***B***	***SE***	***t***	***p***
Distal – Trait Complexes								
Mastery/Intellectual	0.01	0.03	0.39	0.698	-0.04	0.03	-1.211	0.228
Math/Science/Technology	0.03	0.03	1.18	0.242	0.03	0.03	1.03	0.303
Motivation-Related Anxiety	0.04	0.05	0.69	0.492	0.02	0.05	0.477	0.634
Verbal Ability/Openness	0.06	0.06	1.04	0.300	0.04	0.06	0.58	0.563
Competitive Excellence	-0.01	0.05	-0.22	0.828	-0.01	0.05	-0.17	0.866
Proximal								
Self-Regulation Skill					0.21	0.12	1.83	0.069
Technology Attitudes					-0.14	0.10	-1.42	0.159
Prior Experiences					-0.07	0.10	-0.78	0.440
Self-Efficacy					0.09	0.10	0.99	0.323
AI Quality					-0.28	0.10	-2.86	0.005*
AI Speed					0.05	0.10	0.51	0.611
R^2	0.03				0.15			
ΔR^2	-				0.12 ($F = 2.732$, $p = 0.016$*)			

Note: p-values lower than an α value of 0.05 are marked with *.

The set of proximal variables accounted for a statistically significant increase in variance accounted for in within-person variability on the assessment of technological fluency ($F = 2.73$, $p = 0.02$). Model 1 (distal trait complexes) accounted for approximately 3% of total variance ($f^2 = 0.03$). Model 2 (distal trait complexes plus proximal predictors) accounted for 15% of total variance ($f^2 = 0.17$), accounting for a significant increase in variance explained of approximately 12% ($\Delta f^2 = 0.14$).

8.3 Distal and Proximal Predictors of AI Use Frequency

The second outcome variable of interest in this chapter is self-reported frequency of AI use during the assessment, from exclusive use of non-AI resources to exclusive use of AI resources (i.e., AI use frequency categories can be interpreted as essentially ranging from never to always). As shown in **Table 19**, the top frequency category was selected by only three participants. Because such a low number would produce extremely unreliable estimates in a model, the top two levels were combined for further analysis. While there is still considerable unevenness in sample sizes across levels and results of this exploratory regression analysis should therefore be interpreted with caution, this change was intended to account for the most severe sample size disparities (i.e., categories where N < 10).

Table 19 - Distribution of Self-Reported AI Use Frequency Categories

Exclusively non-AI	Majority non-AI	About equal	Majority AI	Exclusively AI
76	25	10	19	3

Because these categories have a meaningful order (i.e., the higher levels represent more use of AI), I used an ordinal logistic regression approach rather than a multinomial logistic regression approach. Whereas multinomial logistic regression ignores order and models the probability (in terms of odds ratios) of each category compared to a reference group, ordinal logistic regression assumes that the categorical outcome variable represents an underlying/latent continuous variable (e.g., frequency of AI use). It estimates the effect of the predictors on the probability of moving to a 'higher' level of the ordered categorical outcome variable.

As in the previous section, I first modeled a regression equation in which the five distal trait complexes were entered simultaneously to explain variance in AI use frequency.

Effect sizes are calculated using Nagelkerke's R^2, a pseudo-R^2 measure designed for use in logistic regression contexts (Nagelkerke, 1991). Next, I added in the proximal predictors (technology experiences, technology attitudes, self-regulatory skill, self-efficacy, as well as the supplemental items which asked about participants' beliefs that AI use would be helpful for solution quality and/or speed) and assessed whether this second step improved prediction. **Table 20** shows the summary results of both models and model comparison, including odds ratios for each predictor. Model fit is evaluated using both Akaike's Information Criterion (AIC) and the Likelihood Ratio Test (LRT).

Table 20 - Summary of Hierarchical Ordinal Logistic Regression Results for Frequency of AI Use

Predictor	Step 1					Step 2				
	B	SE	Odds Ratio	*z*	*p*-value	*B*	SE	Odds Ratio	*z*	*p*-value
Distal – Trait Complexes										
Mastery/Intellectual	-0.01	0.05	0.99	-0.13	0.900	0.01	0.07	1.01	0.07	0.941
Math/Science/Technology	0.05	0.05	1.05	0.89	0.371	0.03	0.06	1.03	0.54	0.590
Motivation-Related Anxiety	0.08	0.10	1.09	0.82	0.413	0.06	0.11	1.06	0.52	0.605
Verbal Ability/Openness	-0.34	0.11	0.72	-2.93	0.003*	-0.32	0.14	0.73	-2.36	0.018*
Competitive Excellence	0.18	0.11	1.20	1.70	0.089	0.08	0.11	1.09	0.75	0.454
Proximal										
Self-Regulation Skill						0.01	0.25	1.01	0.03	0.973
Technology Attitudes						-0.48	0.21	0.62	-2.25	0.025*
Prior Experiences						0.23	0.21	1.26	1.12	0.261
Self-Efficacy						-0.14	0.20	0.87	-0.70	0.484
AI Quality						0.21	0.21	1.24	1.03	0.302
AI Speed						0.85	0.24	2.34	3.54	<0.001*
Nagelkerke's Pseudo-R^2	0.10					0.25				
AIC	303.38					294.11				
Likelihood Ratio Test (LRT)	-					$\chi^2(8, 14) = 21.27$, $p = 0.002$*				

Note: *p*-values lower than an α value of 0.05 are marked with *.

As in the previous section, adding the proximal variables significantly improved prediction of participants' frequency of AI use above and beyond the distal trait complexes alone ($\chi^2(8, 14) = 21.27$, $p = 0.002$), from accounting for approximately a tenth of the variability in self-reported AI use frequency (trait complexes only) to accounting for approximately a quarter of the variability (trait complexes plus proximal variables). Unexpectedly, more positive technology attitudes were associated with lower odds (OR = 0.62) of more frequent AI use – it is possible that this may be an artifact of both smaller group sizes in the more frequent AI categories and somewhat restricted variability (SD = 0.38) in the technology attitudes scale.

8.4 Summary of Process Regression Models

Together, the pattern of results in Chapter 8 suggest that the proximal variables do improve prediction of process variables related to technological fluency, above and beyond the explanatory utility of distal trait complexes. In addition to this general pattern of results, there are two additional notable findings that are more specific in nature. First, although the uneven groups in the AI use frequency variable warrant caution in interpreting the results of the second set of regressions, it is still interesting to note that trait complexes (particularly, the Verbal Ability/Openness to Experience complex) appeared to be more predictive of AI use frequency than of within-person variability on technological fluency item scores (with trait complexes explaining approximately 10% vs. 3% of variability for each criterion, respectively).

One implication is the relationship of distal/proximal predictors with technological fluency processes may vary across specific process aspects. Second, the belief that AI tools are useful in developing higher-quality solutions was negatively correlated with both the Mastery/Intellectual (*r* = -0.16) and Verbal Ability/Openness (*r* = -0.29) trait complexes, and not meaningfully correlated with the Math/Science/Technology complex (*r* = -0.07) – suggesting that beliefs about the utility of specific technological resources may be less related than expected to ability (e.g., spatial/math ability) and non-ability (e.g., spatial/math/technology self-concept; realistic interests) traits typically thought to be predictive of technology use.

CHAPTER 9. DISCUSSION

9.1 Interpreting Technological Fluency Performance Data and Correlates

One of the major goals of the current study was to develop sets of distal and proximal predictor measures that are useful in predicting performance on an assessment of technological fluency. Unfortunately, relatively foundational problems were apparent in the criterion measure which meant the originally proposed analyses (i.e., concerning relative criterion-related validity of distal vs. proximal predictors of performance) would be difficult to meaningfully interpret. The first unexpected finding when conducting preliminary analyses was that the assessment of technological fluency failed to correlate meaningfully with most distal or proximal predictors. Although it is technically possible that the criterion simply did not share variance with these measures, the conceptual overlap between them and the breadth of constructs included in the predictor battery made this highly unlikely. Indeed, further exploration in Chapter 5 revealed specific item-level effects (e.g., negative correlations in the inter-item matrix, item-level floor effects) that suggested some feature of the assessment itself was introducing more complex concerns.

One possible explanation is that as a construct, technological fluency is not well suited for aggregation to a single outcome score (or an aggregation of outcome scores) based on the 'correctness' of a given solution(s), and measurement should instead also consider the extent to which participants tend to engage in (or learn over time to engage in) the particular adaptive strategies explored in Chapter 6 and Chapter 7. In retrospect, theoretical reasoning to support this hypothesis can be found in Ackerman et al.'s (2023) conceptualization of technological fluency as an iterative process, which converges to a solution over time as a function of an individual's multiply-determined strategic approach to a problem (see **Figure 1**).

I originally argued that because search procedures should theoretically be related to solution adequacy/quality, scoring the assessment on solution quality would also account for individual differences in strategic approaches to problem-solving. However, the unexpected lack of predictor-criterion correlations in **Table 9** suggests that other determinants of solution 'correctness' (as assessed in the current study) substantially limit my ability to use performance as a proxy for process. That is, my findings suggest that 'good process' may not always result in 'good performance,' or vice versa. It may be, for example, that someone whose true technological fluency is relatively low might have deceptively high scores on the assessment if they happened to come across an 'ideal' solution early in their search or as part of a larger list (e.g., as provided by tools like ChatGPT), and copied it over without understanding why or how it was correct. This is particularly plausible for assessment procedures which, like the current study, do not require participants to physically enact the solution procedures they develop.

9.2 Individual Differences in Strategic Approaches to Technological Fluency

9.2.1 Ability and Non-Ability Variables Influence Assessment Strategy

Chapter 6 revealed more interesting patterns of differences in strategic approaches to the assessment – particularly, differences in item score patterns between groups whose performance was relatively consistent throughout the assessment (Cluster 2) versus groups with varying rates of performance decline throughout the assessment (Cluster 1 and Cluster 3). Whereas predictor correlates of aggregate performance were mostly weak and somewhat theoretically inconsistent, taking a person-centric approach to categorizing assessment performance provided clearer insight. One illustrative example of what the person-centered approach revealed that was obscured by assessment-level correlations is the relationship between self-reported self-regulatory skill (a key hypothesized proximal influence on technological fluency) and within-person assessment performance. Scores on the self-regulation scale were not significantly related to aggregate assessment score. Closer examination of the cluster analysis results, however, suggest a more meaningful relationship that may have been driven in part by characteristics of participants in the first cluster, who had both (1) the highest average self-regulation scores and (2) low relative item scores and the lowest average number of items completed compared to other groups.

Interestingly, even though Cluster 1 and Cluster 3 participants had more similar overall response pattern structures with each other than with Cluster 2, their ability and non-ability profiles showed clear discrimination (although some effects were relatively small). Cluster 3 generally had higher ability scores, and was more confident (e.g., higher self-efficacy scores), while Cluster 1 scored lower on ability tests and reported higher anxiety (e.g., MTQ – Worry). Taken together with their surprisingly high self-regulation scores, one possible interpretation is that the first cluster cared about doing well, but their performance was hampered by a lack of the self-monitoring skill required to appropriately regulate behavior. The potential impacts of metacognitive deficits (e.g., reduced self-monitoring or inaccurate self-evaluation) on technological fluency are discussed in more detail in the next section. Cluster 2, in contrast, seemed less engaged generally than both other groups, with lower scores on investigative/intellectual traits (although effects were small) and higher scores on technology beliefs/attitudes.

9.2.2 *Qualitative Findings Related to Individual Differences in Problem-Solving Process*

Whereas Chapter 6 provided initial evidence that process differences are a useful source of information on individual differences in technological fluency, the goal of Chapter 7 was to more specifically identify the search procedures which could comprise 'good' vs. 'bad' performance on the technological fluency assessment. Consistent with the relatively weak ability-performance relationship found in Chapter 5, although participants in the high-ability group tended to demonstrate superior reasoning processes and self-management during the search process, these differences were often not reflected in their item-level scores. The participant highlighted in **Figure 12**, for example, did not provide sufficient procedural information in their submitted response to that item to receive a high score despite the fact that they successfully self-corrected from a misguided search strategy and located detailed, appropriate resources. I argue that this (i.e., the fact that the measurement and scoring procedures did not capture adaptive process variables like self-correction) is a core reason that predictor-criterion relationships were so constrained in **Table 9**.

Two more specific findings had important implications for the contributions of the current study as well as directions for future research. First, although participants in both groups chose search terms that were reasonably representative of key terms and phrases in the assessment problems, the high-ability group tended to make better quality inferences during the search. One possible explanation for this finding concerns the role of domain knowledge and knowledge structures in problem-solving, previously studied extensively in well-defined problem contexts such as physics problem-solving (e.g., Chi et al., 1982; Elio & Scharf, 1990; Larkin et al., 1980; Priest & Lindsay, 1992). Although the relatively ill-defined, real-world problem-solving context of the current study is somewhat different from the prior focus on the physics domain, there are notable similarities between my results and previous conclusions about experts' superior ability to make problem-relevant inferences that could apply across contexts. For example, expert problem-solving in physics requires the ability to relate literal problem cues to abstract knowledge about relevant mathematical principles, while expert technological fluency might require participants to relate problem cues to abstract knowledge about various DIY, home, or personal scenarios that could facilitate a more effective solution to the given problem. It may be that participants with more expert knowledge structures concerning a given domain (e.g., cooking) are better able to make appropriate inferences about the quality of a potential solution or resource encountered during search.

The second key finding in Chapter 7 concerned differences in the quality of self-regulatory and metacognitive processes favoring the high-ability group, including self-monitoring and self-evaluation (e.g., double checking or critically considering the appropriateness of some resource before including it in an answer, self-correct after going off course). Prior research in domains including information literacy (Brand-Gruwel et al., 2005) and business management (Güss et al., 2017) has found that self-regulation activities such as self-monitoring are more frequent in expert performance. These processes seem particularly critical for just-in-time, ill-defined learning situations such as those relevant to technological fluency, but more data are needed to understand the mechanisms by which technological resources impact self-regulatory processes during problem-solving. One recent paper, published while data collection was ongoing for this study, suggests that ChatGPT may make it more difficult for learners to self-evaluate their performance on ill-defined problem-solving tasks because access to this tool increases perceived ease of the task (Urban et al., 2024). This is consistent with prior research on metacognitive biases suggesting that experiences of 'fluency' (i.e., ease) during problem-solving or academic tasks may result in maladaptive strategic decisions including the premature cessation of effort (e.g., R. Ackerman, 2019; Bjork et al., 2013).

Ultimately, my results suggest that more direct measurements of the strategic processes involved in technological fluency are needed, either to supplement or replace performance measures. Although I cannot conclude from the current data whether strategic differences between high and low-ability groups stemmed from differences in their *ability* or *willingness* to engage in self-regulatory processes during problem-solving, one potential explanation for the exploratory qualitative findings is that the assessment may have been closer to the zone of tolerable problematicity for high-ability participants – in other words, a range of difficulty within which they are capable of and willing to apply self-regulatory strategies to develop a solution to a novel problem (Elshout, 1987). One implication of Ackerman et al.'s (2023) model of technological fluency (see **Figure 1**) is that the zone of tolerable problematicity impacts whether a participant will choose to engage with a task – it may be that high-ability individuals were more likely to perceive the technological fluency assessment as 'worth the effort'. Unfortunately, the current study design and scoring procedures did not allow me to measure these processes directly. The qualitative analyses generated several interesting hypotheses about the specific processes that comprise 'good' technological fluency, but future work should more directly probe inferences made during problem-solving (e.g., by think-aloud or other related procedures).

9.3 Criterion-Related Validity of Distal Trait Complexes and Proximal Variables for Technological Fluency and Just-in-Time Learning

In Chapter 8, including proximal variables consistently improved prediction of the two process indicators above and beyond distal predictors alone, providing some evidence in support of one of my key predictions (i.e., that distal traits have relatively limited utility for predicting technological fluency). Beyond this general finding, however, I am somewhat limited in the conclusions I can draw from these exploratory analyses. This is because even though the outcome variables used in this chapter were the best options of the currently available data, they are still imperfect proxies of participants' problem-solving processes. For example, although inference quality would be more directly informative of participants' competence on the problem-solving tasks than within-person variability or frequency of AI use, I do not have a quantitative measure of the former available in the current data. If technologically fluent individuals are ultimately more capable of just-in-time learning in real-world settings, including in the workplace (Ackerman et al., 2023), and technological fluency is best assessed by process rather than aggregate performance (as suggested by my collective results), then a lack of direct quantitative process measures is problematic because it limits my ability to determine the individual differences that act as facilitating or inhibiting influences on learning.

Based on the current evidence, it does not appear that trait complexes are highly predictive of technological fluency processes. In fact, even the full (distal + proximal) models had relatively modest explanatory power. However, regardless of whether trait complexes predict process variables directly, they may serve as predictors of a critical variable unassessed in the current study: domain knowledge. Prior research has identified relatively consistent patterns of trait complexes which tend to be either facilitating or inhibiting influences on the acquisition of domain knowledge in both traditional academic contexts (Ackerman, Bowen, et al., 2001; Ackerman, Chamorro-Premuzic, et al., 2011; Ackerman, Kanfer, et al., 2013) and real-world contexts (Beier & Ackerman, 2003). Therefore, one plausible explanation for the relatively modest explanatory power of the models in Chapter 8, which is consistent with key implications of my qualitative findings, is that these models exclude an important mediator (i.e., knowledge). Additional research is needed to determine whether the same patterns of facilitating/inhibiting effects of trait complexes on the acquisition of relevant domain knowledge found in prior research also occur for just-in-time learning in information-rich environments.

9.4 Limitations and Future Research

The current study had several limitations, two primary sets of which (related to measurement of the criterion and measurement of predictors respectively) are discussed below. Throughout the section, directions for future research are proposed that could contribute to the resolution of these issues.

9.4.1 Limitations and Future Directions Related to the Criterion

9.4.1.1 Assessment Procedures

Two procedural limitations of the technological fluency assessment were identified (1) before data collection and (2) after scoring/during analysis. First, the original assessment included only text-based items, and did not require participants to link to any sources. After the first three participants had completed the second session of the study, I revised the assessment to partially mitigate this concern. Unlike in previous piloting with undergraduate research assistants, the first participant exclusively used verbatim ChatGPT-generated answers, and seemed confident in a post-debriefing discussion that a large majority of other participants would behave similarly. The second and third participants similarly relied mostly on AI tools. Although I was concerned that consistent use of AI in the entire sample would lead to problematic reductions in performance variance, I decided against banning ChatGPT in the assessment instructions because generative AI represents an increasingly common informational resource and removing it as an option for participants would have reduced the 'real-world' fidelity of the assessment.

Instead, I made two revisions to the assessment which were intended to reduce the utility of typing problems verbatim into ChatGPT. Because in the real world people encounter these events (e.g., home repair) visually, I replaced text descriptions with visual representations of problems where possible, requiring participants to engineer their own prompts by inferring the problem and how to describe it. I also required participants to provide linked 'shopping lists' of tools and supplies in several items, because real-life problem-solving doesn't stop at identifying what is needed but also requires people to choose among alternative types of tools/supplies and locate where they can be purchased. Interestingly, the results of Chapter 7 suggested that individual differences in process quality are most apparent when the prompt required interpreting visual cues (e.g., as in the first two items on the assessment) – suggesting that future research may benefit from focusing on problem scenarios which can be represented in multimedia formats.

A related, unforeseen factor that may have influenced results is that the AI landscape is undergoing rapid change. As a result, the quality of tools publicly available to participants changed over the course of seven months of data collection. For example, the standard (free) version of ChatGPT was upgraded during the course of data collection from version 3.0 to version 4o mini, the latter of which had substantial upgrades in reasoning capacity, multimedia processing, and search capability (OpenAI, 2024a, 2024b). Additionally, beginning in May 2024 and continuing to roll out in larger waves through the end of data collection, Google implemented AI Overview at the top of search results. This was a key resource among later participants who engaged with AI, and the uneven rollout combined with two large beginning/ending waves of data collection in Spring (February-April) and Fall (August-September) 2024 means that for large subsets of participants, there was a confounding variable affecting highly visible features of the search environment.

The second assessment limitation, which was identified after scoring, was that treating the assessment as a speeded test may have made it more difficult to interpret individual differences in aggregate performance. Because participants were time limited with a substantial number of items (21), participants with lower quality answers on average may still have received relatively high scores if they moved through the assessment very quickly and responded to most or all items. Not only did this likely contribute to some of the psychometric issues noted in Chapter 5, it also doesn’t fully reflect the nature of experiencing these types of tasks in real life. Although the logistical issues associated with a laboratory study informed my decision to treat the assessment as a speeded test, this may not be the optimal procedure for assessing technological fluency.

9.4.1.2 Improving Measurement of Inferences and Metacognitive Processes

As mentioned in the previous section, items which required participants to interpret visual cues seemed to produce the most striking individual differences in process quality. I suspect that this is because in these situations, participants could not simply match terms from the problem to terms encountered during search and were therefore required to make more independent inferences about the relevance of various informational resources. Future research should attempt not only to emphasize items that force participants to make these types of inferences, but also to incorporate study designs which make inferences generally more directly observable. Think-aloud protocols are one type of procedure previously used in problem-solving research which serve to make participants' inferences explicit to the researcher (e.g., Chi et al., 1982; Larkin et al., 1980). Problem-solving tasks might also be followed by brief interviews, where researchers return to a recording of the participant's actions and ask them to guide the researcher through their thought processes.

A related direction for future research concerns more direct evaluation of the self-regulatory beliefs that might explain why some individuals are more likely than others to self-correct when engaged in technological problem-solving. Not only could this provide more insight into whether poor inferences are better explained by individual differences in ability or non-ability variables (which has implications for the trainability of technological fluency), but it could also address current calls in the literature to better understand the nature of metacognitive heuristics and other self-beliefs (e.g., self-efficacy) in the context of problem-solving with generative AI. Urban et al. (2024) found that: (1) participants assigned to use ChatGPT for a creative problem-solving task had lower difficulty perceptions of the task, and (2) lower perceptions of task difficulty were associated with greater over-estimation of performance – but also noted that these findings were based on a single measurement taken at the completion of the task. Experience sampling approaches might support more dynamic characterizations of the development of metacognitive heuristics/beliefs during episodes of technology-supported problem-solving. Patterns of or differences in metacognitive heuristics or specific monitoring/control judgments (Bjork et al., 2013) could then be compared across participants engaging in certain types of strategic behavior (e.g., those who do vs. do not successfully self-correct after beginning down an incorrect search path).

9.4.2 Limitations and Future Directions Related to Predictor Measurement

Although it was not the case for most predictors in the current study, a small number of theoretically important variables did exhibit either range restriction (e.g., technology attitudes) or skew (e.g., self-efficacy), potentially impacting predictor-criterion correlations. Self-efficacy in particular showed negative skew, which given the relatively poor assessment performance in the sample suggests that there was some overconfidence present. It is not clear from the current data what the reason for this may have been, but two possibilities are (1) that confidence in identifying/describing a solution and confidence in having/developing the procedural skill to enact it are related but distinct concepts, and/or (2) that participants were unable to provide valid judgments of their confidence, given a lack of prior first-hand experiences with the exact items included on the assessment. Although I did include a measure of prior experiences, those items were focused more on participants' experiences with leveraging technology for problem-solving, rather than the actual hands-on experience itself. Future researchers should consider including experience measures more directly tied to hands-on procedural skill mastery in the assessed domains.

As delineated in investment-based frameworks of intellectual development like PPIK (i.e., Intelligence-as-Process, Personality, Interest, and Intelligence-as-Knowledge; Ackerman, 1996), I would expect that these revised experience measures would also be correlated with knowledge in the relevant domains.

The current study's lack of knowledge tests was an important limitation. While the original justification for this decision was to establish a foundation of evidence related to the included distal/proximal predictors before expanding the range of individual difference predictors to include domain knowledge as a mediator in future studies, my results suggest that individual differences in relevant domain knowledge may be key to making sense of individual differences in technological fluency, especially as it relates to individuals' ability to make high-quality inferences during search (see Chapter 7). One reasonable question might be whether the most important knowledge tests to include would be those related to knowledge of the directly assessed domains (e.g., home repair, cooking, cleaning/organization) or those related to knowledge of content-agnostic, optimal search procedures (e.g., how to formulate search terms, select/evaluate resource quality). Given prior research on the relevance of domain-specific knowledge structures in expert problem-solving (e.g., Chi et al., 1982), my expectation is that the former would be more informative.

Specifically, I expect that prior domain knowledge should be positively correlated with raw performance on the items relevant to those domains, as well as negatively correlated with time allocated to the same items. Regarding the negative relationship between knowledge and time to solution, I expect that prior knowledge would serve as a partial replacement to search to the extent that individuals do not need to search for information or procedures with which they already familiar. These expectations are based on prior findings from Tidler et al. (2022), who found that scores on a knowledge test of computer components were negatively correlated with completion times on a computer hardware upgrade task ($r_{knowledge,\ time}$ = -0.55) as well as with references to the Internet during the task ($r_{knowledge,\ Internet\ references}$ = -0.52). A follow-on effect of faster search in the context of the current study is that reducing the need for search on some items leaves additional time to devote to other items where the participant may have had less relevant prior knowledge, thereby potentially improving aggregate performance. This interaction between prior knowledge and search strategy also highlights the importance of an assessment that covers multiple domains. That is, because my main focus is on technological fluency itself (i.e., domain-general problem-solving competencies), it would be problematic to assess only one domain (e.g., DIY repairs) where prior knowledge could reduce the need to engage in search throughout the entirety of the assessment.

9.5 Conclusions

The current study made three primary contributions to the literature. First, I developed and implemented an assessment of technological fluency that could be distinguished from similar concepts including digital/information literacy. In contrast to computer and digital literacy assessments, which attempted to assess whether individuals are capable of completing specific, often self-contained tasks (e.g., respectively, basic software use; Hoffman & Blake, 2003; evaluating the credibility of information found online; Bhatt & MacKenzie, 2019), the technological fluency complex accounts for both ‘can do’ and ‘will do’ components influencing the likelihood that individuals will choose to leverage technological aids to solve a novel, external problem (Ackerman et al., 2023).

The construct of technological fluency is expected to have several important practical implications related to changes in the nature of work tasks. Nonroutine problem-solving is increasingly viewed as a desirable competency by organizations (Neubert et al., 2015), and technological aids such as the Internet are widely used by individuals seeking to engage in just-in-time learning or decision-making contexts (Duran-Nelson et al., 2013), suggesting that individuals' propensity to engage these aids effectively will have practical implications for job performance. Two additional contributions were related to (1) the identification and organization of relevant ability and non-ability correlates of technological fluency, and (2) methodological recommendations focused on assessing iterative problem-solving processes. Conclusions derived concerning these topics are presented in turn below.

9.5.1 Conclusions About Ability and Non-Ability Correlates of Technological Fluency

Originally, I sought to provide conclusions about the extent to which the inclusion of proximal strategic variables improved prediction of technological fluency performance above and beyond the contribution of distal trait variables alone. Although criterion measurement issues made the originally planned analyses problematic, the results of alternative and exploratory analyses did provide insight about ability and non-ability distal/proximal variables in relation to the technological fluency assessment. In Chapter 8, for example, I found incremental validity of proximal variables above and beyond distal trait complexes for predicting two aspects of process: (1) within-person variability and (2) frequency of AI use. This provides some evidence in favor of a general expectation of the current study as well as a central proposition of Ackerman et al. (2023). From a person-centered perspective, mean scores on both ability and non-ability variables differed between three clusters of prototypical item response patterns, suggesting that there may be meaningful psychological differences associated with strategic approaches participants took while completing the assessment that are not captured by correlations with aggregate performance.

9.5.2 Conclusions About Process Variables and Assessment Methodology

The qualitative findings, which emphasize the importance of processes like inference making and the ability to engage in self-correction, are consistent with prior complex problem-solving research (Chi et al., 1982; Elio & Scharf, 1990; Larkin et al., 1980; Priest & Lindsay, 1992). What is perhaps unique about the current context is that process quality does not seem to be as directly linked to solution quality as it might be in more well-defined domains (e.g., physics problem-solving). One important question is whether this is fundamentally related to participants' open access to technological tools (including recent developments such as generative AI), or whether it was simply a limitation of the scoring procedures. The answer is likely to be a combination of both.

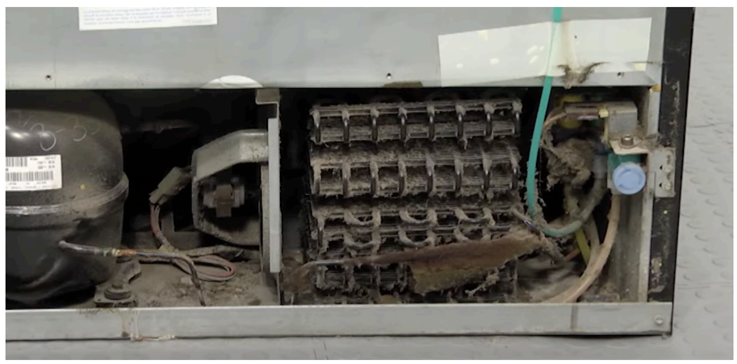
Concerning the latter, my observation of participants during data collection and exploratory qualitative analysis of search processes suggest that some high item scores may have been 'accidental' in that participants simply selected the first suggested answer online without much thought, while some lower item scores may have been assigned to participants who thoughtfully selected resources and corrected misinterpretations, but just failed to provide sufficient details in their written answer. Together, this suggests that revisions to the scoring procedure which take process variables into greater account are warranted.

Concerning the former, there are probably mixed effects of access to search resources on search quality. As discussed by Ackerman et al. (2023), this question has strong relevance to longstanding debates on the effects of calculators in classrooms on student learning – on the one hand, enabling students to conduct operations more quickly or to perform more complicated procedures, while on the other, potentially inhibiting their competence with foundational computational skills. Outside of the classroom, similar arguments have been made regarding the impact of GPS on spatial navigation and cognition more generally (e.g., Gillett & Heersmink, 2019). It is hard to dispute the fact that online search can, *if used effectively*, help to mitigate a lack of direct prior relevant experience with a particular problem scenario. Nonetheless, that condition is critical, and some recent evidence (Urban et al., 2024) does suggest that tools like generative AI may increase the risk of common metacognitive biases/heuristics like experience-based fluency cues (i.e., reductions of perceived difficulty) which risk premature cessation of effort (Bjork et al., 2013). My findings provide initial evidence that the quality of metacognitive processing in technology-supported problem-solving may be correlated with individual differences in ability – this is a critical direction for future research on technological fluency.

APPENDIX A. ASSESSMENT OF TECHNOLOGICAL FLUENCY

Items were presented to participants in a paper packet, with one item per page. All answers were entered online in the Qualtrics questionnaire.

1. You notice a bad odor coming from your refrigerator, which persists even after you have thoroughly cleaned it and checked that no food has rotted. Upon inspecting the fridge, you identify a potential cause (shown below).
    a. Describe the best temporary solutions to mask the odor.
    b. Identify and describe the problem shown in the picture below.
    c. Describe procedures to address the issue in Part b.



2. When you moved into your current apartment, you caused some damage (pictured below) while mounting electronics and hiding cables. You want to repair this damage before moving out to avoid being charged part of your security deposit.
    a. What tools and supplies will be needed to fix the damage? Provide links to all items you would need for the repair.
    b. What procedures are involved to fix the damage?

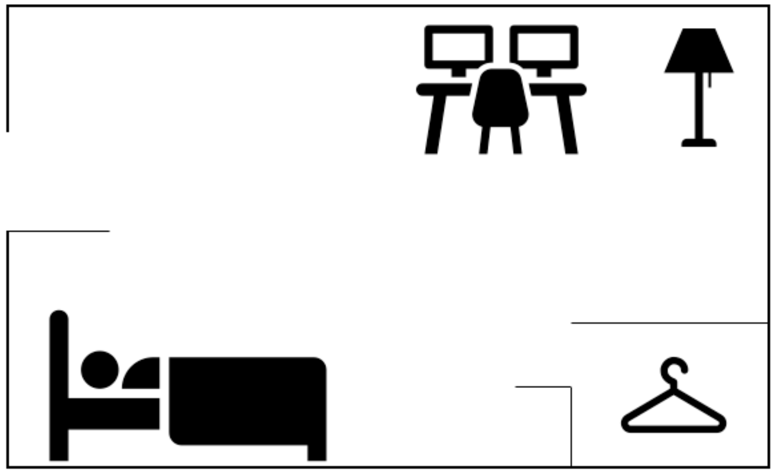
3. You notice that lately, your handheld steamer has been leaving faint stains on clothes as you de-wrinkle them. Explain the most likely cause of this issue, and describe the procedures you should take to stop the steamer from leaving stains.

4. One of the hinges on your kitchen cabinet has become loose, and screws will no longer hold. Find the most appropriate temporary fix that will secure the hinge without leaving permanent damage until you can replace it.

5. You just noticed that the sweater you were planning on wearing out tonight has bad pilling throughout both arms and the neck. You don't have time to pick up a fabric shaver before you need to leave. What is the best way to quickly remove the pilling without damaging the sweater, using household items?

6. You enjoy using cast iron pans, but in your new apartment kitchen (pictured below), you are having trouble heating food consistently. What is the most likely reason this might be happening, and how could you fix the problem?

7. Based on the bedroom layout shown below, how could you make it possible (without moving furniture) to turn the lamp on and off from your bed?



8. The zipper on your favorite pair of jeans won't stay up. Identify a way to keep it closed when you're out until you can replace the zipper, and describe how to make this fix.

9. You want to hang a 25-pound mirror in your bedroom. There are no studs in your preferred location, and the mirror's frame does not have any installation hardware included. What specific equipment (including hardware) will you need to make sure it is hung correctly? Provide links to each item which will be necessary to hang the mirror.

10. You accidentally stripped a screw when assembling a piece of furniture and are not able to remove it with a screwdriver. You do not own a screw extractor. What is the best way to remove the screw without damaging the hole or further stripping the screw? Explain why this method is appropriate to the problem.

11. Many of your favorite items of clothing cannot be put into a drying machine because they will shrink. You typically hang them to dry in your shower, but want a dedicated drying rack which is adjustable, includes hooks, and can be stored compactly/easily. Your budget is $75. What is the best option for a drying rack that meets these criteria? Provide a link to the best option.

12. Your electricity bill has been higher than usual in the apartment you rent. List as many realistic ways as you can to reduce your energy consumption costs and create an action plan to implement these strategies.

13. Your friend is moving into a new apartment, and you have decided to gift them a do-it-yourself set containing tools and supplies they might need for a wide variety of common rental problems. Your budget is $100. List the contents that should be included. Provide a link to each item.

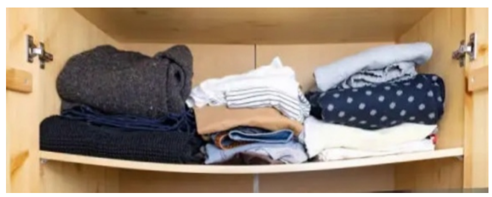
14. Your only storage space is in your bedroom closet, but it is small, and you own several bulky coats and other items which you only wear a few months out of the year. Find a storage system solution which will free up hanging space without creating floor clutter or damaging your bulky clothing items. Provide a link to the item(s) you would purchase in order to create the storage solution.

15. After one of your friends was involved in a not-at-fault car accident, you decided to install a dashcam in your car. You have a $250 budget and would like a dashcam which can report both front and rear road angles at a minimum, but are interested in other popular features as well. What is your best option for a dashcam with the most comprehensive set of included features? Provide a link to the dashcam you select, report its price, and list each of its features.

16. The temperature in your apartment seems to fluctuate a lot regardless of what you set the thermostat to. You need to demonstrate to the apartment manager what the problem is, by tracking the temperature over a few weeks, with objective documentation. Describe (a) what information you would need to collect and (b) how you would collect this information to provide to the apartment manager?

17. You recently received a bouquet of flowers, and want to preserve several of the larger flowers to display in your apartment. However, they are too bulky to press flat without ruining their appearance, and you don't want to display them in a vase.
    a. Describe an alternate way that you can preserve the flowers, including all materials you will need to do so.
    b. Describe how you would display them. Provide links to the items you would need to purchase in order to solve this problem.

18. When you open your dresser, you notice the problem shown below, and are worried about it breaking if it gets any worse. Describe (a) how you could permanently repair this damage, and (b) how you can temporarily reinforce it to prevent further damage in the meantime.



19. You have recently gotten into hiking, and decide that you want to train to hike one of Colorado's 'fourteener' mountains in July (four months from now). You typically go for jogs about three to four times a week, but have little experience hiking or exercising at high altitudes. Excluding travel time to and from the mountain, you would like to complete the hike in a single day.
    a. Choose the mountain to climb, and explain your choice.
    b. Plan a physical training schedule that you will follow to prepare for the hike.

20. You have decided that you want to pursue the Project Management Professional certification, to improve your chances of a promotion at your current job. The certification requires you to pass a written exam which will take place three months from now. Write out a personalized study plan based on the content of the exam that will maximize your chances of passing.

21. You are interested in taking up vegetable gardening as a part of the program in the flyer shown below. Based on the information in the flyer:
    a. List the vegetables you would plant over the course of the rental, keeping in mind planting seasons.
    b. List the equipment you will purchase to get started using the provided credit, not counting seeds or soil. Provide links to all items.

APPENDIX B. ITEMS FOR NON-COPYRIGHTED SELF-REPORT SCALES

Table B1. Technology Attitudes

Item
1. Automobile 'black boxes' (e.g., devices which record and transit driving data to your insurance company or car manufacturer)
2. Epoxy resin
3. Auto-complete search assistance
4. Content moderation on Internet forums
5. Voice assistants (e.g., Amazon Echo, Apple Homepod)
6. Facebook
7. Virtual reality
8. Twitter (X)
9. Restaurant rating apps (e.g., Yelp)
10. Spotify
11. Web browser activity tracking (e.g., Cookies)
12. Calculators
13. GPS (Global Positioning Systems)
14. Electric stoves
15. Chatbots
16. Computers in the K-12 (Kindergarten - 12th Grade) classroom
17. Online whistleblowing databases (e.g., Wikileaks, DDoSecrets)
18. Stud finders
19. Artificial Intelligence
20. Targeted Internet advertising
21. WikiHow
22. Neurological mind-reading devices
23. AI-based evaluations of job applications
24. Speed detection cameras
25. Wikipedia
26. Amazon.com
27. Content recommendation algorithms
28. Dashcams
29. Smart thermostats
30. TikTok
31. YouTube
32. Search engines (e.g., Google, Bing)

Table B2. Prior Technology Experiences

Scale/Item	Wording
Qualitative	1. When I encounter a technical issue (e.g., a problem with my laptop or phone), I try to solve it myself before seeking professional help.
	2. I often use online forums or communities to seek advice for personal problems or challenges.
	3. When new technology or software is released, I often spend time testing its capabilities.
	4. I am more likely to research how to make simple home/auto repairs (e.g., repairing drywall, changing a car taillight) myself, rather than calling a maintenance worker or mechanic.
	5. I watch/follow particular YouTube channels to be more informed about new technologies.
	6. For work-related tasks, I frequently find and use online resources or software to increase my efficiency.
	7. I often find myself looking up tips and tricks to improve my cooking.
	8. My hobbies require me to look up information, assistance, or other resources online.
Quantitative	1. Fix household items using instructions I found online
	2. Use online resources (e.g., YouTube tutorials) to teach myself new skills
	3. Watch educational videos online to understand complex topics
	4. Take online courses or workshops to develop professional skills
	5. Compare features and/or prices online before making a significant purchase
	6. Post my own how-to videos to the Internet
	7. Use online resources to research and develop personalized exercise routines to accomplish fitness goals (e.g., build muscle, lose weight)
	8. Look-up something online in the middle of a discussion, to answer a question or get supporting information
	9. Surf the Internet to get ideas of new activities to try or places to visit
	10. Keep up on new technology by reading relevant blogs or dedicated websites (e.g., Gizmodo, Endaget, Popular Mechanics)
	11. Keep up on current clothing trends by reading relevant blogs or dedicated websites (e.g., my favorite influencers or the websites of my favorite brands)
	12. Take apart computers/other technology to see how these items work

Table B3. Self-Regulatory Skills

Item
1. I determine how to solve a task before I begin.
2. I organize my time to best accomplish my goals.
3. I carefully plan my course of action.
4. I ask myself questions about the material before I begin.
5. I try to understand tasks before I attempt to solve them.
6. I set specific goals before I begin a task.
7. I try to understand the goal of a task before I attempt to answer.
8. I think of several ways to solve a problem and choose the best one.
9. I figure out my goals and what I need to do to accomplish them.
10. I read instructions carefully before I begin a task.
11. I imagine the parts of a task I have to complete.
12. I think about what I really need to learn before I begin a task.
13. I make sure I understand just what has to be done and how to do it.
14. I pace myself while learning in order to have enough time.
15. I try to determine what the task requires.
16. Before I read through information thoroughly, I often skim it to see how it is organized.
17. When I try to solve a problem, I set goals for myself in order to direct my activities.

Table B4. Self-Efficacy/Overclaiming

Scenarios and Increasing Levels of Solution Quality
Scenario 1: You notice a bad odor coming from your refrigerator, which persists even after you have thoroughly cleaned it and checked that no food has rotted. Upon inspecting the fridge, you identify something that looks like a potential cause. (a) Describe the best temporary solutions to mask the odor, (b) identify and describe the most likely source of the problem, and (c) describe procedures to address the issue identified in part B. • I could develop a partly adequate solution to this problem. (C-) • I could develop an adequate solution to this problem. (C) • I could develop a slightly better than adequate solution to this problem. (C+) • I could develop a moderately good solution to this problem. (B-) • I could develop a good solution to this problem. (B) • I could develop a very good solution to this problem. (B+) • I could develop an excellent solution to this problem. (A-) • I could develop an optimal solution to this problem. (A/A+)
Scenario 2: When you moved into your current apartment, you caused some damage while mounting electronics and hiding cables. You want to repair this damage before moving out to avoid being charged part of your security deposit. Given a picture of the damage, (a) list what equipment will be needed and (b) describe the procedures are involved to fix the damage. • I could develop a partly adequate solution to this problem. (C-) • I could develop an adequate solution to this problem. (C) • I could develop a slightly better than adequate solution to this problem. (C+) • I could develop a moderately good solution to this problem. (B-) • I could develop a good solution to this problem. (B) • I could develop a very good solution to this problem. (B+) • I could develop an excellent solution to this problem. (A-) • I could develop an optimal solution to this problem. (A/A+)
Scenario 3: You notice that lately, your handheld steamer has been leaving faint stains on clothes as you de-wrinkle them. Explain the most likely cause of this issue, and describe the procedures you should take to fix it. • I could develop a partly adequate solution to this problem. (C-) • I could develop an adequate solution to this problem. (C) • I could develop a slightly better than adequate solution to this problem. (C+) • I could develop a moderately good solution to this problem. (B-) • I could develop a good solution to this problem. (B) • I could develop a very good solution to this problem. (B+) • I could develop an excellent solution to this problem. (A-) • I could develop an optimal solution to this problem. (A/A+)

Scenario 4: One of the hinges on your kitchen cabinet has become loose, and screws will no longer hold. Find the most appropriate temporary fix that will secure the hinge without leaving permanent damage until you can replace it.

- I could develop a partly adequate solution to this problem. (C-)
- I could develop an adequate solution to this problem. (C)
- I could develop a slightly better than adequate solution to this problem. (C+)
- I could develop a moderately good solution to this problem. (B-)
- I could develop a good solution to this problem. (B)
- I could develop a very good solution to this problem. (B+)
- I could develop an excellent solution to this problem. (A-)
- I could develop an optimal solution to this problem. (A/A+)

Scenario 5: You just noticed that the sweater you were planning on wearing out tonight has bad pilling throughout both arms and the neck. You don't have time to pick up a fabric shaver before you need to leave. What is the best way to quickly remove the pilling without damaging the sweater, using household items?

- I could develop a partly adequate solution to this problem. (C-)
- I could develop an adequate solution to this problem. (C)
- I could develop a slightly better than adequate solution to this problem. (C+)
- I could develop a moderately good solution to this problem. (B-)
- I could develop a good solution to this problem. (B)
- I could develop a very good solution to this problem. (B+)
- I could develop an excellent solution to this problem. (A-)
- I could develop an optimal solution to this problem. (A/A+)

Scenario 6: You enjoy using cast iron pans, but in your new apartment kitchen, you are having trouble heating food consistently. What is the most likely reason this might be happening, and how could you fix the problem?

- I could develop a partly adequate solution to this problem. (C-)
- I could develop an adequate solution to this problem. (C)
- I could develop a slightly better than adequate solution to this problem. (C+)
- I could develop a moderately good solution to this problem. (B-)
- I could develop a good solution to this problem. (B)
- I could develop a very good solution to this problem. (B+)
- I could develop an excellent solution to this problem. (A-)
- I could develop an optimal solution to this problem. (A/A+)

Note: Scores are summed within scenarios, then averaged across scenarios to calculate an aggregate score.

Table B5. Domain Interest and Curiosity

Scale/Item	Wording
Interest	
On a scale of 1 (very boring) to 6 (very interesting), how interesting do you find each of the following domains?	
1.	Furniture and home maintenance/repairs
2.	Making simple appliance repairs
3.	Home and personal technology (e.g., cameras, smart home tools)
4.	Personal care and aesthetics (e.g., home decoration, mending clothes)
5.	Home organization and space optimization
6.	Planning self-improvement projects (e.g., medium- to long-term education or fitness goals)
Curiosity	
On a scale of 1 (not at all curious) to 6 (very curious), how curious would you be to learn more about each of the following domains?	
1.	Furniture and home maintenance/repairs
2.	Making simple appliance repairs
3.	Home and personal technology (e.g., cameras, smart home tools)
4.	Personal care and aesthetics (e.g., home decoration, mending clothes)
5.	Home organization and space optimization
6.	Planning self-improvement projects (e.g., medium- to long-term education or fitness goals)

APPENDIX C. FORMAL HYPOTHESIS TESTING

C.1 Hypotheses Related to Differential Criterion Relationships

Hypotheses 1-7 dealt with assessing differences in the magnitude of correlations between distal and proximal predictors respectively with performance on the assessment of technological fluency. To test these hypotheses, I conducted a series of Steiger's z-tests for dependent correlations (Steiger, 1980). Because conducting a relatively large number of such tests on the same dataset is associated with an increased risk of Type I error (i.e., erroneously rejecting the null), I applied corrections for multiple comparisons. Specifically, I calculated adjusted p-values for each test using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 2000), which I selected over other corrections the family-wise error rate (e.g., the Bonferroni method) because it retains greater statistical power than family-wise error controls while still reducing risk of Type I errors – which is a valuable feature in exploratory contexts with a large number of tests.

The results of the Steiger's z-tests are shown in Table C1. The rows list distal predictors included in each test, while columns list proximal predictors. In all cases, scores on the technological fluency criterion measure were the shared (dependent) third variable. For each cell in the table, pairwise correlations are shown between: 1) the distal predictor and the criterion, 2) the proximal predictor and the criterion, and 3) the distal and proximal predictors. Additionally, the test's z-score, uncorrected p-value, and corrected p-value are reported. Hypotheses 1-7 were evaluated using the corrected p-values in the relevant cells, against an α of 0.05.

Table C1. Summary of results: Steiger's z-tests of dependent correlations

	Self-Regulation (9)	**Tech. Attitudes (10)**	**Prior Experiences (11)**	**Self-Efficacy (12)**
Verbal Ability (1)	$r_{1,criterion}$ = 0.103 $r_{9,criterion}$ = -0.096 $r_{1,9}$ = -0.212 z = 1.461 p = 0.144 $p_{corrected}$ = 0.384	$r_{1,criterion}$ = 0.103 $r_{10,criterion}$ = 0.228 $r_{1,10}$ = -0.151 z = -0.964 p = 0.335 $p_{corrected}$ = 0.564	$r_{1,criterion}$ = 0.103 $r_{11,criterion}$ = 0.164 $r_{1,11}$ = -0.029 z = -0.496 p = 0.620 $p_{corrected}$ = 0.169	$r_{1,criterion}$ = 0.103 $r_{12,criterion}$ = -0.055 $r_{1,12}$ = 0.147 z = 1.376 p = 0.793 $p_{corrected}$ = 0.386
Spatial Ability (2)	$r_{2,criterion}$ = 0.162 $r_{9,criterion}$ = -0.096 $r_{2,9}$ = -0.046 z = 2.046 p = 0.041* $p_{corrected}$ = 0.284	$r_{2,criterion}$ = 0.162 $r_{10,criterion}$ = 0.228 $r_{2,10}$ = -0.065 z = -0.533 p = 0.594 $p_{corrected}$ = 0.792	$r_{2,criterion}$ = 0.162 $r_{11,criterion}$ = 0.164 $r_{2,11}$ = 0.091 z = -0.022 p = 0.983 $p_{corrected}$ = 0.983	$r_{2,criterion}$ = 0.162 $r_{12,criterion}$ = -0.055 $r_{2,12}$ = 0.203 z = 1.965 p = 0.049* $p_{corrected}$ = 0.284
Numerical Ability (3)	$r_{3,criterion}$ = 0.135 $r_{9,criterion}$ = -0.096 $r_{3,9}$ = -0.131 z = 1.764 p = 0.078 $p_{corrected}$ = 0.284	$r_{3,criterion}$ = 0.135 $r_{10,criterion}$ = 0.228 $r_{3,10}$ = -0.061 z = -0.744 p = 0.457 $p_{corrected}$ = 0.665	$r_{3,criterion}$ = 0.135 $r_{11,criterion}$ = 0.164 $r_{3,11}$ = -0.053 z = -0.231 p = 0.818 $p_{corrected}$ = 0.902	$r_{3,criterion}$ = 0.135 $r_{12,criterion}$ = -0.055 $r_{3,12}$ = 0.183 z = 1.701 p = 0.089 $p_{corrected}$ = 0.284
Conscientiousness (4)	$r_{4,criterion}$ = -0.010 $r_{9,criterion}$ = -0.096 $r_{4,9}$ = 0.567 z = 1.063 p = 0.288 $p_{corrected}$ = 0.511	$r_{4,criterion}$ = -0.010 $r_{10,criterion}$ = 0.228 $r_{4,10}$ = 0.083 z = -2.027 p = 0.043* $p_{corrected}$ = 0.284	$r_{4,criterion}$ = -0.010 $r_{11,criterion}$ = 0.164 $r_{4,11}$ = -0.026 z = -1.393 p = 0.164 $p_{corrected}$ = 0.386	$r_{4,criterion}$ = -0.010 $r_{12,criterion}$ = -0.055 $r_{4,12}$ = 0.224 z = 0.412 p = 0.680 $p_{corrected}$ = 0.806
Openness to Experience (5)	$r_{5,criterion}$ = 0.013 $r_{9,criterion}$ = -0.096 $r_{5,9}$ = -0.094 z = 0.844 p = 0.399 $p_{corrected}$ = 0.638	$r_{5,criterion}$ = 0.013 $r_{10,criterion}$ = 0.228 $r_{5,10}$ = -0.209 z = -1.601 p = 0.109 $p_{corrected}$ = 0.318	$r_{5,criterion}$ = 0.013 $r_{11,criterion}$ = 0.164 $r_{5,11}$ = 0.059 z = -1.266 p = 0.205 $p_{corrected}$ = 0.438	$r_{5,criterion}$ = 0.013 $r_{12,criterion}$ = -0.055 $r_{5,12}$ = 0.142 z = 0.589 p = 0.556 $p_{corrected}$ = 0.773
Investigative Interests (6)	$r_{6,criterion}$ = -0.080 $r_{9,criterion}$ = -0.096 $r_{6,9}$ = 0.172 z = 0.149 p = 0.882 $p_{corrected}$ = 0.936	$r_{6,criterion}$ = -0.080 $r_{10,criterion}$ = 0.228 $r_{6,10}$ = -0.226 z = -2.267 p = 0.023* $p_{corrected}$ = 0.284	$r_{6,criterion}$ = -0.080 $r_{11,criterion}$ = 0.164 $r_{6,11}$ = -0.049 z = -1.931 p = 0.054 $p_{corrected}$ = 0.284	$r_{6,criterion}$ = -0.080 $r_{12,criterion}$ = -0.055 $r_{6,12}$ = 0.341 z = -0.248 p = 0.804 $p_{corrected}$ = 0.902
Realistic Interests (7)	$r_{7,criterion}$ = -0.042 $r_{9,criterion}$ = -0.096 $r_{7,9}$ = 0.001 z = 0.441 p = 0.659 $p_{corrected}$ = 0.806	$r_{7,criterion}$ = -0.042 $r_{10,criterion}$ = 0.228 $r_{7,10}$ = 0.012 z = -2.214 p = 0.027* $p_{corrected}$ = 0.284	$r_{7,criterion}$ = -0.042 $r_{11,criterion}$ = 0.164 $r_{7,11}$ = 0.167 z = -1.830 p = 0.067 $p_{corrected}$ = 0.284	$r_{7,criterion}$ = -0.042 $r_{12,criterion}$ = -0.055 $r_{7,12}$ = 0.207 z = 0.117 p = 0.907 $p_{corrected}$ = 0.936
Conventional Interests (8)	$r_{8,criterion}$ = 0.039 $r_{9,criterion}$ = -0.096 $r_{8,9}$ = 0.139 z = 1.182 p = 0.237 $p_{corrected}$ = 0.475	$r_{8,criterion}$ = 0.039 $r_{10,criterion}$ = 0.228 $r_{8,10}$ = 0.213 z = -1.739 p = 0.082 $p_{corrected}$ = 0.284	$r_{8,criterion}$ = 0.039 $r_{11,criterion}$ = 0.164 $r_{8,11}$ = 0.093 z = -1.064 p = 0.287 $p_{corrected}$ = 0.511	$r_{8,criterion}$ = 0.039 $r_{12,criterion}$ = -0.055 $r_{8,12}$ = -0.038 z = 0.746 p = 0.456 $p_{corrected}$ = 0.665

Note: p-values lower than an α value of 0.05 are marked with *.

After correcting to account for the potential of Type 1 error, no significant effects remained (lack of support for H1-H7). However, because some tests of dependent correlations were significant before corrections, and the large number of tests meant that corrections were relatively strict, some relationships emerged which merit further mention. Most notably, technology attitudes were positively correlated with technological fluency ($r = 0.23$), and as predicted (H5), this effect was significantly different from the relationship between technological fluency and distal traits including conscientiousness, investigative interests, or realistic interests (see **Table C1** for exact effect magnitudes). The correlation of spatial ability with performance on the criterion measure ($r = 0.17$) was also significantly different from those of proximal variables including self-regulation and self-efficacy with the criterion measure (see **Table C1** for exact effect magnitudes), but in the opposite direction as predicted (H1).

C.2 Hypotheses Related to Incremental Prediction

Hypotheses 8 and 9 dealt with incremental criterion-related validity of the proximal predictors above and beyond distal traits. To test these hypotheses, I conducted a series of hierarchical regression analyses. All predictor variables were standardized before analysis. To test Hypothesis 8, I first modeled a regression equation in which distal traits (cognitive ability, personality, vocational interests) were entered simultaneously. To assess the magnitude of the distal predictors' ability to predict technological fluency, I calculated R^2 as well as Cohen's f^2. Next, I tested Hypothesis 9 by adding in the proximal predictors (technology experiences, technology attitudes, self-regulatory skill, self-efficacy) and assessed whether this second step improves prediction of technological fluency by the predicted magnitude in R^2 and f^2. **Table C2** shows the summary results of both models and model comparison.

Table C2. Summary of hierarchical multiple regression results.

	Step 1				Step 2			
	B	***SE***	***t***	***p***	***B***	***SE***	***t***	***p***
Distal – Ability								
Verbal Ability	0.47	1.09	0.43	0.670	0.39	1.09	0.36	0.720
Spatial Ability	1.33	1.22	1.09	0.278	1.21	1.21	1.00	0.319
Numerical Ability	0.72	1.20	0.60	0.550	0.89	1.19	0.75	0.455
Distal – Non-Ability								
Conscientiousness	-0.15	1.00	-0.15	0.884	0.73	1.20	0.61	0.543
Openness to Experience	0.07	1.06	0.06	0.950	0.18	1.05	0.18	0.861
Investigative Interests	-0.93	1.02	-0.92	0.361	-0.09	1.04	-0.09	0.931
Realistic Interests	-0.91	0.99	-0.92	0.359	-0.99	0.99	-1.00	0.320
Conventional Interests	0.67	0.98	0.68	0.496	0.15	0.97	0.15	0.882
Proximal								
Self-Regulation Skill					-1.56	1.16	-1.34	0.182
Technology Attitudes					2.11	1.02	2.07	0.041*
Technology Experiences					1.40	1.03	1.37	0.175
Self-Efficacy					-0.69	1.01	-0.68	0.498
R^2	0.05				0.13			
ΔR^2	-				0.08 ($F = 2.644$, $p = 0.037$*)			

Note: p-values lower than an α value of 0.05 are marked with *.

As predicted, the set of proximal variables accounted for a statistically significant increase in variance accounted for in participants' performance on the assessment of technological fluency ($F = 2.644$, $p = 0.037$) – however, the proportion of variance explained was less than expected. Model 1 (distal traits only) accounted for approximately 5% of total variance rather than the predicted 10-15% ($f^2 = 0.05$, rather than the predicted range of 0.11-0.18). Model 2 (distal traits plus proximal predictors) accounted for 13% of total variance ($f^2 = 0.14$), accounting for a significant increase in variance explained of approximately 8% ($\Delta f^2 = 0.09$, rather than the predicted improvement of at least 0.15). Therefore, the evidence for H8 and H9 is mixed.

REFERENCES

Ackerman, P. L. (1989). Within-task intercorrelations of skilled performance: Implications for predicting individual differences? A comment on Henry & Hulin, 1987. *Journal of Applied Psychology, 74*(2), 360-364. https://doi.org/10.1037/0021-9010.74.2.360

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22*(2), 227–257. https://doi.org/10.1016/S0160-2896(96)90016-1

Ackerman, P. L. (2003). Cognitive ability and non-ability trait determinants of expertise. *Educational Researcher*, *32*(8), 15-20. https://doi.org/10.3102/0013189X032008015

Ackerman, P. L. (2021). *New aptitude and attitude assessments study* [Unpublished study].

Ackerman, P. L., et al. (2023). *Technological fluency and individual differences in just-in-time learning* (Manuscript in progress).

Ackerman, P. L., Bowen, K. R., Beier, M. E., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology*, *93*(4), 797–825. https://doi.org/10.1037/0022-0663.93.4.797

Ackerman, P. L., Chamorro-Premuzic, T., & Furnham, A. (2011). Trait complexes and academic achievement: Old and new ways of examining personality in educational contexts. *British Journal of Educational Psychology, 81*(1), 27–40. https://doi.org/10.1348/000709910X522564

Ackerman, P. L., & Ellingsen, V. J. (2014). Vocabulary overclaiming — A complete approach: Ability, personality, self-concept correlates, and gender differences. *Intelligence*, *46*, 216–227. https://doi.org/10.1016/j.intell.2014.07.003

Ackerman, P. L., & Kanfer, R. (2004). Cognitive, affective, and conative aspects of adult intellect within a typical and maximal performance framework. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives* *on intellectual functioning and development* (pp. 119–141). Lawrence Erlbaum Associates Publishers.

Ackerman, P. L., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, *105*(3), 911–927. https://doi.org/10.1037/a0032338

Ackerman, R. (2019). Heuristic cues for meta-reasoning judgments: Review and methodology. *Psychological Topics, 28*(1), 1–20. https://doi.org/10.31820/pt.28.1.1

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl, & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11-39). Springer Berlin Heidelberg.

Ajzen, I., & Cote, N. G. (2008). Attitudes and the prediction of behavior. In W. D. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 289–311). Psychology Press.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Lawrence Erlbaum Associates Publishers.

Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science*, *291*(5509), 1748-1753. https://doi.org/10.1126/science.1059487

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, 84, 191–215. https://psycnet.apa.org/doi/10.1037/0033-295X.84.2.191

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bandura, A. (2013). The role of self-efficacy in goal-based motivation. In E. A. Locke & G. P. Latham (Eds.), *New developments in goal setting and task performance* (pp. 147-157). Routledge.

Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78*(5), 715–722. https://doi.org/10.1037/0021-9010.78.5.715

Basalla, G. (1988). *The evolution of technology*. Cambridge University Press.

Bawden, D. (2008). Origins and concepts of digital literacy. In C. Lankshear & M. Knobel (Eds.), *Digital literacies: Concepts, policies and practices* (pp. 17-32). Peter Lang Publishing.

Beier, M. E., & Ackerman, P. L. (2003). Determinants of health knowledge: an investigation of age, gender, abilities, personality, and interests. *Journal of Personality and Social Psychology*, *84*(2), 439. https://doi.org/10.1037/0022-3514.84.2.439

Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics, 25*(1), 60–83. https://doi.org/10.2307/1165312

Bernaards, C.A., & Jennrich, R.I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement, 65*, 676-696. https://doi.org/10.1177/0013164404272507

Bhatt, I., & MacKenzie, A. (2019). Just Google it! Digital literacy and the epistemology of ignorance. *Teaching in Higher Education*, *24*(3), 302-317. https://doi.org/10.1080/13562517.2018.1547276

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior, 21*(3), 487–508. https://doi.org/10.1016/j.chb.2004.10.005

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using Internet. *Computers & Education, 53*(4), 1207–1217. https://doi.org/10.1016/j.compedu.2009.06.004

Canty, A. & Ripley, B. (2024). boot: Bootstrap R (S-Plus) functions. R package version 1.3-31. https://cran.r-project.org/web/packages/boot/index.html

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, *127*(4), 355–376. https://doi.org/10.1037/0096-3445.127.4.355

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (pp. 7-75). Erlbaum.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Costa, P. T., Jr., & McRae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI–R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: Examples from conscientiousness. *European Journal of Personality, 12*(2), 117–134. https://doi.org/10.1002/(SICI)1099-0984(199803/04)12:2<117::AID-PER295>3.0.CO;2-C

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). Harper Collins.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* Irvington.

D'Zurilla, T. J., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2011). Predicting social problem solving using personality traits. *Personality and Individual Differences, 50*(2), 142–147. https://doi.org/10.1016/j.paid.2010.09.015

D'Zurilla, T. J., & Nezu, A. M. (1990). Development and preliminary evaluation of the Social Problem-Solving Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*(2), 156–163. https://doi.org/10.1037/1040-3590.2.2.156

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Davison, A. C. & Hinkley, D. V. (1997) *Bootstrap methods and their applications.* Cambridge University Press.

Desjarlais, M., & Willoughby, T. (2007). Supporting learners with low domain knowledge when using the internet. *Journal of Educational Computing Research, 37*(1), 1–17. https://doi.org/10.2190/K788-MK86-2342-3600

Diedenhofen, B. & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE, 10*(4): e0121945. http://dx.doi.org/10.1371/journal.pone.0121945

Duran-Nelson, A., Gladding, S., Beattie, J., & Nixon, L. J. (2013). Should we Google it? Resource use by internal medicine residents for point-of-care clinical decision making. *Academic Medicine, 88*(6), 788-794. https://doi.org/10.1097/ACM.0b013e31828ffdb7

Elio, R., & Scharf, P. B. (1990). Modeling novice-to-expert shifts in problem-solving strategy and knowledge organization. *Cognitive Science*, *14*(4), 579-639. https://doi.org/10.1207/s15516709cog1404_4

Elshout, J. J. (1987). Problem solving and education. In E. De Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), *Learning and Instruction* (pp. 259–273). Pergamon Books/University Press.

Evans, T., Thomas, M. O., & Klymchuk, S. (2021). Non-routine problem solving through the lens of self-efficacy. *Higher Education Research & Development*, *40*(7), 1403-1420. https://doi.org/10.1080/07294360.2020.1818061

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Erlbaum.

Firke, S. (2023). janitor: Simple tools for examining and cleaning dirty data. R package version 2.2.0. https://CRAN.R-project.org/package=janitor

Fishbein, M., & Ajzen, I. (1977). *Belief, attitude, intention, and behavior: An introduction to theory and research.* Addison-Wesley.

Fraillon, J., Schulz, W., & Ainley, J. (2013). *International Computer and Information Literacy Study: Assessment Framework.* International Association for the Evaluation of Educational Achievement.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning, 7*(1), 69–89. https://doi.org/10.1080/13546780042000046

Gamer, M., & Lemon, J. (2019). irr: Various coefficients of interrater reliability and agreement. R package version 0.84.1. https://CRAN.R-project.org/package=irr

Gick, M. L. (1986). Problem-solving strategies. *Educational Psychologist, 21*(1-2), 99–120. https://doi.org/10.1080/00461520.1986.9653026

Gillett, A. J., & Heersmink, R. (2019). How navigation systems transform epistemic virtues: Knowledge, issues and solutions. *Cognitive Systems Research, 56,* 36–49. https://doi.org/10.1016/j.cogsys.2019.03.004

Goff, M. (1994). *Understanding vocational self-concept and related domains: A hierarchical factor-analytic approach.* Unpublished doctoral thesis, University of Minnesota.

Graesser, A. C., Sabatini, J. P., & Li, H. (2022). Educational psychology is evolving to accommodate technology, multiple disciplines, and twenty-first-century skills.

Annual Review of Psychology, *73*(1), 547–574. https://doi.org/10.1146/annurev-psych-020821-113042

Gravill, J. I., Compeau, D. R., & Marcolin, B. L. (2006). Experience effects on the accuracy of self-assessed user competence. *Information & Management*, *43*(3), 378–394. https://doi.org/10.1016/j.im.2005.10.001

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new measurement perspective. *Applied Psychological Measurement, 36*(3), 189–213. https://doi.org/10.1177/0146621612439620

Greiff, S., Wüstenberg, S., Holt, D. V., Goldhammer, F., & Funke, J. (2013). Computer-based assessment of complex problem solving: Concept, implementation, and application. *Educational Technology Research and Development*, *61*(3), 407–421. https://doi.org/10.1007/s11423-013-9301-x

Güss, C. D., Devore Edelstein, H., Badibanga, A., & Bartow, S. (2017). Comparing business experts and novices in complex problem solving. *Journal of Intelligence*, *5*(2), Article 20. https://doi.org/10.3390/jintelligence5020020

Guttman, L. (1955). A generalized simplex for factor analysis. *Psychometrika, 20,* 173–192. https://doi.org/10.1007/BF02289015

Herl, H. E., O'Neil Jr, H. F., Chung, G. K. W. K., Bianchi, C., Wang, S. L., Mayer, R., Lee, C. Y., Choi, A., Suen, T., & Tu, A. (1999). *Final report for validation of problem-solving measures* (Technical Report #501). National Center for Research on Evaluation, Standards, and Student Testing. https://cresst.org/wp-content/uploads/TECH501.pdf

Hoffman, M., & Blake, J. (2003). Computer literacy: Today and tomorrow. *Journal of Computing Sciences in Colleges*, *18*(5), 221-233.

Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology, 6*(1), 35–45. https://doi.org/10.1037/h0040767

Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement: II. A four-year prediction study. *Psychological Reports, 12*(2), 547–594. https://doi.org/10.2466/pr0.1963.12.2.547

Holland, J. (1973). *Making vocational choices: A theory of careers.* Prentice-Hall.

Horn, J. L. (1973). On extension analysis and its relation to correlations between variables and factor scores. *Multivariate Behavioral Research, 8*(4), 477-489. https://doi.org/10.1207/s15327906mbr0804_5

Humphreys, L. G., & Ilgen, D. R. (1969). Note on a criterion for the number of common factors. *Educational and Psychological Measurement, 29*(3), 571–578. https://doi.org/10.1177/001316446902900303

Jackman, S. (2024). pscl: Classes and methods for R developed in the political science computational laboratory. R package version 1.5.9. https://github.com/atahk/pscl/

Kanfer, R. (1987). Task-specific motivation: An integrative approach to issues of measurement, mechanisms, processes, and determinants. *Journal of Social and Clinical Psychology, 5*(2), 237-264. https://doi.org/10.1521/jscp.1987.5.2.237

Kanfer, R., & Ackerman, P. L. (2000). Individual differences in work motivation: Further explorations of a trait framework. *Applied Psychology: An International Review, 49*(3), 470–482. https://doi.org/10.1111/1464-0597.00026

Krippendorff, K. (1970). Bivariate agreement coefficients for reliability of data. *Sociological Methodology, 2*, 139–150. https://doi.org/10.2307/270787

Krippendorff, K. (2004), Reliability in content analysis. *Human Communication Research, 30*(3), 411-433. https://doi.org/10.1111/j.1468-2958.2004.tb00738.x

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208*(4450), 1335–1342. https://doi.org/10.1126/science.208.4450.1335

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance.* Prentice-Hall, Inc.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K.(2023). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.6. https://cran.r-project.org/package=cluster

Meyer, F., & Perrier, V. (2024). esquisse: Explore and visualize your data interactively. R package version 2.0.0. https://CRAN.R-project.org/package=esquisse

Moehring, A., Schroeders, U., Leichtmann, B., & Wilhelm, O. (2016). Ecological momentary assessment of digital literacy: Influence of fluid and crystallized intelligence, domain-specific knowledge, and computer usage. *Intelligence*, *59*, 170–180. https://doi.org/10.1016/j.intell.2016.10.003

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30–38. https://doi.org/10.1037/0022-0167.38.1.30

Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691-692. http://links.jstor.org/sici?sici=0006-3444%28199109%2978%3A3%3C691%3AANOAGD%3E2.0.CO%3B2-V

Neubert, J. C., Mainert, J., Kretzschmar, A., & Greiff, S. (2015). The assessment of 21st century skills in industrial and organizational psychology: Complex and collaborative problem solving. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*(2), 238–268. https://doi.org/10.1017/iop.2015.14

Neuwirth, E. (2022). RColorBrewer: ColorBrewer palettes. R package version 1.1-3. https://CRAN.R-project.org/package=RColorBrewer

Nye, C. D. (2022). Assessing interests in the twenty-first-century workforce: Building on a century of interest measurement. *Annual Review of Organizational Psychology and Organizational Behavior, 9,* 415–440. https://doi.org/10.1146/annurev-orgpsych-012420-083120

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior, 98,* 138–151. https://doi.org/10.1016/j.jvb.2016.11.002

OpenAI. (2024a, May 13). *Hello GPT-4o*. https://openai.com/index/hello-gpt-4o/

OpenAI. (2024b, May 13). *Introducing GPT-4o and more tools to ChatGPT free users*. https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/

Organisation for Economic Co-operation and Development. (2010). PISA 2012 problem solving framework [draft for field trial]. OECD.

Owens, W. A., & Schoenfeldt, L. F. (1979). Toward a classification of persons. *Journal of Applied Psychology, 64*(5), 569–607. https://doi.org/10.1037/0021-9010.64.5.569

Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, *84*(4), 890–904. https://doi.org/10.1037/0022-3514.84.4.890

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, *88*(2), 185–200. https://doi.org/10.1111/jopy.12476

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83–105). Lawrence Erlbaum Associates, Inc.

Pressley, M., Yokoi, L., Van Etten, S., & Freebern, G. (1997). Some of the reasons why preparing for exams is so hard: What can be done to make it easier? *Educational Psychology Review, 9*(1), 1–38. https://doi.org/10.1023/A:1024796622045

Priest, A. G., & Lindsay, R. O. (1992). New light on novice-expert differences in physics problem solving. *British Journal of Psychology, 83*(3), 389–405. https://doi.org/10.1111/j.2044-8295.1992.tb02449.x

Revelle, W. (2023). psych: Procedures for psychological, psychometric, and personality research. R package version 2.3.12. https://CRAN.R-project.org/package=psych

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353-387. https://doi.org/10.1037/a0026838

Rounds, J., Su, R., Lewis, P., & Rivkin, D. (2010). *O*NET® interest profiler short form psychometric characteristics: Summary*. The National Center for O*NET Development. https://www.onetcenter.org/reports/IPSF_Psychometric.html

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94*(1), 18–38. https://doi.org/10.1037/0033-2909.94.1.18

Schauberger, P., & Walker, A. (2023). openxlsx: read, write and edit xlsx files. R package version 4.2.5.2. https://CRAN.R-project.org/package=openxlsx

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupation attainment and job performance. *Journal of Personality and Social Psychology, 86*, 162–173. https://doi.org/10.1037/0022-3514.86.1.162

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schunk, D. H., & DiBenedetto, M. K. (2021). Self-efficacy and human motivation. In *Advances in Motivation Science* (Vol. 8, pp. 153–179). Elsevier. https://doi.org/10.1016/bs.adms.2020.10.001

Schunk, D. H., & Pajares, F. (2009). Self-efficacy theory. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 35–53). Routledge/Taylor & Francis Group.

Sharit, J., Taha, J., Berkowsky, R. W., Profita, H., & Czaja, S. J. (2015). Online information search performance and search strategies in a health problem-solving scenario. *Journal of Cognitive Engineering and Decision Making*, *9*(3), 211–228. https://doi.org/10.1177/1555343415583747

Shih, H. P. (2004). Extended technology acceptance model of Internet utilization behavior. *Information & Management*, *41*(6), 719-729. https://doi.org/10.1016/j.im.2003.08.009

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems, 33*(2), 111–126. https://doi.org/10.1016/S0167-9236(01)00139-7

Sievert, C. (2020). *Interactive web-based data visualization with R, plotly, and shiny*. Chapman and Hall/CRC Florida.

Simon, H. A., & Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist, 26*(2), 145–159. https://doi.org/10.1037/h0030806

Snow, R., E. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman, R. J. Sternberg, & Glaser, R. (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 13–59). W.H. Freeman.

Sonnentag, S., Volmer, J., & Spychala, A. (2008). Job performance. In J. Barling & C. L. Cooper (Eds.), *The Sage Handbook of Organizational Behavior* (pp. 427-447). SAGE Publications.

Spacey, R., Goulding, A., & Murray, I. (2004). Exploring the attitudes of public library staff to the Internet using the TAM. *Journal of Documentation, 60*(5), 550-564. https://doi.org/10.1108/00220410410560618

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15*, 201–293. https://doi.org/10.2307/1412107

Stark, S., Chernyshenko, O. S., Drasgow, F., Nye, C. D., White, L. A., Heffner, T., & Farmer, W. L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology, 26*(3), 153–164. https://doi.org/10.1037/mil0000044

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*(2), 245–251. https://doi.org/10.1037/0033-2909.87.2.245

Stone, D. N. (1994). Overconfidence in initial self-efficacy judgments: Effects on decision processes and performance. *Organizational Behavior and Human Decision Processes, 59*(3), 452–474. https://doi.org/10.1006/obhd.1994.1069

Tesluk, P. E., & Jacobs, R. R. (1998). Toward an integrated model of work experience. *Personnel Psychology, 51*(2), 321–355. https://doi.org/10.1111/j.1744-6570.1998.tb00728.x

Tidler, Z. R., Tatel, C. E., & Ackerman, P. L. (2022, September). The Internet and manuals: A use-oriented study of the acquisition of novel procedural knowledge. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 66, No. 1, pp. 1834-1838). SAGE Publications.

Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education, 215*, 1–15. https://doi.org/10.1016/j.compedu.2024.105031

Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., & Peters, O. (2011). Rethinking Internet skills: The contribution of gender, age, education, Internet experience, and hours online to medium- and content-related Internet skills. *Poetics*, *39*(2), 125–144. https://doi.org/10.1016/j.poetic.2011.02.001

Van Iddekinge, C. H., Putka, D. J., & Campbell, J. P. (2011). Reconsidering vocational interests for personnel selection: The validity of an interest-based selection test in relation to job knowledge, job performance, and continuance intentions. *Journal of Applied Psychology, 96*(1), 13–33. https://doi.org/10.1037/a0021193

van Laar, E., Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., & De Haan, J. (2019). Determinants of 21st-century digital skills: A large-scale survey among working professionals. *Computers in Human Behavior*, *100*, 93–104. https://doi.org/10.1016/j.chb.2019.06.017

von Stumm, S., Chamorro-Premuzic, T., & Ackerman, P. L. (2011). Re-visiting intelligence-personality associations: Vindicating intellectual investment. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 217–241). Wiley Blackwell.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds.), *Learning Issues for Intelligent Tutoring Systems* (pp. 19–41). Springer. https://doi.org/10.1007/978-1-4684-6350-7_2

Venables, W. N. & Ripley, B. D. (2002). *Modern applied statistics with S (Fourth Edition).* Springer.

Vonkova, H., Papajoanu, O., Stipek, J., & Kralova, K. (2021). Identifying the accuracy of and exaggeration in self-reports of ICT knowledge among different groups of students: The use of the overclaiming technique. *Computers & Education*, *164*, 104112. https://doi.org/10.1016/j.compedu.2020.104112

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software, 21*(12), 1-20. http://www.jstatsoft.org/v21/i12/

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2022). readxl: Read excel files. R package version 1.4.1. https://CRAN.R-project.org/package=readxl

Willoughby, T., Anderson, S. A., Wood, E., Mueller, J., & Ross, C. (2009). Fast searching for information on the Internet to use in a learning context: The impact of domain knowledge. *Computers & Education, 52*(3), 640–648. https://doi.org/10.1016/j.compedu.2008.11.009

Wittmann, W.W., & Suß, H.-M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik Symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and Individual Differences: Process, Trait, and Content Determinants* (pp. 77-104). Washington, DC: American Psychological Association.

Witten, D. M., & Tibshirani, R. (2010). A framework for feature selection in clustering. *Journal of the American Statistical Association*, *105*(490), 713-726. https://doi.org/10.1198/jasa.2010.tm09415

Wüstenberg, S., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The role of strategy knowledge for the application of strategies in complex problem-solving tasks. *Technology, Knowledge and Learning*, *19*(1–2), 127–146. https://doi.org/10.1007/s10758-014-9222-8

Zeileis, A., & Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News 2*(3), 7-10. https://CRAN.R-project.org/doc/Rnews/

Zimmerman, B. J., & Kitsantas, A. (2005). The hidden dimension of personal competence: Self-regulated learning and practice. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 509–526). Guilford Publications.