Six Popular Financial Literacy Assessment Items and their Resulting Uses and Inferences: An Item Bias Study Utilizing DIF Analyses

by

Diana G. Simpson

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 6, 2022

Keywords: Financial education, financial literacy, "Big Five", differential item functioning, item bias, item response theory

Approved by

Jill Salisbury-Glennon, Associate Professor, Department of Educational Foundations, Leadership, and Technology
Joni Lakin, Associate Professor at the University of Alabama
Paris Strom, Professor, Department of Educational Foundations, Leadership, and Technology
Elisha Wohleb, Clinical Professor, Department of Educational Foundations, Leadership, and Technology

Abstract

Financial education has become the popular antidote to the widespread financial literacy "problem" in the United States. Certain demographic groups (i.e., young adults, the elderly, women, and people of color) have been consistently identified as "financially illiterate." However, substantial increases in financial education – many targeting these specific groups - have achieved little, as research consistently finds that improvement to financial literacy in the U.S. remains elusive. The current study examined the validity and fairness of financial literacy assessments: specifically, whether test score interpretations and uses may be unfair due to measurement bias. Validity and reliability measures were analyzed, and Differential Item Functioning (DIF) analysis was conducted to identify potential measurement bias for the six financial literacy questions included in the 2018 National Financial Capability Study. Item responses were examined by participant age, race, and gender. The study found a lack of validity evidence regarding internal item consistency and item correlation, suggesting that the items measure different or multiple constructs. Furthermore, Differential Item Functioning (DIF) was indicated for three of the six items: Items two, four, and five. Large DIF was indicated for item two for the 65 and older focal group (β = -.13). Moderate DIF was found for item two in favor of the 18-24 age group (β = .06). Item four indicated large DIF for the 18-24 age group (β = -.10) and in favor of the 65+ focal group (β = .11). Items two and five indicated DIF for the Black/African American focal group (β = -.08 and -.06, respectively). Large DIF was found for item five in favor of the 18-24 age group (β = 0.09) and moderate DIF was found in favor of females in the gender group (β = 0.07). Test developers and users may find these items unsuitable for their target population(s).

Dedication

This dissertation is dedicated to my late father, Dr. Philip Grossman, who *also* spent ten years working on his PhD. I always wondered why it took him so long; now I know. It is also dedicated to my late mother, Shirley B. Grossman, who, along with my father, instilled in me a love for learning.

Acknowledgments

So many individuals have played a part in this achievement, several of whom I could not have done this without. However, to God be the Glory, for I could not have done it without His guidance. The endeavor has taught me a great deal about patience, and without God's peace I don't believe I could have persevered.

The Lord provided me with a wonderful committee who exemplified tremendous patience and grace. Dr. Salisbury-Glennon, my committee chair, is genuinely a saint. Her positive attitude, professionalism, and unwavering belief in my ability to finish this process were instrumental in maintaining my momentum. Dr. Strom has provided tremendous support to me since the very beginning as well; I thank you both for sticking with me! Dr. Lakin crafted excellent comps questions and offered thoughtful feedback on research methods, analysis, and the interpretation of data results which was critical to the success of the study. Dr. Wohleb, a relative newcomer to the committee, provided insight as only one from within the discipline is able. Thank you both for giving so freely of your time. Dr. Duke, my outside reader, added so much more to this experience than simply reading my completed draft. As my supervisor and mentor during the (almost) three years that I spent as her graduate research assistant, Dr. Duke has given me the freedom to grow as a researcher and provided many forms of support to guide me graciously in the right direction. Thank you so much, Dr. Duke!

My family, of course, has lived the day-to-day ups and downs of this experience with me. My husband, Scott, has been my most staunch supporter; his grace and mercy are a gift and I thank God for him. My two children, Scott and Claire, began this process with me when they were in elementary school and are now entering their junior years in college! They are my reason for doing this and everything else that I do. I hope I have made them as proud as they have made me.

Disclaimer

The opinions set forth in this dissertation are those of the author and not ETS.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

This chapter provides an overview of the current study. It begins with background information about the study and explains its purpose and significance. The research questions and hypotheses are then presented, followed by definitions of the relevant terms used. Finally, the chapter concludes by providing the limitations and assumptions of the study.

Financial Literacy in America

The President's Advisory Council on Financial Literacy (2008) defined financial education as "the process by which people improve their understanding of financial products, services, and concepts, so they are empowered to make informed choices, avoid pitfalls, know where to go for help and take other actions to improve their present and long-term financial well-being (p. 37)."

Woodyard and Robb (2012) state that "… Americans, on the whole, [are] inadequately prepared to handle their own finances and financial futures (p. 1)." Other studies reveal that Americans have been found to have "managed their finances poorly" due to a lack of financial literacy (Chen & Volpe, 2002, p. 289) and that low financial literacy is related to low self-efficacy (Danes & Haberman, 2007) and lack of confidence (Bucher-Koenen et al., 2017). The subsequent reaction to this problem has been an attempt to increase the consumer's financial knowledge under the assumption that better-informed individuals make better financial choices (Alsemgeest, 2015).

Many financial literacy assessments utilize five core items: the "Big Five" developed by Annamarie Lusardi and Olivia Mitchell (Anderson et al., 2017), which are:

1. Suppose you had $100 in a savings account, and the interest rate was 2% per year. After five years, how much do you think you would have in the account if you left the money to grow?

   A) More than $102

   B) Exactly $102

   C) Less than $102

   D) Don't know

   E) Prefer not to say

2. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

   A) More than today

   B) Exactly the same

   C) Less than today

   D) Don't know

   E) Prefer not to say

3. If interest rates rise, what will typically happen to bond prices?

   A) They will rise

   B) They will fall

   C) They will stay the same

   D) There is no relationship between bond prices and the interest rate

   E) Don't know

   F) Prefer not to say

4. A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.

   A) True

   B) False

   C) Don't know

   D) Prefer not to say

5. Buying a single company's stock usually provides a safer return than a stock mutual fund.

   A) True

   B) False

   C) Don't know

   D) Prefer not to say

The "Big Five" set the standard in financial literacy assessment (Gupta et al., 2018) and have served as foundational questions in many survey instruments intended to measure the construct in the U.S. and other countries (Al-Bahrani et al., 2018; Allgood & Walstad, 2013; Knoll & Houts, 2012; Skimmyhorn, 2016). However, some have doubted whether such a simple test of five questions could accurately measure the construct they purport to measure (Huston, 2010). For example, although all five questions measure knowledge only, the literature identifies both attitudes and behaviors as essential measures of financial literacy. For example, many studies point to a strong relationship between knowledge and behavior in personal finance, as more knowledgeable individuals generally display more responsible or effective financial behaviors (Hilgert et al., 2003; Robb, 2011; Robb & Woodyard, 2011; Xiao et al., 2012). These behaviors include managing one's expenses by using a budget, minimizing debt, and increasing

savings. A recent review of financial literacy education programs revealed that students who receive financial education in school score higher on financial knowledge tests and show more positive money-related attitudes than students who do not receive financial education. Furthermore, these students appear more likely to save (Lee et al., 2019; Yakoboski et al., 2019).

Nonetheless, there is an ongoing debate about how financial education impacts financial behavior (Fernandes et al., 2014; Hastings et al., 2013; Kaiser & Menkhoff, 2020). Studies have found that most individuals cite personal experience as an essential source of their financial learning (CFPB, 2015; Hilgert et al., 2003). In addition, studies indicate that a student's financial understanding is higher if they have had previous personal financial experiences such as opening a bank account or earning money through work or an allowance (Lyons et al., 2006; Sherraden et al., 2011). Lyons et al. (2006) evaluated a financial education program and found that financial education may be correlated with improved financial behaviors, but their findings also indicated that experience might be a more potent factor in behavior change. The authors cautioned that: "researchers may want to re-examine the indicators currently being used to show program impact and whether financial knowledge is the appropriate catalyst to foster behavior change" (p. 27).

Conversely, several research studies found no relationship between financial education and household financial behavior (Allgood & Walstad, 2013; Bernheim et al., 2001; Fernandes et al., 2014). For example, Fernandes et al. (2014) found that "…financial education interventions studied explained only about 0.1% of the variance in the financial behaviors studied, with even weaker average effects of interventions directed at low-income rather than general population samples (p. 1872)." Overall, the existing evidence regarding the relationship between financial knowledge and behavior is scarce and inconclusive (Amagir et al., 2018).

Further confusing the construct of "financial literacy" is the fact that current assessments focus on the underlying message within financial education that financial behaviors are entirely endogenous to the individual. The extant research includes little emphasis on the attitudinal and belief domains of the social contexts where individuals are socialized about finances (Gudmunson & Danes, 2011). Danes et al. (2016) suggest that in the U.S., students are primarily taught about "the benefits of having financial knowledge and healthy financial behaviors" and "the social desirability and impact of being a financially responsible person" (p. 429).

Davies (2015) astutely identified a single question on an assessment posed to students who participated in the US Jump$tart survey (Klein & Mandell, 2007, p.110), which highlights this issue:

Which of the following do you feel is the greatest cause of serious financial difficulty, where families can't pay their bills? The answer choices are as follows:

a) Bad luck, such as unexpected illness or job loss.

b) Not enough savings.

c) Buying too much on credit.

d) Not following a financial plan.

e) Not being able to earn enough money.

Davies (2015) notes:

> Mandell and Klein interpret the responses as indicating that respondents believe that families experience financial difficulties as a consequence of their own financial mismanagement. This is an interesting conclusion because the question only allows respondents to attribute financial difficulty to individual responsibility or serendipity…it is surprising that definitions of financial literacy

do not refer to citizens' understanding of the roles of banks or governments in the conduct of financial systems (p. 304-305).

There is little research on sociocultural factors that may influence financial behaviors. Indeed, the assessment tools upon which most financial literacy studies are based seek only to measure individual knowledge (particularly numeracy) and behavior, based on the assumption that only internal factors are relevant. As a result, few researchers have explored the possibility of factors external to the individual affecting financial outcomes.

In addition, there have been numerous inferences made based on the resulting scores of the "Big Five" financial literacy questions. For example, several studies have concluded that individuals who score higher on financial literacy tests also indicate better financial management practices (Lusardi & Mitchell, 2014; Lyons et al., 2006). In addition, Hilgert et al. (2003) identified a strong correlation between financial literacy and basic financial management skills. In terms of intelligence, researchers have concluded that people with higher education levels have higher scores on financial literacy tests and that intelligence can predict financial literacy. However, while Lachance (2014) showed that level of education is indeed a predictor of financial literacy, the study also determined that the participant's zip code (as a proxy for local quality of education) was an even stronger predictor.

Other research has indicated that African American and Hispanic individuals have lower financial literacy scores than those who identify as white (Lee et al., 2019; Shapiro et al., 2013; Yakoboski, 2019). For example, a recent research study utilizing the 2018 National Financial Capability Study (NFCS) data found that African American homeowners are over three times more likely to have been late with a mortgage payment or taken a loan or hardship withdrawal from their retirement account in the past year (Lin et al., 2019). Similarly, other research

indicates that women often have lower financial literacy scores than men. For example, the 2014 GFLEC and the TIAA Institute, reporting on findings from the 2012 NFCS, found that working women have low levels of financial literacy (Scheresberg et al., 2014).

Age has also been shown to be a factor in financial literacy. For example, Yakoboski et al. (2019) found that financial knowledge was correlated with age group. The youngest group (ages 18-29) was the least financially literate. Those ages 30 to 44 showed improved financial knowledge, but it was not statistically significant. In contrast, participants ages 45 to 59 did show a statistically significant increase in their *P-Fin Index* score, and adults over age 60 showed even higher, statistically significant levels of financial literacy.

Earlier findings, however, indicate that financial literacy scores tend to increase with age until age 60 (Lusardi & Mitchell, 2011) at which time they begin decreasing. Finke et al., 2017 estimate this decline to be about 1% per year. In addition, Gamble et al. (2015) show that cognitive decline related to aging impacts numeracy ability in elderly individuals, yet confidence in their own financial ability does not wane.

These are examples of studies that have been conducted to establish the extent to which Americans are "financially literate." Certain groups have been consistently found to be less financially literate than others. The recommended solution is more financial education, based on the assumption that poor financial literacy results from the individual's lack of knowledge. There is no mention of external constraints (over which the individual may have no control) that may affect these groups' opportuni*ties* to earn more, work more hours, further their education, or save money. This study examines how group membership (specifically, age, race, and gender) may confound the findings of the numerous studies that have identified specific age, gender, or racial groups as being less financially literate.

Validity in Financial Literacy Assessment

Cook & Beckman (2006) state, "because validity is a property of inferences, not instruments, validity must be established for each intended interpretation (p.166)." In addition, the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014), hereafter referred to as "*The Standards*," explicitly states that test instruments must measure what they intend to measure to ensure valid inferences are drawn from test scores.

Modern test theory conceptualizes validity as a process through which test developers establish an evidence-based argument for score interpretations (Kane, 1992; Messick, 1995). In this argument-based approach, multiple forms of evidence are used to substantiate claims within the validity argument, thereby determining whether the scores reflect an accurate measure of the underlying construct. Kane identified two steps to the process: the development of the *interpretive argument*, which is then rated based on its ability to present a *validity argument*. According to Kane, the interpretive argument identifies the "most relevant" assumptions and inferences used to interpret the results of an assessment and "provides a clear basis for choosing the kinds of evidence to be included in the validity argument" (M. T. Kane, 1992, pp. 9–10). In addition, the validity argument involves analyzing and assessing the relative strength of the identified assumptions and inferences, thereby exposing weaknesses.

Many inferences have been made based on the National Financial Capabilities Study conducted in 2009, 2012, 2015, and 2018 as well as many other studies. Inferences made based on such studies influence public policy and public opinion. Additionally, the current lack of consensus concerning the definition of financial literacy has been made evident in the literature, causing the validity of the associated assessment instruments to be questioned. For example,

Fernandes et al. (2014) conducted a meta-analysis of financial literacy studies and found the extant literature lacking in measurement constructs that have been "subjected to standard psychometric tests to establish one-dimensionality, reliability, and validity" and have failed to "establish that the measures of financial literacy represented a construct that was distinct from other correlated traits (p. 1868)." Others have arrived at similar conclusions (Goyal & Kumar, 2021; Huston, 2010)

The current study provides additional research into the validity of financial literacy assessments to improve future program design and, ultimately, provide more accurate feedback to inform educational and social policy.

Theoretical Framework

This study is situated within the theoretical framework of Constructivism and the philosophy of Pragmatism, each of which provides a basis for examining social influences on individual behavior. Based on a theoretical conceptualization that encompasses both individual and sociocultural factors as drivers of financial decision-making and behavior, this study seeks to shed insight on whether America's most widely accepted financial literacy measures may be biased due to a focus on individual knowledge and the omission of consideration for social and cultural influences that inherently affect "financial literacy."

Constructivist theory in education focuses on the learner. Constructivists maintain that our understanding – what we "learn"- is based on our context: culture, prior experience, worldview, etc. The central theme is that individual knowledge is constructed through the learner's experiences instead of memorizing "facts" or being told what is "true." Students construct their understanding of the world by being challenged to support their assertions in the face of new knowledge and experience (Leidner & Jarvenpaa, 2016; Vygotsky, 1978).

Jean Piaget (1896–1980) and Lev Vygotsky (1896-1934) were contemporaries whose work was fundamental to the development of Constructivism. Piaget established the theory of Cognitive Constructivism, which proposed that knowledge is constructed cognitively, within the individual, by creating connections between prior knowledge and new experiences. Vygotsky's Social Constructivist Theory, in contrast, suggested that social learning precedes development in children and that social interaction leads to increased knowledge; that is, that external factors - societal and cultural - dictate what is important to learn, rather than internal (cognitive) functions (Richardson, 1997). Although their suppositions were divergent, Piaget and Vygotsky were both constructivists; they both believed that learning is grounded in the interaction between mind and environment. Learning does not solely occur passively by memorizing "facts" or revealing "truth."

John Dewey contributed much to the theory of Constructivism. Furthermore, he proposed that if it is from this experience that knowledge is conferred, then "knowledge" and "reality" are nothing but constructions in one's mind. In his early work (Dewey, 1896), Dewey introduced the concepts of transaction and experience as components of knowledge construction. He used the concepts of reflective thinking and reciprocal processes to support his belief that both the subjective (individual) and intersubjective (sociocultural) dimensions of the construction of knowledge should be considered within a constructivist framework.

More recently, Self-regulated Learning (SRL) theory has built on constructivist thought to expand on this concept. SRL involves the individual taking an active role in the learning process by exercising control over cognitive, metacognitive, motivational, and behavioral functions to achieve established goals (Panadero & Alonso-Tapia, 2014;

Zimmerman, 1990). More recent contributions to the SRL literature include sociocultural factors. For example, co- regulated learning (CoRL) derives from a Vygotskian perspective to explain transitional processes toward self-regulation. It is grounded in Vygotskian views of higher psychological processes being socially embedded or contextualized and then internalized through social interaction (Mccaslin & Hickey, 2001).

Constructivism provides a theoretical basis upon which an argument may be made that "financial literacy" is vital to measure, considering both individual knowledge and behavior and the contextual and sociocultural factors that affect knowledge and behavior. This framework provides a foundation upon which this study was developed.

Statement of the Problem

Although the lack of financial literacy in the United States has become a widely recognized topic of concern, the definition of the construct within the extant literature is inconsistent, which has led to a myriad of operationalizations and, therefore, dubious measurement instruments. The validity of measurement items in financial literacy requires further analysis to identify best practices in financial literacy assessment. This study hypothesized that one or more of the six most ubiquitous questions used to measure the construct might lack validity evidence due to construct-irrelevant sources of variance. A Product-moment Correlation Coefficient (Pearson's r) and an Item Reliability Coefficient (Cronbach's α) were calculated to measure the item characteristics for the six items. The study utilized Differential Item Functioning (DIF) analysis to identify potential item bias due to construct-irrelevant sources of variance. DIF occurs when individuals from different examinee groups have different probabilities of answering an item correctly after controlling for ability. Holland & Thayer

(1988) introduced the term "differential item functioning" (DIF) to convey the concept of test questions that have a varying degree of difficulty for groups of otherwise comparable proficiency. DIF can result from differences in socioeconomic status, lifestyle, culture, educational experience, and other confounding factors with group membership. It is, therefore, a useful statistical tool for detecting potential construct-irrelevant factors contributing to item performance**.**

Purpose of the Study

The purpose of this quantitative research study was to examine the extent to which financial literacy assessment in America is affected by bias due to group membership. Item bias was analyzed as a potential indicator of construct irrelevant variance (CIV) in a national financial literacy study. Specifically, through Differential Item Functioning, the study analyzed the validity of six commonly used survey questions in the 2018 National Financial Capability Study (NFCS) and, thereby, the cogency of the resulting outcomes and interpretations for groups identified by age, race, and gender.

Research Questions

The current study sought to identify potential threats to the validity of current measures of financial literacy and the resulting consequences of their interpretation and use. The study used differential item functioning (DIF) to measure item bias in financial literacy. Specifically, the study analyzed the six items used to measure financial literacy on the 2018 NFCS survey and addressed the following research questions and hypotheses:

RQ. 1. To what extent do the financial literacy items on the 2018 NFCS exhibit reliability?

RQ. 2. To what extent do the financial literacy items on the 2018 NFCS exhibit validity (measure a common construct)?

RQ. 3. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by age?

RQ. 4. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by gender?

RQ. 5. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by race?

Hypotheses

The researcher posed the following hypotheses:

H1. The items on the 2018 NFCS do not exhibit appropriate levels of reliability.

H2. The items on the 2018 NFCS do not measure a common construct.

H3. The financial literacy items on the 2018 NFCS function differently across groups identified by age.

H4. The financial literacy items on the 2018 NFCS function differently across groups identified by gender, and

H5. The financial literacy items on the 2018 NFCS function differently across groups identified by race.

Significance of the Study

Results derived from this study contribute to the literature by providing additional insight into the validity of current financial literacy assessment practices. The stakeholders who will benefit include financial education program leaders, policymakers, local resource councils, and families throughout the U.S. Proper assessment leads to better decision-making; inadequate

assessment measures and the resulting outcomes are costly and misleading. This study provides stakeholders an opportunity to make better-informed decisions moving forward.

Assumptions

This study assumed that the data retrieved from the 2018 NFCS database was accurate: that the respondents answered the questions truthfully and that the data was recorded and processed correctly. Also, the current study used standard guidelines for interpreting statistical findings (i.e., Pearson's *r*, Cronbach's α). However, financial education and literacy assessment is a relatively new discipline within the social sciences, and therefore standard guidelines may not be appropriate for research findings in this field.

Limitations

Using Differential Item Functioning analyses, the current study sought to identify systematic measurement bias in the six financial literacy items on the 2018 NFCS. There are limitations within the study that should be noted. The National Financial Capability Study was conducted in 2018; therefore, the results reflect that period's data. Also, the statistical model used in this study, differential item functioning (DIF), can provide evidence of potentially biased tests or items but cannot establish why differences occur.

The small number of overall test items studied (six) may have limited the usefulness of some DIF statistical methods. Research has indicated that assessments with only a few items tend to produce a more significant percentage of items flagged for DIF (Buzick & Stone, 2011). Similarly, larger sample sizes increase statistical power (Chalmers et al., 2016), which can be an advantage. However, in a sizeable survey like the one used in the current research, analysis power is substantially increased, leading to "too much statistical power" (Zumbo & Wild, 1999, p. 27). Too much power can create a greater tendency to reject null hypotheses with negligible

differences. The current study focused on measures of effect size and statistical significance to offset these limitations.

Definition of Terms

The following are significant terms used throughout this research study:

***Construct*:** "A psychological process or characteristic believed to account for individual or group differences in behavior (Strauss & Smith, 2009, p. 1)."

***Differential Item functioning (DIF)*:** DIF occurs when individuals from separate examinee groups have different probabilities of answering an item correctly after controlling for ability. (Holland & Thayer, 1988)

***Financial Education*:** A process through which an individual "is to gain an understanding of financial terms, products and systems (financial literacy), as well as to facilitate financial empowerment through newly acquired knowledge (Alsemgeest, 2015, p. 155)."

***Financial Socialization*: "**Financial knowledge, attitudes, and behaviors acquired from formal and informal relationships with family members and other significant influencers during childhood and young adulthood" (Danes, 1994, p. 128)

Summary

Research findings implicating internal factors inherent to the individual's lack of financial literacy have underscored a need to increase financial education efforts. However, even as financial education efforts have burgeoned over the past quarter-century, little change has been found to have occurred; research studies continue to show a "lack of financial literacy" among Americans. A broader view of the potential factors that affect financial literacy may improve financial education efforts.

CHAPTER II: REVIEW OF THE LITERATURE

This chapter begins with a review of the literature pertinent to the history of financial education and literacy in the United States. The chapter continues with an overview of the status of financial literacy in the United States today, including current efforts to define the term and operationalize the construct. Next, a summary of the current assessment tools most commonly used to measure financial literacy today and the outcomes of studies that have used these measurements is given. A discussion concerning the operationalization of the construct of financial literacy and relevant studies is then provided. Next, the theoretical framework underlying the current study and an explanation of the methods used are given. Finally, the chapter concludes with a summary.

A Brief History of Financial Literacy and Education in the U.S.

> All the perplexities, confusions, and distresses in America arise, not from defects in their constitution or confederation, not from a want of honor or virtue, so much as from downright ignorance of the nature of coin, credit, and circulation.
>
> John Adams, 1787

In America, prior to the 1920s, "saving for a rainy day" was a common phrase used to refer to having cash available for emergencies and opportunities. Credit for most middle- and low-income families was difficult to obtain, and buying a home often meant saving up an amount at least equal to 50% value of the home to qualify for a mortgage. At that time, mortgage loans were generally no longer than ten years and required a "balloon" payoff at the end of the term (Snowden, 1987). Saving money to make purchases was the only way for most Americans to obtain goods and services, and as a result, household financial management was generally straightforward. If the savings were not kept at home, they were likely on deposit at the local

bank. Few working-class families held stocks or bonds. After the stock market crash of 1929 and the subsequent economic downturn of the Great Depression across America until 1939, millions of bank patrons lost their savings, unemployment rose to almost 25% of the workforce, and families were displaced or separated.

In 1933, the U.S. Congress passed the Glass-Steagall Act to effect greater control over financial markets. The Act prohibited banks from engaging in investment transactions; by doing so, it allowed for increased regulation of securities markets.

Economic prosperity returned to the United States during and after World War II. America enjoyed unprecedented economic growth, and a sizeable middle class developed. In the 1960s, the United States experienced its most prolonged uninterrupted period of economic expansion in its history. American households' consumption expenditures more than tripled in the two decades between 1950 and 1969. The selection of goods and services available to the middle- and lower-classes had improved substantially due to increased wages and new, enhanced forms of credit. "Buying on time" became a popular way to afford expensive items such as automobiles and fine jewelry. Credit cards became popular.

The steady increase in consumer spending strongly impacted the growth of the American economy, creating a cycle of economic stimulation that continued into the 1970s and 1980s. The combination of increases in consumer demand, available credit, and product marketing fueled an economy based on "consumerism."

An increase in economic disruptions plagued the U.S. during the last quarter of the 20th Century. In the literature, the economic turbulence was attributed to the fact that financial responsibility was systematically transferred to individuals during that time (Hastings et al., 2013). Individuals had greater control and responsibility for their retirement investments due to a

massive conversion from defined benefit pension plans to 401k-type defined contribution plans. Also, due to advances in technology, consumers had far more options to choose from in savings, investing, and lending markets. The World Wide Web made financial information readily available, but consumers often found the abundance of data overwhelming and confusing. In 1999, The Gramm-Leach-Bliley Act repealed the Glass-Steagall Act, allowing banks and investment companies to merge, further mystifying financial consumers. As the 20th Century ended, the inadequacy of Americans' personal and collective financial acumen was as evident as it was widespread and the call for financial education was ubiquitous.

One of the earliest and most well-established organizations in financial education is the National Endowment for Financial Education (NEFE®), a self-endowed, non-profit organization created in 1984. Another nationally acclaimed advocate for financial literacy is the non-profit organization, the National Jump$tart Coalition for Personal Financial Literacy, founded in 1995. Besides serving as a clearinghouse for financial literacy information and curriculum, Jump$tart took a lead role in developing the first national standards for financial education in K-12 classrooms and updated the standards each year. Additionally, Jump$tart has been a leader in conducting research studies to gain insight into American high school students' financial literacy.

The U.S. Federal Government became directly involved in financial education after the 2000-2002 recession when Congress passed the Financial Literacy and Education Improvement Act of 2003. The Act established the Financial Literacy and Education Commission (FLEC), which was appointed to build public awareness of available financial education resources, develop directed materials and distribution strategies, integrate materials within public schools, and establish public-private partnerships to expand resources and outreach.

In the wake of the 2008-2009 financial crisis, the FINRA Foundation and the U.S. Department of the Treasury launched the inaugural National Financial Capability Study (NFCS) to assess Americans' financial capabilities, which they defined in terms of "Making Ends Meet," "Planning Ahead," "Managing Financial Products," "Financial Knowledge," and "Decision-Making" (Woodyard & Robb, 2012). In addition, the Consumer Financial Protection Bureau (CFPB) was created in 2010 by the Dodd-Wall Street Reform and Protection Consumer Act to provide further support to financial education, among other roles.

Most adults in the U.S. are frequent consumers of financial products and services. However, research indicates that efforts toward personal financial education have been inconsistent and ineffective: studies show that financial literacy in the United States has been stagnant for the past 25 years.

Definitions and Constructs of Financial Literacy

The term "Financial Literacy" was introduced by the Jump$tart Coalition for Personal Financial Literacy in its 1997 study of high school students. The Coalition defined the term as "the ability to use knowledge and skills to effectively manage one's financial resources for lifetime financial security" (Jump$tart Coalition, 2017). As operationalized in the early research, the construct was not concisely defined and was generally measured by a person's financial knowledge (Aren & Aydemir, 2014; Hilgert et al., 2003; Huston, 2010).

For example, Huston (2010) conducted a review of financial literacy research that had occurred during the first decade of the 21st Century and found that "lack of clarity is a barrier to a common or general understanding of the financial literacy construct" (p. 305). The study, which examined research from fifty-two unique data sets**,** found that 72% of the studies did not define the term "financial literacy" and that 88% of the studies provided no guidelines for interpreting

the assessment measures. Furthermore, the study determined that nearly one-half of the research studies used the terms "financial literacy" and "financial knowledge" interchangeably. Fernandes et al. (2014) conducted both a meta-analysis and original research regarding financial literacy measurement. The authors found the extant literature lacking measurement constructs that were "subjected to standard psychometric tests to establish one-dimensionality, reliability, and validity" (pg. 1868).

Most recently, in a meta-analysis of over 500 articles published between 2000 and 2019, Goyal & Kumar (2021) found that "very few studies specifically focus on the conceptualization of financial literacy. The construct of financial knowledge seems to be a shortsighted way of defining financial literacy (p. 96)."

The research literature has evolved over the past two decades, and the construct of "financial literacy" has developed through many iterations. These include such concepts as financial wellness (Shim et al., 2009; Yakoboski, 2019), financial well-being (CFPB, 2015; Gudmunson & Danes, 2011; Lee et al., 2019), financial capability (Danes & Yang, 2014; Lusardi, 2015; Lusardi & Mitchell, 2014) and financial fragility (Gupta et al., 2018; Scheresberg, 2013). These conceptualizations present a subset of the domain with a slight variation in the term's operationalization.

For example, the construct of "financial well-being" includes both objective indicators such as income and savings levels, assets, and debt levels, which are both measurable and accessible through multiple sources, and subjective indicators like self-reported financial data and personal values, beliefs, and perspectives (Gudmunson & Danes, 2011).

The construct of "financial wellness," as operationalized by the P-Fin Index (Yakoboski, 2019), is an individual who "has control over day-to-day, month-to-month finances; has the

capacity to absorb a financial shock; is on track to meet his or her financial goals and has the financial freedom to make choices that allow enjoying life (p. 12)."

The construct of "financial capability" also includes the *opportunity* to apply one's knowledge and abilities as an essential part of financial literacy (Johnson & Sherraden, 2007). The U.S. President's Advisory Council on Financial Capability (2014) identified financial capability as having three components: financial literacy to help provide knowledge, access to appropriate products to provide choice, and industry regulation to protect financial consumers (Delgadillo, 2014). The plethora of constructs in the discipline of financial literacy indicates that knowledge alone may not be an adequate measure for "financial literacy." Furthermore, a cohesive, agreed-upon definition of the construct continues to elude researchers in the discipline.

Setting the Standard: The "Big Three" and "Big Five"

Although the literature continued to add additional characteristics to the construct of financial literacy, research studies have continued to predominantly use the "Big Three" and the "Big Five" financial literacy survey questions (Lusardi & Mitchell, 2007, 2014). These survey questions have been utilized in numerous studies, including the 2007–2008 National Longitudinal Survey of Youth; de Bassa Scheresberg (2013); the Federal Reserve Bank's Survey of Consumer Finances (2016, 2019); Gerardi et al. (2010); Hilgert (2003); Lusardi et al. (2020); the National Financial Capability Study (2009; 2012; 2015; 2018); and the RAND American Life Panel (2011). The "Big Five" are inclusive of the "Big Three" and include the following questions:

1. Suppose you had $100 in a savings account and the interest rate was 2% per year. After five years, how much do you think you would have in the account if you left the money to grow?

A) More than $102

B) Exactly $102

C) Less than $102

D) Don't know

E) Prefer not to say

2. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

A) More than today

B) Exactly the same

C) Less than today

D) Don't know

E) Prefer not to say

3. If interest rates rise, what will typically happen to bond prices?

A) They will rise

B) They will fall

C) They will stay the same

D) There is no relationship between bond prices and the interest rate

E) Don't know

F) Prefer not to say

4. A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.

A) True

B) False

C) Don't know

D) Prefer not to say

5. Buying a single company's stock usually provides a safer return than a stock mutual fund.

A) True

B) False

C) Don't know

D) Prefer not to say

The "Big Five" questions have set the standard in financial literacy assessment. They have served as foundational questions in many survey instruments intended to measure the construct in the U.S. and other countries since 2009. Yet, some have doubted whether such a simple test of three or five questions could accurately measure the complex construct they purport to measure. According to Hastings et al. (2013), regarding "the Big Three": "There is remarkably little evidence on whether this set of survey questions is the best approach, or even a superior approach, to measuring financial literacy" (p.353). Stolper and Walter (2017) also argue that although the "Big Three" has become the standard in financial literacy assessment models, there is a lack of evidence concerning the efficacy of the approach.

In addition to the "Big Five" questions, the 2018 NFCS included a sixth question that was developed by Lusardi and Tufano (2015) as part of a survey designed to "… assess key debt literacy concepts, such as the power of interest compounding" (p. 4). The question is:

Suppose you owe $1,000 on a loan and the interest rate you are charged is 20% per year compounded annually. If you didn't pay anything off, at this interest rate, how many years would it take for the amount you owe to double?

A) 2 years

B) Less than 5 years

C) 5 to 10 years

D) More than 10 years

E) Don't know

F) Prefer not to say

This additional question also measures numeracy, as the answer can only be achieved using the compound interest equation or the heuristic, "Rule of 72." Although financial literacy does require numeracy, the overall construct is significantly more complicated, as evidenced by its range of characterizations.

The two authors who developed these questions are considered experts in the field and are widely published. Goyal & Kumar (2021) compiled a list of "prolific" authors in the financial planning discipline; Annamaria Lusardi was first on the list with 26 publications. Olivia S. Mitchell was listed third with 12 peer-reviewed articles. According to Goyal and Kumar (2021), "Annamaria Lusardi and Olivia S. Mitchell also received the highest number of citations - 2,652 and 1,959, respectively" (p. 85). The Financial Industry Regulatory Authority (FINRA) offers these six questions on its website for the public to measure personal financial literacy (https://www.usfinancialcapability.org/quiz.php). These are the same six questions included in the 2018 NFCS to measure the same construct and are the items evaluated for bias in the current study.

A Common Focus on Intrinsic Factors

Financial literacy research and assessment in the U.S. have primarily focused on personal factors and have largely ignored the reasons for social, cultural, or macroeconomic differences in financial outcomes (Danes & Yang, 2014; Gudmunson & Danes, 2011). U.S. government guidelines have moved in this direction as well, encouraging an increased focus on individual responsibility for financial literacy. For example, in 2008, the newly formed Financial Literacy and Education Commission (FLEC) published their initial recommended *Research Priorities,* which suggested that both personal and social factors be included in financial literacy and education assessments. Personal factors included "the core principles of personal finance that every consumer needs to know" (FLEC, 2010, p. 4). Social factors included economic conditions and context and "socialization factors, including conflicting messages, and how those factors vary by gender, life stage, race, socio-economic status, education and ethnicity" (FLEC, 2010, p. 4). Surprisingly, the 2012 update to the *Research Priorities* outlined by FLEC included no social factors. It did, however, recommend reviewing "the economic and financial contexts in individuals' financial decision making" (FLEC, 2012, p. 3).

In 2017, however, five principles were identified in a study by the CFPB:

- Know the Individuals and Families to be Served.
- Provide Actionable, Relevant, and Timely Information
- Improve Key Financial Skills.
- Build on Motivation.
- Make It Easy to Make Good Decisions and Follow Through.

The first four principles focus on improving the individual; the last recognizes that good decisions and implementation are predicated on external factors. Still, within the literature, the predominant focus remains on the individual for improving financial knowledge and behavior.

Theoretical Framework

Many financial literacy research studies include no theoretical basis. Given the overall lack of consistency in the construct's operational definition (see Hung et al., 2009), an absence of a theoretical basis is not surprising. However, there have been a few notable exceptions. The literature indicates that three primary theoretical models have been used within financial literacy: The Family Financial Socialization Theory, the Theory of Planned Behavior, and the Transtheoretical Model of Change.

Danes (1994) developed the Family Financial Socialization Theory, which characterized financial socialization as "…much more inclusive than learning to effectively function in the marketplace. It is acquiring and developing values, attitudes, standards, norms, knowledge, and behaviors (p.128)," which shape the development of financial well-being. In this context, socialization factors may include sociocultural influences (such as media influence and religious affiliation), individual influences (such as peers, family, friends, teachers, and mentors), and the financial institutions, systems, regulations, and government policies within which the individual operates. Gudmunson and Danes (2011) developed a conceptual model showing family financial socialization as a mediator between demographic differences and financial outcomes.

Shim et al. (2009), Vyvyan et al. (2014), and Xiao (2008) each utilized aspects of the Theory of Planned Behavior (TPB) to study financial education and literacy. Shim and his colleagues (2009) developed the Student Financial Well-being Model using, in part, the TPB as a basis. Xiao et al. (2011) found that each of the three elements of TPB (financial attitude,

perceived behavioral control, and subjective norms) was significantly related to financial behavior. Finally, Vyvyan et al. (2014) adapted categories from TPB to assess the relationship between specific socialization factors and the financial capability of people experiencing financial trouble. The factors were categorized as background, attitudes, normative influences, perceived control and skills, and knowledge.

Shockey & Seiling (2004) and Xiao et al. (2004) each applied the Transtheoretical Model of Behavior Change (TTM) to their financial education research. TTM theorizes that behavior change requires an individual to progress through five distinct stages (*pre-contemplation*, *contemplation*, *preparation*, *action*, *and maintenance*) and that each stage requires different strategies for success.

Notably, two of the three theoretical frameworks applied to financial literacy have social factors as a basis. The current study applies conceptual theories from the discipline of educational psychology: specifically, Constructivism and the epistemological underpinnings of Pragmatism, to the discipline of financial education to bridge the gap between personal and social factors as determinants of one's financial literacy. The following section will discuss the constructivist theories of Jean Piaget and Lev Vygotsky, whose foundational work in educational psychology provide a theoretical basis for the current study and the method(s) used. Next, the work of John Dewey will be applied to further ground the study philosophically.

Constructivism

This study is theoretically grounded in the constructivist theory of learning: that individuals learn through interaction with others and within the context of historical and cultural norms that individuals experience throughout their lives. Constructivism is the theory that learners come into an educational setting with prior knowledge, skills, beliefs, and concepts that

facilitate how they construct solutions to problems and lead them to construct new knowledge (Lam, 2011). Constructivism involves a bilateral transfer of knowledge, both internally within an individual and externally between individuals. The social context of one's experiences facilitates the learning process by providing alternative views and supplemental information against which we can measure our understanding and reflectively adjust our propositions (Morgan, 2014). Therefore, for constructivists, there is no "reality" except for the one we self-construct, and there cannot be universal "Truth," for both reality and truth ("facts") are in a constant state of flux.

Constructivist theory in education focuses on the learner. Constructivism's central theme is that individual knowledge is constructed through the learner's experiences instead of memorizing "facts" or being told what is "true." Instead, students construct their understanding of the world by being challenged to support their assertions in the face of new knowledge and experience (Leidner & Jarvenpaa, 2016).

Constructivists maintain that our understanding – what we "learn"- is based on our context: culture, prior experience, worldview, etc. Thus, every learning outcome is unique. The works of two seminal theorists in Constructivist thought, Jean Piaget and Lev Vygotsky exemplify Constructivism's overarching philosophy and differences of opinion between its proponents.

Piaget

Jean Piaget developed a theory of Cognitive Constructivism, which suggests that knowledge is something individuals actively construct by modifying their existing cognitive structure. That is, learning is relative to one's stage of cognitive development. The two main principles underlying Piagetian theory are: 1) the theory of adaptation and concept of cognitive schemes and 2) his theory of cognitive developmental stages (Huitt & Hummel, 2006).

Piaget (1896-1980) believed that knowledge is constructed cognitively, within the individual, by creating connections between prior knowledge and new experiences. His contributions to the domain include the concepts of schemes, assimilation, accommodation, and equilibrium. While interacting with their environments, Piaget thought that children cognitively process information into clusters of related ideas called "schemes." A child's experiences within his environment must either fit (assimilate) into his existing mental structures (schemes), or the existing structures must be changed to make them fit (accommodate). Piaget asserted that this process of assimilation to accommodation – the resolution of the conflict between one's current understanding and the new information that threatens it - continues to repeat until the child's understanding is brought back into a state of "equilibrium." Piaget asserts that this process, which he called "adaptation," is the key to cognitive growth.

Piaget applied his research to educational strategies. He believed that the role of the teacher is to facilitate learning rather than to impart knowledge. His constructivist teaching sought to help students assimilate new information to existing knowledge and make the appropriate cognitive adjustments to accommodate that information (Driscoll, 1994). Piaget suggested that children should not be taught certain concepts until they have reached an appropriate cognitive development stage and advocated student-centered learning and active discovery learning.

Vygotsky

Vygotsky (2012) states:

> … direct teaching of concepts is impossible and fruitless. A teacher who tries to do this usually accomplishes nothing but empty verbalism, a parrotlike

repetition of words by the child, simulating a knowledge of the corresponding concepts but actually covering up a vacuum (p. 150).

Vygotsky's Social Constructivism Theory was foundational to the Constructivist movement (Liu, 2010). Vygotsky (1896-1934) was a prominent Russian psychologist who expressed views that were quite unconventional in his time; he proposed that social learning precedes development in children and that social interaction leads to increased knowledge – concepts not yet even considered in his day. Vygotsky's Social Constructivism theory directly opposed the Piagetian theory. He believed that external factors - societal and cultural - dictate what is important to learn rather than internal (cognitive) functions (Driscoll, 1994).

In his general law of genetic development, Vygotsky stated, "…every function in the child's…development appears twice: first, on the social level, and later, on the individual level; first between people (interpsychological), and then inside the child (intrapsychological)…" (Vygotsky, 1978, p. 57, as quoted in Daniels, 2016, p. 19).

His theory asserted that a child's development results from cultural factors such as dialog with others, inner speech, the institutional environment in which the child is schooled, and the availability of adult mentors for whom the child can serve as an "apprentice." Vygotsky contributed several concepts to the domain of Constructivism, including scaffolding, the Zone of Proximal Development (ZPD), and the More Knowledgeable Other (MKO).

These three concepts are interrelated. Vygotsky believed that children "construct" meaning from the world around them, developing an understanding through experience and others' help. The child mentally absorbs information and relies on the More Knowledgeable Other to bridge the gap between what he already knows and what he has yet to understand. The Zone of Proximal Development is the gap filled, and scaffolding is the "bridge" to

understanding. For Vygotsky, mental life foremost expresses itself through interactions with others.

Societal and cultural influences are integral components of social development theory. Vygotsky emphasized the cultural "tools" that create meaning in a child's mind. He was interested in how a culture's symbolic tools manage, through social interaction, to be internalized by the child into the mental construct that he makes of the world. Vygotsky believed that language was the highest form of human social interaction and cognitive development and that all higher mental functions are initially a response to external (social) occurrences processed internally through language. He termed this process "dialectical discovery," a continuous feedback loop that facilitates learning (Driscoll, 1994).

John Dewey's Pragmatist Philosophy and Contribution to Constructivism

John Dewey helped coordinate the underlying premises of these two theorists by developing a cohesive philosophy of education that addressed a central dispute in the discipline: In summary, the disagreement concerned whether Constructivism, as a school of thought, primarily focused on individual constructions of meaning (as Piaget believed) or social constructions of meaning (as Vygotsky purported).

Dewey's framework was drawn on his philosophical stance as a Pragmatist. In agreement with William James, his philosophical predecessor and "Father of American Psychology," Dewey believed that "we should conceive of all our knowledge as hypotheses to be tested in experience" (Kloppenberg, 1996, p. 104). To a pragmatist, experience is of primary importance; it gives our lives meaning; equally as importantly, each experience is, by definition, unique to its circumstances. If it is from this experience that knowledge is conferred, then these factors preclude "knowledge" and "reality" from being anything but constructions in one's mind.

Collective meanings then develop within societies as individuals act on their knowledge as a community.

In his early work (Dewey, 1896), Dewey introduced the concepts of transaction and experience as components of knowledge construction (Dewey, 1896). In response to the Behaviorist theories popular in his day, Dewey suggested a more transactional, reciprocal approach to “stimulus and response” as the basis for transferring knowledge. Dewey asserted that *knowing* always requires action on the learner's part (Bredo, 1998). That action is the “mediation of an experience” between a stimulus and the subsequent response (Dewey, 1896, p. 2). For example, in *Experience and Education*, Dewey (1938) suggests that it is problems (the "problematic") that identify and structure learning. Dewey (1922) proposed that the way to solve the discrepancies between our beliefs and our experience is to experiment inside our heads, using reflective thinking to perform a "dramatic rehearsal (in imagination) of various competing lines of action" (p. 190). Dewey (1910) called this "reflective practice" (p. 6): "…the scientific reasoning of developing a working hypothesis…to guide the investigation", moving to "inductive discovery" and testing conclusions by "deductive proof" (p. 81). According to Dewey, through reflection, knowledge and action come full circle.

Dewey and Bentley (1949) built on the concept of reflection with the introduction of transactional learning in education. They described learning as a reciprocal, triadic process: Inquiry leads to action, which leads to knowing and further inquiry. Throughout his career, Dewey maintained that both the subjective (individual) and intersubjective (sociocultural) dimensions of the construction of knowledge could – and should - be considered within the same constructivist framework. By bringing both dimensions together, Dewey offered a new

perspective to address the dispute: Constructivism could, and should, focus on individual constructions of meaning *and* social constructions of meaning.

Self-Regulated and Co-Regulated Learning Theories

Self-regulated learning (SRL) theory incorporates both individual and social factors in the learning process. SRL involves the individual taking an active role in the learning process by exercising control over cognitive, metacognitive, motivational, and behavioral functions to achieve established goals (Panadero & Alonso-Tapia, 2014; Zimmerman, 1990). Zimmerman (1989) developed one of the earliest models to explain SRL dynamics. The model identified three interactive factors within SRL: environment, behavior, and person. Zimmerman identified reciprocal relationships between these factors, known as "self-oriented feedback loops," as integral to the self-regulation process because they allow the learner to continue to strive to achieve higher goals (Zimmerman, 1990)

More recent contributions to the SRL literature include sociocultural factors. For example, co-regulated learning (CoRL) derives from a Vygotskian perspective to explain transitional processes toward self-regulation. It is grounded in Vygotskian views of higher psychological processes being socially embedded or contextualized and then internalized through social interaction (Mccaslin & Hickey, 2001). For example, McCaslin and Hickey (2001) defined co-regulation in terms of the interaction between teacher and learner within a zone of proximal development.

As discussed previously, Vygotsky's views were based on the belief that the role of speech within a social context is integral to one's learning and that dialogue with More Knowledgeable Others (MKOs) allows one to develop both cognitively and socially. He believed that as speech becomes more developed, it becomes internalized and allows one to effectively

navigate and act upon the environment. Effective self-regulation, in Vygotskian terms, utilizes inner speech to control one's behavior, plan and execute tasks efficiently, and reign in emotional responses.

To the author's knowledge, SRL theory has not been empirically studied in the financial education domain. However, related studies have been conducted. For example, Howlett et al (2008) found that an individual's self-regulation, future orientation, and financial knowledge were related to their decision to participate in a 401(k) plan. The study points to research that has "investigated how various societal problems are explained by consumers' struggles with self-regulation" and identifies financial decision-making as "a classic decision context that requires self-regulation (p. 226)." As defined in this study, self-regulation referred only to psychological factors like those used to define self-control.

Baumeister et al. (2008) use the terms self-regulation and self-control interchangeably and consider this trait a "limited resource" that, when depleted, will impair decision-making. This concept is referred to as "ego-depletion." Maranges et al. (2017) cite numerous studies on self-control that suggest that ego depletion is related to poor health (e.g., obesity), substance abuse and addiction, risky behavior (e.g., impulsive behaviors such as gambling, cheating, criminal acts, unsafe sex, and, within the realm of personal finance, overspending).

As a theoretical foundation for this study, Constructivism provides a basis for understanding the construction of knowledge as stemming from individual and sociocultural factors. SRL and CoRL provide additional insight into the processes used to construct knowledge. In addition, Dewey & Bentley's (1949) Transactional Learning Theory supports SRL theory in its proposal that learning is a reciprocal, triadic process and that both individual and sociocultural factors are integral to the development of knowledge.

Constructivism and Financial Literacy Assessment

As Constructivists, Piaget and Vygotsky both believed that learning is grounded in the interaction between mind and environment; learning does not solely occur passively by memorizing "facts" or by the revelation of "truth." Constructivism provides a theoretical basis upon which an argument may be made that an individual's "financial literacy" is based on what they have learned through formal and informal methods, both of which may have been influenced by sociocultural factors as well as individual factors. As discussed previously, the research literature indicates that formal financial education has thus far had little to no effect on individuals' financial knowledge, attitudes, and behaviors.

Based on a theoretical conceptualization that encompasses both endogenous and exogenous factors as drivers of financial decision-making and behavior, this study argued that America's current most popular financial literacy measures may be biased due to a lack of consideration for social and cultural factors that inherently affect individuals' financial knowledge, attitudes, and behavior. The theories of Self-Regulated Learning and Co-Regulated Learning provide an additional framework within which financial literacy can be understood as the result of both internal and external influences in financial education.

Assessment in Financial Literacy

Based on the literature, the current practice of financial education does not meet Dewey's standards for "knowledge as experience" (Dewey, 1938). Research has shown that most individuals cite personal experience as the primary source of their financial learning (CFPB, 2015; Hilgert et al., 2003), yet financial literacy is most often assessed using the “Big Three” and “Big Five” knowledge measures. Potential sociocultural factors, such as one’s level (and quality) of education, household financial status during childhood, parental influence, generational

experiences, media influences, and the political environment, are not considered. The following are specific outcomes identified and inferences made in the financial literacy literature that have been widely accepted within the discipline.

Group Membership and Financial Literacy

There is consensus in the literature that particular groups are susceptible to "low financial literacy." For example, Gupta et al. (2018) reviewed the results of the National Financial Capability Study (NFCS, 2016) to measure the construct of "Financial Fragility." The authors defined the construct using answers to the question: "How confident are you that you could come up with $2,000 if the need arose within the next month?" Those who "definitely could not" or "probably could not" come up with the money were considered financially fragile. In contrast, those who "definitely could" or "probably could" find the resources were considered not financially fragile. The study found that 46% of African American participants were financially fragile, while only 37% of white, non-Hispanic participants were considered such. Furthermore, the authors suggested that demographic factors such as age, education, and gender - among others - impact an individual's financial situation. For example, working-aged women (those between the ages of 25 and 60) were found to be more financially fragile than working-aged men; those with a bachelor's degree or higher were less financially fragile than those with lower educational attainment, and high levels of debt and low asset levels were correlated with financial fragility.

Furthermore, outcomes such as "financial fragility" are widely attributed to consumers' individual "choices" in the literature. However, factors such as income and education levels result from many things, including "financial misfortune – circumstances that might be well beyond an individual's control, and particularly systemic factors that affect marginalized groups"

(Pinto & Coulson, 2011, p. 63). The following are examples of research studies that have implicated individual factors as potential predictors of financial behaviors while potentially overlooking socioeconomic, demographic, and cultural factors contributing to participants' financial circumstances.

Age and Financial Literacy

Studies have indicated that financial literacy scores tend to increase with age until retirement when they begin decreasing (Lusardi & Mitchell, 2011). In addition, studies that have addressed age as a factor in financial literacy have suggested that older Americans suffer from a lack of financial savvy. For example, Gamble et al. (2015) show that cognitive decline related to aging affects numeracy ability in elderly individuals. However, subsequent research indicated that confidence in one's own financial ability does not wane after age 60 (Finke et al., 2017).

Lusardi & Mitchell (2014) refer to the results of the 2004 Health and Retirement Study (HRS), which utilized the "Big Three" to measure financial literacy, to emphasize the lack of financial knowledge among older Americans (65+):

> This older U.S. population is quite financially illiterate: only about half could answer the simple 2 percent calculation and knew about inflation, and only one-third could answer all three questions correctly (Lusardi & Mitchell, 2011). This poor showing is notwithstanding the fact that people in this age group had made many financial decisions and engaged in numerous financial transactions over their lifetimes. Moreover, these respondents had experienced two or three periods of high inflation (depending on their age) and witnessed numerous economic and stock market shocks (including the demise of Enron), which should have provided them with information about investment risk (p. 11).

Conversely, a study conducted by Yakoboski et al. (2019) found that financial knowledge correlated with age group. The youngest group (ages 18-29) was the least financially literate. Those ages 30 to 44 showed improved financial knowledge, but the difference was not statistically significant. In contrast, participants ages 45 to 59 did show a statistically significant increase in their *P-Fin Index* score, and adults over age 60 showed an even higher, statistically significant level of financial literacy.

Race, Ethnicity, and Financial Literacy

The financial literacy literature indicates that wealth is demarcated by race. For example, the results of the Jumpstart Coalition for Personal Financial Literacy in the USA (Mandell, 2008) showed lower financial knowledge levels among African American university students compared to European-American university students. Furthermore, much of the literature -- based on studies using the "Big Three" or "Big Five" assessment measures -- suggests that lower financial literacy scores among African American and Hispanic individuals are due to a lack of knowledge and that more financial education is needed to solve the problem (Al-bahrani et al., 2018; Hilgert et al., 2003; Lusardi & Mitchell, 2007).

The literature also indicates that African American individuals exhibit a disproportionate amount of "unfavorable" financial behaviors. For example, research has shown that African American adults are over three times more likely to have been late with a mortgage payment over the past year, and more African American consumers engage in expensive credit card behaviors than white consumers (Yakoboski, 2019). Furthermore, twice as many African American individuals carry student loan debt than their white counterparts and are more likely to have been late on a student loan payment in the past year (Fan & Chatterjee, 2019).

Level of Education, Intelligence, and Numeracy

Studies have found that individuals with higher general cognitive abilities, better aptitude with numbers and numerical calculations (numeracy), or mathematical ability tend to have higher financial literacy levels (Banks & Oldfield, 2007; Cameron et al., 2014; Gerardi et al., 2010). For example, Grinblatt et al. (2009) observe that investors who show higher levels of intelligence on average earn 11% higher returns than those with lower intelligence levels. Difficulty with simple mathematical calculations has been correlated with lower savings levels (Banks & Oldfield, 2007; Smith et al., 2010) and poor planning for retirement (Lusardi & Mitchell, 2011). Gerardi et al. (2010) report "robust" findings of correlations between numerical ability and mortgage delinquency (p. 4). Lachance (2014) similarly found education levels to be a predictor of financial literacy.

Gender Factors

The literature has also found that men indicate higher levels of financial literacy than women (Bucher-Koenen et al., 2017; Chen & Volpe, 2002; Klapper et al., 2015; Lusardi & Mitchell, 2007; Woodyard & Robb, 2012). Research has further indicated that women self-report lower confidence and less interest in learning about personal finance topics than men (Bucher-Koenen et al., 2017; Chen & Volpe, 2002; Lusardi & Mitchell, 2011) and are more conservative and security-conscious (Edwards et al., 2007). However, the relationship between gender and financial literacy is not clear. For example, Brown & Graf (2013) found that gender differences are not due to the level of interest in finance and financial matters between women and men. Furthermore, while some authors (Danes & Yang, 2014; Fonseca et al., 2012; Hsu, 2016; Woodyard & Robb, 2012) have suggested that gender differences are due to the traditional roles of women and men in households, this theory is incomplete, as it does not explain why single

women or high-school age females score lower on financial literacy assessments (Chen & Volpe, 2002).

Validity and Fairness in Financial Literacy Assessment

Validity

Validity in research speaks to an instrument's ability to answer the question it purports to ask and the strength of its conclusions. For example, studies that measure the construct of financial literacy using surveys or questionnaires should be validated based on how well the assessment tool measures the concept of interest (Messick, 1989). Furthermore, the validity of these assessment instruments requires multiple evidence sources to show that the instrument measures what it is supposed to measure. The Standards (American Educational Research Association et al., 2014) present five sources of validity evidence. Specifically, evidence should be collected which relates to (1) test content, (2) response processes, (3) internal structure, (4) relations to other variables, and (5) consequences of testing. The current study focuses on the consequences of testing as a potential source of invalidity.

Messick (1989) identifies two threats to validity based on test consequences: construct under-representation and construct irrelevant variance. Construct underrepresentation is "the degree to which a test fails to capture important aspects of the construct" (American Educational Research Association et al., 2014, p. 12), indicating that an assessment is too narrow. Conversely, construct-irrelevant variance indicates that an assessment is too broad. The *Standards* define construct-irrelevant variance as "the degree to which test scores are affected by processes that are extraneous to the test's intended purpose" (American Educational Research Association et al., 2014, p. 12). Construct-irrelevant variance refers to a difference in test results influenced by factors that should, in theory, be unrelated to test performance, such as native

language, race/ethnicity, gender of the test-taker, or other distinct constructs. If these factors significantly influence performance on a test, the validity of the test results is questionable. Although construct-irrelevant variance and construct underrepresentation cannot be entirely removed from an assessment, the goal is to minimize the presence of both (Hubley & Zumbo, 2011). Construct underrepresentation or construct-irrelevant sources of invalidity can adversely impact groups of test-takers if the results lead to invalid consequences.

Consequences refer to the interpretation and use of test scores to draw specific inferences. Validity evidence based on the consequences of testing reflects the "soundness of proposed interpretations [of test scores] for their intended uses" (American Educational Research Association et al., 2014, p. 19). Consequences may refer to intended or unintended effects and any value implications made by the theory underlying the construct or in the development or implementation of the test and social consequences that may arise from the test (Hubley & Zumbo, 2011).

Numerous inferences have been made based on data from studies using the "Big Three" and "Big Five" financial literacy questions, including responses from the National Financial Capability Studies conducted in 2009, 2012, 2015, and 2018. For example, using the results of the NFCS, researchers have indicated that those who score well on the six financial literacy questions are more likely to have an emergency savings account (de Bassa Scheresberg, 2013) and retirement savings (Lusardi & Mitchell, 2007, 2011) Conversely, those whose responses indicate low financial literacy are more likely to have costly mortgages, less likely to refinance their mortgages during a period of falling interest rates (Campbell, 2006), and are more likely to engage in costly borrowing behavior through the use of credit cards and non-mainstream lenders such as pawnshops, payday loans, and title loans (de Bassa Scheresberg, 2013; Mottola, 2013).

However, other studies have attempted to measure the construct (often using alternative instruments rather than the “Big Five”) and have found little or no relationship to outcomes or behavior. For example, several research studies could not find any relationship between financial education and household financial behavior (Bernheim et al., 2001; Fernandes et al., 2014). Fernandes et al. (2014) studied the relationship between financial education and behavior. The study found that "…financial education interventions studied explained only about 0.1% of the variance in the financial behaviors studied, with even weaker average effects of interventions directed at low-income rather than general population samples (p. 1872)."

Regarding intelligence, researchers have concluded that people with higher education levels have higher scores on financial literacy tests and that intelligence can predict financial literacy. However, this finding is confounded with age, as those with more education are also invariably older. Furthermore, Lachance (2014) showed that level of education is indeed a predictor of financial literacy but also determined that zip code education level was a more significant factor.

Bias and Fairness

The *Standards* state that "a fair test does not advantage or disadvantage some individuals because of characteristics irrelevant to the intended construct" (American Educational Research Association et al., 2014, p. 50). The *Standards* identify procedural fairness, measurement bias, access to the construct being measured, and valid test score interpretation as specific components of test fairness. The current study addresses potential measurement bias in the six financial literacy questions used on the 2018 NFCS.

Measurement bias may be present if the assessment systematically gives one group a relative advantage. Examples include using leading questions or culture-specific terminology or

concepts in test items. Similarly, unequal access to the measured construct would put certain test-takers at a disadvantage (American Educational Research Association et al., 2014. Valid test score interpretation involves an appraisal of proposed individual score use and interpretation (American Educational Research Association et al., 2014; Messick, 1989). The *Standards* require developers and score interpreters to be aware of differences among individuals and not generalize about individuals from the performance of subgroups to which they belong (American Educational Research Association et al., 2014). Financial literacy assessments may violate the valid test score directive for fairness in testing if the results are used to generalize about individuals based on the subgroups to which they belong, and inferences are made that all individuals in a group are "more" or "less" financially literate based solely on their group membership.

However, this potentially unfair interpretation does not necessarily make such assessments biased (i.e., invalid). In educational measurement, the terms bias and fairness are associated but remain different concepts. Bias denotes group differences in scores by examinees who are otherwise equally capable. Alternatively, fairness involves using the test scores and their consequences. In addition, a test can demonstrate no statistical evidence of bias yet still be judged as unfair (Warne et al., 2014). Conversely, a test or item can be biased and still be fair if the explanation for group-based performance differences is directly related to the primary construct(s) and not another irrelevant construct. If this occurs, the observed bias indicates *item impact* but not unfairness (Camilli, 2013).

The distinction between fairness and bias can be a fine line, yet it is an essential delineator between what is valid in assessment and what is not. Kane's argument-based

validation model, upon which modern validity theory is partially based, helps make this distinction**.**

Kane's Argument-based Validation Model

In Kane's words, "the evaluation of test score uses requires an evaluation of the consequences of the proposed uses, and negative consequences can render a score use unacceptable" (Kane, 2013, p. 46). Kane (2013) proposed an "argument-based" validation model that explicitly addressed the assumptions inherent in test scores' proposed interpretations and uses. He identified two steps to the process: developing the interpretive argument (later identified as the interpretive/use argument), which is then rated based on its ability to present a validity argument. According to Kane (2013), the interpretive argument considers the end-use of the assessment first: it answers the question, "how will the results be interpreted?" The interpretive argument identifies the "most relevant" assumptions and inferences used to interpret the results of an assessment and "provides a clear basis for choosing the kinds of evidence to be included in the validity argument" (M. T. Kane, 1990, pp. 9–10).

The validity argument involves analyzing and assessing the relative strength of the inferences identified in the interpretive argument and exposing any weaknesses. Evidence given by the interpretive argument may be empirical (such as referencing extant literature or performing statistical analysis) or judgmental (i.e., based on expert evaluation) but should be thorough enough to make a convincing argument (Chalhoub-Deville, 2016; Zumbo & Wild, 1999).

Kane's (2013) renaming of "Interpretive Argument (I.A.)" to "Interpretation and Use Argument (IUA)" and extensive contribution to the theoretical discussion of external forms of validity evidence were significant. Disagreement among validity theorists regarding the value of

consequences – specifically, construct irrelevant variance and construct underrepresentation – as part of validity had been an ongoing debate; many theorists believed that the scope of validity should be limited to construct-related validity evidence and fundamental investigations of bias (Chalhoub-Deville, 2016). Kane's Argument Based Validation Model reinforced the importance of test use and consequences to validity.

To the author's knowledge, there has never been such a process initiated for the "Big Three" or "Big Five" assessment items or any consideration of the uses and consequences of financial literacy assessments. Indeed, there is little support of validity evidence for financial literacy assessments based on the inferences made about test results. However, some effort has been made to review the consequences of using these questions to measure "financial literacy"; these authors propose that the inferences made from the research are questionable (Hastings et al., 2013; Pinto & Coulson, 2011; Willis, 2017).

Using Differential Item Functioning (DIF) Analysis to Identify Potential Bias

Differential item functioning (DIF) analysis can be used to identify bias (Mitchell & Cawthon, 2015). Holland & Thayer (1988) pioneered the term "differential item functioning" to convey the concept of test questions with varying degrees of difficulty for groups of otherwise comparable proficiency. Originally called "item bias" research, DIF can result from differences in socioeconomic status, lifestyle, culture, educational experience, and other confounding factors with group membership. DIF occurs when an item measures more than one underlying latent trait and when cognitive differences occur between groups on one of these extraneous traits (Martinková et al., 2017). In addition, DIF analysis provides insight into the cognitive and psychological processes involved in item response and test performance between defined groups (Oliveri et al., 2013). Gómez-Benito et al. (2018) propose that DIF studies provide a more

comprehensive measure of validity evidence than the *Standards* (2014) indicate and that DIF studies measure the fundamental fairness of all sources of validity evidence.

The extant literature regarding DIF indicates that it is an evolving measurement tool. For example, Zumbo (2007) identifies three “generations” of DIF. The first generation of DIF used the term “item bias” and failed to distinguish impact from DIF. This distinction was resolved in the second generation, and new statistical DIF methodologies were introduced. Zumbo (2007) describes a subtle but significant shift during the third generation, as DIF was believed to occur "because of some characteristic of the test item and/or testing situation that is not relevant to the underlying ability of interest" (p. 229). He describes the third generation as seeking to explore contextual variables (e.g., cultural environment, socioeconomic status, opportunity to learn) as causes of DIF. Such factors have typically been ignored in DIF analyses (Zumbo, 2007).

Gómez-Benito et al. (2018) argue that item bias and DIF are not distinct concepts. Instead, they assert that third generation DIF analysis involves "combining statistical findings with substantive explanation regarding the construct underrepresentation and/or construct-irrelevant cause of the differential item performance" (p. 107).

Commonly used statistical methods for measuring DIF in educational assessment are the Mantel-Haenszel Chi-Square Test (Mantel & Haenszel, 1959), logistic regression procedures, and Item Response Theory (IRT) models (Gómez-Benito et al., 2018). The Mantel-Haenszel procedure is easy to calculate (no software required) and helps detect DIF in small sample sizes. However, the Mantel-Haenszel procedure cannot distinguish between uniform and non-uniform DIF. Logistic regression procedures can capture both uniform and non-uniform DIF but require statistical analysis software. The Mantel-Haenszel and logistic regression procedures use the total test score to measure the participant's overall ability within the construct. Although no DIF

analysis can determine, with certainty, the specific cause of the DIF, IRT methods do have an advantage in that they measure latent traits (as item difficulty or person ability) rather than using the total test score to estimate an ability level. IRT methods are therefore considered "parametric," and the Mantel-Haenszel and logistic regression procedures "nonparametric" based on this difference. IRT methods have the additional advantage of measuring both uniform and non-uniform DIF (Martinková et al., 2017).

The current study sought to identify DIF using an Item Response Theory (IRT) model to detect potential measurement bias at the item level. The study analyzed the extent to which the scores on each financial literacy item on the 2018 National Financial Capability Study differ based on the participant's group membership (i.e., age, gender, and race).

Item Response Theory Models

IRT models are particularly suitable in educational research because they provide a system of measurement that considers both the test's difficulty and the respondent's level of ability. Item Response Theory posits that an individual's response to an individual test item is related to their level of the latent variable (e.g., financial literacy) and the level of difficulty of the item. The IRT measurement standardizes raw scores to identify ability levels (known as the *item response function*) and then uses this latent variable to measure an item's difficulty. The model measures the probability that an individual with a particular trait will correctly answer an item with a certain difficulty level.

An IRT model requires two assumptions to be proven. Those assumptions are unidimensionality and local independence. Unidimensionality assumes that the test items measure only one latent trait (for this study, financial literacy) and that no other trait influenced the item response. The local independence assumption requires that the respondent's latent trait

measurement is independent of group membership (for this study, the groups are based on age, gender, and race). When the assumptions of IRT are proven, its estimates of the item parameters are independent of the sample. A respondent should show the same ability, independent of the set of items adopted, and conversely, a given item should have the same difficulty, independent of the respondents.

IRT provides a theoretical framework for DIF detection that is aligned with Kane's Argument-based Validation Model. When IRT models are used in DIF detection, differences in the item parameters across groups are examined in relation to the ability parameters. The relationship between the probability of a correct response and the item and ability parameters is measured. The DIF analysis involves both a *reference* group (the ‘‘baseline" group, assumed to be the advantaged group) and a *focal* group (assumed to be the disadvantaged group). The members of each group are "matched" on a measure of ability derived from the test. Then, further statistical procedures are implemented to identify group score differences in item difficulty within those matched groups (NFCS, 2016).

Determining a proper method of DIF analysis to use requires identifying whether the underlying scores are reported using two (dichotomous) or several (polytomous) responses. The current study will analyze the responses as either "correct" or "not correct" and, therefore, will use dichotomous data.

DIF further distinguishes between primary and secondary dimensions as factors to explain why DIF occurs. Each item in a test is designed to measure the central construct, known as the *primary dimension*. DIF items are identified as measuring one or more *secondary dimensions* in addition to the primary dimension (Roussos & Stout, 1996; Shealy & Stout, 1993).

DIF caused by secondary dimensions is either *benign* (reflecting impact, or true differences) or *nuisance* (reflecting bias due to a factor irrelevant to the construct).

Multidimensional IRT Models

Multidimensional IRT models extend the assumption of unidimensionality. The underlying theory is that although there is generally only one primary dimension of interest in a test, there are likely other dimensions within that test that produce construct-irrelevant variance. For example, in financial literacy assessment, reading comprehension and vocabulary may be measured along with the underlying construct. Shealy & Stout (1993) introduced a new DIF test statistic, the simultaneous item bias test (SIBTEST), based on the multidimensional IRT framework. SIBTEST analysis provides an alternative theoretical, hypothesis-driven approach to unidimensional IRT for DIF analysis. Unidimensional DIF requires that DIF analysis be conducted first. Then, if potential DIF is identified, those items are further studied by content specialists and other experts to determine whether item impact or bias is present. Roussos & Stout (1996) suggested that DIF analysis could be implemented after developing a hypothesis (based on pertinent literature and consultation with specialists) regarding how, why, or for whom DIF may be influencing outcomes.

Given the needs of the current study, the SIBTEST software was chosen to measure DIF. The SIBTEST program was chosen due to its multidimensional approach, and because it has been shown to produce minimal Type I error rates in larger sample sizes with fewer test items relative to logistic regression models and the Mantel-Haenszel method (Jiang & Stout, 1998; Magis et al., 2010).

Summary

The theoretical framework of this study and the research questions were significant factors in the choice of research method. There was no precedent in the literature to guide the choice of research method for this study, as no studies have been conducted along a similar path. However, the current study is built on the foundation of the many empirical studies that have indicated specific group differences in financial literacy scores based on the "Big Three" and "Big Five" assessment items. Earlier studies have empirically identified existing correlations between specific groups (i.e., age, gender, race/ethnicity, numeracy/mathematical ability, intelligence) and financial literacy as operationalized by the "Big Three" and "Big Five" assessment items developed by Lusardi, or Lusardi and Mitchell. Many of these studies have been identified previously in this chapter (Banks & Oldfield, 2007; Bucher-Koenen et al., 2017; Cameron et al., 2014; Gerardi et al., 2010; Grinblatt et al., 2009; Gupta et al., 2018; Lusardi & Mitchell, 2007, 2011, 2014; Woodyard & Robb, 2012). Furthermore, several of these studies have based their findings on data from the National Financial Capabilities Study (NFCS). DIF analysis was chosen for this study to measure the extent to which the six financial literacy items on the 2018 NFCS show potential measurement bias: specifically, construct irrelevant variance related to group differences (i.e., age, gender, and race). Such bias may indicate that one or more items is invalid based on the inferences and uses of the results.

Gaps in the Literature

This review of the relevant research literature suggests that over the past 20 years or more: 1.) The construct of "financial literacy" in the United States has not been well defined and is difficult to measure; 2.) The most widely-used assessment items – the "Big Five" – have failed to establish validity evidence to support the inferences made by studies that use them as

measures of "financial literacy," and 3.) the vast amounts of resources allocated to financial education – likely based on the inferences as mentioned above - have failed to improve financial literacy in America.

The present study advances the existing literature by (1) offering a theoretical basis from which to identify extraneous factors which may be affecting documented differences in financial literacy outcomes; (2) providing a statistical analysis of responses given on a recent national assessment of financial literacy to identify potential item bias; and (3) providing further empirical basis to support the development of a consistent and valid measurement assessment for financial literacy.

CHAPTER III: METHODOLOGY

Purpose of the Study

The purpose of this quantitative research study was to examine the extent to which financial literacy assessment in America may be affected by bias due to group membership, specifically measurement bias in the form of construct-irrelevant variance. According to Messick (1989), construct-irrelevant variances threaten the consequential validity of items on a test, thereby undermining the appropriateness and usefulness of the test scores for decision-making. Within a test validation framework – specifically, Kane's (2013) "interpretive-use argument" - the study sought to identify such group bias and, if found, offer "plausible explanations for the differences…to promote sound score interpretations for all examinees" (American Educational Research Association et al., 2014, p.82).

Research Questions

The research questions which guided this study were:

R.Q. 1. To what extent do the financial literacy items on the 2018 NFCS exhibit reliability?

R.Q. 2. To what extent do the financial literacy items on the 2018 NFCS exhibit validity (measure a common construct)?

R.Q. 3. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by age?

R.Q. 4. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by gender?

R.Q. 5. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by race?

Data Source

The current study utilized data from the 2018 National Financial Capability Study. The study is often used to compare the financial behaviors of different demographic groups (Al-Bahrani et al., 2018; Allgood & Walstad, 2016; Knoll & Houts, 2012; Skimmyhorn, 2016). The survey was conducted in 2009, 2012, 2015, and 2018 by FINRA, the Financial Industry Regulatory Authority, through the FINRA Investor Education Foundation. The 2018 dataset was retrieved online at www.https://www.usfinancialcapability.org, where it is available for public use.

Participants

Participants for the 2018 NFCS study were a national sample of 27,091 adults ages 18 and older, approximately 500 from each of the 50 states. According to FINRA, the survey results are "weighted to be representative of the national population in terms of age, gender, ethnicity, education and Census Division" (Lin et al., 2019, p. 42). The survey was self-administered by respondents on a website and conducted between June and October 2018. A copy of the 2018 survey methodology is shown in Appendix I.

The data are publicly available on the data page (Financial Capability Study: Data and Downloads (usfinancialcapability.org). However, the format of the ethnicity data within the dataset had been combined to reflect only "white" and "non-white" participants. Through correspondence with FINRA, the organization that conducted the study, the author obtained the original ethnicity data, which delineated six ethnic identities. The correspondence noted above is shown in Appendix II.

Instrumentation

The codebook for the survey instrument used in the 2018 NFCS is shown in Appendix IV. The survey captured the following data for use in this study:

Section A, Demographics and Classification: Gender, age, ethnicity, education, household income, employment status, and parent's education level.

Section M, Self-assessment and literacy: The six "financial literacy" questions.

The measure of financial literacy is determined by respondents' performance on the six questions that assess the participants' understanding of the relationship between bonds and interest rates, investment risk, real rates of return (inflation), loan amortization, and compound interest.

Data Analysis

The national dataset was downloaded into Excel. The data was free of missing values and outliers. Descriptive statistics were gathered in Excel, and the data were transferred to R version 4.1.2. for further analyses.

Item Characteristics

The "ltm" package for R (version 4.1.2) was used to analyze the six items' characteristics. This analysis included calculating a Product-moment Correlation Coefficient (Pearson's *r*) and an Item Reliability Coefficient (Cronbach's α). The ltm package was accessed online on 2/6/22 at https://cran.r-project.org/web/packages/ltm/ltm.pdf.

Differential Item Functioning (DIF)

This study used a nonparametric approach to assess DIF using the SIBTEST procedure (Shealy & Stout, 1993). A modified “mirt” package for R (Chalmers, 2012) was used to conduct

the analysis. The specific package used in the analysis (difR/difSIBTEST) was accessed online on 1/2/22 at https://rdrr.io/cran/difR/man/difSIBTEST.html.

The difSIBTEST (Magis et al., 2010) estimates respondents' financial literacy levels by first counting the number of questions answered correctly, then identifying an estimate of the probability that an observed difference occurred by chance (*p*-value), and finally, describing the size of the difference (β). Initially, the test performs a procedure to "purify" the data (M. J. Gierl & Khaliq, 2001). The purification process involves screening each item for DIF to establish a subset of items that exhibit no statistically significant DIF. These items then became the comparison against which each of the remaining items, through sequential iterations, is analyzed. This procedure establishes the control for overall ability (the latent variable).

In this study, the dependent variable is the participant's score on each of the six questions. The scores are given as 0 (right) and 1 (wrong), and therefore the dependent variable is dichotomous. The independent variable is group membership. For the race/ethnicity group, the focal groups were *Black or African American, Hispanic or Latino/a,* and *Asian;* responses were minimal from the *Native Hawaiian or Pacific Islander* and *American Indian or Alaskan Native* groups. The individual test scores of each participant in the focal group were compared to those of the reference group, which was *White or Caucasian.* For gender, the focal group was *Females,* and the reference group was *Males*. Age ranges were used to classify participants' ages. The youngest age group (*18 to 24)* and the oldest age group (*65+*) were the focal groups; those from age *25 to 64* were the reference group.

The difSIBTEST analysis was performed to analyze DIF for each focal group. The modified ''mirt" package for R conducts a significance test of a two-tailed hypothesis to detect uniform DIF for each item's focal and reference groups. The DIF analysis compares the

probability of a correct response on an item for the reference group at a measured latent ability level (θ) with the same probability on the item for the focal group at the same ability level (Ibrahim et al., 2021). To correct for ability differences between the focal and reference groups, the observed scores are adjusted using a regression equation based on classical test theory to estimate true scores, TR_s and TF_s for members of each group. If the expected scores are identical, the reference and focal groups are equally matched on θ, and uniform DIF is not present. The level of DIF at θ can be measured as:

$$\beta\ 0\ (\theta) = E_R\ [Y/\ \theta] - E_F\ [Y/\ \theta]$$

If the ability level (θ) is given, this difference can be shown as:

$$\beta\ (\theta) = P_R\ (\theta) - P_F\ (\theta)$$

The SIBTEST statistic β is calculated as the average difference in the probability of a correct response for the two groups and represents the item difficulty parameter. E denotes the expected response given, and Y denotes the actual response given on the item. P represents the response probability for each group, where R is the reference group and F is the focal group (Ibrahim et al., 2021).

SIBTEST utilizes a range of values for the interpretation of β when the null hypothesis is rejected and $p < .05$. Based on guidelines offered by Roussos & Stout (1996), when $\beta < 0.059$, DIF is considered negligible; when $0.059 < \beta < 0.088$, DIF is considered moderate; and when $\beta \geq 0.088$, DIF is considered large.

Limitations of the Study Methods

Through DIF analysis, the current study seeks to identify potential measurement bias in the six financial literacy items on the 2018 NFCS. However, there are limitations within the study that should be noted. First, the larger sample size of the study may lead to "too much

statistical power" (Zumbo & Wild, 1999, p. 27). Too much power can create a greater tendency to reject null hypotheses with negligible differences (a Type I error). The current study focused on measures of effect size and statistical significance to offset these limitations. Second, Differential Item Functioning (DIF) can provide evidence of potentially biased test items but cannot establish why the unexpected differences occur between groups.

Summary

This study used a SIBTEST analysis to examine whether one or more of the six financial literacy items on the 2018 NFCS function differently between matched sub-groups delineated by age, gender, and race. The study hypothesized that the items measure construct-irrelevant factors along with the target latent trait variable and therefore exhibit DIF when they function differently across the subgroups. In the next chapter, the findings of the analysis are reported.

CHAPTER IV: RESULTS

This chapter provides a discussion of the statistical results of this research study. Data analysis results, including descriptive statistics and the results based on the research questions, are presented. The chapter ends with a summary of the findings.

Research Questions

R.Q. 1. To what extent do the financial literacy items on the 2018 NFCS exhibit reliability?

R.Q. 2. To what extent do the financial literacy items on the 2018 NFCS exhibit validity (measure a common construct)?

R.Q. 3. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by age?

R.Q. 4. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by gender?

R.Q. 5. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by race?

Data

The data for this study were obtained from a national database provided by the Financial Industry Regulatory Authority (FINRA). The data were free of missing values and outliers. Analyses for this research study were conducted using R version 4.1.2.

The difR package was used to compute the SIBTEST beta coefficients and relative DIF stat effects (Shealy & Stout, 1993). The difSIBTEST function provides a wrapper to the SIBTEST function from the "mirt" package (Chalmers, 2012) to fit within

the difR framework (Magis et al., 2010) The ‘‘mirt” package identifies potential DIF using Item Response Theory analysis (Chalmers, 2012).

Variables

In this study, the dependent variable is the participant's score on each of the six financial literacy items included in the 2018 NFCS (see Appendix II). The scores were coded as "1" (right) and "0" (wrong), and therefore the dependent variable is dichotomous.

The independent variable is group membership. The focal groups were *Black or African American, Hispanic or Latino/a, and Asian* for the race group. The individual test scores of each focal group were compared to those of the reference group, which was *White or Caucasian.* For gender, the reference group was *Males*, and the focal group was *Females*. Age ranges were used to classify participants' ages. The youngest age group (18 to 25) and the oldest age group (65+) were the focal groups; those aged 26 to 64 were the reference group.

Descriptive Statistics

Participant Characteristics

Twenty-seven thousand ninety-one (27,091) participants across the United States provided complete survey data on all variables. Overall, participant ages ranged from 18 to 101; the data for participants across age groups were normally distributed. The sample had more female than male respondents: 44.1% of participants identified as male (*N*= 11,956), and 55.9% identified as female (*N*=15,135). The percentage of participants who identified as White or Caucasian was 74.2% (*N*=20,099); 9.5 percent as Black or African American *(N*=2,576); 8.6 percent as Hispanic or Latino (*N*=2,338), and 4.5 percent identified as Asian (*N*=1,210). Four percent of participants self-reported their race as "Other" (*N*=868). Table 1 provides a summary of participant characteristics.

Table 1

Descriptive Statistics for the Independent Variables

Items		*n*	%
Gender		11,956	44.1
	Male	15,135	55.9
	Female		
Ethnicity			
	White/Caucasian	20,099	74.2
	Black/African American	2,576	9.5
	Hispanic/Latino	2,338	8.6
	Asian	1,210	4.5
	Combined Other[a]	868	3.2
Age			
	18 – 24	2,795	10.3
	25 – 64	18,784	69.3
	65+	5,512	20.3

Note. The "Combined Other" category includes all Native Hawaiian/Pacific Islander, Alaskan Native/American Indian, and "Other" responses.

Data Analysis Results

R.Q. 1: To what extent do the financial literacy items on the 2018 NFCS exhibit reliability?

Cronbach's Alpha (α) was computed to measure the internal consistency of the six items examined in the current study. By calculating the average inter-correlation of the items, Cronbach's α is a reliability coefficient that measures the relationship between items. A reliability coefficient of greater than .70 is generally considered “acceptable” in social science research (Dogra et al., 2021). However, an alpha between 0.80 and 0.95 has been suggested as preferable for the psychometric quality of scales (Nunnally & Bernstein, 1994). The Cronbach's α for the current items was calculated as .66. This result suggests somewhat low internal consistency and indicates that the items fail to measure similar constructs at an appropriate level to provide validity evidence (Taber, 2018). An analysis was also performed to measure the items’

reliability when individual items were excluded from the calculation. The Cronbach's α for this analysis resulted in reliability coefficients that ranged between 0.29 to 0.49, indicating even lower item consistency.

R.Q. 2: To what extent do the financial literacy items on the 2018 NFCS exhibit validity (measure a common construct)?

A Pearson's product-moment correlation coefficient (*r*) was calculated to measure the strength of the linear relationship between the test items. Only weak correlations were identified between the items indicating that the items are quite independent of one another. This finding also suggests that the items measure different constructs (Dunst & Hamby, 2012). Table 2 summarizes these findings.

Table 2

Correlation Coefficients (Pearson's *r*)

Item	Interest Inflation/ Rate Interest		Int Rate/ Bond Val.	Compound Interest	Interest Expense	Investment Risk
1	1.00					
2	.32*	1.00				
3	.11	.22*	1 .00			
4	.14	.20*	.15	1.00		
5	.27*	.33*	.16	.16	1.00	
6	.21*	.38*	.29*	.23*	.28*	1.00

Note. r = 0.0 – 0.19 (very weak); * r = .20 - .39 (weak); ** r = .40 - .59 (medium); *** r = .60 - .79 (strong)

Differential Item Functioning Analysis

For research questions 3, 4, and 5 regarding the potential bias of the financial literacy questions by group, the difSIBTEST analysis was performed for uniform DIF, and the Benjamini-Hochberg adjustment of p-values was applied. The effect size score (β) can be interpreted as follows: If the effect size is less than 0.059, it is considered negligible; if it falls at

or between 0.059 and 0.088, it is a moderate effect size, and if it is greater than 0.088, the effect size is considered large (Gierl et al., 1999).

R.Q. 3: To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by age?

Two focal groups were analyzed: ages 18 to 24 and ages 65+. The reference group included those ages 25 to 64.

Frequencies

The percent of correct responses by age group was consistent among all six items: Participants ages 65 and older consistently scored higher than those 25 to 64. Respondents ages 25 to 64 likewise scored higher than those 18 to 24. The results are shown in Table 3.

Table 3

Percent Correct Responses to Financial Literacy Questions by Age Group

Item	Age Group		
	18-24	25-64	65+
Q.1: Interest Rates	64.0	73.3	82.9
Q.2: Inflation	35.5	54.9	78.3
Q.3: Bonds/Interest Rates	20.2	24.8	37.6
Q.4: Compound Interest	29.5	31.0	32.3
Q.5: Interest Expense	56.1	76.0	85.4
Q.6: Investment Risk	29.2	44.3	59.0

Ages 18-24 as Focal Group

The results for this age group indicated negligible DIF for items one, three, and six. However, items two, four, and five were indicative of DIF. Item two indicated a moderate effect size (β = .06), while items four and five were found to have large effect sizes (β = -.10 and .09, respectively). The large positive effect size for item five shows that this particular item favored the participants in the focal group (ages 18-24). Table 4 shows the DIF results by age group.

Ages 65+ as Focal Group

The analysis indicated negligible DIF for items one, three, five, and six for this age group. Items two and four indicated large effect sizes (β = -.13 and 0.11, respectively). The large positive effect size for item four indicates that this item favored the participants in the focal group. Table 4 provides a summary of these findings.

Table 4

Differential Item Functioning Analysis by Age Group

	18-24	**65+**
	β	***β***
Item # 1	-0.02	0.02
Item # 2	0.06*	-0.13**
Item # 3	-0.03	-0.02
Item #4	-0.10**	0.11**
Item #5	0.09**	0.01
Item #6	0.02	0.01

Note. β <.059 (negligible) *β = .059-.088 (moderate) **β >.088 (large)

R.Q. 4: To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by gender?

The two gender groups were *Males* and *Females*. Males were the reference group, and females were the focal group.

Frequencies

The percent of correct responses by gender was consistent among all six items: Males consistently scored higher than females. Table 5 summarizes these findings by gender.

Table 5

Percent Correct Responses to Financial Literacy Questions by Gender

Item	Male	Female
Q.1: Interest Rates	79.3	70.4
Q.2: Inflation	65.9	51.1
Q.3: Bonds/Interest Rates	33.9	21.4
Q.4: Compound Interest	38.9	25.0
Q.5: Interest Expense	78.7	73.6
Q.6: Investment Risk	56.7	37.1

Differential Item Functioning Analysis by Gender

Within the gender category, the results indicated negligible effect sizes for all items except item five, which showed a moderate effect size (β = .07). The positive effect size for item four indicates that this item favored the participants in the focal group (females). A summary of the results is given in Table 6.

Table 6

Differential Item Functioning Analysis by Gender

	β
Item #1	0.05
Item #2	0.03
Item #3	-0.02
Item #4	-0.04
Item #5	0.07*
Item #6	-0.04

*Note. β <.059 (negligible) *β = .059-.088 (moderate) **β >.088 (large)*

R.Q. 5: To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by race?

Frequencies

The percent of correct responses by race was somewhat consistent among all six items: Generally, respondents who identified as Asian consistently scored the highest; those who identified as White performed better than those who identified as Black/African American or Hispanic; and Hispanic participants performed better than Black/African American participants. The results are shown in Table 7 and Figure 1.

Table 7

Percent Correct Responses to Financial Literacy Questions– By Race/Ethnicity

Item	White/ Caucasian	Black/ African American	Hispanic/ Latino	Asian
Q.1: Interest Rates	76.6	61.9	67.5	76.6
Q.2: Inflation	57.4	35.9	44.4	61.8
Q.3: Bonds/Interest Rates	31.2	18.6	21.1	28.0
Q.4: Compound Interest	36.2	26.6	28.6	31.8
Q.5: Interest Expense	70.8	58.9	67.0	79.1
Q.6: Investment Risk	48.7	28.2	34.1	51.0

Differential Item Functioning Analysis by Race/Ethnicity

To examine the potential bias of the financial literacy questions by race/ethnicity, three focal groups were identified: Asian, Hispanic/Latino, and Black/African American participants. The reference group included the White/Caucasian respondents.

The analysis indicated negligible effect sizes for all items across races, except two items. Specifically, items two and five showed moderate effect sizes for the Black/African American focal group (β = -.08 and -.06, respectively). A summary of the results is given in Table 8.

Table 8

Differential Item Functioning Analysis by Race/Ethnicity

	Black/African American (β)	Hispanic/Latino (β)	Asian (β)
Item #1	-0.01	-0.01	-0.02
Item #2	-0.08*	0.00	-0.02
Item #3	0.03	0.00	0.00
Item #4	0.05	0.00	0.01
Item #5	-0.06*	0.03	0.03
Item #6	0.01	0.01	0.01

Note. β <..059 (negligible) *β = .059-.088 (moderate) **β >.088 (large)

The following chapter will provide further comments on the study's findings and potential implications within the framework of the literature review and the research questions.

CHAPTER 5: DISCUSSION

This study examined the validity and fairness of financial literacy assessments, specifically whether test score interpretations and uses may be unfair due to measurement bias. This quantitative research study sought to examine the extent to which financial literacy assessment in America is affected by bias due to group membership. In a national financial literacy study, item bias was analyzed as a potential indicator of construct irrelevant variance (CIV).

The general research question addressed by this study was: To what extent does item bias affect the validity of the six financial literacy assessment items used on the 2018 National Financial Capability Study (NFCS)? Using Differential Item Functioning (DIF) analysis, the study analyzed the items in the 2018 NFCS for indications of item bias and, thereby, the cogency of the interpretations and uses for groups identified by age, race, and gender. The national survey data included responses from a representative sample of 27,092 adults ages 18 to 101 from all fifty states. The study investigated the following research questions:

R.Q. 1. To what extent do the financial literacy items on the 2018 NFCS exhibit reliability?

R.Q. 2. To what extent do the financial literacy items on the 2018 NFCS exhibit validity (measure a common construct)?

R.Q. 3. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by age?

R.Q. 4. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by gender?

R.Q. 5. To what extent do the financial literacy items on the 2018 NFCS function differently across groups identified by race?

Based on data analysis, the researcher concluded the following regarding the hypotheses set forth in the study:

HO(1). The items on the 2018 NFCS exhibit appropriate levels of reliability.

HO(2). The items on the 2018 NFCS measure a common construct.

HO(3). The financial literacy items on the 2018 NFCS do not function differently across groups identified by age.

HO(4). The financial literacy items on the 2018 NFCS do not function differently across groups identified by gender, and

HO(5). The financial literacy items on the 2018 NFCS do not function differently across groups identified by race.

Summary of Findings

HO(1) and HO(2): The items on the 2018 NFCS exhibit appropriate levels of reliability and measure a common construct.

The Cronbach's α for the current items was calculated as .661. This result suggests relatively low internal consistency and indicates that the items fail to measure similar constructs at an appropriate level of reliability to provide validity evidence (Taber, 2018). Furthermore, the Pearson's *r* results indicated that the items had a relatively low correlation with one another, further suggesting that these six items may not be measuring "financial literacy" as a singular construct. These findings support rejecting the null hypotheses for Research Questions one and two.

HO(3). The financial literacy items on the 2018 NFCS do not function differently across groups identified by age.

For the 18-24 age group, item two indicated a moderate DIF effect size (β = .06), while items four and five were found to have large effect sizes (β = -.10 and .09, respectively). In the 65+ age group, items two and four indicated large DIF effect sizes (β = -.13 and 0.11, respectively). Large effect sizes for these age groups provide compelling evidence that these items are biased against these groups and should be reconsidered for use in future assessments. Therefore, these findings support the rejection of the null hypothesis for Research Question three.

HO(4). The financial literacy items on the 2018 NFCS do not function differently across groups identified by gender.

A moderate DIF effect size was indicated for item five (β = .07) in favor of the participants in the focal group (females). While this finding does indicate item bias within the gender group, it alone may not be sufficient evidence that the item is invalid. However, it does suggest that a review of the item is warranted. The null hypothesis is therefore supported.

HO(5). The financial literacy items on the 2018 NFCS do not function differently across groups identified by race.

Items two and five showed moderate effect sizes for the Black/African American focal group (β = -.08 and -.06, respectively); again, this finding may not be sufficient evidence that these items are invalid, but it does lend credence to the idea that these items should be reviewed. The null hypothesis is therefore supported.

Discussion

The results of this study indicate that the items that were used to assess financial literacy on the 2018 NFCS were lacking in their ability to measure a common construct, and that three of the six items were flagged for DIF, indicating potential item bias.

However, these three items – although they may have been flagged for DIF in this study - are not necessarily *invalid*. To fully establish item validity using DIF analysis, a review of the items by a subject-matter expert is required. The review is intended to provide professional insight into the possible explanations for why the items may be functioning differently across groups.

Items that indicate DIF can still be valid if the bias is determined to be the result of true differences between the groups. As a subject-matter expert in personal finance who has been teaching financial literacy concepts for almost three decades, the author offers the following review regarding the study results. My approach will be to look at the individual items flagged for DIF, the magnitude of DIF for each group, and potential explanations for each.

As mentioned previously, DIF analysis was conducted on each of the six items to identify potential bias by participant age, race, and gender. Three of the six items were flagged for DIF: Items two, four, and five. The flagged items and their respective results are given below.

Item 2: Interest and Inflation

Q: Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

A. More than today

B. Exactly the same

C. Less than today

D. Don't know

E. Prefer not to say

Item 4: ***Interest Rates and Bond Prices***

Q: If interest rates rise, what will typically happen to bond prices?

A) They will rise

B) They will fall

C) They will stay the same

D) There is no relationship between bond prices and the interest rate

E) Don't know

F) Prefer not to say

Item 5: Mortgage Rates and Interest Expense

Q: A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.

A) True

B) False

C) Don't know

D) Prefer not to say

Item Two

Item two sought to measure the respondent's level of knowledge regarding the relationship between interest rates and inflation. Within the age group, the item exhibited large DIF for those aged 65+ (β = -013) yet also indicated moderate DIF in favor of those ages 18-24

(β = 0.06). This finding seems counterintuitive. Why would the youngest group perform better on this item?

The literature is mixed about the level of financial literacy between age groups. For example, a study by Yakoboski et al. (2019) found that financial knowledge was correlated with age group. The youngest group (ages 18-29) was the least financially literate. Those ages 30 to 44 showed improved financial knowledge, but it was not statistically significant. In contrast, participants ages 45 to 59 did show a statistically significant increase in their financial literacy score, and adults over age 60 showed even higher, statistically significant levels of financial literacy. However, earlier findings indicate that financial literacy scores tend to increase with age until age 60, when they begin decreasing (Lusardi & Mitchell, 2011). Nonetheless, the literature does not suggest that the 18 – 24-year-old age group, given their general aptitude in financial literacy, is more likely to answer this question correctly than other age groups. Yet this is what the DIF analysis suggests.

Conflicting results within the literature and the findings of the current study may suggest that a review of this item is needed, especially when combined with the additional finding that moderate DIF was indicated for this item for those identifying as Black or African American (β= -0.08).

Item Four

Item four sought to measure the respondent's level of knowledge regarding the relationship between interest rates and bond prices. The responses to item four in the current study reflected opposing results to those of item two for the focal groups within the age category. Item four indicated large DIF in favor of the 65+ age group (β=*0.11*) but against the 18-24 age

group (β= *-0.10*). This finding further underscores the need to look at both item two and item four as potentially biased questions, particularly for the age groups.

Item four and item two both require an understanding of interest rates. However, item two, which favored the younger group, required additional understanding regarding inflation, and item four, which favored the oldest group, required knowledge of bond rates. One simple explanation for this finding is that the concept of inflation is being taught in schools more often than it was one or two generations ago. That would indicate that financial education is working and identifies a true difference between age groups.

Item four, which favored the oldest group, required additional knowledge of bond rates. In this case, a plausible explanation is that life experience, which comes with age, is a determining factor in this result. That is, the older respondents were more likely to have had experience with bonds as investments. Again, this reasoning identifies a true difference between age groups. Generally speaking, older individuals have more experience in financial matters. And younger individuals are likely to have been taught about inflation in school.

Item Five

Item five sought to measure the respondent's level of knowledge regarding the relationship between mortgage rates and interest expenses. This item exhibited large DIF in favor of respondents aged 18-24 (β = .09). Moderate DIF was found to favor women (β = 0.07) and against respondents who identified as Black or African American (β = -0.06).

Given that this item was flagged for DIF in all three categories, it is reasonable to suggest that this item be closely examined as an appropriate measure of financial literacy. However, an explanation for the bias is less straightforward. The large DIF indicated in favor of the 18-24 focal group is another counterintuitive finding. Perhaps, again, this result can be explained by the

fact that this concept has recently begun to be taught in schools (high school or college). Alternatively, perhaps this group of young adults learned about this topic through experience, as they were a cohort that came of age during and after the Mortgage Crisis that began in 2008. The literature does not offer an explanation, because the actual score results of this item are clear: younger and older Americans do not perform as well as those ages 25 – 64 on this question. Yet the DIF findings indicate bias within this item; other factors may be involved.

In addition, a moderate level of DIF was found in item five favoring the female focal group. However, no other items in the gender category were flagged for DIF, and females were found to perform less well than males on all items. Therefore, item five also stands out as an anomaly: what is different about this question that would lead to these results?

Item five sought to measure the respondent's understanding of the cost of interest to a borrower over time – specifically, mortgage interest. The literature on this topic is mixed as well. Research has found that men indicate higher levels of financial literacy than women (Bucher-Koenen et al., 2017; Chen & Volpe, 2002; Klapper et al., 2015; Lusardi & Mitchell, 2007; Woodyard & Robb, 2012). The literature has further indicated that women self-report lower confidence and less interest in learning about personal finance topics than men (Bucher-Koenen et al., 2017; Chen & Volpe, 2002; Lusardi & Mitchell, 2011). However, a conflicting study by Brown & Graf (2013) found that gender differences were not due to the level of interest in finance and financial matters between women and men.

Several researchers (Danes & Yang, 2014; Fonseca et al., 2012; Hsu, 2016; Woodyard & Robb, 2012) have suggested that gender differences are due to the traditional roles of women and men in households. This theory is plausible, at least to the extent that it relates to this item,

as women are often the ones to pay the day-to-day bills in a traditional household and may therefore be more aware of mortgage expenses.

However, this explanation cannot account for the full extent of bias that is indicated for this item. This is because the demographic makeup of American households has changed drastically over the past several decades, and the traditional family model as a percentage of US households continues to decline rapidly. In this case, other factors are most certainly at play, yet the literature offers little insight.

Moderate DIF was also indicated on item 5 for respondents who identified as Black/African American (β = -0.06). Though this level of DIF is relatively small, it is the second item to indicate DIF (along with item 2) within this focal group. Also, no other focal group in the race category was flagged for DIF on any item. These results suggest that a reasonable explanation for potential item bias related to this group should be considered.

Studies have consistently indicated that African American adults have lower financial literacy scores than white adults. As noted previously, much of the financial literacy literature -- based on studies using the "Big Three" or "Big Five" assessment measures -- suggests that lower financial literacy scores among African American individuals are due to a lack of knowledge and that more financial education is needed to solve the problem (Al-bahrani et al., 2018; Hilgert et al., 2003; Lusardi & Mitchell, 2007).

The current study's theoretical framework would suggest that the focus on a single individual characteristic – knowledge – fails to consider external factors that may have affected their response to this item, which measured knowledge about mortgages and interest expense. This study was grounded in educational theory, specifically Constructivist theory and self-regulated learning, which hold that learning is based on both internal and external factors. The

study further draws on Dewey's Pragmatist theories, which propose that transactions between teacher and learner, and the experiences between an individual and their world, are critical to educational development. The premise of this study is that individuals learn about finances both cognitively *and* socially, and that knowledge is *contextual*.

The research literature on the topic of 'mortgage interest rates' in general suggests that disturbing external factors do exist and can have a negative effect on individuals' financial knowledge and behaviors. The Black/African American community has been identified as one group that has experienced negative external influences in the mortgage lending process.

For example, Perry (2019) asserted that the 2008 mortgage crisis resulted from inappropriate investment risk in subprime products by mortgage lenders and by borrowers who were victimized by "the fraudulent, irresponsible, and predatory lending practices in this market (p. 175)," particularly African American homebuyers. An example of this is reflected in a report by the State of Illinois, which studied over 82,000 mortgage loans that originated between 2005 and 2007 and found that African American borrowers were three times as likely to be put into a subprime loan than white borrowers in a comparable credit situation. In addition, in 2013, the US Justice Department found several mortgage lending companies (including Wells Fargo, Bank of America, and Countrywide) liable for illegally putting certain groups of borrowers into subprime mortgages or charging higher fees and rates than they did white borrowers because of their race or national origin.

Another example is Bayer et al. (2016), which had similar findings in a study that reviewed the "borrowing side" of the 2008 housing crisis. The study found that African American households were "more likely to become delinquent and default on their mortgages than white households with similar credit scores, house type, neighborhood, and loan

characteristics, especially for mortgages originated for new home purchases in 2005 -2006 (p. 5). In addition, the study found that while participation in subprime mortgages may have impacted these findings to some degree, the more salient factors included assuming a mortgage contract under a high debt-to-income ratio and having greater vulnerability to volatile employment markets and declining home prices.

Unscrupulous mortgage lending and systemic vulnerability to economic loss may be the contextual, socioeconomic factors affecting this group's "knowledge" of financial literacy concepts. Indeed, the St. Louis Federal Reserve published a study in 2015 that found that specific demographic characteristics—being young, less educated, and non-white – were significantly correlated with delinquent mortgage payments and concluded that these individuals "did not choose and cannot readily change these characteristics, so we should refrain from adding insult to injury by suggesting that they simply have brought financial problems on themselves by making risky choices" (Emmons et al., 2016)

Unfortunately, public sentiment toward financial literacy in America is underscored by a general belief that poor financial circumstances reflect individual shortcomings and that anyone can be "successful" if they try hard enough. External factors (e.g., educational opportunities and socioeconomic status) are not considered relevant.

The reporting of test results based on the items examined in this study continues to implicate certain groups as lacking in financial ability. Consequently, public policy towards - and perhaps more importantly, the public opinion of – these groups is greatly affected.

Conclusions and Recommendations for Future Research

This study hypothesized that one or more of the six most ubiquitous questions used to measure the construct might lack validity evidence due to construct-irrelevant sources

of variance (item bias). Specifically, through Differential Item Functioning (DIF), the study analyzed the validity of six commonly used survey questions in the 2018 National Financial Capability Study (NFCS) and, thereby, the cogency of the resulting outcomes and interpretations for groups identified by age, race, and gender.

The *Standards* (American Educational Research Association et al., 2014) explicitly state that test instruments must measure what they intend to measure to ensure valid inferences are drawn from test scores. Modern test theory conceptualizes validity as a process through which test developers establish an evidence-based argument for score interpretations (M. Kane, 2013; Messick, 1995). In this argument-based approach, multiple forms of evidence are used to substantiate claims within the validity argument, thereby determining whether the scores reflect an accurate measure of the underlying construct. However, the literature on financial literacy assessments does not speak to validity evidence.

The current study employed Differential Item Functioning (DIF) analysis to establish validity evidence. DIF may indicate construct-irrelevant variance (item bias), suggesting that a question is invalid. Based on Kane's Interpretive Use model (Kane, 2013), when such construct- irrelevant variance leads to inaccurate inferences or inappropriate score uses, the item is considered invalid.

Lusardi & Mitchell (2011, 2014) developed these six questions, which measure only financial knowledge. As discussed, numerous inferences have been made regarding the results of studies that have used these items, and national policies and resource appropriations have been made based on these inferences. Lusardi & Mitchell (2014) reviewed the financial literacy literature and compared studies based on their own "Big

Three" assessment items. The authors stated that "endogenizing financial knowledge has important implications for welfare as well as policies intended to enhance levels of financial knowledge in the larger population." This statement not only underscores the role of these studies as determinants of public policy but also the significance of knowledge – an endogenous factor – having a primary role in financial literacy. This belief has prevailed within the discipline for decades.

As I put my final thoughts on paper to conclude the current study, I have found that the winds may be shifting in a new direction. In a very recent working paper, Mitchell & Lusardi (2022) indicate a concern for the accuracy of financial literacy assessments and the inferences made from them, particularly in terms of exogenous as well as endogenous factors:

> Examining the effects of financial literacy on behavior has several potential limitations, particularly when examining older respondents. One important issue is whether financial literacy is an endogenously determined variable…Addressing causality is difficult because many data sets often do not provide information on, for example, exogenous sources of variation in financial literacy such as access to financial literacy in school or the workplace, or the opening of a new university where young people could learn about financial literacy or changes in the law affecting financial decisions. A second concern is whether people simply learn by doing: that is, by making many financial decisions about saving and debt, they may also accumulate some financial savvy. A third concern is whether and how well financial literacy is measured, how many questions are needed to

> measure financial literacy…and how important is measurement error. A fourth concern is that psychological biases may affect financial behavior including confidence, self- awareness, and so on (p. 11).
>
> The authors also state that:
>
> We recognize that…people may also suffer from behavioral biases, face complex financial decisions when information is "shrouded," and [can be] subject to sophisticated scams, just to mention a few additional problems related to the economic environment or the way people approach financial decisions (p. 12).

This change of perspective from the authors who developed and promoted the knowledge-as-literacy argument may do as much – or more - to support this study's original premise as the findings herein and additional insights from the literature.

Indeed, Mitchell & Lusardi (2022) also provide further support to the suggestion made in this paper that a theoretical context would provide valuable structure to a financial literacy study. As an example, Mitchell and Lusardi cite Lusardi et al. (2017), which examined the impact of financial literacy on a specific behavioral outcome using an "intemporal model of saving" in which "financial literacy is characterized as a form of human capital that people accumulate to be able to invest in high-return assets." It is notable, however, that the theoretical application presumes an endogenous basis for the level of financial literacy and subsequent behavioral outcome.

Nonetheless, the application of theory can be helpful by offering explanations for results – especially anomalous results, as in the current study. For example, the responses to item four in the current study reflected opposing results to those of item two for the

focal groups within the age category. These items required an understanding of interest rates. However, item two, which favored the younger group, required additional understanding regarding inflation, and item four, which favored the oldest group, required knowledge of the relationship between bond values and interest rates. As Mitchell & Lusardi (2022) have suggested, this may indicate that experience – and not just age – is a confounding factor for these two questions.

Aligned with a framework based on constructivist thought and self-regulated learning (SRL) theory, this study has argued that America's current most popular financial literacy measures may be biased due to an omission of consideration for social and cultural factors that inherently affect individuals' financial knowledge, attitudes, and behavior. The premise of this study is that financial education and assessment should incorporate both factors because individuals learn about money both cognitively and socially. Unfortunately, financial literacy research and assessment in the U.S. has primarily focused on personal factors and have largely ignored the reasons for socio-cultural differences in financial outcomes (Danes & Yang, 2014; Gudmunson & Danes, 2011). Using the current theoretical context, fewer correct responses given by certain groups may reflect contextual, socio-cultural factors irrelevant to the construct of financial literacy, such as the attitudinal and belief domains of the social contexts within which individuals are socialized about finances (Gudmunson & Danes, 2011). The results of this study indicate that three of the six items analyzed in the current study may be biased. This suggests that the items are questionable and should be reviewed by test developers and users before they are included in an assessment instrument. In addition, a potentially more significant problem may be the use of the results of financial literacy

studies that have included these items as a basis to make inferences, influence public policy, and perhaps contribute to systemic biases and stigma regarding financial education and literacy in the United States.

Additional research is needed to establish validity evidence for these six ubiquitous items and financial literacy assessment instruments in general. Such evidence will require a purposeful effort to 1.) standardize the definition of financial literacy, 2.) establish guidelines for the interpretation and use of the results, and 3.) develop assessment instruments that accurately measure the construct as it has been defined.

In addition, applying a theoretical framework to guide financial literacy research is a prerequisite to establishing Financial Education and Literacy as an academic discipline. The inclusion of theoretical underpinnings could also provide a basis for creating a cohesive definition of the construct of "financial literacy" and developing guidelines for identifying appropriate interpretations and uses of research findings. Finally, identifying one or more theoretical frameworks, coupled with a standard definition of "financial literacy," could provide a strong foundation for developing effective assessment instruments.

REFERENCES

Al-bahrani, A., Buser, W., & Patel, D. (2018). Does Math Confidence Matter? How Student Perceptions Create Barriers to Success in Economics Classes. *Journal of Economics and Finance Education*, *17*(1), 61–77. https://econ.gatech.edu/publications/pub/6239

Al-Bahrani, A., Weathers, J., & Patel, D. (2018). Racial Differences in the Returns to Financial Literacy Education. *Journal of Consumer Affairs*, *53*(2), 572–599. https://doi.org/10.1111/joca.12205

Allgood, S., & Walstad, W. (2013). Financial Literacy and Credit Card Behaviors: A Cross-Sectional Analysis by Age. *Numeracy*, *6*(2). https://doi.org/10.5038/1936-4660.6.2.3

Allgood, S., & Walstad, W. B. (2016). The effects of perceived and actual financial literacy on financial behaviors. *Economic Inquiry*, *54*(1), 675–697. https://doi.org/10.1111/ecin.12255

Alsemgeest, L. (2015). Arguments for and against financial literacy education: Where to go from here? *International Journal of Consumer Studies*, *39*(2), 155–161. https://doi.org/10.1111/ijcs.12163

Amagir, A., Groot, W., Maassen van den Brink, H., & Wilschut, A. (2018). A review of financial-literacy education programs for children and adolescents. In *Citizenship, Social and Economics Education* (Vol. 17, Issue 1, pp. 56–80). https://doi.org/10.1177/2047173417719555

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing.* American Educational Research Association.

Anderson, A., Baker, F., & Robinson, D. T. (2017). Precautionary savings, retirement planning and misperceptions of financial literacy. *Journal of Financial Economics*, *126*(2), 383–398. https://doi.org/10.1016/j.jfineco.2017.07.008

Aren, S., & Aydemir, S. D. (2014). A Literature Review on Financial Literacy. *Finansal Araştırmalar ve Çalışmalar Dergisi*, *11*(5), 33–49. https://www.researchgate.net/publication/267641595

Banks, J., & Oldfield, Z. (2007). Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving. *Fiscal Studies*, *28*(2), 143–170. https://doi.org/10.1111/j.1475-5890.2007.00052.x

Baumeister, R. F., Sparks, E. A., Stillman, T. F., & Vohs, K. D. (2008). Free will in consumer behavior: Self-control, ego depletion, and choice. *Journal of Consumer Psychology*, *18*(1), 4–13. https://doi.org/10.1016/j.jcps.2007.10.002

Bayer, P., Ferreira, F., & Ross, S. L. (2016). The Vulnerability of Minority Homeowners in the Housing Boom and Bust. *American Economic Journal: Economic Policy*, *8*(1), 1–27. https://doi.org/10.1257/pol.20140074

Bernheim, B. D., Garrett, D. M., & Maki, D. M. (2001). Education and saving: The long-term effects of high school financial curriculum mandates. *Journal of Public Economics*, *80*(3), 435–465. https://doi.org/10.1016/S0047-2727(00)00120-1

Bredo, E. (1998). Evolution, Psychology, and John Dewey's Critique of the Reflex Arc Concept. *The Elementary School Journal*, *98*(5), 447–466. https://doi.org/10.1086/461908

Brown, M., & Graf, R. (2013). Financial literacy and retirement planning in Switzerland. *Numeracy*, *6*(2). https://doi.org/10.5038/1936-4660.6.2.6

Bucher-Koenen, T., Lusardi, A., Alessie, R., van Rooij, V., & Maarten, C. J. (2017). How financially literate are women? An overview and new insights. *Journal of Consumer Affairs*, *51*(2), 255–283. https://doi.org/10.1111/joca.12121

Buzick, H., & Stone, E. (2011). *Recommendations for Conducting Differential Item Functioning (DIF) Analyses for Students With Disabilities Based on Previous DIF Studies* (Issue August). https://doi.org/10.1002/j.2333-8504.2011.tb02270.x

Cameron, M. P., Calderwood, R., Cox, A., Lim, S., & Yamaoka, M. (2014). Factors associated with financial literacy among high school students in New Zealand. *International Review of Economics Education*, *16*(PA), 12–21. https://doi.org/10.1016/j.iree.2014.07.006

Camilli, G. (2013). Ongoing issues in test fairness. *Educational Research and Evaluation*, *19*(2–3), 104–120. https://doi.org/10.1080/13803611.2013.767602

Campbell, J. Y. (2006). Household finance. *Journal of Finance*, *61*(4), 1553–1604. https://scholar.harvard.edu/campbell/publications/household-finance-presidential-address-american-finance-association

CFPB. (2015). *Financial well-being: The goal of financial education* (Issue January). https://files.consumerfinance.gov/f/201501_cfpb_report_financial-well-being.pdf

CFPB. (2017). *Effective financial education: Five principles and how to use them*. https://files.consumerfinance.gov/f/documents/201706_cfpb_five-principles-financial-well-being.pdf

Chalhoub-Deville, M. (2016). Validity theory: Reform policies, accountability testing, and consequences. *Language Testing*, *33*(4), 453–472. https://doi.org/10.1177/0265532215593312

Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*. https://doi.org/10.18637/jss.v048.i06

Chalmers, R. P., Counsell, A., & Flora, D. B. (2016). It Might Not Make a Big DIF: Improved Differential Test Functioning Statistics That Account for Sampling Variability. *Educational and Psychological Measurement*, *76*(1), 114–140. https://doi.org/10.1177/0013164415584576

Chen, H., & Volpe, R. (2002). Gender Differences in Personal Financial Literacy among College Students. *Financial Services Review*, *11*(3), 289. https://www.researchgate.net/publication/285358406_Gender_Differences_in_Personal_Financial_Literacy_Among_College_Students

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *American Journal of Medicine*, *119*(2), 166.e7-166.e16. https://doi.org/10.1016/j.amjmed.2005.10.036

Danes, S. M. (1994). Parental perceptions of children's financial socialization. *Journal of Financial Counseling and Planning*, *5*(1), 127–149. https://www.afcpe.org/news-and-publications/journal-of-financial-counseling-and-planning/volume-5/%EF%BB%BFparental-perceptions-of-childrens-financial-socialization/

Danes, S. M., & Haberman, H. R. (2007). Teen Financial Knowledge, Self-Efficacy, and Behavior: A Gendered View. *Journal of Financial Counseling and Planning*, *18*(2), 48–60. https://ssrn.com/abstract=2228406

Danes, S. M., & Yang, Y. (2014). Assessment of the use of theories within the journal of financial counseling and planning and the contribution of the family financial socialization conceptual model. *Journal of Financial Counseling and Planning*, *25*(1), 53–68. https://ssrn.com/abstract=2466560

Danes, S.M., Deenanath, V., & Yang, Y. (2016). Evaluation of Financial Literacy Development of High School Students: A Longitudinal Mixed Methods Study. In E. Aprea & & J. S. L. Wuttke, B. Klaus, N. K. Koh, P. Davies, B. Greimel-Fuhrmann (Eds.), *International Handbook of Financial Literacy* (pp. 425–447). Springer.

Daniels, H. (2016). Vygotsky and dialogic pedagogy. In *Dialogic pedagogy: The importance of dialogue in teaching and learning* (Vol. 10, Issue 3, pp. 48–67). https://doi.org/10.21832/9781783096220-004

Davies, P. (2015). Towards a framework for financial literacy in the context of democracy. *Journal of Curriculum Studies*, *47*(2), 300–316. https://doi.org/10.1080/00220272.2014.934717

Delgadillo, L. M. (2014). Financial Clarity: Education, literacy, capability, counseling, planning, and coaching. *Family and Consumer Sciences Research Journal*, *43*(1), 18–28. https://doi.org/10.1111/fcsr.12078

Dewey, J. (1896). The Reflex Arc Concept in Psychology. *Psychological Review*, *3*, 357–370. https://brocku.ca/MeadProject/Dewey/Dewey_1896.html

Dewey, J. (1922). *HUMAN NATURE AND CONDUCT An Introduction to Social Psychology*. Henry Holt and Company. https://www.gutenberg.org/files/41386/41386-h/41386-h.htm

Dewey, J. (1938). *EXPERIENCE & EDUCATION* (A. L. Hall-Quest, Ed.; 1997th ed.). Kappa Delta Pi.

Dewey, J., & Bentley, A. F. (1949). *Knowing and the Known*. Praeger.

Dogra, P., Kaushal, A., & Sharma, R. R. (2021). Antecedents of the Youngster's Awareness About Financial Literacy: A Structure Equation Modelling Approach. *Vision*. https://doi.org/10.1177/0972262921996560

Driscoll, M. P. (1994). Psychology of learning for instruction. In *Psychology of learning for instruction.* Allyn & Bacon.

Dunst, C. J., & Hamby, D. W. (2012). Guide for calculating and interpreting effect sizes and confidence intervals in intellectual and developmental disability research studies. *Journal of Intellectual and Developmental Disability*, *37*(2), 89–99. https://doi.org/10.3109/13668250.2012.673575

Edwards, R., Allen, M. W., & Hayhoe, C. R. (2007). Financial attitudes and family communication about students' finances: The role of sex differences. *Communication Reports*, *20*(2), 90–100. https://doi.org/10.1080/08934210701643719

Emmons, B. W. R., Ricketts, L. R., & Pettigrew, T. (2016, August). Choosing to Fail or Lack of Choice? The Demographics of Loan Delinquency. *In The Balance*, *15*, 1–3. https://www.stlouisfed.org/publications/in-the-balance/2016/choosing-to-fail-or-lack-of-choice-the-demographics-of-loan-delinquency

Fan, L., & Chatterjee, S. (2019). Financial Socialization, Financial Education, and Student Loan Debt. *Journal of Family and Economic Issues*, *40*(1), 74–85. https://doi.org/10.1007/s10834-018-9589-0

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, *60*(8), 1861–1883. https://doi.org/10.1287/mnsc.2013.1849

Finke, M. S., Howe, J. S., & Huston, S. J. (2017). Old age and the decline in financial literacy. *Management Science*, *63*(1), 213–230. https://doi.org/10.1287/mnsc.2015.2293

FLEC. (2010). *Why and How: Background Report Developing the 2011 National Strategy* (Issue November 2010). https://home.treasury.gov/system/files/231/National-Startegy-Background.pdf

FLEC. (2012). *2012 Research Priorities and Research Questions*. https://home.treasury.gov/system/files/231/2012-Research-Priorities-May-12.pdf

Fonseca, R., Mullen, K. J., Zamarro, G., & Zissimopoulos, J. (2012). What Explains the Gender Gap in Financial Literacy? The Role of Household Decision Making. *Journal of Consumer Affairs*, *46*(1), 90–106. https://doi.org/10.1111/j.1745-6606.2011.01221.x

Gamble, K. J., Boyle, P. A., Yu, L., & Bennett, D. A. (2015). *How does aging affect financial decision making?* (Issue 15). http://crr.bc.edu/wp-content/uploads/2015/01/IB_15-1-508.pdf

Gerardi, K., Goette, L., & Meier, S. (2010). *Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data*. https://www.atlantafed.org/-/media/documents/research/publications/wp/2010/wp1010.pdf

Gierl, M. J., & Khaliq, N. S. (2001). Identifying Sources of Differential Item and Bundle Functioning on Translated Achievement Tests: A Confirmatory Analysis. *Journal of Educational Measurement*, *38*(2), 164–187. https://www.jstor.org/stable/1435261

Gierl, M., Rogers, W. T., & Klinger, D. A. (1999). Using Statistical and Judgmental Reviews to Identify and Interpret Translation Differential Item Functioning. *The Alberta Journal of Educational Research*, *45*(4), 353–376. https://doi.org/10.11575/ajer.v45i4.54709

Gómez-Benito, J., Sireci, S., Padilla, J.-L., Hidalgo, M. D., & Benítez, I. (2018). Differential Item Functioning: Beyond validity evidence based on internal structure. *Psicothema*, *30*(1), 104–109. https://doi.org/10.7334/psicothema2017.183

Goyal, K., & Kumar, S. (2021). Financial literacy: A systematic review and bibliometric analysis. In *International Journal of Consumer Studies* (Vol. 45, Issue 1, pp. 80–105). Blackwell Publishing Ltd. https://doi.org/10.1111/ijcs.12605

Grinblatt, M., Keloharju, M., & Linnainmaa, J. (2009). Do smart investors outperform dumb investors? In *Chicago Booth Research Paper* (Vol. 9, Issue 33). http://www.econ.yale.edu/~shiller/behfin/2009_11/grinblatt-keloharju-linnainmaa.pdf

Gudmunson, C. G., & Danes, S. M. (2011). Family Financial Socialization: Theory and Critical Review. *Journal of Family and Economic Issues*, *32*(4), 644–667. https://doi.org/10.1007/s10834-011-9275-y

Gupta, R., Hasler, A., Lusardi, A., & Oggero, N. (2018). Financial fragility in the US: evidence and implications. *Global Financial Literacy Excellence Center, The George Washington University School of Business: Washington, DC*. https://www.nefe.org/_images/research/Financial-Fragility/Financial-Fragility-Final-Report.pdf

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013). FINANCIAL LITERACY, FINANCIAL EDUCATION AND ECONOMIC OUTCOMES. *Annu Rev Econom.*, *May*(5), 347–373. https://doi.org/10.1146/annurev-economics-082312-125807.FINANCIAL

Hilgert, M., Hogart, J., & Beverly, S. (2003). *Household Financial Management: The Connection between Knowledge and Behavior*. https://www.federalreserve.gov/pubs/bulletin/2003/0703lead.pdf

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129–145). Lawrence Erlbaum Associates, Inc.

Howlett, E., Kees, J., & Kemp, E. (2008). The role of self-regulation, future orientation, and financial knowledge in long-term financial decisions. *Journal of Consumer Affairs*, *42*(2), 223–242. https://doi.org/10.1111/j.1745-6606.2008.00106.x

Hsu, J. W. (2016). Aging and Strategic Learning: The Impact of Spousal Incentives on Financial Literacy. *Journal of Human Resources*, *51*(4), 1036–1067. https://doi.org/10.3368/jhr.51.4.1014-6712R

Hubley, A. M., & Zumbo, B. D. (2011). Validity and the Consequences of Test Interpretation and Use. In *Social Indicators Research* (Vol. 103, Issue 2, pp. 219–230). https://doi.org/10.1007/s11205-011-9843-4

Huitt, W., & Hummel, J. (2006). Piaget's Theory of Cognitive Development. In *Educational Psychology Interactive*. http://www.edpsycinteractive.org/topics/cognition/piaget.html

Hung, A. A., Parker, A. M., Yoong, J. K., & Yoong, J. (2009). *Defining and Measuring Financial Literacy* (Labor and Population). https://www.rand.org/content/dam/rand/pubs/working_papers/2009/RAND_WR708.pdf

Huston, S. J. (2010). Measuring Financial Literacy. *Journal of Consumer Affairs*, *44*(2), 296–316. https://doi.org/10.1111/j.1745-6606.2010.01170.x

Ibrahim, A.-W., Rufus, E., & Afolabi, I. (2021). A Comparison of Performance of Two Statistics for Detecting Differential Item Functioning in Two Test Formats. *Journal of Applied Research in Education*, *7*(1). https://www.researchgate.net/publication/350950796

Jiang, H., & Stout, W. (1998). Improved Type I Error Control and Reduced Estimation Bias for DIF Detection Using SIBTEST. *Journal of Educational and Behavioral Statistics*, *23*(4), 291–322. https://doi.org/10.2307/1165279

John Dewey. (1910). *How We Think* (D. C. Heath, Ed.).

Johnson, E., & Sherraden, M. S. (2007). From financial literacy to financial capability among youth. *Journal of Sociology and Social Welfare*, *34*(3), 119–145. https://www.researchgate.net/publication/261174620

Jump$tart Coalition. (2017). *National Standards in K-12 Personal Finance Education*. https://www.jumpstart.org/wp-content/uploads/2018/01/2017_NationalStandardsBook.pdf

Kaiser, T., & Menkhoff, L. (2020). Financial education in schools: A meta-analysis of experimental studies. *Economics of Education Review*, *78*(101930). https://doi.org/10.1016/J.ECONEDUREV.2019.101930

Kane, M. (2013). The Argument-Based Approach to Validation. *School Psychology Review*, *42*(4), 448–457. https://doi.org/10.1080/02796015.2013.12087465

Kane, M. T. (1990). *An Argument-based Approach to Validation*. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR90-13.pdf

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, *112*(3), 527–535. https://doi.org/10.1037/0033-2909.112.3.527

Klapper, L., Lusardi, A., & van Oudheusden, P. (2015). *Financial Literacy Around the World: Insights from the Standard & Poor's Ratings Services Global Financial Literacy Survey*. https://gflec.org/wp-content/uploads/2015/11/Finlit_paper_16_F2_singles.pdf

Klein, L., & Mandell, L. (2007). Motivation and financial literacy. *Financial Services Review*, *16*(2), 105.

Kloppenberg, J. (1996). Pragmatism: An old name for some new ways of thinking? *The Journal of American History*, *83*(1), 100–138. https://www.jstor.org/stable/2945476

Knoll, M. A. Z., & Houts, C. R. (2012). The Financial Knowledge Scale: An Application of Item Response Theory to the Assessment of Financial Literacy. *Journal of Consumer Affairs*, *46*(3), 381–410. https://doi.org/10.1111/j.1745-6606.2012.01241.x

Lachance, M. E. (2014). Financial literacy and neighborhood effects. *Journal of Consumer Affairs*, *48*(2), 251–273. https://doi.org/10.1111/joca.12024

Lam, F. (2011). The Socratic Method as an approach to learning and its benefits [Carnegie Mellon University]. In *Dietrich College Honors Theses*. https://doi.org/10.1184/R1/6686570.v1

Lee, J. M., Lee, J., & Kim, K. T. (2019). Consumer Financial Well-Being: Knowledge is Not Enough. *Journal of Family and Economic Issues*, 218–228. https://doi.org/10.1007/s10834-019-09649-9

Leidner, D., & Jarvenpaa, S. (2016). *The Use of Information Technology to Enhance Management School Education: A Theoretical View*. *19*(3), 265–291. https://www.jstor.org/stable/249596

Lin, J. T., Bumcrot, C., Ulicny, T., Mottola, G., Walsh, G., Ganem, R., Kieffer, C., & Lusardi, A. (2019). The state of U.S. financial capability: The 2018 National Financial Capability Study. In *FINRA Investor Education Foundation* (Issue June). https://gflec.org/initiatives/national-financial-capability-study/

Liu, C. C. (2010). Evolution of Constructivism. *Contem Issues in Ed Research*, *3*(4), 63–66. https://doi.org/10.19030/cier.v3i4.199

Lusardi, A. (2015). Financial literacy: Do people know the ABCs of finance? *Public Understanding of Science*, *24*(3), 260–271. https://doi.org/10.1177/0963662514564516

Lusardi, A., Michaud, P. C., & Mitchell, O. S. (2017). Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, *125*(2), 431–477. https://doi.org/10.1086/690950

Lusardi, A., & Mitchell, O. (2007). Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth. *Journal of Monetary Economics*, *54*(1), 205–224. https://doi.org/10.1016/j.jmoneco.2006.12.001

Lusardi, A., & Mitchell, O. S. (2011). Financial Literacy: Implication for Retirement Wellbeing. In *National Bureau of Economic Research*. https://doi.org/10.3386/w17078

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, *52*(1), 5–44. https://doi.org/10.1257/jel.52.1.5

Lusardi, A., Mitchell, O. S., & Oggero, N. (2020). Debt Close to Retirement and Its Implications for Retirement Well-being. In *Remaking Retirement: Debt in an Aging Economy* (pp. 15–34). https://pensionresearchcouncil.wharton.upenn.edu/remaking-retirement-debt-in-an-aging-economy/

Lusardi, A., & Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics and Finance*, *14*(4), 332–368. https://doi.org/10.1017/S1474747215000232

Lyons, A. C., Chang, Y., & Scherpf, E. M. (2006). Translating financial education into behavior change for low-income populations. *Journal of Financial Counseling and Planning*, *17*(2), 27–45. https://ssrn.com/abstract=2232122

Magis, D., Béland, S., Tuerlinckx, F., & de Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods*, *42*(3), 847–862. https://doi.org/10.3758/BRM.42.3.847

Mandell, L. (2008). *THE FINANCIAL LITERACY OF YOUNG AMERICAN ADULTS: Results of the 2008 National Jump$tart Coalition Survey of High School Seniors and College Students*. https://views.smgww.org/assets/pdf/2008%20JumpStart%20Financial%20Literacy%20Survey.pdf

Mantel, N., & Haenszel, W. (1959). Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease. *JNCI: Journal of the National Cancer Institute*, *22*(4), 719–748. https://doi.org/10.1093/jnci/22.4.719

Maranges, H. M., Schmeichel, B. J., & Baumeister, R. F. (2017). Comparing cognitive load and self-regulatory depletion: Effects on emotions and cognitions. *Learning and Instruction*, *51*, 74–84. https://doi.org/10.1016/j.learninstruc.2016.10.010

Martinková, P., Drabinová, A., Liaw, Y.-L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking Equity: Why Differential Item Functioning Analysis Should Be a Routine Part of Developing Conceptual Assessments. *CBE—Life Sciences Education*, *16*(2). https://doi.org/10.1187/cbe.16-10-0307

Mccaslin, M., & Hickey, D. (2001). Self-regulated learning and academic achievement: A Vygotskian view. In *Self-regulated learning and academic achievement: Theory, research, and practice* (pp. 227–252). https://www.researchgate.net/publication/257927378

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, *18*(2), 5–11. https://doi.org/10.3102/0013189X018002005

Messick, S. (1995). Validity of Psychological Assessment. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Mitchell, O. S., & Lusardi, A. (2022). *Financial Literacy and Financial Behavior at Older Ages*. https://repository.upenn.edu/prc_papers

Mitchell, R. E., & Cawthon, S. W. (2015). Academic Achievement: Test Bias and Fairness. In G. Gertz & P. Boudreault (Eds.), *Deaf Studies Encyclopedia*. Sage.

Morgan, D. L. (2014). Pragmatism as a Paradigm for Social Research. *Qualitative Inquiry*, *20*(8), 1045–1053. https://doi.org/10.1177/1077800413513733

Mottola, G. (2013). In Our Best Interest: Women, Financial Literacy, and Credit Card Behavior. *Numeracy*, *6*(2). https://doi.org/10.5038/1936-4660.6.2.4

NFCS. (2016). *Financial Capability in the United States 2016*. https://www.usfinancialcapability.org/downloads/NFCS_2015_Report_Natl_Findings.pdf

Nunnally, J. C., & Bernstein, I. H. (1994). The Assessment of Reliability. In *Psychometric Theory* (pp. 248–292). McGraw-Hill, Inc.

Oliveri, M. E., Ercikan, K., & Zumbo, B. (2013). Analysis of Sources of Latent Class Differential Item Functioning in International Assessments. *International Journal of Testing*, *13*(3), 272–293. https://doi.org/10.1080/15305058.2012.738266

Panadero, E., & Alonso-Tapia, J. (2014). How do students self-regulate? Review of Zimmerman's cyclical model of self-regulated learning. *Anales de Psicologia*, *30*(2), 450–462. https://doi.org/10.6018/analesps.30.2.167221

Perry, V. G. (2019). A Loan at Last? Race and Racism in Mortgage Lending. In *Race in the Marketplace* (pp. 173–192). Springer International Publishing. https://doi.org/10.1007/978-3-030-11711-5_11

Pinto, L., & Coulson, E. (2011). Social Justice and the Gender Politics of Financial Literacy Education. *Journal of the Canadian Association for Curriculum Studies*, *9*(2), 54–85. https://jcacs.journals.yorku.ca/index.php/jcacs/article/view/34299

President's Advisory Council on Financial Literacy. (2008). *2008 Annual Report to the President*. https://www.treasury.gov/about/organizational-structure/offices/Domestic-Finance/Documents/PACFL_Draft-AR-0109.pdf

Richardson, V. (1997). *Constructivist Teacher Education: Building New Understandings*. Falmer Press.

Robb, C. A. (2011). Financial Knowledge and Credit Card Behavior of College Students. *Journal of Family and Economic Issues*, *32*(4), 690–698. https://doi.org/10.1007/s10834-011-9259-y

Robb, C. A., & Woodyard, A. S. (2011). Financial Knowledge and Best Practice Behavior. *Journal of Financial Counseling and Planning*, *22*(1), 60–70. https://files.eric.ed.gov/fulltext/EJ941903.pdf

Roussos, L., & Stout, W. (1996). *A Multidimensionality-Based DIF Analysis Paradigm* (Vol. 20, Issue 4). http://www.copyright.com/

Scheresberg, C. de B. (2013). Financial Literacy and Financial Behavior among Young Adults: Evidence and Implications. *Numeracy*, *6*(2). https://doi.org/10.5038/1936-4660.6.2.5

Scheresberg, C. de B., Lusardi, A., & Yakoboski, P. (2014). Working Women's Financial Capability: An Analysis across Family Status and Career Stages. In *TIAA Institute* (Issue May). https://www.tiaainstitute.org/publication/working-womens-financial-capability-analysis

Shapiro, T., Meschede, T., & Osoro, S. (2013). *The roots of the widening racial wealth gap: explaining the Black-White economic divide* (Issue February). https://heller.brandeis.edu/iere/pdfs/racial-wealth-equity/racial-wealth-gap/roots-widening-racial-wealth-gap.pdf

Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika*, *58*(2), 159–194. https://EconPapers.repec.org/RePEc:spr:psycho:v:58:y:1993:i:2:p:159-194

Sherraden, M. S., Johnson, L., Guo, B., & Elliott, W. (2011). Financial Capability in Children: Effects of Participation in a School-Based Financial Education and Savings Program. *Journal of Family and Economic Issues*, *32*(3), 385–399. https://doi.org/10.1007/s10834-010-9220-5

Shim, S., Xiao, J. J., Barber, B. L., & Lyons, A. C. (2009). Pathways to life success: A conceptual model of financial well-being for young adults. *Journal of Applied Developmental Psychology*, *30*(6), 708–723. https://doi.org/10.1016/j.appdev.2009.02.003

Shockey, S. S., & Seiling, S. B. (2004). Moving Into Action Moving Into Action: Application Of The Transtheoretical Model Of Behavior Change To Financial Education. *Journal of Financial Counseling and Planning*, *15*(1). https://www.afcpe.org/wp-content/uploads/2018/10/vol1515.pdf

Skimmyhorn, W. L. (2016). Comparing Military and Civilian Household Finances: Descriptive Evidence from Recent Surveys. *Journal of Consumer Affairs*, *50*(2), 471–483. https://doi.org/10.1111/joca.12109

Smith, J., McArdle, J., & Willis, R. (2010). Financial Decision Making and Cognition in a Family Context. *The Economic Journal*, *110*(548), F363–F380. https://doi.org/10.1111/j.1468-0297.2010.02394.x

Snowden, K. A. (1987). Mortgage Rates and American Capital Market Development in the Late Nineteenth Century. *Source: The Journal of Economic History*, *47*(3), 671–691. https://www.jstor.org/stable/2121335

Stolper, O. A., & Walter, A. (2017). Financial literacy, financial advice, and financial behavior. *Journal of Business Economics*, *87*(5), 581–643. https://doi.org/10.1007/s11573-017-0853-9

Stoneberg, B. D. (2004). A study of gender-based and ethnic-base differential item functioning (DIF) in the Spring 2003 Idaho Standards Achievement Tests applying the Simulataneous Bias Test (SIBTEST) and the Mantel-Haenszel Chi Square Test. In *Measurement* (Issue March). http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED489949&ERICExtSearch_SearchType_0=eric_accno&accno=ED489949

Strauss, M. E., & Smith, G. T. (2009). Construct Validity: Advances in Theory and Methodology. *Annu Rev Clin Psychol.*, *5*(1), 1–25. https://doi.org/10.1146/annurev.clinpsy.032408.153639

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

U.S. President's Advisory Council on Financial Capability for Young Americans, (2014). https://www.loc.gov/item/lcwaN0016852/

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Vygotsky, L. S. (2012). *Thought and Language*. MIT Press.

Vyvyan, V., Blue, L., & Brimble, M. (2014). Factors that influence financial capability and effectiveness: Exploring financial counsellors' perspectives. *Australasian Accounting, Business and Finance Journal*, *8*(4), 3–22. https://doi.org/10.14453/aabfj.v8i4.2

Warne, R. T., Yoon, M., & Price, C. J. (2014). Exploring the various interpretations of "test bias." *Cultural Diversity and Ethnic Minority Psychology*, *20*(4), 570–582. https://doi.org/10.1037/a0036503

Willis, L. E. (2017). Finance-informed citizens, citizen-informed finance: An essay occasioned by the international handbook of financial literacy. *Journal of Social Science Education*, *16*(4), 16–27. https://doi.org/10.4119/UNIBI/jsse-v16-i4-1761

Woodyard, A., & Robb, C. (2012). Financial Knowledge and the Gender Gap. *Journal of Financial Therapy*, *3*(1), 1–16. https://doi.org/10.4148/jft.v3i1.1453

Xiao, J. J. (2008). Applying Behavior Theories to Financial Behavior. In *Handbook of Consumer Finance Research* (Issue August, pp. 69–81). https://doi.org/10.1007/978-0-387-75734-6_5

Xiao, J. J., O'neill, B., Prochaska, J. M., Kerbel, C. M., Brennan, P., & Bristow, B. J. (2004). A consumer education programme based on the Transtheoretical Model of Change. *International Journal of Consumer Studies*, *28*(January, 2004), 55–65. https://doi.org/10.1111/j.1470-6431.2004.00334.x

Xiao, J. J., Serido, J., & Shim, S. (2012). Financial Education, Financial Knowledge, and Risky Credit Behavior of College Students. In D. J. Lamdin (Ed.), *Consumer Knowledge and Financial Decisions: Lifespan Perspectives* (pp. 113–128). Springer New York. https://doi.org/10.1007/978-1-4614-0475-0_8

Xiao, J. J., Tang, C., Serido, J., & Shim, S. (2011). Antecedents and Consequences of Risky Credit Behavior Among College Students: Application and Extension of the Theory of Planned Behavior. *Journal of Public Policy and Marketing*, *30*(2), 239–245. https://doi.org/10.1509/jppm.30.2.239

Yakoboski, P. J. (2019). *Financial Literacy and Wellness among African-Americans New Insights from the Personal Finance (P-Fin) Index*. https://gflec.org/wp-content/uploads/2019/11/TIAA-Institute-GFLEC_African-American-P-Fin-Index_November-2019.pdf

Yakoboski, P. J., Lusardi, A., & Hasler, A. (2019). *Financial Literacy in the United States and Its Link to Financial Wellness*. https://www.tiaainstitute.org/sites/default/files/presentations/2019-03/TIAA%20Institute-GFLEC_P-Fin%20Index%20Report_April%202019_FINAL.pdf

Zimmerman, B. J. (1989). A Social Cognitive View of Self-Regulated Academic Learning. *Journal of Educational Psychology*, *81*(3), 329–339. https://doi.org/10.1037/0022-0663.81.3.329

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

Zumbo, B. D. (2007). Three Generations of DIF Analyses: Considering Where It Has Been, Where It Is Now, and Where It Is Going. *Language Assessment Quarterly*, *4*(2), 223–233. https://doi.org/10.1080/15434300701375832

Zumbo, B. D., & Wild, W. R. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF) LOGISTIC REGRESSION MODELING AS A UNITARY FRAMEWORK FOR BINARY AND LIKERT-TYPE (ORDINAL) ITEM SCORES.*

APPENDIX I: Author's Correspondence with FINRA

ARC RESEARCH
An SVC Company

FINRA Investor Education FOUNDATION

2018 NATIONAL FINANCIAL CAPABILITY STUDY

STATE-BY-STATE SURVEY METHODOLOGY

The 2018 National Financial Capability Study (NFCS) was funded by the FINRA Investor Education Foundation and conducted by ARC Research (formerly known as Applied Research & Consulting).

Survey Instrument

The survey instrument used in the 2018 NFCS was based on the 2015 questionnaire, which was updated and modified to include input from academics, policy-makers, and researchers who have used the NFCS data from previous years.

Sample

The sample consisted of 27,091 adults (18+) across the U.S., with approximately 500 respondents per state, plus the District of Columbia. To provide additional utility for researchers working with the data, the 2018 NFCS included oversamples in two states, for a total of 1,250 respondents each in OR and WA.

- Respondents were drawn using non-probability quota sampling from established online panels consisting of millions of individuals who have been recruited to join, and who are offered incentives in exchange for participating in online surveys. Specifically, the panels used for this survey were provided by SSI (Survey Sampling International), EMI Online Research Solutions, & Research Now. These panels use industry-standard techniques to verify the identities of their panel members and to ensure that their demographic characteristics are valid and up-to-date. For additional details on sample and response statistics, click here.

- Within each state, quotas were set to approximate Census distributions for age by gender, ethnicity, education level, and income based on data from the Census Bureau's American Community Survey.

- A pure probability sample of over 25,000 observations would have an estimated margin of error of half a percentage point (i.e., plus or minus 0.5 percent), and the margin of error would increase somewhat for sub-groupings of the sample. As in all survey research, there are possible sources of error—such as coverage, nonresponse and measurement error—that could affect the results.

- Note: As with previous NFCS State-by-State surveys, we did not specifically target heads of households or primary financial decision-makers.

630 Third Ave. | 9th Floor | New York, NY 10017 | 212.226.1007 | www.arcllc.com

MG

Mottola, Gary <Gary.Mottola@finra.org>

To: dianasimpson@auburn.edu

Fri 5/15/2020 1:25 PM

NFCS NDA template 10.2019.... 49 KB

Hi Diana,

Nice chatting with you.

As we discussed, we aggregate some variables and pull out zip code to protect the identity of the respondents. However, if you sign this one-page data usage agreement I can send you the full file.

The first step is for you to fill out the form and send it to me--but don't sign it. We will sign it first and send it back to you for your signature, which you will provide electronically via an application called DocuSign. The DocuSign email will come to you from tiffany.bolden@finra.org.

Once we have both signatures, I can send you the data.

And here's a link to the NFCS Research Award.

https://www.usfinancialcapability.org/research-award.php

Best regards,

Gary

APPENDIX II

Financial Literacy Questions on the 2018 NFCS

1. Suppose you had $100 in a savings account and the interest rate was 2% per year. After five years, how much do you think you would have in the account if you left the money to grow?
   a. More than $102
   b. Exactly $102
   c. Less than $102
   d. Don't know
   e. Prefer not to say
2. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?
   a. More than today
   b. Exactly the same
   c. Less than today
   d. Don't know
   e. Prefer not to say
3. If interest rates rise, what will typically happen to bond prices?
   a. They will rise
   b. They will fall
   c. They will stay the same
   d. There is no relationship between bond prices and the interest rate

e. Don't know

f. Prefer not to say

4. Suppose you owe $1.000 on a loan and the interest rate you are charged is 20% per year compounded annually. If you didn't pay anything off, at this interest rate, how many years would it take for the amount you owe to double?

   a. Less than two years

   b. At least 2 years but less than 5 years

   c. At least 5 years but less than 10 years

   d. At least 10 years

   e. Don't know

   f. Prefer not to say

5. A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.

   a. True

   b. False

   c. Don't know

   d. Prefer not to say

6. Buying a single company's stock usually provides a safer return than a mutual fund.

   a. True

   b. False

   c. Don't know

   d. Prefer not to say

APPENDIX III

2018 NFCS Questionnaire/Codebook

ARC RESEARCH
An SVC Company

FINRA Investor Education FOUNDATION

2018 National Financial Capability Study
State-by-State Survey Instrument

Note:

- Changes to the NFCS State-by-State survey are footnoted in this document. Footnotes are labeled with the year that the change was implemented (2018, 2015 or 2012).

Sample Characteristics:

- N ≈ 500 respondents per state (plus D.C.)
  - Oversamples in Oregon and Washington (total N ≈ 1,250 in each of these two states)[1]
- Quotas within each state by:
  - Age
  - Gender
  - Income
  - Ethnicity
  - Education

Coding Notes:

- For all questions in the survey except A3a:
  - Code 98 = Don't know
  - Code 99 = Prefer not to say

- For A3a:
  - Code 999 = Prefer not to say

- For questions that have been modified, 2018, 2015, 2012 and 2009 codes may differ from each other.

[1] 2018: Added oversamples of two states.

Z) Thank you very much for participating in this research.

- Please be assured that **all of your answers will be completely ANONYMOUS and CONFIDENTIAL**. Therefore, please try to answer these questions as openly and honestly as possible.

A1a) [SECTION A: DEMOGRAPHICS & CLASSIFICATION QUESTIONS]

A2) Please enter your 5 digit home zip code.

[_____]
[EDIT: 00001-99998]

[LOAD ALL GEO INFORMATION TO DATA]
[CHECK TOTAL STATE QUOTA, IF FULL, TERMINATE & SKIP TO QTERM]

A3) What is your gender?

Male 1
Female.....2

A3a)[2] What is your age?

[DROP DOWN MENU; PUNCH MATCHES AGE]

[13 13
14 14
15 15
16 16
17 17
18 18
19 19
20 20
…etc. *…etc.*
97 97
98 98
99 99
100 100
101 or older.....101
Prefer not to say ..... 999]

[IF Q.A3a = 13-17, 999 (REF), TERMINATE & SKIP TO QTERM]

[2] 2012: Changed from age ranges in 2009 to continuous years in 2012. Tracking comparisons can be made by coding individual years into the age ranges used in 2009.

A3b) [BUILDER: CREATE GENDER/AGE NET FROM Q's A3 & A3a:

Male 18-24 1
Male 25-34 2
Male 35-44 3
Male 45-54 4
Male 55-64 5
Male 65+ 6
Female 18-24 7
Female 25-34 8
Female 35-44 9
Female 45-54 10
Female 55-64 11
Female 65+ 12

CHECK GENDER/AGE QUOTA BY STATE, IF FULL, TERMINATE & SKIP TO QTERM]

A4)[3] Which of the following best describes your race or ethnicity?

Select all that apply.

[M]

White or Caucasian 1
Black or African-American 2
Hispanic or Latino/a 3
Asian 4
Native Hawaiian or other Pacific Islander 7
American Indian or Alaska Native 5
Other 6
Prefer not to say 99
[IF Q.A4 = 99 (REF), TERMINATE & SKIP TO QTERM][CODE 99 EXCLUSIVE]

[3] 2015: Changed "Asian/Pacific Islander" into two separate categories. Tracking comparisons can be made by coding into 2012 categories. Minor wording changes (from "Native American" in 2012 to "American Indian" in 2015).

A4a)[4] [BUILDER: PUNCH ETHNICITY

IF SINGLE RESPONSE:
IF Q.A4 = 1, PUNCH 1
IF Q.A4 = 2, PUNCH 2
IF Q.A4 = 3, PUNCH 3
IF Q.A4 = 4, PUNCH 4
IF Q.A4 = 7, PUNCH 4
IF Q.A4 = 5 or 6, PUNCH 5
IF MULTIPLE RESPONSES:
IF Q.A4 = 3, PUNCH 3
IF Q.A4 = 4 AND 7 ONLY, PUNCH 4
IF Q.A4 NE 3 OR (NE 4 AND 7 ONLY), PUNCH 5

White non-Hispanic 1
Black non-Hispanic.....2
Hispanic (any race) 3
Asian non-Hispanic.....4
Other non-Hispanic (American Indian, Other, 2+ ethnicities) 5

CHECK ETHNICITY QUOTA BY STATE, IF FULL, TERMINATE & SKIP TO QTERM]

A5)[5,6] What was the highest level of education that you completed?

Did not complete high school 1
High school graduate – regular high school diploma 2
High school graduate – GED or alternative credential 3
Some college, no degree 4
Associate's degree.....5
Bachelor's degree.....6
Post graduate degree 7
Prefer not to say 99

[IF Q.A5 = 99 (REF), TERMINATE & SKIP TO QTERM]
[CHECK EDUCATION QUOTA BY STATE, IF FULL, TERMINATE & SKIP TO QTERM]

[4] 2015: Programming logic updated to correspond to changes to A4 (ethnicity).
[5] 2012: Changed "high school graduate" into two separate categories (regular diploma and GED). Tracking comparisons between 2012 and 2009 can be made by coding into 2009 categories.
[6] 2015: Changed "some college" and "college graduate" into three separate categories ("some college, no degree," "associate's degree," and "bachelor's degree"). Minor wording changes (from "last year of education" and "post graduate education" in 2012 to "highest level of education" and "post graduate degree" in 2015).

A6) What is your marital status?

Married.....1
Single.....2
Separated.....3
Divorced.....4
Widowed/widower.....5
Prefer not to say.....99

[IF Q.A6 = 99 (REF), TERMINATE & SKIP TO QTERM]

A7) Which of the following describes your current living arrangements?

I am the only adult in the household.....1
I live with my spouse/partner/significant other.....2
I live in my parents' home.....3
I live with other family, friends, or roommates.....4
Prefer not to say.....99

[IF Q.A7 = 99 (REF), TERMINATE & SKIP TO QTERM]

A7a) [BUILDER: PUNCH MARITAL STATUS VARIABLE:

If Q.A6 = 1, PUNCH MARRIED
If Q.A6 = 2 – 5 AND Q.A7 = 2, PUNCH LIVING WITH PARTNER
If Q.A6 = 2 – 5 AND Q.A7 = 1, 3, or 4, PUNCH SINGLE

Married.....1
Living with partner.....2
Single.....3

If Q.A7a = 1, CVAR "spouse"
If Q.A7a = 2, CVAR "partner"

IF Q.A7a = 1 OR 2, CVAR "Does your household"
IF Q.A7a = 3, CVAR "Do you"]

A11)[7] How many children do you have who are financially dependent on you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]]? Please include children not living at home, and step-children as well.

1.....1
2.....2
3.....3
4 or more.....4
No financially dependent children.....5
Do not have any children.....6
Prefer not to say.....99

[IF Q.A11 = 99, TERMINATE & SKIP TO QTERM]

[7] 2012: Changed question order (appears earlier in the survey than in 2009).

A8) What is your [IF Q.A7a = 1 OR 2) INSERT: household's] approximate annual income, including wages, tips, investment income, public assistance, income from retirement plans, etc.? Would you say it is…

Less than $15,000 1
At least $15,000 but less than $25,000 2
At least $25,000 but less than $35,000 3
At least $35,000 but less than $50,000 4
At least $50,000 but less than $75,000 5
At least $75,000 but less than $100,000 6
At least $100,000 but less than $150,000 7
$150,000 or more 8
Don't know 98
Prefer not to say 99

[IF Q.A8 = 98 (DK) OR 99 (REF), TERMINATE & SKIP TO QTERM]
[CHECK INCOME QUOTA BY STATE, IF FULL, TERMINATE & SKIP TO QTERM]

AM21)[8] Have you ever been a member of the U.S. Armed Services, either in the active or reserve component?

Currently a member of the U.S. Armed Services 1
Previously a member of the U.S. Armed Services 2
Never a member of the U.S. Armed Services 3
Prefer not to say 99

[IF Q.AM21 = 2 (PREVIOUSLY), ASK; OTHERWISE SKIP TO Q.AM22]

AM30)[9] When did you complete your service in the military?

Within the past year 1
1 to 3 years ago 2
4 to 10 years ago 3
More than 10 years ago 4
Prefer not to say 99

AM31)[10] Did you retire from the military?

Yes 1
No 2
Don't know 98
Prefer not to say 99

[8] 2012: Military question added to State-by-State survey. See note that follows question X3.
[9] 2015: New question.
[10] 2015: New question.

AM32)[11] What was your most recent military service branch and component?

[DISPLAY WITH BREAKS ON THE LIST]

Army 1
Army National Guard (full-time, activated, or non-activated) 2
Army Reserve (full-time, activated, or non-activated) 3

Navy 4
Navy Reserve (full-time, activated, or non-activated) 5

Air Force 6
Air National Guard (full-time, activated, or non-activated) 7
Air Force Reserve (full-time, activated, or non-activated) 8

Marine Corps 9
Marine Corps Reserve (full-time, activated, or non-activated) 10

Coast Guard 11
Coast Guard Reserve (full-time, activated, or non-activated) 12

Don't know 98
Prefer not to say 99

[IF Q.A6 = 1 (MARRIED), ASK; OTHERWISE SKIP TO Q.X3]

AM22)[12] Has your spouse ever been a member of the U.S. Armed Services, either in the active or reserve component?

Currently a member of the U.S. Armed Services 1
Previously a member of the U.S. Armed Services 2
Never a member of the U.S. Armed Services 3
Prefer not to say 99

X3) [BUILDER: PUNCH QUESTIONNAIRE VERSION:

If Q.AM21 = 1 OR Q.AM22 = 1, PUNCH 2 (MILITARY)
ALL OTHERS, PUNCH 1 (CORE)

Core questions 1
Military ..... 2]

Note on Military Questions:

- Depending on their military status, respondents were also shown military-specific wording for several questions, as indicated in the programming instructions in this document.

[11] 2015: New question.
[12] 2012: Military question added to State-by-State survey.

[IF Q.X3 = 2 (MILITARY), ASK; OTHERWISE SKIP TO Q.A9]

X4) [BUILDER: PUNCH MILITARY STATUS VARIABLE:

If Q.AM21 = 1 (CURRENT MEMBER), PUNCH RESPONDENT IN SERVICE
If Q.AM21 = 2, 3, 99 (PREV, NEVER, REF) AND Q.AM22 = 1 (SPOUSE CURRENT MEMBER), PUNCH SPOUSE IN SERVICE

Respondent in service 1
Spouse in service ..... 2]

A9) Which of the following best describes your current employment or work status?

Self-employed 1
Work full-time for an employer [IF Q.AM21 = 1 INSERT: or the military] 2
Work part-time for an employer [IF Q.AM21 = 1 INSERT: or the military] 3
Homemaker 4
Full-time student 5
Permanently sick, disabled, or unable to work 6
Unemployed or temporarily laid off 7
Retired 8
Prefer not to say 99

[IF Q.A9 = 99, TERMINATE & SKIP TO QTERM]

A40)[13] [IF Q.A9 = 1-3 (EMPLOYED), INSERT: In addition to your main employment, did you also do other work for pay in the <u>past 12 months</u>?]

[IF Q.A9 = 4-8 (NOT EMPLOYED), INSERT: Did you do any work for pay in the <u>past 12 months</u>?]

Yes 1
No 2
Don't know 98
Prefer not to say 99

[IF Q.A7a = 1 OR 2, ASK; OTHERWISE SKIP TO Q.A10a]

A10) Which of the following best describes your [spouse/partner]'s current employment or work status?

Self-employed 1
Work full-time for an employer [IF Q.AM22 = 1 INSERT: or the military] 2
Work part-time for an employer [IF Q.AM22 = 1 INSERT: or the military] 3
Homemaker 4
Full-time student 5
Permanently sick, disabled, or unable to work 6
Unemployed or temporarily laid off 7
Retired 8
Prefer not to say 99

[IF Q.A10 = 99, TERMINATE & SKIP TO QTERM]

[13] 2018: New question.

A10a) [BUILDER: HOUSEHOLD RETIREMENT STATUS:

IF Q.A9 = 1 – 3, PUNCH NON-RETIRED HOUSEHOLD
IF ((Q.A7a = 3 AND Q.A9 = 4 – 7) OR (Q.A7a = 1, 2 AND Q.A9 = 4 – 7 AND Q.A10 = 1 – 7)), PUNCH NON-RETIRED HOUSEHOLD
IF Q.A9 = 8, PUNCH RETIRED-HOUSEHOLD – RESPONDENT RETIRED
IF Q.A7a = 1, 2 AND Q.A9 = 4 – 7 AND Q.A10 = 8, PUNCH RETIRED HOUSEHOLD – RESPONDENT NOT WORKING AND SPOUSE RETIRED

Non-retired household 1
Retired household--Respondent retired 2
Retired household--Respondent not working and spouse retired ..... 3]

[IF Q.X3 = 2 (MILITARY), ASK; OTHERWISE SKIP TO Q.A21]

AM7) What is your [IF Q.X4 = 2 INSERT: spouse's] military service branch and component?

[DISPLAY WITH BREAKS ON THE LIST]

Army 1
Army National Guard (full-time, activated, or non-activated) 2
Army Reserve (full-time, activated, or non-activated) 3

Navy.....4
Navy Reserve (full-time, activated, or non-activated).....5

Air Force.....6
Air National Guard (full-time, activated, or non-activated) 7
Air Force Reserve (full-time, activated, or non-activated).....8

Marine Corps 9
Marine Corps Reserve (full-time, activated, or non-activated) 10

Coast Guard 11
Coast Guard Reserve (full-time, activated, or non-activated).....12

Don't know 98
Prefer not to say.....99

[IF Q.A5 = 2, 3, 4, 5 (HS GRAD, SOME COLLEGE, ASSOCIATE'S DEGREE) AND Q.A9 NE 5 (NOT FT STUDENT), ASK; OTHERWISE SKIP TO Q.A22]

\# A21)[14,15] Are you a part-time student taking courses for credit?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.A5 = 2, 3, 4, 5 (HS GRAD, SOME COLLEGE, ASSOCIATE'S DEGREE) AND ((Q.A9 = 5 OR Q.A21 = 1) (FT OR PT STUDENT)), ASK; OTHERWISE SKIP TO Q.A14]

\# A22)[16,17] Which of the following best describes the school you are attending?

Four-year college or university.....1
Two-year community college 2
Vocational, technical, or trade school 3
Other 4
Don't know 98
Prefer not to say 99

[IF Q.A7a = 1, 2, ASK; OTHERWISE SKIP TO Q.A41]

\# A14) Who in the household is most knowledgeable about saving, investing and debt?

You.....1
Someone else 2
You and someone else are equally knowledgeable 3
Don't know 98
Prefer not to say 99

\# A41)[18] What was the highest level of education completed by the person or any of the people who raised you?

Did not complete high school 1
High school graduate/GED.....2
Some college, no degree 3
Associate's degree.....4
Bachelor's degree.....5
Post graduate degree 6
Don't know 98
Prefer not to say 99

\# A16) [END OF SCREENER]

[14] 2012: New question.
[15] 2015: Question base updated to correspond to changes to A5 (education).
[16] 2012: New question.
[17] 2015: Question base updated to correspond to changes to A5 (education).
[18] 2018: New question.

J) [SECTION J: FINANCIAL ATTITUDES & BEHAVIORS]

Ja) These days, a lot of people are thinking about financial issues. We are interested in your opinions on some of these issues.

J1) Overall, thinking of your assets, debts and savings, how satisfied are you with your current personal financial condition?

Please use a 10-point scale, where 1 means "Not At All Satisfied" and 10 means "Extremely Satisfied."

Not At All Satisfied 1	2	3	4	5	6	7	8	9	Extremely Satisfied 10	Don't know	Prefer not to say
1	2	3	4	5	6	7	8	9	10	98	99

J2) When thinking of your financial investments, how willing are you to take risks?

Please use a 10-point scale, where 1 means "Not At All Willing" and 10 means "Very Willing."

Not At All Willing 1	2	3	4	5	6	7	8	9	Very Willing 10	Don't know	Prefer not to say
1	2	3	4	5	6	7	8	9	10	98	99

J3) Over the <u>past year</u>, would you say your [IF Q.A7a = 1 OR 2 INSERT: household's] spending was less than, more than, or about equal to your [IF Q.A7a = 1 OR 2 INSERT: household's] income? Please do not include the purchase of a new house or car, or other big investments you may have made.

Spending **less** than income.....1
Spending **more** than income.....2
Spending about **equal** to income.....3
Don't know.....98
Prefer not to say.....99

J4) In a <u>typical month</u>, how difficult is it for you to cover your expenses and pay all your bills?

Very difficult.....1
Somewhat difficult.....2
Not at all difficult.....3
Don't know.....98
Prefer not to say.....99

J40)[19] In the past 12 months, which one of the following best describes your [IF Q.A7a = 1 OR 2 INSERT: and your [spouse/partner]'s] income?

Roughly the same amount each month 1
Occasionally varies from month to month 2
Varies quite often from month to month 3
Don't know 98
Prefer not to say 99

J5) Have you set aside emergency or rainy day funds that would cover your expenses for 3 months, in case of sickness, job loss, economic downturn, or other emergencies?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.A11 = 1, 2, 3, 4 (FINANCIALLY DEPENDENT CHILD), ASK; OTHERWISE SKIP TO Q.J8]

J6) Are you setting aside any money for your children's college education?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.A10a = 1 (NOT RETIRED), ASK; OTHERWISE SKIP TO Q.J9]

J8) Have you ever tried to figure out how much you need to save for retirement?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.A10a = 2, 3 (RETIRED), ASK; OTHERWISE SKIP TO Q.J10]

J9) [IF Q.A10a = 2 INSERT: Before you retired, did you try to figure out how much you needed to save for retirement?]

[IF Q.A10a = 3 INSERT: Before your [spouse/partner] retired, did you try to figure out how much you needed to save for retirement?]

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[19] 2018: New question.

J10) In the past 12 months, [IF Q.A7a = 3 INSERT: have you/ IF Q.A7a = 1 OR 2 INSERT: has your household] experienced a large drop in income which you did not expect?

Yes....1
No....2
Don't know....98
Prefer not to say....99

J20)[20] How confident are you that you could come up with $2,000 if an unexpected need arose within the next month?

I am certain I could come up with the full $2,000....1
I could probably come up with $2,000....2
I could probably not come up with $2,000....3
I am certain I could not come up with $2,000....4
Don't know....98
Prefer not to say....99

J32)[21] How would you rate your current credit record?

Very bad....1
Bad....2
About average....3
Good....4
Very good....5
Don't know....98
Prefer not to say....99

[20] 2012: New question.
[21] 2015: New question.

J33)[22] How strongly do you agree or disagree with the following statements?

Please give your answer on a scale of 1 to 7, where 1 = "Strongly Disagree," 7 = "Strongly Agree," and 4 = "Neither Agree Nor Disagree". You can use any number from 1 to 7. (Select an answer for each)

[RANDOMIZE]

		Strongly Disagree 1	2	3	Neither Agree nor Disagree 4	5	6	Strongly Agree 7	Don't Know	Prefer not to Say
J33_1)	I worry about running out of money in retirement	1	2	3	4	5	6	7	98	99
J33_40)[23]	Thinking about my personal finances can make me feel anxious	1	2	3	4	5	6	7	98	99
J33_41)[24]	Discussing my finances can make my heart race or make me feel stressed	1	2	3	4	5	6	7	98	99

J41)[25] How well do these statements describe you or your situation?

[RANDOMIZE]

		Does not describe me at all	Describes me very little	Describes me somewhat	Describes me very well	Describes me completely	Don't know	Prefer not to say
J41_1)	Because of my money situation, I feel like I will never have the things I want in life	1	2	3	4	5	98	99
J41_2)	I am just getting by financially	1	2	3	4	5	98	99
J41_3)	I am concerned that the money I have or will save won't last	1	2	3	4	5	98	99

J42)[26] How often do these statements apply to you?

[RANDOMIZE]

		Never	Rarely	Sometimes	Often	Always	Don't know	Prefer not to say
J42_1)	I have money left over at the end of the month	1	2	3	4	5	98	99
J42_2)	My finances control my life	1	2	3	4	5	98	99

[22] 2015: New question.
[23] 2018: New question.
[24] 2018: New question.
[25] 2018: New questions.
[26] 2018: New questions.

J43)[27] If you were to set a financial goal for yourself today, how confident are you in your ability to achieve it?

Not at all confident 1
Not very confident 2
Somewhat confident 3
Very confident 4
Don't know 98
Prefer not to say 99

J14) [END OF SECTION J]

[27] 2018: New question.

B) [SECTION B: BANKING]

[DISPLAY Q'S B1 AND B2 ON SAME SCREEN]

B1) [IF Q.A7a = 3 INSERT: Do you/ IF Q.A7a = 1 OR 2 INSERT: Does your household] have a checking account?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

B2) [IF Q.A7a = 3 INSERT: Do you/ IF Q.A7a = 1 OR 2 INSERT: Does your household] have a savings account, money market account, or CDs?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.B1 = 1 (YES), ASK; OTHERWISE SKIP TO Q.B40]

B4) Do you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] overdraw your checking account occasionally?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

B40)[28] How strongly do you agree or disagree with the following statement?

Please give your answer on a scale of 1 to 7, where 1 = "Strongly Disagree," 7 = "Strongly Agree," and 4 = "Neither Agree Nor Disagree". You can use any number from 1 to 7.

	Strongly Disagree 1	2	3	Neither Agree nor Disagree 4	5	6	Strongly Agree 7	Don't know	Prefer not to say
I would feel comfortable going to a bank or credit union branch to ask a question about a product or service	1	2	3	4	5	6	7	98	99

[28] 2018: New question.

[IF Q.B1 OR Q.B2 = 1 (YES), ASK; OTHERWISE SKIP TO Q.B31]

B41)[29] How often do you access your checking or savings account in the following ways?

		Never	Sometimes	Frequently	Don't know	Prefer not to say
B41_1)	Online banking with a laptop or desktop computer	1	2	3	98	99
B41_2)	Mobile banking with text messaging, mobile app, or Internet browser or email on a mobile phone	1	2	3	98	99

B14) [MOVED TO END OF SECTION C]

B31)[30] How often do you use your mobile phone to pay for a product or service <u>in person</u> at a store, gas station, or restaurant (e.g., by waving/tapping your mobile phone over a sensor at checkout, scanning a barcode or QR code using your mobile phone, or using some other mobile app at checkout)?

Frequently 1
Sometimes.....2
Never.....3
Don't know 98
Prefer not to say 99

B42)[31] How often do you use your mobile phone to transfer money to another person?

Frequently 1
Sometimes.....2
Never.....3
Don't know 98
Prefer not to say 99

B43)[32] How often do you use websites or apps to help with financial tasks such as budgeting, saving, or credit management (e.g., GoodBudget, Mint, Credit Karma, etc.)? Please do not include websites or apps for making payments or money transfers.

Frequently 1
Sometimes.....2
Never.....3
Don't know 98
Prefer not to say 99

[29] 2018: New questions.
[30] 2015: New question. Replaces B22_8 from 2012.
[31] 2018: New question.
[32] 2018: New question.

B44)[33] In the past 12 months, how often have you taken on a work assignment through a website or mobile app, such as Uber, Task Rabbit, Care.com, etc.?

Frequently 1
Sometimes 2
Never 3
Don't know 98
Prefer not to say 99

B16) [END OF SECTION B]

[33] 2018: New question.

C) [SECTION C: RETIREMENT ACCOUNTS]

Ca) The following are questions about retirement accounts and pensions. Please answer to the best of your knowledge. If you really do not know the answer, please select "don't know."

C1)[34] Do you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] have any retirement plans through a current or previous employer, like a pension plan [IF Q.X3 = 2 INSERT: , a Thrift Savings Plan (TSP),] or a 401(k)?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.C1 = 1 (YES) AND Q.A7a = 1 OR 2, ASK; OTHERWISE SKIP TO Q.C3]

C2) Were these plans provided by your employer or your [spouse/partner]'s employer, or both?

Your employer.....1
Your [spouse's/partner's] employer 2
Both your employer and your [spouse's/partner's] employer.....3
Don't know 98
Prefer not to say 99

[IF Q.C1 = 1 (YES), ASK; OTHERWISE SKIP TO Q.C4]

C3) Are any of these retirement plans the kind where you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] get to choose how the money is invested?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

C4)[35,36] Do you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] have any other retirement accounts NOT through an employer, like an IRA, Keogh, SEP, or any other type of retirement account that you have set up yourself?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[34] 2012: The base for all questions in this section (C1 through C11) changed from *non-retired households* in 2009 to *all respondents* in 2012. Tracking comparisons to 2009 can be made by looking at the responses of only non-retired households (A10a = 1) in 2012 and 2015.
[35] 2015: Minor wording changes (added "myRA" to list of examples).
[36] 2018: Minor wording changes (removed "myRA" from list of examples).

[IF Q.C3 = 1 OR Q.C4 = 1 (YES), ASK; OTHERWISE, SKIP TO Q.B14]

C5) Do you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] regularly contribute to a retirement account like a [IF Q.X3 = 2 INSERT: Thrift Savings Plan (TSP),] 401(k) or IRA?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[DISPLAY Q'S C10 & C11 ON SAME SCREEN]

C10) In the last 12 months, have you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] taken a loan from your retirement account(s)?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

C11) In the last 12 months, have you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] taken a hardship withdrawal from your retirement account(s)?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.B1 OR B2 = 1, 98, 99 (YES, DK, REF), ASK; OTHERWISE SKIP TO Q.C40]

B14)[37] Not including retirement accounts, [IF Q.A7a = 1 OR 2 INSERT does your household/IF Q.7a = 3 INSERT: do you] have any investments in stocks, bonds, mutual funds, or other securities?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

C40)[38] Over the past 12 months, approximately how often did you buy a lottery ticket?

Almost every day 1
A few times a week 2
About once a week 3
About once a month 4
Less than once a month 5
Never 6
Don't know 98
Prefer not to say 99

[37] 2015: Changed question order (appears later in the survey than in 2012).
[38] 2018: New question.

C41)[39] Do you currently have a will?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

C16) [END OF SECTION C]

[39] 2018: New question.

D) [SECTION D: GOVERNMENT BENEFITS]

D40)[40] Do you receive either Medicaid benefits or food stamps/SNAP?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

D17) [END OF SECTION D]

[40] 2018: New question.

E) [SECTION E: HOME & MORTGAGES]

Ea_1)[41] Do you [IF Q.A7a = 1 OR 2 INSERT: or your [spouse/partner]] currently own your home?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.Ea_1 = 1 (YES OWN HOME), ASK; OTHERWISE SKIP TO Q.E17]

E3a) Following are some questions about your home. If you own more than one home, please refer to your primary residence.

[DISPLAY Q'S E7 & E8 ON SAME SCREEN]

E7)[42] Do you currently have any mortgages on your home?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

E8)[43] Do you have any home equity loans?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[IF Q.E7 = 1 OR Q.E8 = 1 (YES), ASK; OTHERWISE SKIP TO Q.E15]

E20)[44] Do you currently owe more on your home than you think you could sell it for today?

Yes, owe more.....1
No 2
Don't know 98
Prefer not to say 99

[41] 2015: Changed format of question from a grid ("Do you currently own any of the following? – Your home") in 2012 to a single question in 2015. Minor wording changes to accommodate new question format.
[42] 2012: Minor wording changes (from "a mortgage" in 2009 to "any mortgages" in 2012).
[43] 2012: Minor wording changes (from "a home equity loan" in 2009 to "any home equity loans" in 2012).
[44] 2012: New question.

[IF Q.E7 = 1 (YES), ASK; OTHERWISE SKIP TO Q.E17]

E15)[45] How many times have you been late with your mortgage payments in the past 12 months? (If you have more than one mortgage on your home(s), please consider them all.)

Never.....1
Once.....2
More than once.....3
Don't know.....98
Prefer not to say.....99

E17) [END OF SECTION E]

[45] 2015: Changed time frame of question (from *2 years* in 2012 to *12 months* in 2015). Tracking comparisons are not possible.

F) [SECTION F: CREDIT CARDS]

F1) How many credit cards do you have? Please include store and gas station credit cards but NOT debit cards.

1 1
2-3 2
4-8 3
9-12 4
13-20 5
More than 20 6
No credit cards 7
Don't know 98
Prefer not to say 99

[IF Q.F1 = 7 (None), 98 (DK), 99 (REF), SKIP TO F12]

F2) In the past 12 months, which of the following describes your experience with credit cards? (Select an answer for each)

[DO NOT RANDOMIZE]

		Yes	No	Don't Know	Prefer not to Say
F2_1)	I always paid my credit cards in full	1	2	98	99
F2_2)	In some months, I carried over a balance and was charged interest	1	2	98	99
F2_3)	In some months, I paid the minimum payment only	1	2	98	99
F2_4)	In some months, I was charged a late fee for late payment	1	2	98	99
F2_5)	In some months, I was charged an over the limit fee for exceeding my credit line	1	2	98	99
F2_6)	In some months, I used the cards for a cash advance	1	2	98	99

F10) Thinking about when you obtained your most recent credit card, did you collect information about different cards from more than one company in order to compare them?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

F12) [END OF SECTION F]

G) [SECTION G: OTHER DEBT]

G1) [IF Q.A7a = 3 INSERT: Do you/ IF Q.A7a = 1 OR 2 INSERT: Does your household] currently have an auto loan? (This does not refer to an auto lease).

Yes.....1
No 2
Don't know 98
Prefer not to say 99

G20)[46] Do you currently have any unpaid bills from a health care or medical service provider (e.g., a hospital, a doctor's office, or a testing lab) that are past due?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

G30)[47] Do you currently have any student loans? If so, for whose education was this/were these loan(s) taken out?

Select all that apply.
[CODES 97, 98, 99 EXCLUSIVE]
[DISPLAY WITH BREAK ON THE LIST]

[M]

Yes, have student loan(s) for:
- Yourself.....1
- Your spouse/partner 2
- Your child(ren).....3
- Your grandchild(ren).....4
- Other person.....5

No, do not currently have any student loans.....97
Don't know 98
Prefer not to say 99

[IF Q.G30 = 1-5 (HAVE STUDENT LOAN), ASK; OTHERWISE SKIP TO Q.G40]

G33)[48] Before you got your most recent student loan, did you try to figure out how much your monthly payments would be?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

[46] 2012: New question.
[47] 2015: New question. Replaces G21 from 2012.
[48] 2015: New question.

G35)[49] How many times have you been late with a student loan payment in the past 12 months? (If you have more than one student loan, please consider them all.)

Never, payments are not due on my loans at this time.....1
Never, I have been repaying on time each month.....2
Once.....3
More than once.....4
Don't know.....98
Prefer not to say.....99

G22)[50,51] Are you concerned that you might not be able to pay off your student loans?

Yes.....1
No.....2
Don't know.....98
Prefer not to say.....99

[[IF Q.A5 = 4, 5, 6, 7 (SOME COLLEGE OR MORE), ASK; OTHERWISE SKIP TO Q.G25]

G40)[52] Do you wish you had chosen to go to a less expensive college?

Yes.....1
No.....2
Don't know.....98
Prefer not to say.....99

[49] 2015: New question.
[50] 2012: New question.
[51] 2015: Question base changed to correspond to G30 (new student loan question).
[52] 2018: New question.

G25)[53] In the past 5 years, how many times have you… (Select an answer for each)

[RANDOMIZE]

		Never	1 time	2 times	3 times	4 or more times	Don't Know	Prefer not to Say
G25_1)[54]	Taken out an auto title loan? Auto title loans are loans where a car title is used to borrow money for a short period of time. They are NOT loans used to purchase an automobile.	1	2	3	4	5	98	99
G25_2)	Taken out a short term "payday" loan?	1	2	3	4	5	98	99
G25_3)[55, 56]	Gotten an advance on your tax refund? This is sometimes called a "refund anticipation check" or "Rapid Refund" (Not the same as e-filing)	1	2	3	4	5	98	99
G25_4)	Used a pawn shop?	1	2	3	4	5	98	99
G25_5)	Used a rent-to-own store?	1	2	3	4	5	98	99

G38)[57] Have you been contacted by a debt collection agency in the past 12 months?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

G23)[58] How strongly do you agree or disagree with the following statement?

Please give your answer on a scale of 1 to 7, where 1 = "Strongly Disagree," 7 = "Strongly Agree," and 4 = "Neither Agree Nor Disagree". You can use any number from 1 to 7.

	Strongly Disagree 1	2	3	Neither Agree nor Disagree 4	5	6	Strongly Agree 7	Don't Know	Prefer not to Say
I have too much debt right now	1	2	3	4	5	6	7	98	99

G10) [END OF SECTION G]

[53] 2012: New questions. The "Yes/No" scale in 2009 was replaced by a frequency scale ("How many times") in 2012, therefore tracking comparisons to 2009 are not possible. Minor wording changes to individual items (not documented here) to accommodate the new scale.
[54] 2012: Description of auto title loans added in 2012.
[55] 2012: Minor wording changes (from refund anticipation "loan" in 2009 to refund anticipation "check" in 2012).
[56] 2018: Question was removed in 2015 and reinstated in 2018.
[57] 2015: New question.
[58] 2012: New question.

H) [SECTION H: INSURANCE]

H1) Are you covered by health insurance?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

H30)[59] In the last 12 months, was there any time when you…

[RANDOMIZE]

		Yes	No	Don't Know	Prefer not to Say
H30_1)	Did NOT fill a prescription for medicine because of the cost	1	2	98	99
H30_2)	SKIPPED a medical test, treatment or follow-up recommended by a doctor because of the cost	1	2	98	99
H30_3)	Had a medical problem but DID NOT go to a doctor or clinic because of the cost	1	2	98	99

H8) [END OF SECTION H]

[59] 2015: New questions.

M) [SECTION M: SELF-ASSESSMENT & LITERACY]

M1) How strongly do you agree or disagree with the following statements?

Please give your answer on a scale of 1 to 7, where 1 = "Strongly Disagree," 7 = "Strongly Agree," and 4 = "Neither Agree Nor Disagree". You can use any number from 1 to 7. (Select an answer for each)

[RANDOMIZE]

		Strongly Disagree 1	2	3	Neither Agree nor Disagree 4	5	6	Strongly Agree 7	Don't Know	Prefer not to Say
M1_1)	I am good at dealing with day-to-day financial matters, such as checking accounts, credit and debit cards, and tracking expenses	1	2	3	4	5	6	7	98	99
M1_2)	I am pretty good at math	1	2	3	4	5	6	7	98	99

M4) On a scale from 1 to 7, where 1 means very low and 7 means very high, how would you assess your overall financial knowledge?

Very Low 1	2	3	4	5	6	Very High 7	Don't know	Prefer not to say
1	2	3	4	5	6	7	98	99

M40)[60] Were you ever required to take financial education?

Yes.....1
No 2
Don't know 98
Prefer not to say 99

M20)[61] Was financial education offered by a school or college you attended, or a workplace where you were employed?

Yes, but I did not participate in the financial education offered 1
Yes, and I did participate in the financial education 2
No 3
Don't know 98
Prefer not to say 99

[60] 2018: New question.
[61] 2012: New question.

[IF Q.M20 = 2 (PARTICIPATED), ASK, OTHERWISE SKIP TO Q.M5a]

M21)[62] When did you receive that financial education?

[DO NOT RANDOMIZE]

		Yes	No	Don't Know	Prefer not to Say
M21_1)	In high school	1	2	98	99
M21_2)[63]	[IF Q.A5 = 4, 5, 6, 7] In college	1	2	98	99
M21_3)	From an employer	1	2	98	99
M21_4)	[IF Q.AM21 = 1 or 2] From the military	1	2	98	99

M41)[64] In total, about how many hours of financial education did you receive?

1-2 hours 1
3-10 hours 2
More than 10 hours 3
Don't know 98
Prefer not to say 99

M42)[65] Overall, how would you rate the quality of the financial education you received?

Please give your answer on a scale of 1 to 7, where 1 means "very low" and 7 means "very high."

Very Low 1	2	3	4	5	6	Very High 7	Don't know	Prefer not to say
1	2	3	4	5	6	7	98	99

M5a) Following are some multiple choice questions. If you don't know the answer, just select "don't know."

M6) Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?

More than $102 1
Exactly $102 2
Less than $102 3
Don't know 98
Prefer not to say 99

[62] 2012: New questions.
[63] 2015: Question base updated to correspond to changes to A5 (education).
[64] 2018: New question.
[65] 2018: New question.

M7) Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

More than today 1
Exactly the same 2
Less than today 3
Don't know 98
Prefer not to say 99

M8) If interest rates rise, what will typically happen to bond prices?

They will rise 1
They will fall 2
They will stay the same 3
There is no relationship between bond prices and the interest rate 4
Don't know 98
Prefer not to say 99

M31)[66] Suppose you owe $1,000 on a loan and the interest rate you are charged is 20% per year compounded annually. If you didn't pay anything off, at this interest rate, how many years would it take for the amount you owe to double?

Less than 2 years 1
At least 2 years but less than 5 years 2
At least 5 years but less than 10 years 3
At least 10 years 4
Don't know 98
Prefer not to say 99

M9a)[67] There are a few questions left, and the survey will be complete.

Following are two statements. Please indicate whether each statement is true or false. If you don't know, just select "don't know."

[RANDOMIZE Q.M9 AND Q.M10]

M9) A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.

True 1
False 2
Don't know 98
Prefer not to say 99

[66] 2015: New question.
[67] 2015: Minor wording changes (from "two questions" in 2012 to "a few questions" in 2015).

M10) Buying a single company's stock usually provides a safer return than a stock mutual fund.

True.....1
False.....2
Don't know.....98
Prefer not to say.....99

M11) [END OF SECTION M]

999) [POINT OF COMPLETE]