Centre for Social and Behaviour Change

Pause and Practice: Does it help build confidence?

WORKING PAPER 2412

DECEMBER 2024

Sneha Shashidhara

Pavan Mamidi

Pause and Practice: Does it help build confidence?

Sneha Shashidhara[a]*, Pavan Mamidi[a],

[a] *Center for Social and Behaviour Change, Ashoka University, Sonipat, India*

Corresponding Author: Dr. Sneha Shashidhara, Centre for Social and Behaviour Change, Ashoka University. Email: sneha.shashidhara@ashoka.edu.in

Acknowledgement

This research has been funded by a BMGF core grant to CSBC (INV 026871). We thank Bijoyetri Samaddar, Aditya Laumas, Pooja Haldea at CSBC, and Dr. Sharon Barnhardt for her guidance. We would also like to thank Akhilesh Singh, Vaibhav, and Sonika for their help with data collection.

Declaration of Conflict

The authors declare no conflict of interest.

Data Availability

All data, analysis code, and research materials are available in a masked link [https://osf.io/m527c/?view_only=c97e39ee58214e2a8a8aab6dfcd5c131].

Pause and Practice: Does it help build confidence?

Financial inclusion is the availability of financial resources, including transactions, payments, savings, credit and insurance, and equal opportunities to access these financial services. Digital financial systems enable financial inclusion by providing accessible, affordable, and secure financial services to underserved populations, reducing barriers such as geographic distance and high transaction costs. However, despite the growing usage of digital payments, many people in India prefer cash-based transactions due to various factors, one of which is likely the low confidence level in navigating the payment apps or the smartphone itself. Here, in a lab-in-the-field experiment, we test simple video interventions targeted at smartphone users to perform a QR-code-based money transfer transaction. Participants were asked to practise the transaction while watching the video to learn the task step-by-step. The findings indicate that participants exposed to videos with practice embedded in them demonstrated higher accuracy and confidence in QR code transactions, suggesting the effectiveness of these interventions.

Keywords: **QR code, active learning, confidence, randomised control experiment**

Introduction

Digital financial services (DFS) have emerged as a key component in the global effort to achieve financial inclusion, especially in mid- and low-income countries. The World Bank defines financial inclusion as individuals' and businesses' having access to valuable and affordable financial products and services that meet their needs – transactions, payments, savings, credit and insurance – delivered responsibly and sustainably. The rise of mobile banking, digital wallets, and other forms of DFS has reduced the physical and financial barriers that traditionally excluded vulnerable groups from formal financial systems (Demirgüç-Kunt et al., 2020). Digital platforms such as mobile money and QR code-based payment systems are particularly transformative, offering accessible, affordable, and secure transaction methods for populations previously reliant on cash (Aker & Mbiti, 2010; Jack & Suri, 2014).

Despite these advancements, digital financial inclusion remains a challenge, especially in rural and low-income areas with low technological literacy and limited access to smartphones and the Internet (Asongu et al., 2018). Research in India, for instance, shows that while digital payments have expanded rapidly, many individuals continue to rely on cash due to their discomfort or lack of familiarity with digital systems (Mazzotta et al., 2023). This gap highlights the need for more targeted interventions that not only make technology accessible but also empower users to engage confidently with it. The 'pause and practice' method, examined in this study, seeks to address these barriers by providing users with structured opportunities to build their confidence while interacting with new digital financial tools.

Self-efficacy, or the belief in one's ability to succeed in specific tasks, is a central determinant of behaviour change and skill acquisition (Bandura, 1997). In the context of DFS, confidence plays a critical role in users' willingness to adopt and effectively use these platforms

(Hastings et al., 2013). According to research in a variety of fields, when individuals feel capable of mastering a task, they are more likely to engage with it persistently (Berry & West, 1993; Gist, 1987; Li et al., 2005; Margolis & McCabe, 2004; Vera et al., 2014). Studies have explored the role of psychological traits, such as self-esteem and confidence, that affect financial behaviour after knowledge is accounted for (Abreu & Mendes, 2012; Antonides et al., 2011; Brown & Taylor, 2014; Tang et al., 2015; Tang & Baker, 2016). Thus, building self-efficacy or confidence through structured interventions could be a crucial strategy in increasing the uptake of DFS. Our study explores whether such interventions can enhance users' success in adopting DFS technologies by improving their accuracy and confidence.

Participants in rural Uttar Pradesh, India, with smartphones and without DFS experience, were introduced to QR code transactions via video as part of the field experiment. India has adopted UPI (United Payments Interface), a real-time payment system that allows users to send and receive money, pay bills, and make online purchases using a mobile app. QR codes can be scanned to initiate such payments, where the code contains the receiver's UPI ID, a unique identifier linked to a user's bank account. In treatment arm one, they had to pause and do the task alongside, and in the second treatment, they were also given additional information to handle contingencies. The participants were tested on the QR code task shown in the video and the same task in a new app. They were also tested on both tasks two days later to test retention and learning.

Literature Review

Financial Knowledge

Financial education and knowledge have been shown to affect behaviour across different studies (Billari et al., 2023; Collins, 2013; Duflo & Saez, 2003; Kaiser et al., 2022). Kaiser et al., 2022 review experimental evidence and show the mechanism of financial education leading to better knowledge and behaviour. The effectiveness of such financial education in low-income and low-education settings, such as rural India, is more debatable (Bhutoria & Vignoles, 2018; Fernandes et al., 2014; Holzmann, 2010; Kaiser & Menkhoff, 2017; Walstad et al., 2017). Here, the intensity of education and proving it at the 'right time' becomes more important (Kaiser & Menkhoff, 2017). There is now increasing evidence that behavioural nudges and interventions that combine nudges and education lead to improved financial behaviour (Banerjee & Duflo, 2012; Carpena et al., 2017; Isler et al., 2022; Smith et al., 2020; Thaler & Sunstein, 2008).

Active Learning

Active learning pedagogies have been known to improve learning compared to more traditional lecture-style pedagogies (Ballen et al., 2017; Blázquez et al., 2019; Dewsbury et al., 2022; Freeman et al., 2014; Prince, 2004; Røe et al., 2022). The European Commission's recent Digital Education Action Plan (2021–2027) emphasises that digital education should support teaching that is more personalised, flexible, and centred on the needs of students (European Commission, 2021), which deeply resonates with one of the early definitions of active learning "instructional activities involving students in doing things and thinking about what they are doing" (Bonwell & Eison, 1991). In particular, pausing in between lectures improved learning (Bachhel & Thaman, 2014; Richards et al., 2017; Ruhl et al., 1990), and so has hands-on learning (Hartman et al., 2000; Hearns et al., 2009; Parkinson et al., 2003). Thus, we tested if practising the task while watching the videos step-by-step improved learning.

Confidence

Confidence is a psychological trait known to be separate from one's ability to do the task (Allwood & Montgomery, 1987; Dougherty, 2001; Fischhoff et al., 1977; Jackson & Kleitman, 2014; Lichtenstein & Fischhoff, 1977; Stankov & Lee, 2008). It has been established that confidence is separate from other constructs like intelligence, personality, locus of control, and self-efficacy, albeit correlated with them (Jackson & Kleitman, 2014; Stankov, 2000, 2004; Stankov & Lee, 2008). Overall, there is a trend of overconfidence in the population; people are more likely to overestimate their performance (Camerer & Lovallo, 1999; Harvey, 1997; Juslin, 1994; Klayman et al., 1999; Mcgraw et al., 2004; Svenson, 1981). This overconfidence trend has been known to have positive effects that may be evolutionarily advantageous (Hirshleifer et al., 2012; Johnson & Fowler, 2011; McKay & Dennett, 2009; Trivers, 2000; Weinberg, 2009).

In this study, we incorporate the principles of the pause procedure, hands-on training, and other behaviour science aspects into the intervention to improve DFS learning. We hypothesise that confidence is an important predictor of this learning and study how the intervention and general confidence interact for optimal learning.

Methods

Experimental Design Overview

We conducted a randomised control experiment. Primary owners of smartphones who have never used any form of DFS from the districts of Benaras and Chandauli in Uttar Pradesh (most populous state in India) were eligible. Primary is defined as the phone being with them the majority of the time. Participants were assigned randomly to one of three experimental groups: one control and two treatment arms. The control arm included a video explaining the QR-code-based monetary transaction on the PhonePe app (one of the most used apps for DFS in rural India). The first treatment (Pause & Practice) included pausing and practising the task while watching the control video. The second treatment (Pause & Practice Plus) included additional information to handle contingencies.

In-person baseline surveys included measures of digital literacy, financial literacy, and experience accessing financial products. This was followed by the intervention, an immediate test of the QR code transaction task, and the participants' confidence. Participants were also asked to do the same transaction on a different app, Gpay. The endline survey was conducted 2-4 days later, including the primary outcome measure of completing QR code transactions without assistance on PhonePe and Gpay, standard demographics, and general confidence measures. Figure 1 shows the overall experiment structure.

Treatment Components

Participants were shown a short video explaining the QR code transaction process on the PhonePe App. Each group got slightly different video content, but the overall style was the same. A local content creator made the videos to ensure they were in the proper Hindi dialect, giving participants a sense of familiarity. Videos were designed to be cognitively simple, with relevant visuals of the app shown. The QR code process was described in simple language and narrated by a relatable person who was visible on the screen. Table 1 highlights the differences between the three videos, thus, the three arms. The two treatment arms' videos included an instruction to pause the video and do the task alongside on a separate device. These instructions came five times to demarcate the major steps of this process. The second treatment arm *also* included information on managing contingencies *and* confidence-building messages.

Five steps of pause and practice:

Step 1: Locate the PhonePe app and open it. Pause and practice alongside.

Step 2: Find the scan icon on the app and click it. Pause and practice alongside.

Step 3: Scan the QR code. Pause and practice alongside.

Step 4: Enter the amount and select the account to pay from. Pause and practice alongside.

Step 5: Enter the UPI ID. Pause and practice alongside

Data Collection

The districts were selected through convenience sampling as our field team there had previously interacted with the frontline health workers in that area. We leveraged these relationships for recruitment. We informed frontline workers a few days before to invite eligible participants to Anganwadi Centers (rural child care centres) on the day of the study, and they were recruited on a first-come, first-serve basis. Upon reaching the desired target of 20 participants, recruitment was closed. Participants were asked screening questions to determine if they had used DFS before, and enumerators verified that there were no DFS apps on their phones at baseline. We continued data collection by visiting Anganwadis in different villages in these districts until we reached our desired sample of 100 participants per arm.

The baseline and endline surveys were conducted in person, with on-ground enumerators administering the survey on tablets, using licensed software Qualtrics on their offline app. 10%

of the surveys were recorded and were around 45 minutes long. Only complete surveys were used for analysis, and no participants with partial surveys were contacted again to resume the survey. While the enumerators were familiar with the broad outline of the study, they were not made aware of the details of the treatment groups to reduce potential bias. All participants were compensated with Rs. 100 for baseline and Rs. 150 for Endline. The survey and intervention were administered in Hindi.

The entire questionnaire was piloted in 2 villages. The same enumerators collected the pilot data to ensure the data quality and address any concerns faced in the field. 10% of the participants were chosen for an additional short survey across treatment and control groups and enumerators. A week after the primary data collection, these surveys were conducted by a separate set of enumerators via a phone call. It included questions about the previous survey length and comfort rating. These were matched with the original survey to estimate the data quality.

Treatment was watching QR code transaction explanation videos. The control group watched the video once and could request to watch it again. In the treatment arms, the participants watched it once, and during the second time, they paused and practised alongside. All arms were asked to do the transaction independently at the end of the intervention on the day of baseline on the PhonePe app (Task 1 immediate outcome) and a new Gpay app (Task 2 immediate outcome). At Endline, they again performed the transaction without seeing the video on PhonePe (Task 1 after 2 days) and Gpay (Task 2 after 2 days). The accuracy of the tasks was calculated as a percentage of the total steps (5) completed successfully. For each task, participants were asked to rate their confidence before starting the task (pre-confidence) and after finishing the five steps (those that did not finish the five steps were assigned 0), but before knowing the outcome (post-confidence).

Digital and Financial Literacy

We estimated digital and financial literacy at the Baseline before the intervention. Participants were asked questions about smartphone use through pictures. They had to point to the location they would touch for specific functions. Five such questions probed the use of browsers, WhatsApp, YouTube, torch, and bars showing network signals. A digital score was calculated based on the number of correctly answered questions.

Participants were asked about financial services, and their interest, loans, and insurance knowledge was tested (3 questions). Participants were also asked how they access financial services, withdraw money from the ATM, and deposit cheques. They pointed to specific parts on a relevant image to answer how they perform these tasks. There were three such questions. A composite financial literacy score was computed on the total number of correct answers to these six financial literacy questions.

General Confidence

General Confidence was measured using three general/non-financial tasks, presented in the same order to all participants. These tasks were done at Endline at the end of the QR code tasks and other survey questions. The trials were predefined for each task, and the same set of trials was presented to each participant in the same order. The first was the esoteric analogies test (Jackson & Kleitman, 2014). The task contained 20 items of the following type: CHICK is to HEN as CALF is to __. The options were BULL, COW, COAT, ELEPHANT (answer = COW). The enumerator read out the words and entered the option the participant chose. After each trial, the participants were asked to rate their confidence in their previous answer on a scale of 1-5.

The second task was panamath (Halberda et al., 2008), a non-symbolic task that measures a person's approximate number system abilities. Participants judged whether more green or pink dots were presented on a grey screen. Each image is shown for 1.5 seconds. The difference between dots ranged from 3-10. The probability of the difference being 3 to 12 decreased by half with every increase in difference by 1, until 7. The probability of the difference being 3 was 49%, 4 was 24%, 5 was 12%, 6 was 6%, and 7-10 was 2% each. The number of dots in each colour ranged from 5 to 20. The size of the dots remained constant in all trials, but their position was jittered across a 6X8 grid. There were 24 trials; the respondents answered and gave confidence ratings for each trial.

The last was a vocabulary task (Jackson & Kleitman, 2014). A word was presented on the screen, and four options were shown below. The enumerator read out the word and the options and asked which of the options was the closest in meaning to the target word. The enumerator entered the answer and the confidence rating for each of the 20 items. The analogy and vocabulary words included common words used in everyday language that did not exceed the 7th-grade reading level. Participants were not expected to read, which allowed participants without literacy to participate.

An average reaction time was computed across all the tasks to measure processing speed. Average accuracy across all tasks provided a proxy for IQ (intelligence quotient) or general ability. Note, this was not used to look for IQ effects but to understand what covaried with task-specific measures. General confidence was the average confidence across the three tasks. Bias was computed as the difference between average subjective confidence measures and objective accuracy for each task and averaged across tasks. It indicates whether an individual was able to match their confidence levels with their level of accuracy (Stankov et al., 2012). Bias scores range between +1 to -1. High (positive) and low (negative) scores indicate over- and under-confidence.

Sample Size Determination

Based on pilot data, we arrived at a sample size of 100 per group. With 100 per group and a 74% completion rate in the control group, we can detect a relative improvement of 20%. As sampling was convenience-based, we continued data collection till we completed at least 100 in each group.

Model Specifications

We used three outcomes (accuracy, pre-confidence, and post-confidence) for each of the two tasks (Task 1 – PhonePe; Task 2 – Gpay) and two time points (Immediate outcome – immediately after the intervention; After 2 days – at Endline). Ordered Logistic Regression was used for all outcomes. A dummy variable was created for the treatment with pause & practice, pause & practice plus, and Control levels. A digital literacy score and a financial literacy score were added as covariates. Further, the model controlled for demographic information: age, gender, education, employment, household income, number of people in the household, number of teenagers in the household, religion, caste, and duration of owning the smartphone. We then repeated the regressions for the outcome variables, looking for an effect of average reaction time and accuracy of the three independent tasks used to compute general confidence, as well as general confidence and bias. The same demographic values and digital and financial literacy scores were added as controls.

Lastly, to test if the average measures (reaction time, accuracy, confidence or bias) had a significantly greater effect on one outcome compared to another, we used an interaction method. For example, to test if average accuracy had a greater effect on PhonePe task accuracy

Immediate outcome (Y1) or Gpay task accuracy Immediate outcome (Y2), we constructed a new variable (Y3), a concatenated version of the two outcome variables and a dummy variable for outcome type (1 for Y1 and 2 for Y2). A new interaction term was created: the multiplication of average accuracy and outcome type. Ordered logistic regression was run on the concatenated outcome variable Y3, with treatment levels, average accuracy, digital and financial literacy scores, demographic controls and the two new factors of outcome type and the interaction between average accuracy and outcome type. If the interaction term is significant, it would suggest that the effect of average accuracy on Y1 and Y2 are significantly different. All pairwise differences in its effect on outcome measures were computed for a given average measure.

Transparency and Openness

We described our sampling plan, all data exclusions (there were none), all manipulations, and all measures in the study, and we adhered to the Journal of Applied Psychology methodological checklist. All analyses, including randomisation and data checks, were conducted using custom-made MATLAB (The MathWorks, Inc.) scripts and scripts in StataCorp. (2023). The study hypothesis and analysis plan were pre-registered, but no masked link is available, and they will be shared after the review. All data, analysis code, and research materials are available in a masked link [https://osf.io/m527c/?view_only=c97e39ee58214e2a8a8aab6dfcd5c131].

Results

Sample demographics

We collected a sample of 444 participants who were surveyed on two days, the day of the intervention and two days later, to measure recall. 91% of participants who completed the first session also completed the second. 148 people were part of the control group that watched the video of how to do the QR transactions (33%), 171 were part of pause & practice - P&P (39%), and 125 were part of pause & practice plus - P&P plus (28%).

Three-fourths of our participants were women, and roughly one-third (34%) belonged to the OBC (other backward classes or caste) category. 30% had not completed more than 10th-grade education, 32% listed household duties as their primary employment, and 26% had an average monthly household income of less than INR 5,000 ($60). Respondents' ages ranged from 18-58 years, with a mean age of 26.75 (SD = 7.64). Most participants were Hindus (94%). The mean household size was 8 (SD = 4.64), with 40% of households having more than two teenage members. All demographic characteristics by treatment arms are presented in Table A1. We also evaluate if the treatment groups and their respective controls are different using t-tests.

We compared the proportions attrited across the four groups using a two-tailed Fischer's exact test. There was a marginal increase in the proportion attrited in P&P (12%) compared to the Control group (5.7%) at the 5% level (odds ratio = 0.43, and p = 0.043). No difference existed between P&P plus (9.4%) and other arms.

Accuracy

Our survey included two primary outcomes: accuracy and confidence. Accuracy was measured by the number of correctly completed transaction steps. Enumerators observed and marked steps as the participants attempted the two tasks: QR code transactions in PhonePe and GPay. Both treatment arms had significantly higher accuracy in the PhonePe task immediately after the intervention (P&P: coeff = 1.225, $p < 0.001$; P&P plus: coeff = 1.218, $p < 0.001$). The treatments also showed an increase in the GPay task, the same task in a novel platform (P&P: coeff = 0.606, $p < 0.05$; P&P plus: coeff = 0.626, $p < 0.05$). The average accuracies for pause and practice, pause and practice plus, and control for the PhonePe task were 82%, 82%, and 64%, and for the GPay task, they were 73%, 75%, and 62%. (Table 2).

Participants attempted the same tasks 2 days later (Table 2). Treatment groups were more accurate in the PhonePe task (P&P: coeff = 0.698, $p < 0.001$; P&P plus: coeff = 0.484, $p < 0.1$), but not the GPay task (coeff < 0.09, $p > 0.1$). The PhonePe tasks' accuracies were higher (P&P: 76%, P&P plus: 71%, Control: 62%) than GPay (P&P: 66%, P&P plus: 68%, Control: 63%).

Confidence

Before each task, participants were also asked to rate their confidence in their ability to do the task correctly (pre-confidence). Both treatment arms showed significantly higher confidence before attempting the PhonePe task, immediately after the intervention (P&P: coeff = 0.408, $p < 0.1$; P&P plus: coeff = 0.721, $p < 0.01$), as would be expected given that they had a chance to practice. Interestingly, they also showed higher confidence before attempting the GPay novel task (P&P: coeff = 0.527, $p < 0.05$; P&P plus: coeff = 0.510, $p < 0.05$). Pause and Practice showed some effect in both tasks after two days (PhonePe: coeff = 0.351, $p < 0.1$; GPay: coeff = 0.376, $p < 0.1$), while pause and practice plus did not (coeff < 0.36, $p > 0.1$). The regression values and averages are listed in Table 3.

Participants were also asked about their confidence in having done all the steps correctly before they found out the outcome of the task (post-confidence). Participants in both treatment arms were more confident in having done the PhonePe task accurately at step 5, immediately after the intervention (P&P: coeff = 1.180, $p < 0.001$; P&P plus: coeff = 0.980, $p < 0.001$); as well as the GPay task (P&P: coeff = 0.704, $p < 0.01$; P&P plus: coeff = 0.600, $p < 0.05$). The treatment did not affect post-confidence when the tasks were done two days later (coeff < 0.45, $p > 0.1$). The regression values and averages are listed in Table 4.

General Confidence

The experiment included measures of general confidence. Participants completed three tasks, which gave us a measure of average reaction time, accuracy, confidence, and bias (See Methods). Here, we examined the effects of these independent measures on the task-related outcomes: accuracy, pre-confidence, and post-confidence for PhonePe and GPay tasks at each of the two time points immediately after the intervention and two days later (Table 5).

Average reaction time, a proxy for processing speed, did not affect any accuracies or post-confidences. There was a small negative effect on the pre-confidence in both tasks' immediate outcomes and the GPay task two days later (coeff < -0.044, $p < 0.1$). On the other hand, the average accuracy showed a positive effect on all the outcome variables (coeff > 2.01, $p < 0.05$), with effects being numerically greater for post-confidence than pre-confidence. Average confidence on independent tasks had a positive effect on all confidence outcomes (coeff > 2.3, $p < 0.01$) and none of the accuracies (coeff < 0.85, $p > 0.1$).

Lastly, we investigated the effect of average bias. Bias indicates how well matched a participant's confidence is with their ability. A positive bias suggests overconfidence, and a negative bias suggests underconfidence. Among the accuracies, bias had a small negative effect on the PhonePe task 2 days later (coeff = -1.631, $p < 0.05$), and the rest were insignificant (coeff < 0.08, $p > 0.1$). It showed a positive effect on GPay's post-confidence at both time points (coeff

> 1.6, $p < 0.05$) and not PhonePe's (coeff < 0.8, $p > 0.1$). It positively affected all pre-confidence outcomes (coeff > 1.6, $p < 0.01$).

Pairwise comparison of the effect of average measures on outcomes

We computed whether the effect in one outcome differed from another using all pairwise comparisons for each of the average measures of accuracy, confidence, and bias computed from the three independent tasks. This analysis is not reported for average reaction time as the measure had limited effect on the outcomes. Average accuracy had a greater effect on all Post-confidence outcomes compared to all pre-confidence outcomes (coeff = 2.46, $p < 0.001$) and compared to all accuracy outcomes (coeff > 0.90, $p < 0.01$; except no difference between Gpay accuracy two days later and Gpay post-confidence Immediate outcome). It had a lower effect on all the pre-confidence outcomes than accuracy outcomes (coeff > 1.55, $p < 0.001$). Average confidence had a greater effect on pre- and post-confidence outcomes compared to accuracy outcomes (coeff > 0.96, $p < 0.01$; except no difference between PhonePe pre-confidence Immediate outcome and Gpay accuracy Immediate outcome or two days later); there was no clear difference between pre- and post-confidence outcomes. Average bias had a greater effect on pre- and post-confidence outcomes than accuracy outcomes (coeff > 0.08, $p < 0.01$, except the effect of PhonePe accuracy two days later was greater than the effect on GPay post-confidence two days later coeff = 0.51, $p < 0.001$). Average bias had a greater effect on pre-confidence than post-confidence outcomes (coeff > 0.76, $p < 0.05$, except no difference between PhonePe pre-confidence Immediate outcome and GPay post-confidence Immediate outcome)

Digital and Financial Literacy

Along with the demographic controls, we also controlled for digital and financial. Here, we examine the coefficients for these variables on all outcome variables (Table 6). These constructs were based on objective knowledge questions that participants answered before the intervention. Both digital and financial literacy have positive effects on all outcome variables (Digital: coeff > 1.08, $p < 0.1$; Financial: coeff > 0.95, $p < 0.1$).

Demographic Effects

All regression models were controlled for demographic variables. We report here any demographic effects on the accuracy and confidence outcome variables. There was an overall trend of decrease in all accuracy and post-confidence measures with age (coeff > -0.036, $p < 0.05$; except GPay post-confidence two days later), and a decrease in the same measures in female compared to male participants (coeff > -0.549, $p < 0.1$; except GPay post-confidence two days later). The same effect was seen for age and gender in GPay pre-confidence two days later (coeff> -0.036, $p < 0.05$) and gender only in PhonePe pre-confidence Immediate outcome (coeff = -0.549, $p < 0.05$). All other pre-confidence measures were not significant. We saw a decrease in all accuracy and confidence measures for participants who had completed secondary school or less compared to those who had a diploma or a higher degree of education (coeff > -0.439, $p < 0.1$; except GPay post-confidence two days later). Being employed in agriculture or as a part of the MNREGA employment scheme of the government increased all confidence measures compared to other forms of employment that included self-employment, informal, public, and private sectors in all confidence measures (coeff > 0.450, $p < 0.1$; except GPay pre-confidence Immediate outcome), and both measures of accuracy for PhonePe (coeff > 0.478, $p < 0.1$), but not GPay. Income showed a decreasing trend in all measures for those earning less than Rs. 10,000 per month compared to those earning more than Rs. 10,000 per month, but the effect was consistently seen. Similarly, there was a decrease in outcome measures with increasing household members, which was not very consistent. Interestingly, caste did not significantly affect any measure except a decrease in GPay post-confidence Immediate outcome in OBC (other backward classes) and other reservation categories compared to the General category (coeff < -1.011, p < 0.1). As our data was 94% of Hindus, the majority religion, we were not able to look for the effects of religion.

Self Reports

Of the 444 surveyed two days after the intervention, 121 (27.5%) reported having practised the task at home. Of the 121, 59 reported being successful, 61 unsuccessful, and 1 did not know the outcome. Additionally, participants were asked to report their trust in different financial services. 185 (41.7%) mentioned they trusted banks a lot, 158 (35.6%) to an extent, 88 (19.8%) neither trusted nor distrusted, 5 distrusted to some degree (1.1%), and 8 (1.8%) distrusted a lot. 144 (32.4%) trusted ATMs a lot, 168 (37.8%) to an extent, 85 (19.1%) neither trusted nor distrusted, 8 distrusted to some degree (1.8%), and 39 (8.8%) distrusted a lot. Lastly, 141 (31.8%) trusted digital transactions a lot, 186 (41.9%) trusted to some extent, 74 (16.7%) neither trusted nor distrusted, 13 (2.9%) distrusted to some extent, and 30 (6.8%) distrusted a lot.

Discussion

The above study tested two behavioural interventions to improve the learning of a QR code task. Broadly, we ask if active learning, which includes doing the task alongside a teaching video, can improve the accuracy of the task when performed independently and people's confidence in their ability to do the task. We found that both the treatments of pause & practice and pause & practice plus (with additional information to handle contingencies) increase the accuracy in the task on which people were trained (PhonePe), and this effect transfers to a novel platform (GPay). The interventions also improved the subjective confidence of participants. Additionally, we uncovered specific effects of general confidence and bias. We controlled for digital and financial literacy, as well as demographic variables in all analyses.

Our intervention aimed to improve learning in multiple ways. We used short videos with simple language and a relatable narrator (visible throughout the video) to put people at ease. An external content creator was hired to make the script more engaging. The narrator was a local vlogger, popular in that geography for creating fun, informative videos. Thus, the control video also used behaviour science principles, often lacking in many financial knowledge-building videos. The core intuition of our intervention is that practice, in many ways, is a simulation-like condition where one can attempt learning with no cost of mistakes. The instruction to practice is the experimenter actively encouraging the learner, implicitly saying that the learner has the ability to do the task, an endorsement, or a tacit transfer of confidence. We hypothesised that by adding an instruction to pause and practice at strategic points, we allow for a breakdown of the task (overall goal) into smaller steps (manageable sub-goals), previously known to improve the uptake of behaviour (Delaney et al., 1998; Duncan et al., 2017; Hershfield et al., 2019; Rai et al., 2023). The strategy of breaking into effective sub-goals is woven into the intervention and is given to participants, lessening the cognitive load on them (Duncan et al., 2017). The pause and practice method is a form of active learning instead of passively viewing the information, which is known to improve learning (Blázquez et al., 2019; Freeman et al., 2014). Additionally, information on how to handle contingencies – what to do when the transaction doesn't go according to the ideal scenario usually taught, and what are the crucial benefits with appropriate explanations – can further build people's confidence in doing the task and translate to other such tasks (Inzana et al., 1996; Weisman & Markman, 2017).

Accuracy was increased for both arms in both tasks in the first time point, immediately after intervention. Two days later, the treatment arms still had a small effect on the PhonePe task. This suggests that active learning is superior to traditional learning methods (Ballen et al., 2017; Dewsbury et al., 2022; Freeman et al., 2014; Prince, 2004; Røe et al., 2022). Practising leading to improved performance aligns with previous literature (Calamia et al., 2012; McCaffrey et al., 2000), and even mental practice has been known to improve performance (Driskell et al., 1994). We wanted to further probe the effects of active learning in the domain of confidence. Having practised the task, treatment arms expressed greater confidence before starting to do the task independently. Interestingly, they also showed greater confidence before doing a novel task, and the first treatment arm participants (pause & practice) showed greater confidence before doing both tasks two days later. Practising alongside increases their specific confidence in doing tasks (Jin et al., 2022; Leopold et al., 2005; Mengelkamp & Bannert, 2010) and is retained two days later.

Further, we found that participants who practised alongside finishing the task were more confident in getting the answer correct. Being able to correctly place their confidence, i.e., high confidence when they complete the task, shows better learning overall (Shashidhara et al., 2024).

We then checked if the two treatments differed in any outcome measure, i.e., the effect of the additional information on handling contingencies, and we found no significant differences. This may be due to the length of the video of pause and practice plus, which may have yielded benefits on repeated viewing.

To comprehensively understand confidence's role in learning, we independently measured general confidence (average across three cognitive tasks). We measured average reaction time, a proxy for processing speed, along with general confidence. Reaction time negatively affected both tasks' pre-confidence on the day of the intervention and GPay two days later. Taking more time to do cognitive tasks, or processing in general, is correlated with decreased confidence before attempting the task (Ackerman, 2014) and does not affect accuracy or confidence after having done the task.

Average accuracy across the three tasks, a proxy for ability, affected all accuracy, pre-confidence, and post-confidence measures for both tasks and time points. Much literature suggests that general ability is a good predictor of task performance (Jackson & Kleitman, 2014; Schmidt & Hunter, 2004; Spearman, 1904). Evidence suggests it matters in everyday complex tasks and job performance (Gottfredson, 1997; Schmidt & Hunter, 2004). Interestingly, average accuracy had the most effect on post-confidence, more than even the accuracy outcomes. The participants who are better at the independent tasks report more confidence in getting the task correct at the end of the task, and even more so than the objective task accuracy. We saw the least effect on the subjective measure of pre-confidence before the task, suggesting confidence in one's ability to do the task cannot be completely explained by ability (Jackson & Kleitman, 2014; Stankov, 2004).

Average confidence had an effect on all the confidence outcomes, pre- and post-, without a clear distinction of having a greater effect on one or the other. Average bias, a more objective measure of confidence that examines the relation between subjective confidence and objective accuracy, showed the highest effect on pre-confidence, followed by post-confidence and accuracy. People who overestimate their ability in general also tend to express higher confidence before the task, suggesting a positive effect of overconfidence (Hirshleifer et al., 2012; Johnson & Fowler, 2011; Weinberg, 2009).

We saw an increase in all outcome measures with increased digital and financial literacy (Hastings et al., 2013; Kaiser et al., 2022). We also saw demographic effects. Overall, there was a decrease in accuracy and post-confidence measures with age and gender (a decrease in females compared to males). Our results match much of the literature showing decreased confidence in women and, at times, accuracy too (Boekaerts & Rozendaal, 2010; Lundeberg et al., 1994; McMurran et al., 2023; Stankov et al., 2012). We saw a decrease in all measures with lower education compared to participants with higher educational degrees. The digital technology task could explain the age and education effects (Lundeberg et al., 1994; Zhang, 2023). Interestingly, we did not see much effect of caste on confidence.

Self-reports on trust showed that 77.3% trusted banks, 70.2% trusted ATMs, and 73.7% trusted digital transactions. Trust is an important factor in financial inclusion and has been known to affect bank use (Ghosh, 2021). Our participants already had access to a bank account and thus showed higher trust, while trust in DFS was slightly lower. Lastly, we also looked at how the different outcome measures are related, and we saw a strong correlation between them (See Appendix). The independent tasks showed a strong correlation in reaction time, accuracy, confidence, and bias (See Appendix).

One of the study's limitations is that the follow-up was after only two days. We did not conduct an RCT and evaluate the impact of our interventions in real life. However, we did ascertain if the effects of the intervention were retained two days later to test for learning. Future studies are needed to bridge the gap, i.e., to study if such confidence-building interventions lead to better uptake of the behaviour in the long run. Confidence could predict successful intervention uptake (Stankov et al., 2014).

In development contexts, confidence is frequently cited as a barrier to technology adoption. Research from low-income settings has shown that individuals are less likely to engage with digital financial tools if they lack confidence in navigating them (Demirgüç-Kunt et al., 2020). We add to a small but growing literature on the real-world application of behaviour science and psychology. We find our intervention increases confidence prior to doing the task, which is more likely to make the intervention sustainable (Bandura, 1997), and through a series of independent tasks and confidence judgements, we find both general ability and overconfidence bias a predictor of performance. Confidence-building strategies, like the pause and practice method, may thus play a critical role in bridging the digital divide and enhancing financial inclusion.

Conclusion

While more people join the digital financial infrastructure daily, a digital divide remains. Here, we tested interventions to teach simple QR code transactions to smartphone users through practicing alongside. We found that such intervention improves both performance and confidence in one’s ability to do the tasks.

Based on what we learned from our study, we recommend using active learning rather than passive videos for interventions. Both practising alongside and having practical information increased confidence in one's ability to do the task before attempting it, even in a novel platform. More interventions need to examine participants' confidence, which may be a better indicator of them accepting the behaviour and trying said behaviour on their own.

Bibliography

Abreu, M., & Mendes, V. (2012). Information, overconfidence and trading: Do the sources of information matter? *Journal of Economic Psychology*, *33*(4), 868–881. https://doi.org/10.1016/J.JOEP.2012.04.003

Ackerman, R. A. (2014). The diminishing criterion model for metacognitive regulation of time investment. *Journal of Experimental Psychology: General*, *143*(3), 1349–1368. https://doi.org/10.1037/A0035098

Aker, J. C., & Mbiti, I. M. (2010). Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives*, *24*(3), 207–232. https://doi.org/10.1257/JEP.24.3.207

Allwood, C. M., & Montgomery, H. (1987). Response selection strategies and realism of confidence judgments. *Organizational Behavior and Human Decision Processes*, *39*(3), 365–383. https://doi.org/10.1016/0749-5978(87)90029-X

Antonides, G., Manon de Groot, I., & Fred van Raaij, W. (2011). Mental budgeting and the management of household finance. *Journal of Economic Psychology*, *32*(4), 546–555. https://doi.org/10.1016/J.JOEP.2011.04.001

Asongu, S. A., Nwachukwu, J. C., & Orim, S. M. I. (2018). Mobile phones, institutional quality and entrepreneurship in Sub-Saharan Africa. *Technological Forecasting and Social Change*, *131*, 183–203. https://doi.org/10.1016/J.TECHFORE.2017.08.007

Bachhel, R., & Thaman, R. G. (2014). Effective use of pause procedure to enhance student engagement and learning. *Journal of Clinical and Diagnostic Research : JCDR*, *8*(8). https://doi.org/10.7860/JCDR/2014/8260.4691

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: self-efficacy drives performance gains with active learning. *CBE Life Sciences Education*, *16*(4). https://doi.org/10.1187/CBE.16-12-0344/ASSET/IMAGES/LARGE/AR56FIG2.TIF.GZ.JPEG

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Worth Publishers. https://psycnet.apa.org/record/1997-08589-000

Banerjee, A., & Duflo, E. (2012). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. In *Population and Development Review* (Vol. 37, Issue 4). PublicAffairs. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1728-4457.2011.00462.x

Berry, J. M., & West, R. L. (1993). Cognitive Self-efficacy in Relation to Personal Mastery and Goal Setting across the Life Span. *Http://Dx.Doi.Org/10.1177/016502549301600213*, *16*(2), 351–379. https://doi.org/10.1177/016502549301600213

Bhutoria, A., & Vignoles, A. (2018). Do Financial Education Interventions for Women from Poor Households Impact Their Financial Behaviors? Experimental Evidence from India. *Journal of Research on Educational Effectiveness*, *11*(3), 409–432. https://doi.org/10.1080/19345747.2018.1465317

Billari, F. C., Favero, C. A., & Saita, F. (2023). Online financial and demographic education for workers: Experimental evidence from an Italian Pension Fund. *Journal of Banking & Finance*, *151*, 106849. https://doi.org/10.1016/J.JBANKFIN.2023.106849

Blázquez, B. O., Masluk, B., Gascon, S., Díaz, R. F., Aguilar-Latorre, A., Magallón, I. A., & Botaya, R. M. (2019). The use of flipped classroom as an active learning approach improves academic performance in social work: A randomized trial in a university. *PLOS ONE*, *14*(4), e0214623. https://doi.org/10.1371/JOURNAL.PONE.0214623

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction*, *20*(5), 372–382. https://doi.org/10.1016/J.LEARNINSTRUC.2009.03.002

Bonwell, C. C., & Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom*. School of Education and Human Development, George Washington University.

Brown, S., & Taylor, K. (2014). Household finances and the 'Big Five' personality traits. *Journal of Economic Psychology*, *45*, 197–212. https://doi.org/10.1016/J.JOEP.2014.10.006

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring Higher the Second Time Around: Meta-Analyses of Practice Effects in Neuropsychological Assessment. *The Clinical Neuropsychologist*, *26*(4), 543–570. https://doi.org/10.1080/13854046.2012.680913

Camerer, C., & Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review*, *89*(1), 306–318. https://doi.org/10.1257/AER.89.1.306

Carpena, F., Cole, S., Shapiro, J., & Zia, B. (2017). The ABCs of Financial Education: Experimental Evidence on Attitudes, Behavior, and Cognitive Biases. *Https://Doi.Org/10.1287/Mnsc.2017.2819*, *65*(1), 346–369. https://doi.org/10.1287/MNSC.2017.2819

Collins, J. M. (2013). The impacts of mandatory financial education: Evidence from a randomized field study. *Journal of Economic Behavior & Organization*, *95*, 146–158. https://doi.org/10.1016/J.JEBO.2012.08.011

Delaney, P. F., Reder, L. M., Staszewski, J. J., & Ritter, F. E. (1998). The strategy-specific nature of improvement: The Power Law Applies by Strategy Within Task. *Psychological Science*, *9*(1), 1–7. https://doi.org/10.1111/1467-9280.00001

Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2020). The Global Findex Database 2017: Measuring Financial Inclusion and Opportunities to Expand Access to and Use of Financial Services. *The World Bank Economic Review*, *34*, 2–8. https://doi.org/10.1093/wber/lhz013

Dewsbury, B. M., Swanson, H. J., Moseman-Valtierra, S., & Caulkins, J. (2022). Inclusive and active pedagogies reduce academic outcome gaps and improve long-term performance. *PLOS ONE*, *17*(6), e0268620. https://doi.org/10.1371/JOURNAL.PONE.0268620

Dougherty, M. R. P. (2001). Integration of the Ecological and Error Models of Overconfidence Using a Multiple-Trace Memory Model. *Journal of Experimental Psychology: General*, *130*(4), 579–599. https://doi.org/10.1037/0096-3445.130.4.57

Driskell, J. E., Copper, C., & Moran, A. (1994). Does mental practice enhance performance? *Journal of Applied Psychology*, *79*(4), 481–492. https://doi.org/10.1037//0021-9010.79.4.481

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, *118*(3), 815–842. https://doi.org/10.1162/00335530360698432

Duncan, J., Chylinski, D., Mitchell, D. J., & Bhandari, A. (2017). Complexity and compositionality in fluid intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(20), 5295–5299. https://doi.org/10.1073/PNAS.1621147114/SUPPL_FILE/PNAS.201621147SI.PDF

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Https://Doi.Org/10.1287/Mnsc.2013.1849*, *60*(8), 1861–1883. https://doi.org/10.1287/MNSC.2013.1849

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with Certainty: The Appropriateness of Extreme Confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(4), 552–564. https://doi.org/10.1037/0096-1523.3.4.552

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8410–8415. https://doi.org/10.1073/PNAS.1319030111/SUPPL_FILE/PNAS.1319030111.ST04.DOCX

Ghosh, S. (2021). How important is trust in driving financial inclusion? *Journal of Behavioral and Experimental Finance*, *30*, 100510. https://doi.org/10.1016/J.JBEF.2021.100510

Gist, M. E. (1987). Self-Efficacy: Implications for Organizational Behavior and Human Resource Management. *Https://Doi.Org/10.5465/Amr.1987.4306562*, *12*(3), 472–485. https://doi.org/10.5465/AMR.1987.4306562

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24 1 SPEC.ISS*, 79–132. https://doi.org/10.1016/S0160-2896(97)90014-3

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature 2008 455:7213*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Hartman, B. A., Miller, B. K., & Nelson, D. L. (2000). The Effects of Hands-on Occupation Versus Demonstration on Children's Recall Memory. *The American Journal of Occupational Therapy*, *54*(5), 477–483. https://doi.org/10.5014/AJOT.54.5.477

Harvey, N. (1997). Confidence in judgment. *Trends in Cognitive Sciences*, *1*(2), 78–82. https://doi.org/10.1016/S1364-6613(97)01014-0

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013). Financial literacy, financial education, and economic outcomes. *Annual Review of Economics*, *5*(Volume 5, 2013), 347–373. https://doi.org/10.1146/ANNUREV-ECONOMICS-082312-125807/CITE/REFWORKS

Hearns, M. K., Miller, B. K., & Nelson, D. L. (2009). Hands-On Learning versus Learning by Demonstration at Three Recall Points in University Students. *Http://Dx.Doi.Org/10.3928/15394492-20090825-01*, *30*(4), 169–171. https://doi.org/10.3928/15394492-20090825-01

Hershfield, H., Shu, S., & Benartzi, S. (2019). Temporal Reframing and Participation in a Savings Program: A Field Experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3097468

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are Overconfident CEOs Better Innovators? *The Journal of Finance*, *67*(4), 1457–1498. https://doi.org/10.1111/J.1540-6261.2012.01753.X

Holzmann, R. (2010). Bringing Financial Literacy and Education to Low and Middle Income Countries. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.1707673

Inzana, C. M., Driskell, J. E., Salas, E., & Johnston, J. H. (1996). Effects of preparatory information on enhancing performance under stress. *The Journal of Applied Psychology*, *81*(4), 429–435. https://doi.org/10.1037/0021-9010.81.4.429

Isler, O., Rojas, A., & Dulleck, U. (2022). Easy to shove, difficult to show: Effect of educative and default nudges on financial self-management. *Journal of Behavioral and Experimental Finance*, *34*, 100639. https://doi.org/10.1016/J.JBEF.2022.100639

Jack, W., & Suri, T. (2014). Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. *American Economic Review*, *104*(1), 183–223. https://doi.org/10.1257/AER.104.1.183

Jackson, S. A., & Kleitman, S. (2014). Individual differences in decision-making and confidence: Capturing decision tendencies in a fictitious medical test. *Metacognition and Learning*, *9*(1), 25–49. https://doi.org/10.1007/S11409-013-9110-Y

Jin, S., Verhaeghen, P., & Rahnev, D. (2022). Across-subject correlation between confidence and accuracy: A meta-analysis of the Confidence Database. *Psychonomic Bulletin and Review*, *29*(4), 1405–1413. https://doi.org/10.3758/S13423-022-02063-7/TABLES/1

Johnson, D. D. P., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature 2011 477:7364*, *477*(7364), 317–320. https://doi.org/10.1038/nature10384

Juslin, P. (1994). The Overconfidence Phenomenon as a Consequence of Informal Experimenter-Guided Selection of Almanac Items. *Organizational Behavior and Human Decision Processes*, *57*(2), 226–246. https://doi.org/10.1006/OBHD.1994.1013

Kaiser, T., Lusardi, A., Menkhoff, L., & Urban, C. (2022). Financial education affects financial knowledge and downstream behaviors. *Journal of Financial Economics*, *145*(2), 255–272. https://doi.org/10.1016/J.JFINECO.2021.09.022

Kaiser, T., & Menkhoff, L. (2017). Does Financial Education Impact Financial Literacy and Financial Behavior, and If so, When? *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.2753510

Klayman, J., Soll, J. B., González-Vallejo, C., & Barlas, S. (1999). Overconfidence: It Depends on How, What, and Whom You Ask. *Organizational Behavior and Human Decision Processes*, *79*(3), 216–247. https://doi.org/10.1006/OBHD.1999.2847

Leopold, S. S., Morgan, H. D., Kadel, N. J., Gardner, G. C., Schaad, D. C., & Wolf, F. M. (2005). Impact of educational intervention on confidence and competence in the performance of a simple surgical task. *Journal of Bone and Joint Surgery*, *87*(5), 1031–1037. https://doi.org/10.2106/JBJS.D.02434

Li, W., Lee, A. M., & Solmon, M. A. (2005). Relationships among Dispositional Ability Conceptions, Intrinsic Motivation, Perceived Competence, Experience, and Performance. *Journal of Teaching in Physical Education*, *24*(1), 51–65. https://doi.org/10.1123/JTPE.24.1.51

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, *20*(2), 159–183. https://doi.org/10.1016/0030-5073(77)90001-0

Lundeberg, M. A., Fox, P. W., & Punćochaŕ, J. (1994). Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments. *Journal of Educational Psychology*, *86*(1), 114–121. https://doi.org/10.1037/0022-0663.86.1.114

Margolis, H., & McCabe, P. P. (2004). Self-Efficacy: A Key to Improving the Motivation of Struggling Learners. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *77*(6), 241–249. https://doi.org/10.3200/TCHS.77.6.241-249

Mazzotta, B. D., Chakravorti, B., & Bijapurkar, R. (2023). *THE COST OF CASH IN INDIA* .

McCaffrey, R. J., Duff, K., & Westervelt, H. J. (2000). *Practitioner's Guide to Evaluating Change with Neuropsychological Assessment Instruments* (R. J. McCaffrey, K. Duff, & H. J. Westervelt, Eds.). Springer US. https://doi.org/10.1007/978-1-4615-4233-9

Mcgraw, A. P., Mellers, B. A., & Ritov, I. (2004). The affective costs of overconfidence. *Journal of Behavioral Decision Making*, *17*(4), 281–295. https://doi.org/10.1002/BDM.472

McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, *32*(6), 493–510. https://doi.org/10.1017/S0140525X09990975

McMurran, M., Weisbart, D., & Atit, K. (2023). The relationship between students' gender and their confidence in the correctness of their solutions to complex and difficult mathematics problems. *Learning and Individual Differences*, *107*, 102349. https://doi.org/10.1016/J.LINDIF.2023.102349

Mengelkamp, C., & Bannert, M. (2010). Accuracy of confidence judgments: Stability and generality in the learning process and predictive validity for learning outcome. *Memory and Cognition*, *38*(4), 441–451. https://doi.org/10.3758/MC.38.4.441

Parkinson, T. M., Force, J. E., & Smith, J. K. (2003). Hands-on Learning: Its Effectiveness in Teaching the Public about Wildland Fire. *Journal of Forestry*, *101*(7), 21–26. https://doi.org/10.1093/JOF/101.7.21

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/J.2168-9830.2004.TB00809.X

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Rai, A., Sharif, M. A., Chang, E. H., Milkman, K. L., & Duckworth, A. L. (2023). A field experiment on subgoal framing to boost volunteering: The trade-off between goal granularity and flexibility. *The Journal of Applied Psychology*, *108*(4), 621–634. https://doi.org/10.1037/APL0001040

Richards, L. W., Wang, A. T., Mahapatra, S., Jenkins, S. M., Collins, N. M., Beckman, T. J., & Wittich, C. M. (2017). Use of the pause procedure in continuing medical education: A randomized controlled intervention study. *Medical Teacher*, *39*(1), 74–78. https://doi.org/10.1080/0142159X.2016.1230664

Røe, Y., Wojniusz, S., & Bjerke, A. H. (2022). The Digital Transformation of Higher Education Teaching: Four Pedagogical Prescriptions to Move Active Learning Pedagogy Forward. *Frontiers in Education*, *6*, 784701. https://doi.org/10.3389/FEDUC.2021.784701/BIBTEX

Ruhl, K. L., Hughes, C. A., & Gajar, A. H. (1990). Efficacy of the Pause Procedure for Enhancing Learning Disabled and Nondisabled College Students' Long- and Short-Term Recall of Facts Presented through Lecture. *Http://Dx.Doi.Org/10.2307/1510392*, *13*(1), 55–64. https://doi.org/10.2307/1510392

Schmidt, F. L., & Hunter, J. (2004). General Mental Ability in the World of Work: Occupational Attainment and Job Performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162

Shashidhara, S., Mamidi, P., Gupta, G., & Radhakrishnan, D. (2024). *Improving Parental Engagement in Foundational Literacy and Numeracy Learning* (2307). https://csbc.org.in/upload/Parental_Engagement_in_FLN.pdf

Smith, R. C., Loibl, C., Imam, K., & Jones, L. (2020). Improving Financial Behaviors Through Nudges. *The Journal of Extension*, *58*(1), 2. https://doi.org/10.34068/joe.58.01.02

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201. https://doi.org/10.2307/1412107

Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence*, *28*(2), 121–143. https://doi.org/10.1016/S0160-2896(99)00033-1

Stankov, L. (2004). Mining on the "no man's land" between intelligence and personality. *Learning and Individual Differences: Process, Trait, and Content Determinants.*, 315–337. https://doi.org/10.1037/10315-014

Stankov, L., & Lee, J. (2008). Confidence and Cognitive Test Performance. *Journal of Educational Psychology*, *100*(4), 961–976. https://doi.org/10.1037/A0012546

Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences*, *22*(6), 747–758. https://doi.org/10.1016/J.LINDIF.2012.05.013

Stankov, L., Morony, S., & Lee, Y. P. (2014). Confidence: the best non-cognitive predictor of academic achievement? *Educational Psychology*, *34*(1), 9–28. https://doi.org/10.1080/01443410.2013.814194

StataCorp. (2023). *Stata Statistical Software* (Release 18). StataCorp LLC.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, *47*(2), 143–148. https://doi.org/10.1016/0001-6918(81)90005-6

Tang, N., & Baker, A. (2016). Self-esteem, financial knowledge and financial behavior. *Journal of Economic Psychology*, *54*, 164–176. https://doi.org/10.1016/J.JOEP.2016.04.005

Tang, N., Baker, A., & Peter, P. C. (2015). Investigating the Disconnect between Financial Knowledge and Behavior: The Role of Parental Influence and Psychological Characteristics in Responsible Financial Behaviors among Young Adults. *Journal of Consumer Affairs*, *49*(2), 376–406. https://doi.org/10.1111/JOCA.12069

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge*. Yale University Press.

Trivers, R. (2000). The Elements of a Scientific Theory of Self-Deception. *Annals of the New York Academy of Sciences*, *907*(1), 114–131. https://doi.org/10.1111/J.1749-6632.2000.TB06619.X

Vera, M., Le Blanc, P. M., Taris, T. W., & Salanova, M. (2014). Patterns of engagement: the relationship between efficacy beliefs and task engagement at the individual versus collective level. *Journal of Applied Social Psychology*, *44*(2), 133–144. https://doi.org/10.1111/JASP.12219

Walstad, W., Urban, C., J. Asarta, C., Breitbach, E., Bosshardt, W., Heath, J., O'Neill, B., Wagner, J., & Xiao, J. J. (2017). Perspectives on evaluation in financial education: Landscape, issues, and studies. *The Journal of Economic Education*, *48*(2), 93–112. https://doi.org/10.1080/00220485.2017.1285738

Weinberg, B. A. (2009). A MODEL OF OVERCONFIDENCE. *Pacific Economic Review*, *14*(4), 502–515. https://doi.org/10.1111/J.1468-0106.2009.00466.X

Weisman, K., & Markman, E. M. (2017). Theory-based explanation as intervention. *Psychonomic Bulletin and Review*, *24*(5), 1555–1562. https://doi.org/10.3758/S13423-016-1207-2/METRICS

Zhang, M. (2023). Older people's attitudes towards emerging technologies: A systematic literature review. *Public Understanding of Science (Bristol, England)*, *32*(8), 948. https://doi.org/10.1177/09636625231171677

Tables and Figures

Table 1: A brief description of the Videos in each of the arms.

Video	Description of Message	Duration
Control	The video explains the reasons for using UPI transactions and then explains QR code transactions using the PhonePe app, from finding the app to entering the UPI ID and completing the transaction.	3 min 47 sec
Pause and Practice	The same video as above has additional instructions to pause the video and do the step alongside on a different device at five different intervals.	4 min 25 sec
Pause and Practice with additional info	The same video as T1 with five pause-and-practice instructions and a few additional pointers to make the task easier and inspire confidence and trust in this DFS process. Additional info: 1. Video prompts users to ensure internet and location services are enabled before launching the app 2. Explains what a QR code is and how it's more robust than manually entering an account or phone number 3. Explains that the phone may ask permission to use the camera on the first try and what to press when they see such a window. 4. Mentions that scanning will not happen without the internet 5. Mentions that internet is needed to proceed from the amount page UPI ID page. 6. Same UPI ID is used across apps, and there is a limit of 10 transactions a day with a given UPI ID Motivational/Confidence building: 1. Only the amount specified by you can be transferred 2. How to keep the UPI pin safe (not to share/ write it down) 3. Entering the wrong PIN will not lead to any additional	6 min 55 sec

	problems, just that there won't be any transactions. 4. In case the phone is lost, no one can use your account as they won't have access to the UPI ID. 5. If someone tries to use your account and enters the wrong ID 3 times, the account is locked for 24 hours Grievance Redressal - 1. Mutually agree on the amount sent to the beneficiary and sort out any changes in the amount. In case of extra money is sent, one can also approach the beneficiary's bank.

Table 2: Results from Regression. The outcome variables are accuracy for PhonePe and GPay tasks immediately after intervention and two days later. The treatment effects indicate coefficients and standard errors for ordinal logistic regressions followed by mean and standard deviation for each outcome variable. Regressions include digital and financial literacy and demographic controls (not shown).

	Task 1 Accuracy Immediate outcome		Task 2 Accuracy Immediate outcome		Task 1 Accuracy After 2 days		Task 2 Accuracy After 2 days	
	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]
Pause and Practice	1.225***	0.820	0.606*	0.725	0.698**	0.758	0.0116	0.656
	[0.255]	[0.024]	[0.245]	[0.031]	[0.231]	[0.026]	[0.243]	[0.033]
Pause and Practice Plus	1.218***	0.822	0.629*	0.747	0.484+	0.709	0.093	0.683
	[0.278]	[0.028]	[0.262]	[0.034]	[0.256]	[0.035]	[0.259]	[0.037]
Control	-	0.638	-	0.619	-	0.616	-	0.626
		[0.032]		[0.035]		[0.032]		[0.036]

*+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001*

Table 3: Results from Regression. The outcome variables are pre-confidence (confidence before doing the tasks) for PhonePe and GPay tasks immediately after intervention and two days later. The treatment effects indicate coefficients and standard errors for ordinal logistic regressions followed by mean and standard deviation for each outcome variable. Regressions include, digital and financial literacy, demographic controls (not shown).

	PhonePe Pre-Confidence Immediate outcome		GPay Pre-Confidence Immediate outcome		PhonePe Pre-Confidence After 2 days		GPay Pre-Confidence After 2 days	
	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]
Pause and Practice	0.408+	0.462	0.527*	0.443	0.351+	0.579	0.376+	0.541
	[0.209]	[0.028]	[0.212]	[0.027]	[0.211]	[0.027]	[0.237]	[0.214]
Pause and Practice Plus	0.721**	0.526	0.510*	0.450	0.364	0.596	0.079	0.253
	[0.228]	[0.033]	[0.232]	[0.035]	[0.231]	[0.032]	[0.249]	[0.031]
Control	-	0.340	-	0.361	-	0.527	-	0.226
		[0.030]		[0.029]		[0.031]		[0.032]

*+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001*

Table 4: Results from Regression. The outcome variables are post-confidence (confidence after doing the tasks) for PhonePe and GPay tasks immediately after intervention and two days later. The treatment effects indicate coefficients and standard errors for ordinal logistic regressions followed by mean and standard deviation for each outcome variable. Regressions include digital and financial literacy, and demographic controls (not shown).

	PhonePe Post-Confidence Immediate outcome		GPay Post-Confidence Immediate outcome		PhonePe Post-Confidence After 2 days		GPay Post-Confidence After 2 days	
	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]	OLogit Treatment Effect [SE]	Mean [SD]
Pause and Practice	1.184***	0.496	0.704**	0.326	0.304	0.390	0.448	0.834
	[0.243]	[0.032]	[0.234]	[0.034]	[0.252]	[0.034]	[0.348]	[0.027]
Pause and Practice Plus	0.980***	0.492	0.600*	0.448	0.275	0.394	-0.023	0.773
	[0.260]	[0.040]	[0.255]	[0.040]	[0.272]	[0.040]	[0.366]	[0.040]
Control	-	0.306	-	0.361	-	0.328	-	0.777
		[0.034]		[0.029]		[0.036]		[0.042]

$+ p < 0.10$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$

Table 5: Results from Regression. The outcome variables are accuracy, pre-confidence, and post-confidence for PhonePe and GPay tasks immediately after intervention and two days later. The treatment effects indicate coefficients and standard errors for ordinal logistic regressions. Regressions include experiment group dummy, digital and financial literacy, and demographic controls (not shown). IO refers to immediate outcome, and A2d refers to after two days.

		Average Reaction Time		Average Accuracy		Average Confidence		Average Bias	
		OLogit Treatment Effect	[SE]	OLogit Treatment Effect	[SE]	OLogit Treatment Effect	[SE]	OLogit Treatment Effect	[SE]
Accuracy	PhonePe IO	-0.018	[0.028]	2.728*	[1.083]	0.091	[0.637]	-0.889	[0.654]
	GPay IO	-0.030	[0.027]	2.015*	[1.027]	0.848	[0.631]	0.081	[0.652]
	PhonePe A2d	-0.029	[0.026]	4.219***	[1.046]	-0.010	[0.627]	-1.631*	[0.647]
	GPay A2d	-0.016	[0.026]	4.381***	[1.054]	0.795	[0.620]	-0.832	[0.639]
Pre-Confidence	PhonePe IO	-0.043+	[0.023]	2.707**	[0.929]	2.654***	[0.604]	1.682**	[0.613]
	GPay IO	-0.043+	[0.023]	2.143*	[0.951]	4.283***	[0.701]	3.429***	[0.688]
	PhonePe A2d	-0.030	[0.023]	3.245***	[0.940]	4.648***	[0.642]	3.337***	[0.627]
	GPay A2d	-0.044+	[0.023]	3.645***	[0.928]	4.943***	[0.663]	3.350***	[0.640]
Post-Confidence	PhonePe IO	-0.002	[0.025]	4.331***	[1.113]	2.534***	[0.719]	0.808	[0.700]
	GPay IO	-0.011	[0.025]	4.023***	[1.102]	3.418***	[0.769]	1.623*	[0.708]
	PhonePe A2d	0.013	[0.028]	5.707***	[1.257]	2.335**	[0.780]	0.155	[0.745]
	GPay A2d	0.033	[0.040]	4.833**	[1.689]	5.351***	[1.033]	5.037***	[1.223]

*+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001*

Table 6: Results from Regression. The outcome variables are accuracy, pre-confidence, and post-confidence for PhonePe and GPay tasks immediately after intervention and two days later. The treatment effects indicate coefficients and standard errors for digital and financial literacy of ordinal logistic regressions described in Table 2-4. Regressions include demographic controls (not shown). IO refers to immediate outcome, and A2d refers to after two days.

		Digital Literacy		Financial Literacy	
		OLogit Treatment Effect	[SE]	OLogit Treatment Effect	[SE]
Accuracy	PhonePe IO	1.316*	[0.521]	3.130***	[0.607]
	GPay IO	1.243*	[0.502]	1.612**	[0.551]
	PhonePe A2d	1.778***	[0.485]	1.162*	[0.514]
	GPay A2d	1.964***	[0.498]	0.993+	[0.525]
Pre-Confidence	PhonePe IO	1.477**	[0.454]	0.962*	[0.438]
	GPay IO	1.524**	[0.466]	0.954*	[0.429]
	PhonePe A2d	2.189***	[0.464]	1.801***	[0.452]
	GPay A2d	1.385**	[0.458]	1.378**	[0.443]
Post-Confidence	PhonePe IO	1.906***	[0.519]	1.840***	[0.485]
	GPay IO	1.085*	[0.507]	1.451**	[0.483]
	PhonePe A2d	2.680***	[0.587]	1.465**	[0.512]
	GPay A2d	1.581+	[0.865]	1.416*	[0.696]

*+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

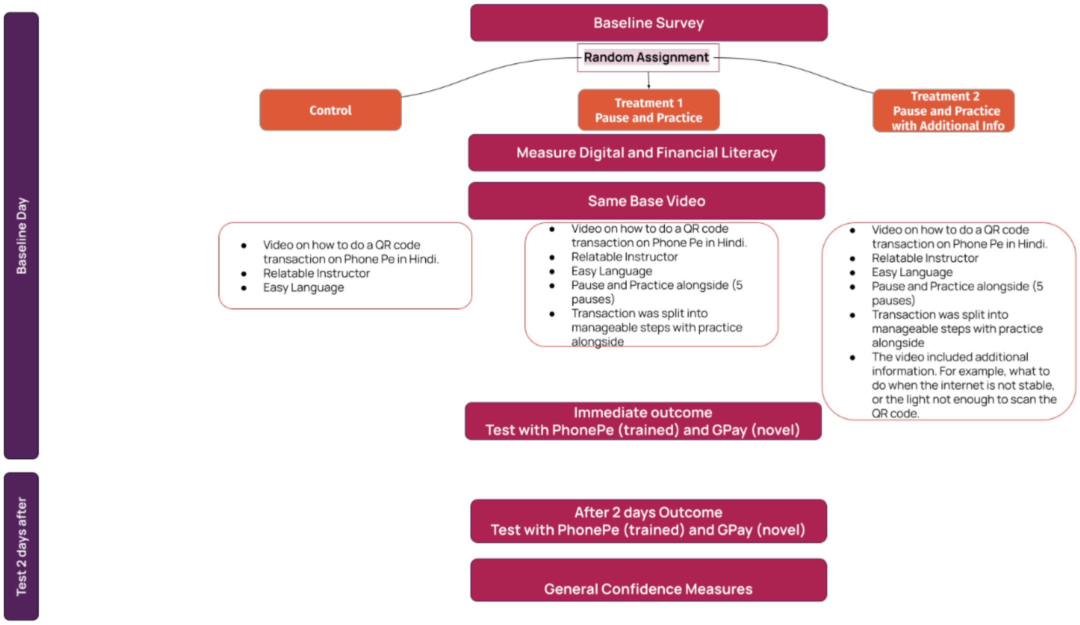

Figure 1: The Figure shows the overall study structure. People were randomly assigned to treatment and control groups. On the first day, they were shown a video explaining QR code transactions in the PhonePe app. Immediately afterwards they were tested on QR code transactions in PhonePe and GPay. Two days later, at endline they were tested again on the same tasks.

Appendix

Table A1: Balance Checks for all the control variables.

	(1)	(2)	(3)	t-test / Fischer	t-test / Fischer
Variable	Control	Pause and Practice	Pause and Practice with FAQ	Difference (1)-(2)	Difference (1)-(3)
Mean age (SE)	27.28 (0.64)	25.98 (0.59)	27.18 (0.66)	-1.5	-0.12
Gender: Prop. Female	0.77	0.70	0.80	0.70	1.19
Household members	7.99 (0.37)	8.18 (0.40)	8.07 (0.37)	0.35	0.16
Teen members: Prop. 1	0.25	0.21	0.30	0.80	1.26
Teen members: Prop. 2	0.32	0.35	0.37	1.13	1.25
Teen members: Prop. more than 2	0.43	0.44	0.34	1.05	0.66
Education: Prop. 10th pass or less	0.34	0.32	0.24	0.88	0.60+
Education: Prop. 12th pass	0.26	0.38	0.33	1.78*	1.41
Education: Prop. Diploma or Higher	0.40	0.30	0.43	0.66+	1.15
Employment: Prop. agriculture and mnrega	0.38	0.40	0.37	1.11	0.96
Employment: Prop. managing their household	0.28	0.28	0.42	0.98	1.80*
Employment: Prop. Other	0.34	0.32	0.22	0.90	0.54*
Income: Prop. less than 5k	0.22	0.32	0.21	1.65+	0.92
Income: Prop. 5k - 10k	0.35	0.36	0.38	1.05	1.11
Income: Prop. more than 10k	0.43	0.32	0.42	0.62*	0.96
Religion: Prop. Hindu	0.91	0.95	0.96	1.73	2.31

Caste: Prop. General	0.27	0.25	0.30	0.88	1.18
Caste: Prop. OBC	0.32	0.35	0.37	1.13	1.25
Caste: Prop. Other	0.41	0.41	0.33	0.99	0.70
Own smartphone duration: Prop. more than year	0.64	0.60	0.63	0.85	0.99

*+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

Relationship between the outcome measures

Along with the independent variable of the treatment arm and the covariates of digital and financial literacy and demographic controls, we ran ordered logit regressions to understand the relationship between the outcome measures. The outcome measures were arranged chronologically, and each measure's impact on all future measures was calculated (Fig A1). For example, the pre-confidence of the PhonePe task immediately after the intervention was the first measure, and thus, its impact on all other measures was computed. Each measure strongly affected all other measures, as expected in related tasks. PhonePe's pre-confidence Immediate outcome did not correlate with the accuracy of both tasks two days later. Post-confidence Immediate outcome measures of both tasks and GPay two days later did not affect their accuracy. PhonePe accuracy two days later did not affect the post-confidence of the Gpay task. Overall, we see a strong correlation between accuracy and confidence measures, which aligns with the literature (Jackson & Kleitman, 2014; Jin et al., 2022).

Relationship between the General Tasks

We computed Pearson correlations of reaction time, accuracy, confidence, and bias across the three general tasks of analogy, panamath and vocabulary. As expected, there is a high correlation between these task variables (Table A2). There was a stronger correlation between analogy and vocabulary tasks than with panamath in both accuracy and confidence. This illustrates the domain effect of analogy and vocabulary as language tasks and panamath as a more perceptual task (Jackson & Kleitman, 2014; Jin et al., 2022). This difference is seen in reaction time. Bias suggests a stronger relationship between analogy and panamath. Both these tasks involved a form of decision-making involving fluid intelligence that may have been absent in vocabulary, which involves only crystallised intelligence (Jackson & Kleitman, 2014).

		Pre-Confidence	Post-Confidence	Accuracy	Pre-Confidence	Post-Confidence	Accuracy	Pre-Confidence	Post-Confidence	Accuracy	Pre-Confidence	Post-Confidence	Accuracy
		PhonePe IO	PhonePe IO	PhonePe IO	GPay IO	GPay IO	GPay IO	PhonePe A2d	PhonePe A2d	PhonePe A2d	GPay A2d	GPay A2d	GPay A2d
		Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]	Logit Treatment Effect [SE]
Pre-Confidence	PhonePe IO		3.308***	1.252***	3.913***	1.736***	0.680*	3.538***	0.581+	0.295	2.373***	2.805***	0.273
			[0.343]	[0.340]	[0.335]	[0.304]	[0.318]	[0.321]	[0.320]	[0.301]	[0.292]	[0.481]	[0.302]
Post-Confidence	PhonePe IO			70.070	2.500***	3.791***	2.352***	2.626***	1.779***	1.544***	2.207***	2.129***	1.310***
				[2288.2]	[0.276]	[0.328]	[0.336]	[0.276]	[0.295]	[0.291]	[0.267]	[0.437]	[0.293]
Accuracy	PhonePe IO				1.163***	3.979***	3.365***	1.650***	3.322***	3.378***	1.663***	1.587+	2.400***
					[0.309]	[0.517]	[0.366]	[0.305]	[0.544]	[0.361]	[0.305]	[0.917]	[0.347]
Pre-Confidence	GPay IO					3.125***	1.320***	3.159***	1.424***	0.617+	3.306***	2.456***	0.779*
						[0.345]	[0.336]	[0.323]	[0.330]	[0.319]	[0.323]	[0.456]	[0.318]
Post-Confidence	GPay IO						66.400	2.527***	2.162***	1.235***	2.722***	2.399***	1.686***
							[1461.6]	[0.270]	[0.293]	[0.285]	[0.273]	[0.413]	[0.302]
Accuracy	GPay IO							1.322***	2.948***	1.715***	1.616***	1.586**	2.437***
								[0.250]	[0.449]	[0.273]	[0.256]	[0.614]	[0.295]
Pre-Confidence	PhonePe A2d								2.484***	0.945**	5.158***	5.438***	0.969**
									[0.377]	[0.307]	[0.389]	[0.665]	[0.316]
Post-Confidence	PhonePe A2d									5.546***	2.838***	5.782***	4.977***
										[0.640]	[0.289]	[0.670]	[0.578]
Accuracy	PhonePe A2d										1.282***	-2.111	3.180***
											[0.267]	[3.804]	[0.336]
Pre-Confidence	GPay A2d											5.765***	1.963***
												[0.666]	[0.331]
Post-Confidence	GPay A2d												1.310
													[0.183]
Accuracy	GPay A2d												

Figure A1: The above figure shows the effect of outcome measures on each other. We see that all accuracy and confidence measures are strongly correlated.

Table A2:

Task Combinations	Reaction Time (r)	Accuracy (r)	Confidence (r)	Bias (r)
Analogy-Panamath	0.523***	0.251***	0.739***	0.584***
Analogy-Vocabulary	0.614***	0.643***	0.823***	0.341***
Panamath-Vocabulary	0.610***	0.289***	0.730***	0.387***

*+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001*