ARTICLE

An item response theory approach to constructing and evaluating brief and in-depth financial literacy scales

Rob Ranyard[1] ⓘ | Simon McNair[2] | Gianni Nicolini[3] |
Darren Duxbury[4]

[1]Centre for Decision Research, Leeds University Business School, Leeds, UK

[2]Consumer Markets Team, Behavioral Insights Team, London, UK

[3]Faculty of Economics, Tor Vergata University of Rome, Rome, Italy

[4]Accounting and Finance Group, Newcastle University Business School, Tyne, UK

Correspondence
Rob Ranyard, Leeds University Business School, Leeds, UK.
Email: r.ranyard@leeds.ac.uk

Abstract

We applied item response theory (IRT) to construct and evaluate new brief and in-depth financial literacy scales. A survey of a UK adult sample ($N = 589$) included 50 questions to assess knowledge about managing financial resources and competence in using personal finance-related information—including five widely used items, on interest rates, inflation, investment diversification, mortgages and bonds. IRT applied to a scale of these items identified some limitations, overcome via further iterations to construct a new brief scale with sound psychometric properties. IRT was then applied iteratively to our pool, resulting in an in-depth, 20-item scale, also psychometrically sound, covering four broad financial domains: everyday money transactions; the concept of money; borrowing; and saving and investment. Parallel 10-item sub-scales were also evaluated. The validity of the new scales was demonstrated by regression analyses which found that, controlling for demographic variables, financial literacy predicted key indicators of financial well-being.

KEYWORDS

financial literacy, item response theory, scale development

1 | INTRODUCTION

Financial literacy, defined by the US President's Advisory Council on Financial Literacy as "the ability to use knowledge and skills to manage financial resources effectively for a lifetime of financial well-being" (PACFL, 2008) is recognized as a global societal issue (Organization for Economic Cooperation and Development [OECD], 2016). Worldwide, it has been estimated that only one in three adults has a level of financial literacy that could be considered necessary for fully informed, rational financial decision making (Klapper *et al.*, 2015). Moreover, much research has implicated financial literacy as important in various financial behaviors including: effective retirement planning (Lusardi and Mitchell, 2007; 2011a; van Rooij *et al.*, 2011a; 2011b); stock investing (Abreu and Mendes, 2010); inflation expectations (de Bruin *et al.*, 2010); and having more savings and lower borrowing (Stango and Zinman, 2009). In addition, it has been found that those who lack basic understanding of fundamental economic concepts are also more likely to be unbanked (Grimes *et al.*, 2010).

Despite the abundance of research on financial literacy that has highlighted its importance for a range of everyday financial decisions, there remain several issues pertaining to exactly *how* financial literacy should be assessed. First, there are issues of definition and operationalization (Hung *et al.*, 2009; Remund, 2010). In a review of 72 studies of financial literacy, Huston (2010) observed that only seven provided clear and formal definitions of "financial literacy" as a construct, while a further eight studies "somewhat" defined the construct. Huston argued for a conceptualization that includes both understanding and application of financial concepts, commensurate with the PACFL definition above: "... Financial literacy could be defined as measuring how well an individual can understand and use personal finance-related information." (306). Huston (2010) further argued that in-depth financial literacy assessment should adequately represent its key domains, which she considered to be: the basic concept of money, including interest rates and inflation; borrowing, including mortgages; saving and investment, including investment diversification; and resource protection, including insurance.

A contrasting approach, pioneered by Annamaria Lusardi and her colleagues (Lusardi, 2011; Lusardi and Mitchell, 2011a; 2011b), aims to assess financial literacy efficiently with brief scales based on a small number of survey questions. This approach has been widely adopted, including by the OECD (2016). Both approaches have advantages and disadvantages. On the one hand, the in-depth approach gives a richer profile of respondents' strengths and weaknesses in different domains of knowledge that can be particularly useful for evaluating financial education programs or the educational needs of individuals. On the other hand, research programs with more general theoretical or policy goals require an efficient, but valid and reliable, brief scale. The main aim of this paper is to develop both a brief, 5-item scale and an in-depth, 20-item scale of financial literacy with internationally relevant items. In addition, we aim to construct and evaluate 10-item scales to provide alternatives to the 13-item scale of Fernandes *et al.* (2014) and the 10-item scale of Houts and Knoll (2019). The new scales should have similarly sound psychometric properties but more coherent content representative of four key domains of financial literacy: everyday money transactions; the concept of money; loans and mortgages; and saving and investment.[1]

Knoll and Houts (2012) have argued that few existing financial literacy scales were created based on rigorous psychometric analysis. They addressed this problem by applying item response theory (IRT), a psychometric technique for the construction of sensitive and reliable measures, to a pool of items from existing US financial literacy surveys. The result was a new, 20-item "financial knowledge scale" consisting of items with a good distribution of difficulty

that discriminated different levels of knowledge well. Informed by this work, we applied IRT to develop new scales. It is important to note, however, that Knoll and Houts' scale was validated against US surveys, and three of their items pertaining to the US financial context would not be relevant to people living in other countries. For instance, items requiring knowledge of 401(k)s, or of IRAs, are not likely to be good candidates by which to assess financial knowledge of non-US adults. A second problem with the content of Knoll and Houts' scale is that it is heavily biased toward the domain of savings and investments, with nine of the 17 internationally relevant items being from that domain. These problems also apply to Fernandes *et al.*'s (2014) 13-item scale, which includes one US-specific item and otherwise is heavily weighted toward investment knowledge. We address these issues by devising an initial pool of items across 10 financial topics that are relevant to people internationally and develop scales with items sampling these topics. The main aim of the present article, then, is to advance the assessment of financial literacy by applying current good practice in scale construction and evaluation, with scale content comprising internationally relevant items that sample across key domains of financial literacy. We build on the work of Lusardi (2011) to construct a brief scale, and also that of Huston (2010) to construct an in-depth scale.

Our initial pool of 50 items sampling 10 topics was selected with an international population in mind (see Appendix A, and Nicolini, 2019). Our starting point was the "big five" items devised by Lusardi (2011) and her colleagues (Lusardi and Mitchell, 2011a; 2011b; 2014). Although these were initially devised for an investigation of preparedness for retirement, they have since been widely used to assess financial literacy, including by Knoll and Houts (2012). These five questions (see Table 6) represent our first five topics: Interest Rates, Inflation, Mortgages, Bonds and Investment Diversification. Five additional topics were included in order to deepen and broaden our coverage: Bank Accounts, Payments, Loans, Stock Investment and Pensions.[2] The authors added one question from an existing survey and devised 44 new ones, varying in difficulty, to construct an initial pool of 50 questions, five per topic. The ten topics cover three of Huston's (2010) four key financial domains: *the concept of money* (Interest rates, Inflation); *borrowing* (Mortgages, Loans); and *saving and investment* (Investment diversification, Bonds, Stock investments and Pensions). In addition, we covered the important domain of *everyday financial transactions* (Bank accounts, Payments) rather than Huston's fourth key domain of *resource protection*, because such everyday transactions comprise a basic domain of financial knowledge common to all countries, whereas resource protection varies internationally, in particular the availability of different insurance products.

Our starting point for a brief scale consisted of Lusardi's (2011) "big five" items. We applied an initial IRT which identified some limitations. We then applied further iterations of IRT to construct a new brief scale with sound psychometric properties. To develop an in-depth scale, IRT was applied iteratively to the pool of 50 questions until a final scale was constructed, consisting of 20 items covering the 10 topics and four broad financial domains described above. This scale was partitioned into parallel 10-item scales in two ways, one covering either easier or harder topics, and the other being parallel 10-item scales of equivalent difficulty broadly sampling the 10 topics.

Turning to the evaluation of the scales, our approach was to first assess their basic scale properties in terms of the IRT criteria explained in the next section. We then examined their construct validity by regressing demographic variables on each scale, since previous research has found that financial literacy varies with age, gender, level of education and income (Lusardi and Mitchell, 2011b; Nicolini, 2019). Finally, we compared how well various scales predicted survey responses to certain questions related to financial well-being. These items relate to three

of the four household financial management activities deemed critical by the US Federal Reserve (Hilgert *et al.*, 2003), namely: cash-flow management, saving, and credit management.

The remainder of this article is organized as follows. We next present an outline of IRT to elucidate the value of the technique in establishing more valid psychometric instruments. This is followed by the details of our survey of UK respondents. Next, the results section presents the two broad sets of analyses summarized earlier, that is, the construction and evaluation of the scales respectively. We conclude with a discussion of the extent to which we have been successful in our main aim of producing brief and in-depth financial literacy scales with sound psychometric properties. We also consider the possible uses of these scales, their limitations, and open questions for future research on the measurement of financial literacy.

1.1 │ An overview of two-parameter IRT

In line with Knoll and Houts' (2012) similar approach for a US sample, we applied two-parameter IRT analysis to a starting battery of 50 items from which to construct our scales. Broadly, two-parameter IRT is a means of assessing the contribution constituent items make to that scale's overall validity in determining one's level of a latent construct, such as ability. It thus serves as a quantitative means of determining how well a scale and the items therein measure the intended construct. Whereas in classical test theory measures of latent ability vary with the average difficulty of the test items, those resulting from IRT do not. IRT has been used to construct scales in several other domains such as numeracy (Weller *et al.*, 2013), reasoning (Primi *et al.*, 2017), and financial wellbeing (Consumer Finance Protection Bureau, 2015). As an in-depth explication of IRT is beyond the scope of this report (see Embretson and Reise, 2000, for a detailed breakdown), we focus our introduction to IRT on detailing the two parameters estimated and outlining the assumptions of two-parameter IRT models. For our purposes two-parameter IRT is preferable to the one-parameter model, which assumes equal item discrimination, and the three-parameter model, which adds a parameter to estimate rate of guessing. Instead of the latter, we reduce noise in the data due to guessing by including "do not know" and "prefer not to say" response options for all items. In fact, we found that the three-parameter model was not a significantly better fit to the data than the two-parameter model, which attests to the effectiveness of these steps to reduce guessing.

Assuming the items reflect the latent construct of interest, the first step of two-parameter IRT is to recode responses to binary form: correct = 1, incorrect = 0. We include "do not know" responses in the incorrect category, whereas "prefer not to say" is counted as missing data. IRT determines the probability of a correct response to an item as a function of one's ability (here, level of financial literacy), contingent upon: (a) the *discriminability* of the item (denoted by a); and (b) the *difficulty* of the item (denoted by b). Ability is assumed to follow a normal distribution with mean of 0 (and a *SD* of 1). As such, the *difficulty* parameter, b is the z-score reflecting the ability level at which the respondent would have a 50% chance of correctly answering the item; higher values for b indicate more difficult items. The *discriminability* parameter, a, reflects how well the item detects differences in ability level, with higher quantifications reflecting better-discriminating items. Figure 1 presents a graph illustrating these parameters, where each line represents an item characteristic curve.

In the figure, all three items share a difficulty parameter of $b = 0$, indicating that a person of average ability would score 50% on these three items. Relative to an item difficulty of $b = 0$, then, those at higher ability levels would thus be predicted to score more than 50%. Each item

FIGURE 1 Example illustrating item response theory parameters of difficulty (*b*), and discriminability (*a*)
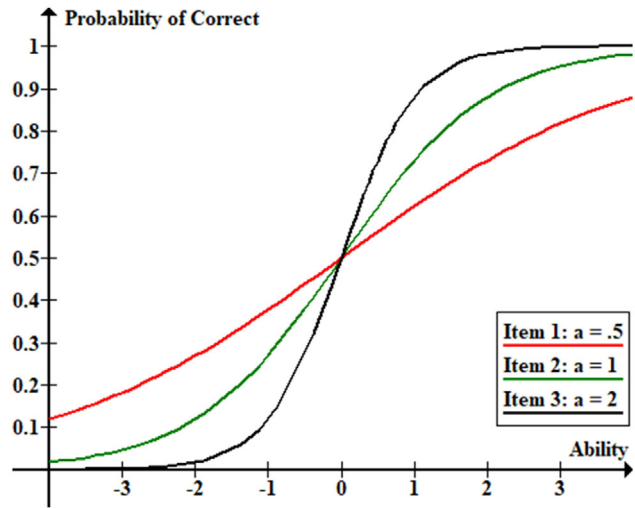
represented in the figure varies in discriminability as evidenced by the differing slopes of the three lines. An item characteristic curve with a less steep slope represents a lower value of *a*, indicating worse discriminability for that item. From the graph, the lowest-discriminating item will yield only marginal increases in probability of correct responding across different ability levels, thus indicating poor discriminability, whereas the item with the steepest slope and highest *a* has very good discriminability. In this work, we assume that values greater than 0.8 have acceptable discriminability (Embretson and Reise, 2000). In addition to the assumptions indicated above, IRT models assume the following: (a) invariance of item parameters across samples varying in characteristics such as age and latent ability; (b) local independence of items, in practice meaning that responses should not be highly correlated; and (c) unidimensionality, usually tested by confirmatory factor analysis and assessing the goodness of fit of the one-factor solution.

Scale construction using IRT is an iterative process, beginning with an initial scale (a scale consists of a set of items, each scored correct or incorrect, with the respondent's score being the total number of correct items). The scale properties that we used to evaluate the psychometric quality of the initial and subsequent scales are the following: (a) all scale *items* should have acceptable *discriminability* (i. e., *a* > 0.8); (b) scale *items* should be distributed in *difficulty* across the relevant ability range (generally from about $b = -1$ to $+1$, where average is 0); (c) *scales* should have satisfactory *overall item reliability* (e.g., marginal reliability, or Cronbach's alpha > .8); and (d) *scales* should have satisfactory *overall discriminability* across an appropriate range of respondent ability (e.g., -2 to $+2$ *SD* units for a general purpose scale). After the IRT analysis of the initial scale, items are dropped and/or replaced iteratively until no further improvements at the item or the scale level are achievable.

2 | METHOD

2.1 | Participants

On our behalf, the survey company Survey Sampling International implemented an online survey in May 2016 targeting a representative UK sample of adults who were paid £5 for their

participation. Prior to this, the study received ethical approval from the Leeds University Business School Research Ethics Committee (Application #: AREA15-103). The resulting sample of $N = 589$ usable responses had the following demographic characteristics: 53% female (female coded 1; male coded 0); mean age = 41–45 years old; 53% educated to university-level; mean monthly income = £1,000–£1,500. As expected from a national UK sample, there were no significant correlations between age and either gender or monthly income. However, consistent with national statistics, younger participants ($r = -.20$, $p < .001$, $n = 582$) and men ($r = -.12$, $p < .01$, $n = 573$) tended to have a higher level of education, and women tended to earn less than men ($r = -.33$. $p < .001$, $n = 519$). Finally, level of education and monthly income were positively related ($r = .36$, $p < .001$, $n = 524$).

2.2 | Questionnaire

The questionnaire began with background information about the research team, the nature and purpose of the survey, respondents' right to withdraw, and other such matters. Respondents were informed that there would be 101 questions in four sections: (a) questions about themselves; (b) 50 multiple choice "quiz" questions to assess how much they knew about financial matters; (c) questions about their financial behavior; and (d) three questions concerning their financial circumstances. As well as their participation fee, respondents were motivated by the promise, after completion, of a summary of their performance on the financial literacy questions.

As outlined earlier, our pool of 50 financial literacy questions comprised five questions of varying expected difficulty in each of 10 topics (see Appendix A). Two of the 10 topics concerned basic money transactions (Payments, Bank Accounts), while another two concerned basic properties of the money concept (Interest Rates, Inflation). A further two others concerned borrowing (Mortgages, Loans) and the remaining four concerned saving and investment: Investment diversification, Bonds, Stock Investment and Pensions. The 10 topics, all relevant to the financial well-being of citizens across contemporary market economies, were identified from a literature review (Nicolini, 2019), building on the five questions devised by Lusardi (2011). To ensure that individual items are similarly relevant, questions concerning specific national financial matters and specific financial products were excluded. The items were piloted and finalized with a sample of financial consumers selected from the Italian adult population with the support of a consumer union organization and have been further tested in five European countries. Decreasing correct response rates from Question 1 through to Question 5 in each of the 10 topics confirmed the increasing level of difficulty of the items across the five questions. All 50 items typically offered four or five response options: two or three specific responses (including one correct response), "Do not know", and "Prefer not to say". As mentioned earlier, for the purpose of IRT analyses, answers were binary coded as correct or incorrect, with "Do not know" responses included as incorrect, while "Prefer not to say" responses were treated as missing values.

Parts 3 and 4 of the survey asked numerous questions concerning respondents' preferences for, and use of, several types of financial products including bank accounts, types of payment used, types of investments and savings held, mortgages and other borrowing, retirement and planning, and about respondents' current financial circumstances. Analysis of most of these items is outside of the scope of the current report. The present study focused on three questions that are indicators of respondents' current financial well-being. As mentioned earlier, these

items speak to three of the four household financial management activities deemed critical by the US Federal Reserve (Hilgert *et al.*, 2003), namely: savings, credit management, and cash-flow management.

1. Precautionary savings
   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/No)
2. Credit management
   What is the total amount of your debts (not considering any mortgages on houses)? (Open response)
3. Cash flow management
   In a typical month, how difficult is it for you to cover your expenses and pay all your bills and obligations on time? (1—Not at all difficult, 4—Very difficult)

3 | RESULTS

We first describe our two new scales and how they were constructed, beginning with the 20-item in-depth scale. Prior to constructing this scale, we carried out an exploratory principal components analysis (PCA) to assess the dimensionality of the 10 subscales that comprise the topic structure of the 50 items. This led to the identification of two clusters of topics that differed in difficulty, which informed development of the in-depth scale. Next, we split the 20-item scale into two, parallel 10-item subscales which provide more efficient alternatives that may be more useful in some circumstances. Finally, we constructed a brief 5-item scale, beginning with Lusardi's (2011) five items, replacing two of them to improve the scale's psychometric properties. Following the construction sections, we present our evaluation of the new scales, in terms of their basic properties, how they relate to demographic variables and how they predict the reported financial well-being of our respondents.

3.1 | A new in-depth financial literacy scale

3.1.1 | Exploratory PCA

For each of the 10 subscales comprising the 50 items overall, we calculated the total number correct (out of 5) for each respondent that yielded complete data (see Table 1). Respondents who failed to answer all five items for a subscale, either due to not providing a response or choosing the "Prefer not to say" response option, in turn did not yield a score for that subscale. Table 1 indicates that on average the Payment subscale was the easiest, with the highest number correct, while the Bonds subscale was the most difficult.

Exploratory PCA was applied to the subscale scores to examine the dimensionality of the 10 financial topics (based on $N = 476$ respondents with no missing data). Initial tests showed that the data was suitable for PCA Varimax rotation: the Kaiser-Meyer-Olkin statistic indicated good sampling adequacy, KMO = .89; and Bartlett's test of sphericity was significant: $X^2(45) = 1{,}642$, $p < .001$. The key findings are presented in Table 1 according to factor loadings > .3. Two components with Eigenvalues >1 were observed, together accounting for 58% of variance in subscale scores. The communalities of the subscales were reasonable, ranging from .45 to .69. It

TABLE 1 Means, *SD*s of number correct, and principal components analysis of the 10 financial literacy subscales, UK adult sample, *N* = 589

| Topic | Mean | *SD* | *N* missing | Rotated coefficients | | Communalities |
				1	2	
Payments	3.65	1.74	27	.83		.69
Bank accounts	3.40	1.40	29	.82		.67
Inflation	3.08	1.72	32	.71		.59
Debt	2.11	1.28	22	.62	.37	.52
Interest rates	2.60	1.47	27	.61		.45
Bonds	1.03	1.27	41		.81	.68
Stock Inv.	1.74	1.13	33		.74	.61
In. Diversification	1.66	1.50	37	.39	.66	.59
Pensions	1.74	1.13	38	.40	.58	.49
Mortgages	2.37	1.50	41	.51	.49	.50

Note: Each subscale comprises five items. Missing responses comprised instances where no response was given, and all "Prefer not to say" responses. Varimax rotation, two-component solution: Factor loadings > .3 shown.

can be seen that the component structure is complex with respect to some subscales, with several loading high on one component and moderate on the other, and the Mortgage subscale loading moderately on both. Also, generally speaking, the more difficult subscales load higher on Component 2, and the easier ones load higher on Component 1.

Taking a factor loading cut-off of ~0.3, the two components can be interpreted loosely as comprising easier and harder domains of financial literacy, with the Mortgage subscale moderately related to both. However, although in PCA terms these are two dimensions, it is clear that conceptually they are clusters of topics at the lower and higher end of a single "difficulty" dimension. Nevertheless, having had this distinction highlighted by the PCA analysis, we decided to first apply IRT to construct easier and harder subscales of the in-depth scale, and then to combine them into the full scale.

3.1.2 | Easier and harder financial literacy subscales

In order to construct the easier and harder subscales, IRT analysis was first applied to each of the ten 5-item subscales pertaining to the discrete financial topics. From this, the three best-discriminating items of each topic were selected for further analysis, as a first step in ensuring that all items of the final scales would have a satisfactory degree of discriminability while retaining a set of items in each topic. The 15 items comprising the five easier topics (i.e., Interest Rates, Inflation, Bank Accounts, Payments and Loans) were then subject to further IRT analysis. On inspecting the results, we removed the item in each topic with the lowest discriminability and checked that the 10 remaining items were well distributed across the range of difficulty. On this basis we selected two items per domain to produce the 10-item easier subscale. The details of this scale are presented in Appendix B.

Selection of the 10 items comprising the harder subscale began as detailed for the easier scale by firstly focusing on the three best-discriminating items from each of the five harder

financial domains (Investment Diversification, Mortgages, Stock Investments, Pensions and Bonds). IRT analysis was applied to the 15-item scale comprising these items and the item in each domain with the lowest discriminability was removed. In this case IRT showed that the 10-item scale included two items with unsatisfactory discriminability. Consequently, these items, one from the Bond, and one from the Pension domain were replaced by Stock investments items to produce the final harder subscale. The details of this scale are also presented in Appendix B.

The in-depth scale was constructed by merging the easier and harder subscales to give a 20-item scale, which was then subjected to further analysis. To begin, we computed a tetrachoric correlation matrix for the 20 in-depth scale items and subjected this to principal component analysis (PCA) using maximum likelihood estimation with Varimax rotation. This analysis was conducted in order to confirm, based on an item-level analysis, the two-factor structure of the in-depth scale (i.e., the "easier" and "harder" components). The Kaiser-Meyer-Olkin statistic indicated good sampling adequacy, KMO = .72; and Bartlett's test of sphericity was significant: $X^2(190) = 7,361$, $p < .001$. Extracting two factors, both cumulatively accounted for 55% of variance. From the factor loadings presented in Table 2, we see that items 1–10—those pertaining to the "easier" scale"—for the most-part load more highly on the second factor, while items 11–20—those pertaining to the "harder" scale—load more highly on the first factor. IRT assumes a unidimensional underlying construct. The two-dimensional solution identified here is consistent with this, since the two dimensions are poles of the single dimension of difficulty.

We further scrutinized the in-depth scale by applying IRT, the results of which can be seen in Table 3. The results of the IRT show that, within this longer scale, the discrimination parameter for the Bond item was relatively low ($a = .72$). Nevertheless, this item was retained in order to preserve a good range of item difficulty, which, as the table shows, was in the range −1.5–2.2. Cronbach's alpha = .86 indicates excellent item reliability overall, and this cannot be improved by removing any items from the scale, which further strengthens the scale's position as targeted and reliable. Table 3 shows that there was a good distribution of easy and difficult items across the range, and the test characteristic curve (see Figure 2) shows that a person of average ability would score about 50% (10/20), and that the test would be most sensitive to differences in ability across a wide range of ability (−2.0–2.0). In terms of content, the new in-depth financial literacy scale assesses all the content areas of the financial topics of the survey, assessing each of the four broader domains with at least four items, as recommended by Huston (2010).

3.1.3 | Parallel 10-item subscales

In order to provide alternatives to the 13-item scale of Fernandes *et al.* (2014) and the 10-item scale of Houts and Knoll (2019) with sound psychometric properties but a more representative content of domains of financial literacy, we identified two parallel subscales of the 20-item scale, consisting of either the even (10a) or the odd numbered items (10b). The items of each subscale can be identified by the item numbers of the in-depth scale in Table 3. We assessed their psychometric properties, beginning with IRT analyses. For subscale 10a, while the bond item's discrimination parameter was low ($a = .60$), removing it only marginally improved the subscale's reliability from .75 to .76. As such, we retained the item to maintain parity with the second subscale and to offer a good range of item difficulty and breadth of coverage. Item

TABLE 2 Item-level factor loadings and communalities for two-factor principal component analysis of the 20-item in-depth financial literacy scale

Item	Rotated component loadings		Communalities
	1	2	
1. Interest1	.33	.61	.71
2. Interest4	.48	.37	.60
3. Inflation1	.45	.67	.83
4. Inflation3	.52	.62	.82
5. BankAccount1	.01	.86	.81
6. BankAccount2	.06	.79	.77
7. Payments1	.01	.85	.92
8. Paymets2	.05	.76	.84
9. Loans2	.44	.43	.56
10. Loans5	.45	.57	.67
11. InvestDiv1	.66	.20	.77
12. InvestDiv2	.74	.34	.88
13. InvestDiv5	.65	.50	.82
14. Mortgage3	.49	.20	.52
15. Mortgage4	.60	.03	.56
16. Stock1	.66	.08	.53
17. Stock3	.58		.61
18. Stock4	.64	.29	.62
19. Pension3	.63	.08	.59
20. Bond3	.54	.07	.46

reliability of subscale 10b was also very good, with Cronbach's alphas of .75. A Guttman's split-half correlation of .68 was observed, demonstrating the internal consistency of the in-depth scale.

3.2 | A new brief financial literacy scale

As a starting point for the construction of a brief, five-item scale, we applied IRT to a scale comprising the five items devised by Lusardi (2011). Table 4 presents the main results of the IRT analysis of this scale. It can be seen that three of the items were relatively easy ($b = -1.17$ to -0.31); one was harder, and the other was very difficult ($b > 9$). The table also shows that although the discriminability of four items was acceptable, that of the most difficult item was not ($a = .15$). This was confirmed with a Cronbach test of item reliability: alpha = .52 for the 5-item scale, which improved to alpha = .56 when the difficult item was removed. However, the test characteristic curve, presented in Figure 3a, shows that the test as a whole exhibited good sensitivity over a reasonable range of ability (-2.0–1.0). It can be seen that the curve is steepest in the ability range, $-.5$–0.0, reflecting a bias toward easier items.

TABLE 3 Discrimination and difficulty parameter estimates for the in-depth scale ($N = 507$)

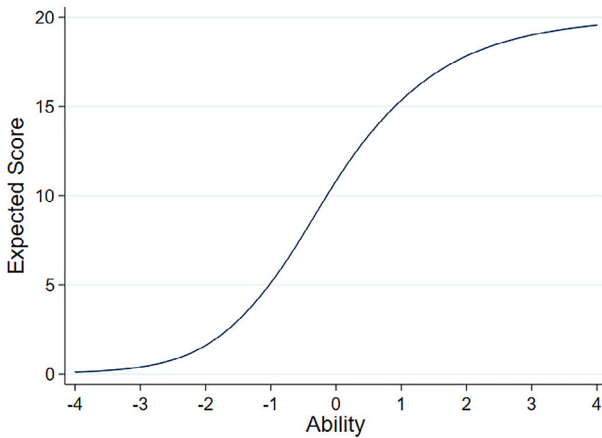| Topic/label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Interest rates *Interest1*	78	Suppose you had £100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? ***More than £102***, *Exactly £102, Less than £102*	−1.16	2.21
	45	Suppose you had £100 in a savings account and the interest rate was 10% per year. After 2 years, how much do you think you would have in the account if you left the money to grow? *£110, £120, **£121***	.09	1.38
Inflation *Inflation1*	58	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, Exactly the same as today,* ***Less than today***	−.38	2.86
Inflation3	57	Imagine depositing £100 in your savings account. If after 1 year the balance on your account (included the interest) is £104 and the inflation rate during the last year was 5%, How rich do you think you would be compared 1 year before? *More rich, Exactly as rich,* ***Less rich***	−.35	2.87
Bank accounts *BankAccount1*	77	You have an "overdraft" in your bank account if... ***You use more money than is in your account***, *You receive interest on your deposit from the bank, You pay by cheques*	−1.25	1.89
BankAccount2	73	What do you NOT need to access to your bank account in the case of e-banking? *A username and password, A device connected to the internet (computer, tablet, etc.),* ***An ID Card***	−1.07	1.71
Payments *Payments1*	83	What kind of card lets you buy something now and pay for it in the future? *Debit card,* ***Credit Card***, *Pre-paid card*	−1.57	2.04

(Continues)

TABLE 3 (Continued)

Topic/label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Payments2	63	If you have no money on your bank account and you cannot overdraw, which of the following payment options do you have to buy something in a shop? *Cheques [without any cheque-guarantee card option], Debit card, **Pre-paid card***	−.74	1.29
Loans Loans2	57	Everything else equal, if the maturity of a mortgage is longer the installments will be... *The same, **Smaller**, Bigger*	−.43	1.47
Loans5	50	Which of the following pay back options for a £100 debts shows the highest APR? *£**102 after 1 week**, £105 after 1 month, £110 after 2 months*	−.17	1.86
Investment div. InvestDiv1	41	Buying a single company's stock usually provides a better return than buying a mutual stock fund *True, **False***	.26	1.31
InvestDiv2	24	Compared with an investment in stocks, the risk in investing in stock mutual funds is... *Equal, Greater, **Smaller***	.87	1.99
InvestDiv5	48	Suppose you invested 1 year ago £1,000 in a well-diversified stock mutual fund. If the performance of the stock market index (the one where the fund invested its money) in the same period has been +5%, how much returns do you expect from your investment? *Less than £5, More than £500, **Around £50***	−.11	2.37
Mortgages Mortgage3	43	In a 15-year mortgage which of the following options will minimize the total interest paid over the life of the loan? *Annual payments, Semi-annual payments, **Monthly Payments***	.20	1.05
Mortgage4	31	Consider a 15-year mortgage where you can choose between monthly payments and 6-month payments. If after 5 years you would like to close the mortgage in advance, which one will make you pay less?***Monthly***	.97	.95

TABLE 3 (Continued)

Topic/label	Percent correct	Item	b (Difficulty)	a (Discrimination)
		payment mortgage, *6-monthly payment mortgage, Both would make people pay the same amount*		
Stock invest. *Stock1*	31	Ignoring the case of default of the issuer, which of the following investment products guarantees the reimbursement of the invested capital? *Stocks,* ***Bonds****, Stock mutual funds*	.91	1.09
Stock3	19	Which is the standard measure for the default risk of a bond issuer?***Rating****, APR, Benchmark*	1.66	.98
Stock4	34	If the same company issues short-term and a long-term bonds, typically the interest rate of the long-term bond is... ***Higher****, The same, Lower*	.51	1.46
Pensions *Pension3*	44	Do you think that the performance of financial markets can affect the performance of a pension fund?*No—they are unrelated****Yes—they are positively related (i.e., when the market value goes up so too does the pension fund value)*** *Yes—they are negatively related (i.e., when the market value goes down so too does the pension fund value)*	.10	1.03
Bonds *Bonds3*	19	If you expect a drop in interest rates what is a good investment strategy? ***Buy a bond,*** *Sell a bond, Hold a bond*	2.21	.72

Note: Response options for each item are provided in italics, with the correct response in bold. All items also provided "Do not know" and "Prefer not to say" response options.

The results of the above IRT analysis suggest that replacing two items might result in an improved short scale, with better item discriminability, better distribution of difficulty and sensitivity over a wider range of ability. The bond item was replaced with a savings item that was relatively difficult but with satisfactory discriminability, and the relatively easy mortgage item was replaced with a more difficult one to give a better distribution of difficulty overall. As with our in-depth scale, we began our analysis of this new brief scale by subjecting tetrachoric correlation coefficients for item responses to principal component analysis (PCA) using maximum likelihood estimation with Varimax rotation. The analysis indicated good sampling adequacy, with a Keyser-Meyer-Olkin statistic = .72; and Bartlett's test of sphericity was significant: $X^2(10) = 604.86$, $p < .001$. A single factor accounted for 48% of variance. We continued our analysis by applying IRT to the new brief scale. Table 5 shows that the difficulty of the five items is reasonably well distributed across the range from −1 to 1, and all items have adequate
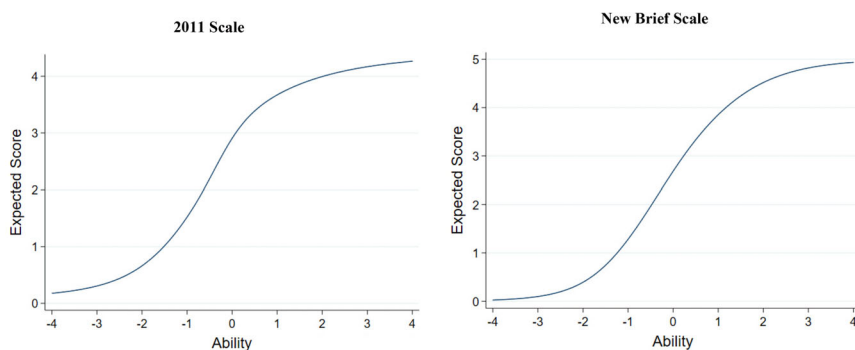
FIGURE 2 Test characteristic curve for the in-depth scale

TABLE 4 Discrimination and difficulty parameter estimates for a 5-item financial literacy scale comprising Lusardi's (2011) Items: ($N = 565$)

Topic/ label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Interest rates *Interest1*	78	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***, *Exactly £102, Less than £102*	−1.17	1.84
Inflation *Inflation1*	58	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, Exactly the same as today,* ***Less than today***	−.31	3.02
Investment div. *InvestDiv1*	41	Buying a single company's stock usually provides a safer return than a stock mutual fund. *True,* ***False***	.39	.90
Mortgages *Mortgage1*	63	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****, False*	−.75	1.01
Bonds *Bond1*	19	If interest rates rise, what will typically happen to bond prices? *They will rise,* ***They will fall****, They will stay the same*	9.13	.15

Note: Response options for each item are provided in italics, with the correct response in bold. All items also provided "Do not know" and "Prefer not to say" response options.

2011 Scale   **New Brief Scale**

FIGURE 3 Test characteristic curves for the 5-item scale of Lusardi and Mitchell's (2011a; 2011b) items, and the new brief scale

discriminability. Furthermore, the test characteristic curve (Figure 3b) shows a more balanced pattern compared to that for the initial scale, with no ceiling or floor effects evident. Also, the figure shows that the scale is sensitive across a wider range of ability (−2.0–2.0). The overall reliability, as assessed by a Cronbach's alpha of .59, is also a minor improvement on the initial scale. In terms of content, the new brief financial literacy scale samples knowledge and use of finance-related information in three broad financial domains: concept of money (Interest Rates, Inflation); savings and investment; and borrowing (Mortgages).

3.3 | Socio-demographic variations in financial literacy

Multiple regression analyses were carried out with the in-depth and brief financial literacy measures as criterion, together with the scale comprising Lusardi's (2011) five items. The predictor variables for each analysis were the socio-demographic variables of age, gender, level of education and monthly income. The analyses were on the subset of participants for whom complete data was available for the variables being analyzed, with sample sizes varying from $N = 437$ to 487. The socio-demographic predictors accounted for a significant proportion of variation on each measure, with adjusted R-square varying from .16 to .20. These and the beta coefficients for each predictor are shown in Table 6. The latter show the independent contribution of each predictor variable to the variance explained, controlling statistically for the other socio-demographic variables, some of which are significantly correlated in this sample.

In all analyses, age, education and monthly income were significant predictors of financial literacy. The strongest predictor was age, with older adults scoring higher on all measures. For example, on the more sensitive in-depth scale, the average number correct rose steadily from 8.25/20 for participants in their twenties, rising by about one item per decade to 13/20 for the over sixties. The next strongest predictor was education, such that financial literacy was higher with higher education levels. Next, higher monthly income was associated with higher financial literacy. Finally, the regression analyses found that gender was a significant predictor on the brief scale and the scale comprising Lusardi's (2011) items, but not on the in-depth scale. However, a univariate analysis identified a significant effect of gender on the in-depth scale, with male participants scoring 11.2/20 on average, compared to 9.7/20 for females. Overall, the

TABLE 5 Discrimination and difficulty parameter estimates for a new brief scale ($N = 563$)

Topic/ label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Interest *Interest1*	78	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**, Exactly £102, Less than £102*	−1.17	1.82
Inflation *Inflation1*	58	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, Exactly the same as today, **Less than today***	−.35	1.94
Investment *InvestDiv1*	41	Buying a single company's stock usually provides a safer return than a stock mutual fund. *True, **False***	.32	1.24
Mortgages *Mortgage3*	43	In a 15-year mortgage which of the following options will minimize the total interest paid over the life of the loan? *Annual payments, Semi-annual payments, **Monthly Payments***	.91	1.15
Stock invest. *Stock1*	31	Ignoring the case of default of the issuer, which of the following investment products guarantees the reimbursement of the invested capital? *Stocks, **Bonds**, Stock mutual funds*	.89	1.01

Note: Response options for each item are provided in italics, with the correct response in bold. All items also provided "Do not know" and "Prefer not to say" response options.

	Financial literacy scale		
	In-depth	**Brief**	**Lusardi 5**
Age	.38***	.35***	.31***
Gender	−.07	−.10*	−.09*
Income	.12*	.12*	.15***
Education	.14**	.16***	.15***
R^2	.18	.20	.17
N	437	487	487

TABLE 6 Linear regression analyses predicting financial literacy scores from four socio-demographic variables

Note: *$p < .05$; **$p < .01$; ***$p < .001$ Standardized coefficients shown. Gender coding, male = 0, female = 1.

adjusted R^2 measure shows that a similar proportion of variance in each measure was accounted for by this set of four socio-demographic variables.

3.4 | Predicting financial well-being from financial literacy

The distributions of responses to the three questions related to financial well-being were as follows (excluding "do not know" responses): 58% ($N = 532$) stated that they had set aside emergency or rainy day funds ("yes" coded 1, "no" coded 0); 57% ($N = 509$) stated that they had no debts (some debts coded 1, no debts coded 0); and for how difficult people felt it was to cover monthly expenses ($N = 554$): 7.6% responded "Very difficult" (coded 1); 11.7% responded "Difficult" (coded 2); 31.9% responded "A bit difficult" (coded 3); and 48.7% responded "Not at all difficult" (coded 4).

As one would expect these indicators are correlated: those with precautionary savings tended to have less difficulty covering monthly expenses ($r = .34$, $N = 520$, $p < .001$); and those with no debts tended to have less difficulty ($r = -.29$, $N = 491$, $p < .001$). Not surprisingly, having precautionary savings and having debts were inversely related ($r = -.16$, $N = 479$, $p < .001$).

In order to assess the extent to which the new financial literacy scales predict the first two of the above indicators of financial well-being (not having debts, and having precautionary savings), we conducted separate hierarchical logistic regressions. For each logistic regression, at Step 1 the four socio-economic variables described earlier were entered, and at Step 2 scores on one of the financial literacy scales was entered. Responses to the third question, how difficult people felt it to cover monthly expenses, were analyzed by hierarchical linear regression. Again, socio-economic variables were entered at Step 1, and financial literacy scores at Step 2. The results of these analyses are summarized in Tables 7 and 8. The regression coefficients indicate the independent contribution of each variable to the prediction in the final model. The R^2 value, or equivalent, estimates the proportion of variance explained after all predictors are included in the regression, while the R^2 change values estimate the additional contribution of the financial literacy measure when added to the socio-demographic variables.

Taking the linear regression first, presented in Table 7, it can be seen that the full set of predictors account for over 10% of variance in ratings of difficulty in covering expenses. The main finding is that higher financial literacy is significantly associated with less difficulty in meeting monthly expenses, after controlling for socio-demographic factors. For the in-depth scale, about half of the variance explained is due to the addition of the financial literacy measure, while this is around a third in the case of the brief scale. With respect to the socio-demographic variables, only personal monthly income is significantly related to difficulty in covering expenses. The table shows that the beta coefficients for the brief scale are a little lower than that for the in-depth scale. For comparison, the results for the scale comprising Lusardi's (2011) items are also shown. It can be seen that this beta coefficient is a little lower still.

Turning to whether respondents report having precautionary savings or debts, the findings are presented in Table 8. Regarding precautionary savings, financial literacy yields significant predictive value when measured on both new scales, while also controlling for socio-demographic factors. Concerning whether people have debts, both scales significantly predict this, indicating that better performance on these scales is associated with not holding debts, controlling for socio-demographic factors. Age and income are the key sociodemographic factors, each of which is significantly predictive of having precautionary savings, and also of holding debt. To the extent that such sociodemographic factors capture an individual's life-stage and economic

TABLE 7 Linear hierarchical regression predicting difficulty in covering monthly expenses from socio-demographic variables and measures of financial literacy (in-depth and brief scales)

	Difficulty covering monthly expenses		
Demographic factors			
Age	−.01	−.06	.07
Gender	.02	.01	.00
Income	−.16**	−.16**	.18***
Education	.08	.07	−.08
Financial literacy			
In-depth	.29***		
Brief		.23***	
LM5			.18**
R^2	.13	.11	.09
R^2 change	.06***	.04***	.03***
N	420	468	465

Note: *$p < .05$; **$p < .01$; ***$p < .001$. Standardized coefficients shown. Gender coding, male = 0, female = 1.

TABLE 8 Binary hierarchical logistic regressions of whether respondents have precautionary savings or debts predicted by socio-demographic variables and measures of financial literacy

	Precautionary savings				Debts			
Demographic factors	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
Age	−.10*	.90	.11**	.91	−.90*	.92	−.13**	.88
Gender	−.24	1.28	−.25	1.27	−.01	1.00	.11	1.11
Income	.19**	0.85	.17***	.83	−.16**	1.18	−.17**	1.19
Education	.14	0.82	.20	.87	−.16	1.17	−.10	1.10
Financial literacy								
In-depth	.07**	.85			−.08**	.93		
Brief			.17*	.93			−.14*	.87
R^2	.10***		.09***		.08***		.08***	
R^2 change	.02**		.01**		.02**		.01	
N	411		456		385		432	

Note: *$p < .05$; **$p < .01$; ***$p < .001$.

circumstances, financial literacy is still found to significantly predict the tendency to hold precautionary savings (positive) and debt (negative).

4 | DISCUSSION

Beginning with a pool of 50 survey questions, we applied IRT to construct and evaluate new brief and in-depth financial literacy scales. The pool included five widely used items devised by

Lusardi and colleagues (Lusardi, 2011) on Interest rates, Inflation, Investment diversification, Mortgages and Bonds. Many important and influential studies of financial literacy have used some, all, or slight variations of these items (e.g., Lusardi and Tufano, 2009; Lusardi and Mitchell, 2011a; 2011b; 2014; van Rooij et al., 2011a; 2011b; Klapper et al., 2015). That they have been shown to predict various financial outcomes is testament to their utility. However, there is relatively little known about how well they perform as accurate assessments of the financial literacy construct. We contributed to this by applying IRT to a UK sample's responses to a scale of these items. This identified some limitations, including rather low item reliability, partly due to one item with poor discriminability. We showed that if this and another item were replaced the resulting scale had improved psychometric properties. The items of this new brief scale were reasonably well distributed in difficulty and the scale was sensitive to ability differences across a wider range. In this instance, then, applying IRT has offered insights as to how Lusardi's (2011) tried-and-trusted items perform, and allowed a means of offering an improved brief scale that retains three of the original items, but which is better able to discriminate between ability levels. We note, however, that although the improved scale has slightly better item reliability, it still yields a relatively low Cronbach's alpha. This is generally the case for brief scales, which, we suggest, should only be used if brevity is really necessary. Notwithstanding this, we have fulfilled the first part of our main aim, to produce a useful, brief financial literacy scale with internationally relevant items and satisfactory psychometric properties.

We also applied IRT iteratively to our pool of items to produce an in-depth, 20-item scale, also psychometrically sound, covering four broad financial domains. Huston (2010), in an extensive review and synthesis of financial literacy assessments, argued that valid scales should seek to include three to five items for each core domain of financial literacy. Our in-depth scale meets this criterion; it presents five items across four domains, three of which are in line with those identified by Huston (2010): the concept of money, including inflation and interest; borrowing; saving and investment. Our scales also include a fourth important domain, everyday money transactions, with items to assess more basic, everyday financial knowledge concerning bank accounts and modes of payment, typically overlooked on such scales. With 1.5 million UK consumers being unbanked (Financial Inclusion Commission, 2015) and 24.2 million US households, underbanked (FDIC, 2017), it would appear apt to include items that tap fundamental, everyday knowledge rather than focus only on topics of knowledge that may tacitly assume that respondents already have such a level of financial engagement; previous research has observed that access to financial services does influence financial literacy (Atkinson et al., 2006).

We showed that the in-depth scale is sensitive across a relatively wide and balanced range of ability. It would be useful for a representative national sample when a longer scale is feasible and a more sensitive measure covering a broad range of topic areas is required. We also evaluated two 10-item subscales of the in-depth scale, both with sound psychometric properties. All these longer scales had very good item reliability (Cronbach's alphas > .75) and were sensitive to changes in ability across a useful range. Furthermore, the items of each scale had at least satisfactory discriminability. This fulfilled the second part of our main aim, which was to apply current good practice in scale construction to produce an in-depth financial literacy scale for a UK population, thereby extending Knoll and Houts' (2012) similar approach for US respondents. Furthermore, we also produced two parallel 10-item subscales that match the efficiency and cover a wider range of topics than both the 10-item scale of Houts and Knoll (2019) and the 13-item scale of Fernandes et al. (2014).

Although our scales cover a wide range of topics across four key domains (the basic concept of money, borrowing, saving and investment, and basic money transactions), our scales do not

cover the other key domain identified by Huston (2010), namely, resource protection, including insurance. Future research on financial literacy measurement should explore this important domain in greater depth.

Turning to the validation of the scales, we first found that age, education and monthly income were predictive of performance on both the brief and in-depth scales. Previous research, stratified by age groups, has indicated financial literacy follows an inverted U-shape, with average performance rising across age ranges until the early sixties when it begins to decline (Finke et al., 2016). Although we also identified a steady rise in average performance across the earlier decades, we did not see a fall after Age 60. However, we did not investigate the specific age of participants, merely age bands with the oldest being greater than 65 years.

Regarding education, while Klapper et al.'s (2015) global study of financial literacy indicates a generally positive trend between these factors, Lusardi and Mitchell (2011b), in another global study of financial literacy, offer some caution that education is "...not a perfect proxy for financial literacy". This may explain our finding that although education level was predictive of scores on all scales, the correlations were low in range. It is perhaps notable that in England—which accounts for 84% of the UK population—financial education in schools only became a compulsory part of the national curriculum in 2014, which might help explain why education yields only minor direct predictive value. The curriculum now specifically teaches teenage children about a range of issues assessed in our developed scales including interest, credit, forms of payment, inflation, and savings and investments.

Regarding income, the hypothesis that this positively related to financial literacy has been supported in several previous studies, for example, Lusardi and Mitchell (2011a), Honekamp (2012), and Knoll and Houts (2012).

We also identified a significant gender gap, although this was not significant on the in-depth scale after controlling for age and education. This may be because this scale includes a more balanced range of topics, some of which may have a smaller financial literacy gender gap. The existence of a gender gap, with females scoring systematically lower than males in financial literacy tests, is a quite common result in financial literacy studies, regardless of the items used to assess financial literacy and the assessment methodology. For instance, females perform significantly worse than males on financial literacy scales as reported by Sekita (2011) in a study on Japan, by Van Rooij et al. (2011a; 2011b) in the Netherlands, and by Honekamp (2012) in Germany. Our findings indicate that future research clarifying the topics where the gender gap is highest would be useful.

Finally, we compared how well the various scales predicted certain indicators of financial well-being. These indicators relate to three of the four household financial management activities deemed critical by the US Federal Reserve (Hilgert et al., 2003), namely: cash-flow management, saving, and credit management. The in-depth scale yielded significant predictive validity of three important financial wellbeing indicators (precautionary saving, debt, and making ends meet on a monthly basis), out-performing the scale comprising the widely used five items devised by Lusardi (2011) for two of these outcomes. This corroborates previous research which has linked lower financial literacy particularly with experiencing financial instability and cash flow management issues (Disney and Gathergood, 2011). This is important, not least because more recent research suggests that financial wellbeing (represented by money management stress and financial stability) is directly predictive of general wellbeing (Netemeyer et al., 2017). Developing more effective assessments of financial literacy, then, is an important endeavor given the significant applied utility that financial literacy holds as a construct.

In sum, as assessed against a UK national sample, our new financial literacy scales exhibited sound psychometric properties at both the scale and item levels. Furthermore, our initial validation tests showed that they are sensitive to differences in financial literacy across key socio-demographic variables. Finally, the main substantive contribution of our study was to show that financial literacy, as measured by both in-depth and brief scales, significantly predicts key aspects of the current financial well-being of a UK national sample.

In establishing a series of scales containing items that have systematically been determined to yield a more rigorous assessment of financial literacy for UK respondents, we hope to encourage similar practice among international colleagues. Global studies have shown that, broadly, levels of financial literacy vary across nations (Atkinson and Messy, 2011; Klapper *et al.*, 2015). However, a key issue remains that the appropriateness of the items used for the samples studied does not form part of these comparisons. To that end, we agree with Nicolini *et al.* (2013) that there "...is a need to co-ordinate, if not standardize, financial literacy assessments across countries" (p. 695) such that scales are more tailored to the national context; in other words, such that scales assess knowledge we might reasonably expect the nationality of the sample to need to know. The use of IRT to construct scales can facilitate precisely such an effort through the shared use of common "anchor" items with known existing parameter values. We concur with Huston (2010) and others (Knoll and Houts, 2012; Hauff *et al.*, 2020) in advocating this approach.

ACKNOWLEDGMENTS

ORCID

Rob Ranyard https://orcid.org/0000-0003-4293-9510

ENDNOTES

[1] As explained later, we replaced Huston's (2010) domain of resource protection with the domain of everyday money transactions.

[2] Initially this topic was labeled "Insurance and retirement" (Nicolini, 2019). However, we excluded the one insurance item, leaving four concerning pension knowledge.

REFERENCES

Abreu, M. and Mendes, V. (2010) Financial literacy and portfolio diversification. *Quantitative Finance*, 10(5), 515–528.

Atkinson, A. and Collard, S. (2006) Levels of financial capability in the UK: results of a baseline survey. *Financial Services Authority Consumer Research Paper*, (47), 1–144. http://www.pfrc.bris.ac.uk/publications/Reports/Fincap_baseline_results_06.pdf.

Atkinson, A. and Messy, F.-A. (2011) Assessing financial literacy in 12 countries: an OECD/INFE international pilot exercise. *Journal of Pension Economics and Finance*, 10(4), 657–665.

de Bruin, B., Wändi, W.V., Downs, J.S., Fischhoff, B., Topa, G. and Armantier, O. (2010) Expectations of inflation: the role of demographic variables, expectation formation, and financial literacy. *Journal of Consumer Affairs*, 44(2), 381–402.

Consumer Finance Protection Bureau. (2015) *CFPB financial wellbeing scale: scale development and technical report*, http://www.sjdm.org/dmidi/files/CFPB_Financial_Well-Being_Scale_Technical_Report.pdf.

Disney, R. and Gathergood, J. (2011) *Financial literacy and indebtedness: new evidence for UK consumers*. University of Nottingham, https://ssrn.com/abstract=1851343.

Embretson, S.E. and Reise, S.P. (2000) *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Federal Deposit Insurance Corporation (FDIC). (2017) *FDIC National Survey of Unbanked and Underbanked Households*, https://www.fdic.gov/householdsurvey/.

Fernandes, D., Lynch, J.G., Jr., Netemeyer, R.G. and G, R. (2014) Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8), 1861–1883.

Financial Inclusion Commission. (2015) *Improving the financial health of the nation*, http://www.financialinclusioncommission.org.uk/pdfs/fic_report_2015.pdf.

Finke, M.S., Howe, J.S. and Huston, S.J. (2016) Old age and the decline in financial literacy. *Management Science*, 63(1), 213–230.

Grimes, P.W., Rogers, K.E., Smith, R.C. and C, R. (2010) High school economic education and access to financial services. *Journal of Consumer Affairs*, 44(20), 317–335.

Hauff, J.C., Carlander, A., Gärling, T. and Nicolini, G. (2020) Retirement financial behaviour: how important is being financially literate? *Journal of Consumer Policy*, 1–22.

Hilgert, M.A., Hogarth, J.M. and Beverly, S.G. (2003) Household financial management: the connection between knowledge and behavior. *Federal Reserve Bulletin*, 89, 309–322.

Honekamp, I. (2012) *Financial literacy and retirement Savings in Germany*. Networks Financial Institute at Indiana State University. Working Paper number: 2012-Wp-03.

Houts, C.R. and Knoll, M.A.Z. (2019) The financial knowledge scale: new analyses, findings, and development of a short form. *Journal of Consumer Affairs*, 54, 775–800. https://doi.org/10.1111/joca.12288.

Hung, A., Parker, A.M. and Yoong, J. (2009) *Defining and measuring financial literacy* (No. 708), 15 February, 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.

Klapper, L., Lusardi, A. and Oudheusden, P.V. (2015) *Financial literacy around the world*. Standard and Poor's Ratings Services Global Financial Literacy Survey, http://media.mhfi.com/documents/2015-Finlit_paper_17_F3_SINGLES. pdf.

Knoll, M.A. and 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.

Lusardi, A. (2011) *Americans' financial capability*. National Bureau of Economic Research. Report number: w17103, https://www.nber.org/papers/w17103.

Lusardi, A. and Mitchell, O.S. (2007) *Financial literacy and retirement planning: new evidence from the rand American life panel*. Michigan Retirement Research Center Research. Paper Number: WP 2007–157, https://ssrn.com/abstract=1095869 or https://doi.org/10.2139/ssrn.1095869.

Lusardi, A. and Mitchell, O.S. (2011a) Financial literacy and retirement planning in the United States. *Journal of Pension Economics and Finance*, 10, 509–525.

Lusardi, A. and Mitchell, O.S. (2011b) Financial literacy around the world: an overview. *Journal of Pension Economics and Finance*, 10(4), 497–508.

Lusardi, A. and Mitchell, O.S. (2014) The economic importance of financial literacy: theory and evidence. *Journal of Economic Literature*, 52(1), 5–44.

Lusardi, A. and Tufano, P. (2009) Teach workers about the perils of debt. *Harvard Business Review*, 87(11), 22–24.

Netemeyer, R.G., Warmath, D., Fernandes, D. and Jr, J.G.L. (2017) How am I doing? Perceived financial well-being, its potential antecedents, and its relation to overall well-being. *Journal of Consumer Research*, 45, 68–89.

Nicolini, G. (2019) *Financial Literacy in Europe: Assessment Methodologies and Evidence from European Countries*. New York: Routledge.

Nicolini, G., Cude, B.J. and Chatterjee, S. (2013) Financial literacy: a comparative study across four countries. *International Journal of Consumer Studies*, 37, 689–705.

Organization for Economic Cooperation and Development (OECD). (2016) *OECD/INFE International Survey of Adult Financial Literacy Competencies*, https://www.oecd.org/daf/fin/financial-education/OECD-INFE-International-Survey-of-Adult-FInancial-Literacy-Competencies.pdf.

President's Advisory Council on Financial Literacy (PACFL). (2008) *Annual Report to the President*. President's Advisory Council on Financial Literacy, http://www.treas.gov/offices/domestic-finance/financial-institution/fineducation/council/index.shtml.

Primi, C., Morsanyi, K., Donati, M.A., Galli, S. and Chiesi, F. (2017) Measuring probabilistic reasoning: the construction of a new scale applying item response theory. *Journal of Behavioral Decision Making*, 30(4), 933–950.

Remund, D.L. (2010) Financial literacy explicated: the case for a clearer definition in an increasingly complex economy. *Journal of Consumer Affairs*, 44(2), 276–295.

van Rooij, M., Lusardi, A. and Alessie, R. (2011a) Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2), 449–472.

van Rooij, M., Lusardi, A. and Alessie, R. (2011b) Financial literacy and retirement planning in The Netherlands. *Journal of Economic Psychology*, 32, 593–608.

Sekita, S. (2011) Financial literacy and retirement planning in Japan. *Journal of Pension Economics and Finance*, 10(4), 637–656.

Stango, V. and Zinman, J. (2009) What do consumers really pay on their checking and credit card accounts? Explicit, implicit, and avoidable costs. *American Economic Review*, 99(2), 424–429.

Weller, J.A., Nathan, F.D., Tusler, M., Mertz, C.K., Burns, W.J. and Peters, E. (2013) Development and testing of an abbreviated numeracy scale: a Rasch analysis approach. *Journal of Behavioral Decision Making*, 26(2), 198–212.

APPENDIX A: 50 FINANCIAL LITERACY QUESTIONS IN 10 TOPICS (RIGHT ANSWERS IN BOLD)

Interest Rates

Interest 1 (from Lusardi, 2011)

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

Exactly £102

Less than £102

Do not know

Prefer not to say

Interest 2

Suppose you borrow £200 for 2 years from a bank that asks you to pay interest at 2% a year. After 2 years, how much do you think you would have to pay to settle your debt?

More than £204

Exactly £204

Less than £204

Do not know

Prefer not to say

Interest 3

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 £110

Exactly £110

Less than £110

Do not know

Prefer not to say

Interest 4

Suppose you had £100 in a savings account and the interest rate was 10% per year. After 2 years, how much do you think you would have in the account if you left the money to grow?

£110

£120

£121

Do not know

Prefer not to say

Interest 5

What is the interest rate (the APR—annual percentage rate) of a loan where a lender gives you £100 and you have to pay back £110 after 1 month?

10%

Around 120%

More than 200%

Do not know

Prefer not to say

Inflation

Inflation 1 (from Lusardi, 2011)

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

Exactly the same amount as today

Less than today

Do not know

Prefer not to say

Inflation 2

Imagine that the interest rate on your savings account was 4% per year and inflation was 4% per year. After 1 year, how much would you be able to buy with the money in this account?

More than today

Exactly the same amount as today

Less than today

Do not know

Prefer not to say

Inflation 3

Imagine you deposit £100 in your savings account. If after 1 year the balance on your account (including the interest) is £104 and the inflation rate during the last year was 5%, how rich do you think you are compared to 1 year before?

More rich

Exactly as rich as 1 year before

Less rich than 1 year before

Do not know

Prefer not to say

Inflation 4

If your bank will pay 4% a year on the £100 balance in your savings account, how much inflation would you expect if you think that you will maintain your purchasing power after 2 years?

0% a year

No more than 4% per year

No more than 8% per year

Do not know

Prefer not to say

Inflation 5

In the last year the inflation rate was 20%. If today the price of a London Underground ticket is £2.40, how much was the price 1 year ago?

£1.92

£2.00

£0.48

Do not know

Prefer not to say

Mortgages

Mortgage 1 (from Lusardi, 2011)

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

False

Do not know

Prefer not to say

Mortgage 2

Compared to a 15-year mortgage, if you want to reduce the total interest paid over the life of a loan, which of the following mortgages would you have to prefer?

A 30-year mortgage

A 20-year mortgage

A 10-year mortgage

Do not know

Prefer not to say

Mortgage 3

For a 15-year mortgage which of the following options will minimize the total interest paid over the life of the loan?

Annual payments

Semi-annual payments

Monthly payments

Do not know

Prefer not to say

Mortgage 4

Consider a 15-year mortgage where you can choose between monthly payments and 6-month payments. If after 5 years you would like to close the mortgage in advance, which one would result in the lowest final payment?

Monthly payment mortgage

6-month payment mortgage

Both of them would make people pay the same amount

Do not know

Prefer not to say

Mortgage 5

To reduce the payments (installments) of a 15-year mortgage with 6-month payments, which of the following is a good option?

Reduce the maturity of the loan, switching to a 10-year mortgage

Increase the frequency of the payments switching to monthly payments

Reduce the collateral

Do not know

Prefer not to say

Investment Diversification

InvestDiv 1 (Lusardi, 2011)

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

True

False

Do not know

Prefer not to say

InvestDiv 2

Compared with an investment in stocks, the risk of investing in stock mutual funds is...

...equal

...greater

...smaller

Do not know

Prefer not to say

InvestDivDiv 3

In a stock mutual fund, when the number of stocks issued by different companies increases, what happens to the investor's risk?

Increase

Decrease

Nothing: it is the same

Do not know

Prefer not to say

InvestDiv 4

Which of the following investment options fits well for investors that want to double their money in a very short term?

Money market mutual fund

Stock mutual fund

Single stock investment

Do not know

Prefer not to say

InvestDiv 5

Suppose you invested £1,000 1 year ago in a well-diversified stock mutual fund. If the performance of the stock market index (in which the fund invested its money) during the same period has been +5%, what return do you expect from your investment?

Less than £5

More than £500

Around £50

Do not know

Prefer not to say

Bonds

Bond 1 (Lusardi, 2011)

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

They will rise

They will fall

They will stay the same

There is no relationship between bond prices and the interest rates

Do not know

Prefer not to say

Bond 2

Is there a relationship between interest rates and bond prices?

Yes, when interest rates fall bond prices fall

Yes, when interest rates fall bond prices rise

No, there is no relationship

Do not know

Prefer not to say

Bond 3

If you expect a drop in the interest rate what is a good investment strategy?

Buy a bond

Sell a bond

Hold a bond

Do not know

Prefer not to say

Bond 4

Buying a bond is a good strategy if you think that...

...stock indices will rise

...interest rates will fall

...inflation will rise

Do not know

Prefer not to say

Bond 5

If you expect a rise in the interest rate what is the worst investment strategy?

Buy gilts that mature in less than a year (new issue)

Buy gilts with two to 10-year maturities (new issue)

Buy gilts that mature between 10 and 30 years in the future (new issue)

Do not know

Prefer not to say

Bank Accounts

BankAccount 1

You have an "overdraft" on your bank account if...

...you use more money than you have in your account

...you receive interest on your deposit from the bank

...you pay by cheques

Do not know

Prefer not to say

BankAccount 2

What do you NOT need to access your bank account in the case of e-banking (or "Internet banking")?

A user ID (or Username) and a password

A device connected to Internet (computer, tablet, etc.)

Your passport

Do not know

Prefer not to say

BankAccount 3

If the balance of your bank account is zero and you issue a cheque...

...your account will be automatically closed by the bank

...the cheque will be paid only if you have an overdraft facility

...your credit score will drop for sure

Do not know

Prefer not to say

BankAccount 4

Which of the following scenarios does NOT apply to e-banking (or "Internet banking")?

Access to bank services and information 24 hr a day

Large use of cash transaction (withdrawals and deposits)

Access to bank services from abroad

Do not know

Prefer not to say

BankAccount 5

If the balance of your account over a full year has been zero, your bank...

...cannot charge you any fee

...will close your account

...will send you a bank statement anyway

Do not know

Prefer not to say

Payments

Payments 1

What kind of card lets you buy something now and pay for it in the future?

Debit card

Credit card

Pre-paid card

Do not know

Prefer not to say

Payments 2

If you have no money in your bank account and you cannot have an overdraft, which of the following payment options do you have to buy something in a shop?

Cheques (without any cheque guarantee card option)

Debit card

Pre-paid card

Do not know

Prefer not to say

Payments 3

Which of the following payment options will affect the balance of your bank account?

Cash

Pre-paid card

Debit card

Do not know

Prefer not to say

Payments 4

If you pay the balance of your credit card in full at the end of the month, do you have to pay interest?

Yes

No

Only if you used your credit card abroad

Do not know

Prefer not to say

Payment 5

Do you think it is possible to use a credit card to withdraw cash from an ATM?

No, you can do it with a debit card, but not with a credit card

Yes, but you will be charged

Yes, and it will be free of charge

Do not know

Prefer not to say

Stock Investments

Stock 1

Ignoring the case of the issuer defaulting, which of the following investment products guarantees the reimbursement of invested capital?

Stocks

Bonds

Stock mutual funds

Do not know

Prefer not to say

Stock 2

Using cash for saving, which of the following risks do you avoid?

Risk of inflation

Risk of theft

Risk of liquidity

Do not know

Prefer not to say

Stock 3

Which is the standard measure of the default risk of a bond issuer?

Rating

APR

Benchmark

Do not know

Prefer not to say

Stock 4

If the same company issues short-term and long-term bonds, typically the interest rate of the long-term bond is...

...higher

...the same

...lower

Do not know

Prefer not to say

Stock 5

You can invest £100 in stock A or in stock B. The value of stock A is £50, the value of stock B is £1. Comparing the risk of buying two £50 A-shares with the risk of buying one hundred £1 B-shares we can say that...

...the risks are the same

...the two £50 A-shares investment is riskier

...the one hundred £1 B-share investment is riskier

Do not know

Prefer not to say

Loans

Loans 1

Typically, if you buy things (mobile-phones, TVs, etc.) using credit you will pay...

...more than paying cash

...the same amount of paying cash

...less than paying cash

Do not know

Prefer not to say

Loans 2

Everything else being equal, if the maturity of a mortgage is longer, the installments will be...

...the same

...smaller

...bigger

Do not know

Prefer not to say

Loans 3

Suppose you need to borrow £100. Bank "A" allows you to repay £10 a month for 12 months. Bank "B" allows you to repay £120 after 12 months. Which is the loan with the higher APR?

Bank "A"

Bank "B"

The APRs of the two loans are the same

Do not know

Prefer not to say

Loans 4

With a mortgage, if the value of the collateral is higher...

...the interest rate is lower

...the interest rate is higher

...the interest rate does not change

Do not know

Prefer not to say

Loans 5

Which of the following pay back options for a £100 debt has the highest APR?

£102 after 1 week

£105 after 1 month

£110 after 2 months

Do not know

Prefer not to say

Pensions*

Pensions 1*

Compared with a non-smoker, the premium for the healthcare insurance plan of a regular smoker...

...is higher

...is lower

...is the same

Do not know

Prefer not to say

Pensions 2

Taxation on pension fund income is...

...higher than taxes on earned income

...equal to taxes on earned income

...**lower than taxes on earned income**

Do not know

Prefer not to say

Pensions 3

Do you think that the performance of financial markets can affect the performance of a pension fund?

No, they have no relationship

Yes, they are positively related (when the market goes up the value of pension funds goes up too)

Yes, they are negatively related (when the market goes up the value of pension funds goes down)

Do not know

Prefer not to say

Pensions 4

Feeding a pension fund with £10,000 a year for 10 years is equal to feeding it £5,000 a year for 20 years

True

False, to feed a pension fund with £10,000 a year for 10 years is better

False, to feed a pension fund with £5,000 a year for 20 years is better

Do not know

Prefer not to say

Pensions 5

Which of the following sentences on retirement investment products is wrong?

They benefit from a tax-shield

The money in these products can be distrained on by creditors

Even a small withdrawal for any reason is prohibited by the law until the retirement of the worker

Do not know

Prefer not to say

*In the early development of the pool of 50 questions this topic was denoted "Retirement and Insurance." In the present study the insurance question was omitted, and the topic relabeled "Pensions."

APPENDIX B: PSYCHOMETRIC PROPERTIES OF THE EASIER AND HARDER 10-ITEM SUBSCALES

The two scales and their properties are presented in Tables B1 and B2 and Figures B1 and B2.

As expected, most items on the easier scale were, in fact, relatively easy, with difficulty parameters ranging from $b = -1.4$ to 0.9. Consistent with these difficulty parameter estimates, the test characteristics curve (Figure B1a)—which estimates expected scale performance contingent on ability—shows that a respondent of average ability 0 would be expected score 80% (8/10) on Scale 1. All 10 items had good discriminability ($a > 1.0$) and Cronbach's alpha confirmed that the scale was reliable (alpha = .83). Figure B1a also shows that the overall scale would be most sensitive to differences in ability across the ability range -2.0–1.0. At higher ability levels, however, the test would be susceptible to ceiling effects. In terms of content, the easier subscale assesses financial literacy with respect to the broader domains of the concept of money (interest rates and inflation), basic money transactions (banking and payment systems) and borrowing (loans and debts).

Turning to the harder scale, Table B2 shows that, as expected, most items were relatively difficult, in the range -0.1–1.6, but well-distributed across that range. Consistent with this, the test characteristics curve (Figure 2b) shows that a respondent of average ability would score about 30% (3/10). All items had good discriminability ($a > 1.0$), and Cronbach's alpha confirmed that the test was reliable (alpha = .79). Figure B2b also shows that the scale would be most sensitive to differences in ability in the higher range (-1.0–3.0). In terms of content, the harder subscale assesses financial literacy mainly with respect to saving and investment (Investments, Savings, Retirement and planning and Bonds) but also in the domain of borrowing (Mortgages).

TABLE B1 Discrimination and difficulty parameter estimates for the easier subscale ($N = 541$)

Topic/label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Interest rates *Interest1*	78	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, Exactly £102, Less than £102*	−1.16	2.14
Interest4	45	Suppose you had £100 in a savings account and the interest rate was 10% per year. After 2 years, how much do you think you would have in the account if you left the money to grow? *£110, £120, **£121***	.09	1.28
Inflation *Inflation1*	58	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, Exactly the same as today, **Less than today***	−.34	2.95
Inflation3	57	Imagine depositing £100 in your savings account. If after 1 year the balance on your account (included the interest) is £104 and the inflation rate during the last year was 5%, How rich do you think you would be compared 1 year before? *More rich, Exactly as rich, **Less rich***	−.33	2.60
Bank accounts *BankAccount1*	77	You have an "overdraft" in your bank account if... ***You use more money than is in your account**, You receive interest on your deposit from the bank, you pay by cheques*	−1.11	2.56
BankAccount2	73	What do you NOT need to access to your bank account in the case of e-banking? *A username and password, An device connected to the internet (computer, tablet, etc.), **An ID Card***	−.95	2.18
Payments *Payments1*	83	What kind of card lets you buy something now and pay for it in the future? *Debit card, **Credit Card**, Pre-paid card*	−1.41	2.72

(Continues)

T A B L E B1   (Continued)

Topic/label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Payments2	63	If you have no money on your bank account and you cannot overdraw, which of the following payment options do you have to buy something in a shop? *Cheques [without any cheque-guarantee card option], Debit card, **Pre-paid card***	−.63	1.71
Loans *Loans2*	57	Everything else equal, if the maturity of a mortgage is longer the installments will be... *The same, **Smaller**, bigger*	−.44	1.27
Loans5	50	Which of the following pay back options for a £100 debts shows the highest APR?£***102 after 1 week**, £105 after 1 month, £110 after 2 months*	−.17	1.59

Note: Response options for each item are provided in italics, with the correct response in bold. All items also provided "Do not know" and "Prefer not to say" response options.

TABLE B2 Discrimination and difficulty parameter estimates for harder subscale ($N = 548$)

Topic/ label	Percent correct	Item	b (Difficulty)	a (Discrimination)
Investment div. *InvestDiv1*	41	Buying a single company's stock usually provides a safer return than a stock mutual fund. *True,* **False**	.26	2.00
InvestDiv2	24	Compared with an investment in stocks, the risk in investing in stock mutual funds is... *Equal, Greater,* **Smaller**	.81	2.71
InvestDiv5	48	Suppose you invested 1 year ago £1,000 in a well-diversified stock mutual fund. If the performance of the stock market index (the one where the fund invested its money) in the same period has been +5%, how much returns do you expect from your investment? *Less than £5, More than £500,* **Around £50**	−.01	2.09
Mortgages *Mortgage3*	43	In a 15-year mortgage which of the following options will minimize the total interest paid over the life of the loan? *Annual payments, Semi-annual payments,* **Monthly Payments**	.27	1.04
Mortgage4	31	Consider a 15-year mortgage where you can choose between monthly payments and 6-month payments. If after 5 years you would like to close the mortgage in advance, which one will make you pay less? **Monthly payment mortgage***, 6-monthly payment mortgage, Both would make people pay the same amount*	.81	1.31
Investment div. *InvestDiv1*	31	Ignoring the case of default of the issuer, which of the following investment products guarantees the reimbursement of the invested capital? *Stocks,* **Bonds***, Stock mutual funds*	.78	1.41
InvestDiv3	19	Which is the standard measure for the default risk of a bond issuer? **Rating***, APR, Benchmark*	1.38	1.33
InvestDiv4	34	If the same company issues short-term and a long-term bonds, typically the interest rate of the long-term bond is... **Higher***, The same, Lower*	.52	1.64
Pensions *Pensions1*	44	Do you think that the performance of financial markets can affect the performance of a pension fund?\ *No—They are unrelated*	.13	1.46

(Continues)

TABLE B2 (Continued)

Topic/ label	Percent correct	Item	b (Difficulty)	a (Discrimination)
		Yes —they are positively related (i.e., when the market value goes up so too does the pension fund value) *Yes—they are negatively related (i.e., when the market value goes down so too does the pension fund value)*		
Bonds *Bonds3*	19	If you expect a drop in the interest rates what is a good investment strategy? ***Buy a bond,*** *Sell a bond, Hold a bond*	1.56	1.06

Note: Response options for each item are provided in italics, with the correct response in bold. All items also provided "Do not know" and "Prefer not to say" response options.
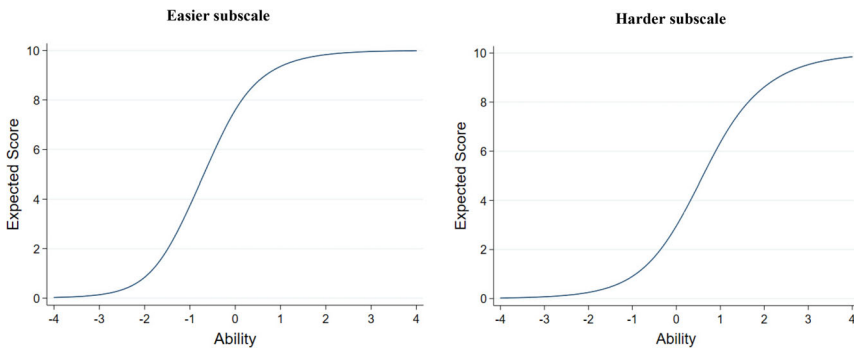

FIGURE B1 Test characteristic curves for the easier and harder subscales