



City Research Online

City St George's, University of London

Citation: Priolo, G., Stablum, F., Vacondio, M., D'Ambrogio, S., Caserotti, M., Conte, B., De Roni, P., du Plooy, H., Grillo, V. D., Mastromatteo, L. Y., et al (2026). The Robustness of Mental Accounting Across 21 Countries. Journal of Consumer Research, doi: 10.1093/jcr/ucag022

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/38475/>

Link to published version: <https://doi.org/10.1093/jcr/ucag022>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

The Robustness of Mental Accounting Across 21 Countries

GIULIA PRIOLO 
FEDERICA STABLUM
MARTINA VACONDIO
SIMONE D'AMBROGIO
MARTA CASEROTTI
BEATRICE CONTE
PRISCA DE RONI
HILDA DU PLOOY
VIVIAN D. GRILLO
LIBERA Y. MASTROMATTEO
ELISA TEDALDI
FILIPPO TOSCANO 
JESÚS AGUILAR-ARMIJO
PARISA AHMADI GHOMROUDI
LUCIAN ALEXA
MATHIAS H. ANDERSEN
PER A. ANDERSSON
KARINE AOUN BARAKAT
CAROLINA BARROS
RUGGERO BASANISI
TARA BEILNER
SERGIU BURLACU
THAI CAO
ALESSANDRA CARELLA
ARIANNA CHIAPPI
ZAFER ÇİFTÇİ
CLAUDIA CİVAI
ALANA DALY
VALDONĖ DARŠKUVIENĖ
MARTA DE PEDIS
EARLE J. DU PLOOY
MOHAMMED EL-MIR
CHRISTIAN T. ELBÆK
SONDOS ELKOT
VALERIA FANGHELLA

EMAN FARAHAT
AMY G. FEHL 
AMA P. FENNY 
PAUL A. G. FORBES
GEMMA GARBI
ANDRÉ GONÇALVES
SEVIAS GUVURIRO
ALI HAJIAN
STEVE HEINKE
AUSTIN HOWARD
SUDHARSANA JAGATHEESH
JAYANAND
PEIRAN JIAO
GABRIELA JIGA-BOY
ALEJANDRA JORDANO DE
CASTRO
TOBIAS KALENSCHER
AUSTĚJA KAŽEMEKAITYTĖ
AFREEN KHALID
KIANA KOTHE
PHILIP KRÜGER
NGÂN LÊ THỊ KIỀU
GINTARĖ LECKĖ
YANINA LEDOVAYA
MENGYU LIM
LUCA M. LÜPKEN
HUONG MAI THI XUAN
LAURA MANGOLD
ALFARISI MAULANA
MAYA MAZE
HAJDI MOCHE
ZAHRA MORADI
ADEL MOUMIN
VALERIA NAVA
MICHELLE J. Y. NEOH

LEONARDO NICOLAO
HAMZA O. K. EL HALLAOUI
SEBASTIAN OLSCHESKI
ADOBEA Y. OWUSU
AHMET F. OZATES
SOFIA PELICA
SONJA PERKOVIC
ANANDA W. M. PUTERI
HAGAI RABINOVITCH
GUILHERME A. RAMOS
NICOLE ROBITAILLE 
CAROLINE ROUX 
BENJAMIN SCHEIBEHENNE
MARTIN SCHOEMANN
MOHAMMAD SEIDISAROEI
SANJAY SINGH
MUSTAFA Z. SÖYÜK
LIZA STEINER
AMIRA TAREK
BERTO USMAN
HANNAH VAN ALEBEEK
MOHAMMAD H. VAZIRIAN
EVGENIYA VEDERNIKOVA
JANET L. WIJAYA
XINXIN ZHU
JICHUAN ZONG
LEAF VAN BOVEN 
STEPHAN DICKERT
LORELLA LOTTO
DAVID J. HARDISTY 
JUSTIN POMERANCE 
BEATRIZ PEREIRA
KAI RUGGERI
ENRICO RUBALTELLI 

Editors: June Cotte and Stephen A. Spiller

Associate Editor: Carey K. Morewedge

Advance Access publication July 10, 2026

Giulia Priolo (gip.msc@cbs.dk) is a Postdoctoral Researcher at the Department of Management, Society and Communication, Copenhagen Business School, Dalgas Have, 15, 2000 Frederiksberg, Denmark. Federica Stablum (federica.stablum@gmail.com) is a Research Fellow at the Department of Mechanical Engineering, University College London, Roberts Engineering Building, Torrington Place, London, WC1E 7JE, United Kingdom. Martina Vacondio (martina.

continued on page 22

© The Author(s) 2026. Published by Oxford University Press on behalf of Journal of Consumer Research, Inc. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com. • Vol. 00 • 2026 <https://doi.org/10.1093/jcr/ucag022>

First introduced four decades ago, the influential concept of mental accounting—how people mentally organize, evaluate, and track financial activities—posits that consumers often defy traditional economic rationality, treating money as non-fungible across discrete mental accounts. The present study provides the first large-scale test of the replicability and generalizability of mental accounting effects, using a sample of 5,589 participants from 21 countries. The results demonstrate that mental accounting effects are replicable, robust, and generalizable. Hierarchical Bayesian meta-analyses revealed a 100% replication rate for all tested scenarios, while unpooled analyses showed a 90.5% replication rate (133/147 effects). Further analyses indicated that mental accounting effects were weaker in lower-income countries than in higher-income countries. Multidimensional scaling suggested that mental accounting effects vary along three interpretable dimensions that reflect social context (individual vs. interactive decisions), decision perspective (deciding for self vs. other), and role in price determination (setting vs. evaluating prices). Across a diverse population and after controlling for multiple factors, the findings show that consumers make decisions based on mentally formed accounts that consistently diverge from their objective financial value.

Keywords: mental accounting, replication, behavioral science, psychology, consumer behavior, non-WEIRD population

Consumers have long been known to arrange spending into different categories to mentally organize financial information used for making decisions (Thaler 1985). Planned allocations and balances for each of these individual “mental accounts” are some of the bases of such decisions. Because of this categorization, the same amount of money can be perceived differently depending on its “account.” For example, consumers perceive \$100 for leisure differently from \$100 for rent. Similarly, a \$100 check received for overtime work is perceived differently from a \$100 check for a tax refund. This earmarking of money is the cornerstone of the mental accounting model of consumer behavior originally developed over the 1980s and 1990s (Kahneman and Tversky 1982; Thaler 1985, 1990, 1999; Tversky and Kahneman 1981).

However, since Mental Accounting was introduced and empirically tested, modern technology has completely changed our relationship with money and the way we spend it (Shah, Khan and Khan 2025). In several financial environments, the norm now would be for both the overtime pay and the tax rebate to be received via direct deposit. Checks are also used less frequently as a mode of payment, with innovations like digital payment platforms being rapidly adopted. Cash use has declined significantly, although its prevalence still varies across countries (Shy 2023; Tarlin 2021).

Considering the substantial changes in consumer financial ecosystems that have occurred in the last few decades, it is appropriate to ask whether classic mental accounting principles continue to describe how consumers organize, evaluate, and allocate money. To address

this issue, the current research presents a large-scale test of the replicability and generalizability of mental accounting effects, using a sample of 5,589 participants from 21 countries.

This research provides three main contributions. First, we show that mental accounting remains a robust construct in modern times across different countries and demographic factors, with only slightly weaker effect sizes than the original studies. Second, we explore the relationship between a country’s economic development—assessed as Gross Domestic Product (GDP) per capita—and between-country differences in mental accounting. Finally, we propose three distinct dimensions (via multidimensional scaling analysis) across which mental accounting effects appear to vary. The three dimensions include social context (interactive vs. individual decisions), decision perspective (deciding for self vs. other), and role in price determination (setting vs. evaluating prices), suggesting a broader structure than the original, largely unidimensional account, and presenting a clear framework for valuable future research.

THE ROLE OF MENTAL ACCOUNTING IN CONSUMER RESEARCH

Mental accounting (Thaler 1985, 1990, 1999) is one of the most important models used in consumer research to understand how individuals manage their expenses and money. The concept was first introduced by Tversky and Kahneman (1981) (see also Kahneman and Tversky 1982) and further developed by Thaler (1999, 183) who

defined it as “the set of cognitive operations used by individuals and households to organize, evaluate, and keep track of financial activities.” The foundational evidence for this model precedes technological innovations that significantly changed consumers’ relationships with money and the way they use it. Despite this, mental accounting is still the topic of numerous experimental scenarios and real-world applications in fields like finance (Barberis and Huang 2001; Choi, Laibson and Madrian 2009; Shefrin and Statman 2000), charitable giving (Sussman, Sharma, and Alter 2015), savings and debts (Thaler 1990), and energy consumption (Hahnel *et al.* 2020). Thaler’s work, in particular, played a fundamental role in the early development of behavioral sciences.

Despite initial resistance (Shiller 2017), mental accounting is now generally accepted as a key framework for studying and explaining human behavior, as evidenced by publications in leading economics (Kőszegi and Matějka 2020; Lian 2021) and marketing journals since its inception (Gourville 1998; Heath and Soll 1996; Prelec and Loewenstein 1998; Reinholdt, Bartels, and Parker 2015; Soman and Cheema 2011; Soman and Gourville 2001). The theory has continued to develop in the last 10 years (Atlas and Bartels 2018; Choe and Kan 2021; Evers, Imas, and Kang 2022; Fei, Bartels, and Zhang 2026; Hossain 2018; Kelting, Robinson, and Lutz 2019; Lee and Morewedge 2023; Tully and Sharma 2018; see also Sussman, Wang, and Apalkova 2023).

To provide a sense of the importance of mental accounting, a Web of Science search conducted in March 2026 returned a total of 712 papers published on the topic between 1984 and 2026, 45 of which appeared in the five top marketing journals (i.e., *Journal of Consumer Research*, *Journal of Marketing Research*, *Journal of Marketing*, *Journal of Consumer Psychology*, and *Marketing Science*). These 712 papers were cited 24,065 times, with the subset of 45 published in top marketing journals being cited 7,113 times. Similarly, the original papers have all been widely cited: Tversky and Kahneman (1981) 13,280 citations, Kahneman and Tversky (1982) 1,235 citations, Thaler (1985) 2,657 citations, Thaler (1999) 2,453 citations, and Shafir and Thaler (2006) 123 citations. The pivotal role played by mental accounting in influencing the study of consumer behavior is showcased by the decision to award the 2017 Nobel Prize in economics to Richard Thaler. Two recent literature reviews (Silva, Moreira, and Bortolon 2023; Skwara 2023) further reflect the current interest in the topic.

One of the most significant contributions of mental accounting to the understanding of consumer behavior is the integration of psychological constructs into the analysis of consumers’ decision-making processes,

which challenges the classical economic view of consumer rationality. Thaler (1985) suggested that the economic theory of the consumer did not lead to widespread applications in marketing because it neglected most of the marketing variables with the exception of price and product characteristics. Instead, Thaler explained how many marketing variables would fall into the category that Tversky and Kahneman (1981) had labeled as framing. Thus, Thaler (1985) proposed a set of new concepts showing that consumers are impacted by the way they code outcomes as gains versus losses, by how they evaluate a purchase (transaction utility), and by budgetary rules.

For instance, Thaler’s early work showed that consumers create distinct mental accounts for specific types of expenses (e.g., paying bills vs. paying for leisure activities) and are unwilling to transfer money across different accounts. This is a violation of the principle of fungibility (Abeler and Marklein 2017; Thaler 1990), which states that different forms of wealth can be substituted for each other, that is, be used interchangeably to meet a person’s economic needs. Individuals may instead prefer to track their budget and spending by splitting their economic activities into different categories based on what they use their money for (Heath and Soll 1996; Soman and Gourville 2001; Soman and Lam 2002) or the medium used to make a purchase (Liu and Chou 2018; Liu and Dewitte 2021; Prelec and Loewenstein 1998; Thomas, Desai, and Seenivasan 2011).

The creation of different mental accounts has thus been considered an intuitive strategy that consumers use to maintain self-control on their expenses (Thaler 1999). For instance, most people associate their investments with savings and are reluctant to use them for exceptional expenses (e.g., remodeling their kitchen). Consumers also use money differently depending on how they categorize it (i.e., depending on the mental account “opened” for a specific purchase). For example, they tend to use a tax refund for exceptional and hedonic expenses rather than everyday purchases (as they would with their salary), or to save it (as they usually do when receiving a bonus from stock investments; Shefrin and Thaler 1988). In addition, mental accounting also influences decisions as consumers rely on contextual cues (e.g., how a specific expense is framed). For instance, the construct of transaction utility suggests that money could be sacrificed in some circumstances but not others (e.g., paying more for a drink in a fancy resort vs. a rundown store; Thaler 1985). Similarly, people’s use of money changes depending on whether it is framed as belonging to a specific mental account (e.g., leisure activities) or not (Thaler 1985, 1990, 1999).

Mental accounting is still very influential in the marketing literature as shown by more recent work on

budgeting decisions (see [Sussman et al. 2023](#) for a review). At times, consumers may be strict with their mental budgets: consumers tend to have a fixed budget to spend on fuel, thus selecting lower grades of gasoline when prices increase and higher grades when prices decrease ([Hastings and Shapiro 2013](#)). At other times, consumers may have more malleable mental accounts: they tend to spend more on purchases labeled as exceptional than on purchases labeled as regular expenses. This is because exceptional purchases are framed more narrowly and are less easily incorporated into future budgets ([Sussman and Alter 2012](#); [Sussman et al. 2015](#)). Similar malleability is observed with the use of promotional credits, where the discount offered by a promotional gift card is deducted twice (“double discount”; [Cheng and Cryder 2018](#)). Recent work has also focused on models to generate predictions of preference outcome timing (hedonic hypothesis; [Evers et al. 2022](#)) and to explain reference price adaptation ([Baucells and Hwang 2017](#)) as well as the likelihood of consumers purchasing add-ons ([Erat and Bhaskaran 2012](#)).

THE NEED FOR REPLICATION IN CONSUMER BEHAVIOR RESEARCH

Considering the pivotal role of mental accounting in consumer research, we deemed it essential to replicate its core findings for three main reasons. First, in the last two decades, psychological research has undergone a “replication crisis” that extended to related fields like marketing and consumer research ([Krishna 2021](#); [Munafò et al. 2017](#)). This crisis called into question the reliability of many (then-considered) classic findings in psychology and marketing (e.g., ego depletion, social priming), and showed that some questionable methodological and statistical practices (e.g., small sample sizes, “p-hacking” and the file-drawer effect) had become widespread ([Ioannidis 2005](#); [Loken and Gelman 2017](#); [Open Science Collaboration 2015](#); [Simmons, Nelson, and Simonsohn 2011](#); [Simonsohn, Nelson, and Simmons 2014](#)). Scholars in the field have thus called for more rigorous and transparent scientific practices ([Wood 2024](#)). As research practices continue to evolve and improve, it has become essential to ensure that foundational models and theories developed in the past, including mental accounting, are robust across time and under new, improved research standards. Researchers have already presented large-scale replications of related subjects, such as prospect theory ([Ruggeri et al. 2020](#)) and heuristics ([Klein et al. 2014](#)), thereby bolstering their reliability and expanding their impact within the behavioral sciences. To the best of our knowledge, such testing has not yet been performed on mental accounting.

Second, another critique of behavioral sciences is the reliance on samples from higher-income, Western

countries for drawing conclusions. This approach has been referred to as WEIRD (Western, Educated, Industrialized, Rich, and Democratic) ([Henrich, Ensminger et al. 2010](#); [Henrich, Heine, and Norenzayan 2010](#)), though the appropriateness of that dichotomization may itself be problematic ([Ghai 2021](#)). Still, if the goal is to robustly understand human behavior, broader and more inclusive approaches to sampling that expand the countries and populations involved are necessary. Otherwise, narrow samples may relate only to behaviors observed within those specific systems.

The original mental accounting studies themselves were conducted with relatively small U.S. samples and often limited to college students, an even less generalizable set of participants. We therefore invested heavily in testing whether the conclusions of these original studies generalize more broadly by involving individuals and populations outside those typically considered WEIRD. Large between-nations studies on intertemporal choice show that cultural dimensions and broader financial conditions systematically shape economic decision-making across countries ([Ruggeri et al. 2022](#); [Wang, Rieger, and Hens 2016](#)), suggesting that such variation is plausible also for mental accounting and that a broader sample will produce a richer understanding of human behavior.

Another important reason to extend the analysis of mental accounting to other countries is that it is related to consumption decisions that may vary depending on specific factors unique to different countries and populations. For instance, mental accounting effects could be lower in countries with lower average incomes than those in which it has traditionally been studied. It may also vary based on countries’ levels of income inequality, depending on the presence of a large welfare state, or when using different currency denominations. Larger income inequality may lead to more mental accounting because a larger number of households need to carefully consider the way they spend money compared to countries where income inequality is low. Relatedly, the welfare state may provide support to citizens leading to a reduced need to keep tight control on the way they spend or save money (e.g., because of public healthcare or pension systems). This would be consistent with findings that inequality and welfare-state orientation shape household financial fragility and planning behavior ([Comelli 2021](#); [Setterfield and Kim 2024](#)). Finally, larger currency denominations may induce people to accept discounts regardless of the base price as the high nominal values of the currency can give the impression that the reduction represents a substantial cut to the original price, a pattern that would be consistent with the denomination effect observed in consumer research ([Raghubir and Srivastava 2009](#)).

Third, much has changed in the way consumers manage their finances compared to the world for which the original scenarios were designed. For instance, when the model was developed in the 1980s, most transactions relied on cash or checks, and credit card use was still limited in many countries. Now, online banking offers an easy way to check one's expenses, whereas in the past, it could only be done by tallying receipts, balancing a checkbook, or receiving a bank statement. At the same time, other developments, such as digital payments and subscription plans, have made it harder for consumers to have a clear idea of their spending (Mack 2023). The rapid increase in efficient digital transactions has been so sudden in some areas and slower in others, presenting both a need and an opportunity to test if or how mental accounting relates in contemporary consumer systems.

KEY AIMS OF THE PRESENT STUDY

Assessing Replicability of Mental Accounting

Building on these motivations, we designed and pre-registered our study to rigorously test the replicability and generalizability of mental accounting over time and across diverse cultural and economic contexts, using a large and statistically reliable sample. Failing to replicate the original results would raise questions about a model at the core of the behavioral sciences and cast doubt on many of its practical applications. Alternatively, partial replication in certain countries could help identify boundary conditions where mental accounting is more likely to impact consumers' decisions, offering important insights for real-world applications. Critically, a successful replication of a majority of the scenarios in most countries would strengthen our confidence in its reliability and robustness.

Examining Heterogeneity by Individual Differences

Using a large, non-student sample allows us to test potential moderating effects on mental accounting. While previous work explored the role of some individual differences, such as cognitive style (Hossain 2018), it remains unclear whether mental accounting constitutes a robust and universal psychological model of economic decisions. To this end, our second objective was to investigate the possible moderation effect of household income and financial literacy (Bai 2023; Deuflhard, Georgarakos, and Inderst 2019; Lusardi and Mitchell 2017), which have only been tested in previous smaller-scale studies (Muehlbacher and Kirchler 2019). Although previous studies have consistently shown that understanding basic economic concepts makes a difference in people's decisions, for example, saving more for

retirement (Deuflhard *et al.* 2019; Lusardi and Mitchell 2017), the possible relationship between these two variables is not trivial and has seldom been investigated in the past (Silva *et al.* 2023). From one perspective, mental accounting is a bias with potentially negative consequences for spending and saving, since it violates fungibility. Conversely, effectively harnessing mental accounting can facilitate self-control over expenses. Thus, reasonable arguments could be made for a relationship in either direction. Relatedly, it is worth assessing possible effects of household income, as wealthier individuals have fewer financial constraints, potentially reducing the need for sensitivity to smaller values compared to those with a lower income.

Examining Heterogeneity by Country-Level Factors

Lastly, by collecting data in a diverse set of countries, including those underrepresented in behavioral research, we could also explore potential effects of socio-economic differences across countries. Considering how nations differ in their economic conditions (e.g., GDP per capita, currency denomination) and broader societal structures (e.g., levels of democracy and inequality), it is plausible that mental accounting may manifest differently across contexts. For example, in lower-GDP countries, consumers may have less discretionary income, making even small discounts economically significant, regardless of whether they apply to low- or high-priced items. As a result, people may focus more on the absolute value of the savings than on relative pricing differences. Additionally, mental accounting strategies developed in higher-income contexts may be less intuitive or ecologically valid in settings where financial decisions are more constrained or less flexible.

Failing to detect effects of individual or socio-economic country characteristics on how mental accounting manifests would reinforce the notion of mental accounting as a universal cognitive mechanism, which operates consistently across individual differences and cultural contexts. Conversely, identifying significant effects across countries would offer valuable insights into the contextual variability of mental accounting. These findings could inform practitioners and policymakers by enabling them to tailor interventions or financial strategies more effectively, leveraging mental accounting principles and considering clients' individual characteristics and geographical or cultural backgrounds.

MENTAL ACCOUNTING OPERATIONALIZATION

We chose to replicate a subset of mental accounting scenarios to capture the theory's core principles while

still reflecting the breadth of contexts examined in the literature. The selected scenarios vary along several important dimensions: whether decisions are made individually or in social settings, whether individuals decide for themselves or on behalf of others, and whether the task involves making a choice or determining a price. Specifically, we selected seven scenarios that represent the foundational elements of the original model and have significantly impacted thinking across scientific and applied domains. To ensure contemporary relevance, we also incorporated more recent scenarios that reflect situations consumers may face today (table 1).

Crucially, all selected scenarios have clear relevance to consumer behavior and are frequently cited. According to a Web of Science search in March 2026, three of the five original papers in which the scenarios first appeared have more than 2,000 citations each (Thaler 1985, 1999; Tversky and Kahneman 1981). Another has been cited more than 1,200 (Kahneman and Tversky 1982), and the most recent was cited more than 100 times (Shafir and Thaler 2006). Beyond their impact, these scenarios enable us to test for the heterogeneity and variety of mental accounting effects while remaining grounded in two of its most important tenets: transaction utility and decision frames. Below, we outline the theoretical underpinnings of each selected scenario and explain why they are relevant for understanding mental accounting in consumer decision-making.

Combined Outcomes—MrAB Scenarios

One of Thaler's initial objectives was to use "the concept of mental accounting to move further towards a behaviorally based theory of consumer choice" (Thaler 1985, 201). In doing so, he extended the value function to analyze compounded outcomes (e.g., two gains or two losses, a mixed gain ($x > 0 > y$; $x > |y|$) or a mixed loss ($x > 0 > y$; $x < |y|$); for additional details see the web appendix). The "MrAB" scenarios test a set of conditions in which participants are asked to report which of two characters is happier (or more upset). In each situation, the two characters obtain either a gain or a loss, but for one of them it is presented as a lump sum while for the other it is presented as two segregated outcomes (table 1). This approach allows for the assessment of when people prefer an integrated (lump sum) versus segregated outcome. Although the analysis of combined outcomes has been studied less frequently than other mental accounting effects, it is particularly relevant for consumer research due to its wide-ranging practical applications (e.g., pricing strategies like bundling product features or add-ons and product framing, Erat and Bhaskaran 2012; Prelec and Loewenstein 1998; Skwara 2023).

Transaction Utility—Game and Drink Scenarios

Transaction utility is defined as the price the consumer pays compared to the reference price (Thaler 1985). According to Thaler, a major factor in determining the reference price is fairness, which is defined as the cost to the seller. The "Game" scenario (Thaler 1985) shows that sellers set a price based on the anchor available (e.g., the price marked on the ticket, market price, or actual cost incurred) and their assessment of a fair price is influenced by the buyers' perceptions of the sellers' costs (table 1). This perception, in turn, impacts the *reference price*. In the "Drink" scenario, participants perceive that the fair price to pay is significantly different depending on where a drink is bought (e.g., a small, run-down grocery store or a fancy resort) although acquisition utility does not change.

Decision Frames—Jacket Scenario

Tversky and Kahneman (1981) (see also Kahneman and Tversky 1982) were the first to investigate how mental accounting can influence people's decisions by modifying contextual factors that have no impact on the economic value of a transaction. In the "Jacket" scenario (Kahneman and Tversky 1982; Thaler 1999; Tversky and Kahneman 1981), consumers are buying a jacket and a Bluetooth speaker (a calculator in the original scenario) for \$140. They are told that either the low price (\$15) or the high price item (\$125) is available at a \$5 discount in another store. If consumers were to simply consider whether it is worth driving 20 minutes to save \$5, they would reach the same decision regardless of the product to which the discount is applied. The same would be true if consumers were to compare their current wealth with the wealth resulting from having both products minus \$140. Instead, they use the cost of the product as a reference price and prefer to drive 20 minutes when the discount is applied to the low-priced item.

Sunk Costs—Play Scenarios

Mental accounting is also related to the impact of sunk costs on consumers' spending decisions. When consumers have already opened a mental account for a specific purchase, they are unwilling to use additional resources from it, even though the previous expense might be unrecoverable. This creates paradoxical effects, as illustrated by the "Play" scenario (Kahneman and Tversky 1982, 1984; Thaler 1999; Tversky and Kahneman 1981). This scenario demonstrates that consumers are less likely to repurchase a ticket for a play after losing it than after losing an equivalent amount of cash. Buying the ticket a second time, when the account is already negative, feels aversive. Attending the show is

TABLE 1

MENTAL ACCOUNTING EFFECT CATEGORIES, SCENARIOS, AND CONDITIONS INCLUDED IN THE STUDY

Combined outcomes: <i>MrAB</i> (Thaler 1985)—DV: A/B/No difference			
<i>MrAB1</i>	<i>MrAB2</i>	<i>MrAB3</i>	<i>MrAB4</i>
Person A was given two tickets to the lottery. They won \$50 with one ticket and \$25 with the other ticket. Person B was given a ticket to one lottery. They won \$75. Who is happier?	Person A received a fine and owed \$100. The same day, they received a second fine and owed \$50. Person B received a fine and owed \$150. There were no other repercussions from this mistake. Who is more upset?	Person A bought a lottery ticket and won \$100. Also, they damaged one door in their apartment in an accident and had to pay the landlord \$80. Person B bought a lottery ticket and won \$20. Who is happier?	Person A damaged their car in a parking lot. They had to spend \$200 to repair the damage. The same day they damaged the car, they won \$25 in a raffle. Person B damaged their car in a parking lot. They had to spend \$175 to repair the damage. Who is more upset?
Transaction utility: <i>Game</i> (Thaler 1985)—DV: Willingness to accept			
<i>Game1</i>	<i>Game2</i>	<i>Game3</i>	
You are going to a sold-out game of your favorite local sports team, and you have an extra ticket. The price marked on the ticket is \$5, but this ticket has been offered to you by a friend for free. You want to get rid of the extra ticket, and you know the ticket's face value price is \$5. You find someone who wants the ticket. There is no law against charging more than the price marked on the ticket. What price do you ask for selling the extra ticket if the person you sell it to is a friend/is a stranger? What price would you have asked for if you found out that the ticket's face value price was \$10 instead? If they are a friend/stranger?	You are going to a sold-out game of your favorite local sports team, and you have an extra ticket. The price marked on the ticket is \$5, which is what you paid for each ticket. You want to get rid of the extra ticket, and you know the ticket's face value price is \$5. You find someone who wants the ticket. There is no law against charging more than the price marked on the ticket. What price do you ask for selling the extra ticket if the person you sell it to is a friend/stranger? What price would you have asked for if you found out that the ticket's face value price was \$10 instead? If they are a friend/stranger?	You are going to a sold-out game of your favorite local sports team, and you have an extra ticket. The ticket's face value price is \$5, but you paid \$10 for each ticket when you bought them from another person. You want to get rid of the extra ticket and you know the ticket's face value price is \$5. You find someone who wants the ticket. There is no law against charging more than the price marked on the ticket. What price do you ask for selling the extra ticket if the person you sell it to is a friend/stranger? What price would you have asked for if you found out that the ticket's face value price was \$10 instead? If they are a friend/stranger?	
Transaction utility: <i>Drink</i> (Thaler 1985)—DV: Willingness to pay			
<i>Drink1</i>	<i>Drink2</i>		
You are on the beach on a hot day, and you would really like a cold drink. A friend offers to go buy the drink from a fancy resort hotel. They ask how much you are willing to pay for the drink: they will only buy it if it costs as much or less than the price you state. You trust your friend, and there is no possibility of bargaining with the bartender. What price do you tell them?	You are on the beach on a hot day and you would really like a cold drink. A friend offers to go buy the drink from a small, run-down grocery store. They ask how much you are willing to pay for the drink: they will only buy it if it costs as much or less than the price you state. You trust your friend, and there is no possibility of bargaining with the store owner. What price do you tell them?		
Mental accounting as outcomes framing: <i>Jacket</i> (Kahneman and Tversky 1982; Tversky and Kahneman 1981)—DV: Yes/No			
<i>Jacket1</i>	<i>Jacket2</i>		
You are about to purchase a jacket for \$125 and a Bluetooth speaker for \$15. The salesman informs you that the speaker you wish to buy is on sale for \$10 at the other branch store, located 20 minutes away. Would you make the trip to the other store?	You are about to purchase a jacket for \$15 and a Bluetooth speaker for \$125. The salesman informs you that the speaker you wish to buy is on sale for \$120 at the other branch store, located 20 minutes away. Would you make the trip to the other store?		

(CONTINUED)

TABLE 1
CONTINUED

Sunk cost and mental accounts with a negative balance: <i>Play</i> (Tversky and Kahneman 1981)—DV: Yes/No	
<i>Play1</i>	<i>Play2</i>
Imagine that you are on your way to a music event with a ticket for which you have paid \$40. On entering the venue, you discover that you have lost the ticket. Would you pay \$40 for another ticket?	Imagine that you are on your way to a music event, and intend to buy a ticket worth \$40. On entering the venue, you realize that you have lost \$40 in cash. Would you still pay \$40 for a ticket?
Local versus global representations: <i>Gym</i> (Shafir and Thaler 2006)	
DVs: "I feel like I wasted \$20," "I feel like I wasted nothing, since my visit had already been paid for," "I feel like I wasted something, but no specific amount or measure comes to mind" (Scale from 1 = "Strongly agree" to 5 = "Strongly disagree")	
<i>Gym1</i>	<i>Gym2</i>
Imagine that you go to the gym on a regular basis. Due to your busy working schedule, you can only go on Monday nights. The membership to the gym allows you to pay \$20 (non-refundable) per session. One Monday, just after you have paid and changed into your gym clothes, you receive a phone call that requires you to leave and forgo your exercise that evening. How well does each of the following statements capture your feelings about the cost of the missed workout?	Imagine that you go to the gym on a regular basis. Due to your busy working schedule, you can only go on Monday nights. The membership costs you \$1000 a year (i.e., roughly \$20 per visit). One Monday, just after you have paid and changed into your gym clothes, you receive a phone call that requires you to leave and forgo your exercise that evening. How well does each of the following statements capture your feelings about the cost of the missed workout?
Historical accounting of purchases and bonuses: <i>Plane</i> (Shafir and Thaler 2006)	
DV: What do you think is the most appropriate thing to do? "Pay your friend \$35 for the coupon"/"Consider it a gift and do not pay for the coupon"/"Pay some, but not the full amount for the coupon, for example, half the price."	
<i>Plane1</i>	<i>Plane2</i>
You are flying with a friend who has upgraded their ticket to business class and has two coupons, each worth \$35, which can be used to upgrade your ticket from economy to business class. They purchased one coupon for the standard price of \$35 and received the other one for free. Just one coupon is required to upgrade your class on the current flight, and your friend offers you the coupon that they purchased so that you can upgrade as well.	You are flying with a friend who has upgraded their ticket to business class and has two coupons, each worth \$35, which could be used to upgrade your ticket from economy to business class. They purchased one coupon for the standard price of \$35 and received the other one for free. Just one coupon is required to upgrade your class on the current flight and your friend offers you the coupon that they got for free so that you can upgrade your ticket as well.

seen by consumers as a transaction in which the ticket price is exchanged for the experience of seeing the play. Thus, buying the ticket twice increases the cost without enhancing the experience.

Local versus Global Representations—Gym Scenario

Decision frames demonstrate that the perceived value of goods is subjective and malleable and can be influenced by various factors. Shafir and Thaler (2006) analyzed how people perceive the value of items depending on their structure and suggested that some construals appear more natural in certain contexts than others. One such case is illustrated by the "Gym scenario." Imagine a consumer who goes to the gym but then has to leave without working out. Mental accounting theory

predicts that if someone pays per session, they will experience the missed workout as a loss, as they paid for a session they did not use. In contrast, someone who purchased a yearly membership may be upset over the missed workout, but they should not attach a specific monetary value to each individual session, thus making the financial loss less painful.

Purchase and Bonuses—Plane Scenario

In the "Plane" scenario, Shafir and Thaler (2006) showed that when people exchange a coupon, they consider how it was acquired in the first place. The authors predicted that the recipient of an offered coupon for a flight upgrade may insist on paying for the coupon if the owner had originally paid for it but not if the owner obtained it for free. Shafir and Thaler (2006) predict that

people are unlikely to charge a friend more than what they originally spent and will thus be more likely to give the coupon away as a gift without asking for any money if they obtained it for free.

Replicating these scenarios is important not only for their practical relevance, but also because they cover a set of core principles in consumer research such as the diminishing marginal sensitivity of gains and losses (Kahneman and Tversky 1979), transaction utility (Thaler 1985), or decision frames (Kahneman and Tversky 1982; Thaler 1999; Tversky and Kahneman 1981). As evidence of their centrality, some of these concepts have been covered extensively in the marketing and consumer research literature on mental accounting (e.g., transaction utility: 45 papers, 4,338 total citations; decision frames: 91 papers, 2,181 citations; source: Web of Science).

Importantly, the scenarios were chosen to represent a heterogeneous set of conditions including different methods to measure consumers' preferences. In some cases, participants were asked to set a price (willingness to pay or to accept); in others, they evaluated fixed prices (e.g., would you buy a product for \$X?) or made a choice (e.g., whether to drive to a second store or not). We deemed it important to show that mental accounting can be replicated in many different countries across a set of different preference elicitation measures. This is especially relevant considering the work on procedural invariance (Tversky, Slovic, and Kahneman 1990). It is possible that some procedures (e.g., willingness to pay or choice between fixed prices) have a different impact on decisions in cultures other than those in which mental accounting has been originally studied. Similarly, the scenarios included interactions with other people (e.g., a friend offering to buy a beer, exchanging a coupon with a friend or selling a ticket to a friend vs. a stranger) as well as individual decisions (e.g., buying the jacket and Bluetooth speaker for oneself). Finally, the scenarios covered both situations in which participants were asked to assess outcomes and events happening to fictional characters (e.g., is Mr. A or Mr. B happier?) versus first-person decisions (e.g., setting the price for a drink or deciding whether to drive to a second store for a discount). We return to these distinctions later when examining heterogeneity across scenarios.

MAJOR FINDINGS AND CONTRIBUTIONS

Our results contribute to the literature on consumer research and marketing in several ways. First, we demonstrate that mental accounting remains robust and replicates across a diverse set of people globally. Across 5,589 participants from 21 countries, we found strong evidence for the robustness of mental accounting, with

all seven selected scenarios replicating in the same direction as the originals and 90.5% of country-level estimates remaining statistically significant (notably higher than the 70% average replication rate observed across social science replication efforts; Altmejd *et al.* 2019). Furthermore, although effect sizes were attenuated, mental accounting effects persisted across diverse populations.

Second, we demonstrate that individual differences in financial literacy and demographic characteristics do not account for the variability in how mental accounting manifests across countries, whereas a country's economic conditions do. Specifically, our analysis identifies a link between GDP per capita and mental accounting. We find a stronger association of mental accounting effects with GDP per capita up to about \$50,000, after which the correlation levels off. We demonstrated that this relationship holds even when controlling for other socio-economic characteristics of countries (i.e., currency denomination, degree of democracy, and inequality), but treat the findings as exploratory and recommend more nuanced analyses of the role of GDP and other potential covariates.

Finally, through an exploratory multidimensional scaling analysis, we identified an informative three-dimensional structure in how mental accounting effects vary across countries. This suggests that the variability of mental accounting between populations depends on scenario features such as whether decisions involve social context, require a particular decision perspective, or demand different levels of engagement in the role of price determination. While mental accounting is indeed a collection of different effects that encompass how consumers make financial decisions, our results add a new layer to the original model by showing that these commonly used scenarios can be structured into a parsimonious set of categories.

Our results offer reassurance for the mental-accounting-based policies already implemented while providing a reliable base to create new ones, such as initiatives to support consumers in managing debt or recognizing fraudulent financial practices. Similarly, home banking apps could leverage mental accounting to help customers better understand their savings and spending patterns, and make more informed financial decisions. We present a summary of our main findings and contributions in [table 2](#).

MATERIALS AND METHODS

The study received approval from the Ethical Committee for Psychological Research at the University of Padova (Protocol number: 4692) and from the University Human Research Ethics Committee at Concordia University (Protocol number: 30016952). In

TABLE 2
OVERVIEW OF RESEARCH QUESTIONS, MAIN RESULTS, AND CONTRIBUTIONS

Research question	Main results	Main contributions
RQ1: Is mental accounting robust and reliable across countries and over time?	Unpooled analysis: 90.5% overall replication Pooled analysis: 100% overall replication	Mental accounting is a robust and reliable phenomenon across a diverse sample from 21 countries and over time
RQ2: Are there moderating factors of mental accounting across countries?	Socio-demographic and Financial literacy: no significant effects. GDP per capita: significant, positive association up to US\$ 50,000 threshold. Currency denomination, Democracy Index, Gini coefficient: no significant effects.	GDP per capita is positively and non-linearly associated with stronger mental accounting effects. Individual differences do not explain variation in mental accounting across countries.
Exploratory Analysis: How are the mental accounting scenarios related?	Multidimensional scaling analysis (MDS) presents a three-dimensional structure: Dimension 1: social context involvement (individual vs. interactive decisions) Dimension 2: decision perspective (self vs. other) Dimension 3: role in price determination (setting vs. evaluating prices)	Mental accounting effects are not uniform across scenario types; between-country effect size patterns cluster along three interpretable dimensions

all countries, participants included in the analyses provided their informed consent at the beginning and the end of the survey, as described in the “Procedure” section. We did not use any forms of deception or hidden purposes, and all information about the study was fully disclosed. The study was pre-registered on the Open Science Framework (OSF) repository (<https://osf.io/u3j9b/overview>). All materials, datasets, and code, as well as any deviations from the pre-registration, are reported below and on our OSF project page (<https://osf.io/kgesy>).

Sample

The lead researchers and collaborators recruited participants from 32 countries, covering 23 languages. We included all countries where we could find appropriate collaborators and employed no other method to systematically include/exclude countries. Each country team was composed of at least two collaborators with direct connections to the country, either as natives or due to work relocation. To ensure consistency, the lead researchers shared guidelines on translation and adaptation of stimuli, as well as data collection. We excluded three countries (Colombia, Hungary, and South Korea) that dropped out during data collection, and eight countries (Ghana, Iran, Israel, Russia, Singapore, South Africa, Turkey, and the UK) that failed to reach the minimum pre-registered sample size (250 participants per country based on a power analysis; see the “Power and Error” section). In the main text, we thus report results from 21 countries (i.e., Austria, Brazil, Canada, China, Denmark, Egypt, France, Germany, India, Indonesia, Italy, Lithuania, Morocco, Netherlands, Portugal,

Romania, Spain, Sweden, Switzerland, the USA, and Vietnam). Participants received no monetary compensation to partake in the study, and were all adults, with the legal age defined within each country. Exclusion criteria at participant-level included a completion rate below 100% on the survey and reporting a value below 4 on a self-reported data quality question (i.e., “How serious have you been about filling in the survey?”, from 1 “Not serious at all” to 5 “Absolutely serious”). We also calculated one-third of the median Length of Interview (LOI) and conducted analyses both including and excluding participants whose completion times fell outside this range. Results in the main text refer to samples not considering the LOI exclusion criteria. A total of 5,589 participants completed the survey and met the pre-registered inclusion criteria, with 57.1% female and a median age of 30 (IQR: 24–40, see [table 1](#) in the [web appendix](#)). Results related to the full dataset without exclusion criteria are reported in the [web appendix](#) (section 3).

Instrument

To test the replicability of the mental accounting effects, we selected seven scenarios ([table 1](#)) with multiple conditions ([Kahneman and Tversky 1982](#); [Shafir and Thaler 2006](#); [Thaler 1985](#); [Tversky and Kahneman 1981](#)). The survey employed a within-subjects design, where each participant saw all seven scenarios and all the conditions for each scenario. We chose a within-subjects design to increase statistical power given a lower sample size. Furthermore, previous replication efforts used a similar methodology to test multiple effects in one study ([Ruggeri et al. 2020](#)). We used a

nested randomized presentation of the scenarios to minimize learning effects and reduce the likelihood of presenting two related scenarios close to each other (see the “Procedure” section). We allowed two modifications to ensure robust and current relevance. In the “Drink” scenario (Thaler 1985), “cold beer” was replaced with “a drink” to include participants from cultures where alcohol is prohibited or discouraged. In the “Jacket” scenario (Kahneman and Tversky 1982; Tversky and Kahneman 1981), the “calculator” was replaced with a “Bluetooth speaker” to reflect a more current and common product. After consulting with the lead researchers, local collaborators could make reasonable adaptations to enhance cultural appropriateness and reliability. Instances of these adaptations are: the substitution of “lottery ticket” in the first and third “MrAB” scenarios (Thaler 1985) with the term “lucky draw” or “raffle” in countries where lotteries are frowned upon or banned, such as Singapore and India; the substitution of “plane” in the “Plane” scenario (Shafir and Thaler 2006) with “train” or “bus” when traveling by plane within a country was deemed unrealistic by the collaborators. All instances where a modification occurred are reported on our OSF project page (<https://osf.io/kgesy>).

Financial Values and Adaptation

To adapt the financial values to current values and respective countries’ currencies, we performed a calculation based on median incomes in two steps, following previously published replications (Ruggeri *et al.* 2020). First, we calculated how much (in percentage) every financial value corresponded to the U.S. household median income in the year of publication of the original selected scenarios (i.e., 1981, 1985, and 2006). We then multiplied the percentage obtained in step one by the current median income in each country. For example, \$50 represents 0.21% of the U.S. median income (\$23,620) in 1985. We then obtained our adapted financial value by multiplying by 0.21% the current U.S. median income (\$74,099), which resulted in \$156. Finally, the resulting values were rounded to ensure that they were easy to process while maintaining the proportions of the financial values within scenarios (in the example above, \$156 was rounded up to \$160). After performing these calculations, the collaborators in each country were consulted to verify the plausibility of the obtained values in their respective currencies. For countries that do not publish an official median income (e.g., Brazil, Egypt, and Iran), we used the average household income. For countries where the local currency’s purchasing value was substantially lower, we applied an upward adaptation to the entire group of values. Some countries adjusted the values in certain specific scenarios to ensure that prices were realistic. For example, the

prices in the Gym scenario were significantly reduced in Canada (yearly membership: CAD\$1,400 after calculation, CAD\$700 final adjusted value) and Singapore (yearly membership: 2,370 SGD after calculation, 1,200 SGD final adjusted value) as the value after calculation did not align with current market prices. On the contrary, in countries where the calculated value was aligned with current market prices, the values remained unvaried, for example in Turkey (yearly membership: 1,785 TRY after calculation, 1,800 TRY final adjusted value) or Austria (yearly membership: 830 € after calculation, 800 € final adjusted value). All instances where a modification occurred are reported on our OSF project page, and a table with all the adjusted and final values, as well as the yearly household income for each scenario in every country (<https://osf.io/kgesy>).

Translation

We translated the survey into all languages using forward-and-backward translation from English. One individual translated the entire survey from English to the target language. Afterwards, a second individual translated it back into English, and the two English versions were compared. In case any discrepancies in the translation could not be resolved through discussion, a third person was involved as a mediator. In countries where collaborators were not native speakers (or if there were not enough collaborators to complete the forward-and-backward translation), they recruited at least one native speaker to aid the process. We implemented this method for all the non-English-speaking countries. Specific issues that arose around appropriateness and cultural adaptation were resolved country by country; for example, in Egypt, the option “non-binary/third gender” was changed to “other.” We aimed to maintain the survey as gender-neutral as possible. However, in some languages, this was not possible because of grammatical rules and/or the social norms of the country.

Demographics and Individual Differences

At the end of the survey, participants reported demographic and socio-economic information, including their age, gender identity, years of schooling (from elementary school to doctoral degree or specialization), country of residence, native language, and last year’s household income. To ensure between-country comparability, we calculated income ranges based on each country’s median monthly household income (or mean, if median data were unavailable). Participants could choose from six income brackets relative to the median: less than 50%, 50–75%, 75–100%, 100–125%, 125–150%, and over 150%, with an option to “Prefer not to answer”. In the survey, we presented actual monetary values calculated on a per-country basis (for instance, participants in the

U.S. were presented with the following seven options: Less than \$35,000; \$35,000–\$52,000; \$52,000–\$70,000; \$70,000–\$87,000; \$87,000–\$105,000; More than \$105,000; Prefer not to answer). When needed, we adapted the income measure based on how this information is typically reported in each country (i.e., median vs. mean, annual vs. monthly, individual vs. household income). Three countries (i.e., Portugal, Russia, and Brazil) created additional response options to account for the severe income inequality among their citizens. We assessed financial literacy (i.e., knowledge of inflation, interest rates, and savings accounts), with four items adapted from previous studies (Lusardi and Mitchell 2014; Skagerlund *et al.* 2018; van Rooij, Lusardi, and Alessie 2012). We assigned the value 1 to each correct response and 0 to each incorrect response. We then summed the number of correct responses to compute an overall score, which ranged from 0 to 4. A higher score indicated a greater degree of financial literacy.

Pilot

We shortened all scenarios to reduce the overall length of the survey and to make the materials easier to process. We used a between-subjects design to pre-test the original versus shortened versions of each scenario on 100 participants located in the United States, using Prolific Academic. The results showed that the two versions did not differ statistically; see “Explanation of existing data” on OSF (<https://osf.io/u3j9b/overview>).

Power and Error

We conducted a simulation-based power analysis using the R software and the Stan probabilistic programming language for Bayesian statistical inference. For each scenario, we defined the target effect size as half of the effect size found in the original scenarios to minimize the risk of performing the analyses on inflated effect sizes. We aimed to obtain 0.8 power to detect the target effect size for each scenario with a standard alpha error probability of 0.05. We estimated the power for each sample from 50 to 300 participants in increments of 50 and found that 250 participants per country were needed to achieve 0.8 power in all scenarios.

Procedure

We collected data exclusively online from September to October 2022, coordinating the process through a single institutional Qualtrics account. We shared an advertising message inviting participants to complete a short questionnaire on “how people make decisions” via social media (e.g., country-based platforms and online forums), following the Demić-Večkalov method

described in previous large-scale studies (Ruggeri *et al.* 2020, 2022). We deliberately excluded direct references to mental accounting in the recruitment messages and the informed consent. In cases where the Demić-Večkalov method proved insufficiently effective, we supplemented recruitment with student samples and collaborator networks. This approach has reliably yielded large, heterogeneous samples of independent participants while minimizing recruitment bias in similar large-scale studies (Ruggeri *et al.* 2020, 2022). Participants could complete the survey in either their local language or English, with both versions adapted to local currency and monetary values (see the “Instrument” section). Before starting the survey, they saw an informed consent statement explaining that the study focused on decision-making. Those who did not consent were automatically redirected to the end of the survey, and we did not collect any data from them. Participants who provided consent proceeded to answer 17 questions about the seven scenarios (see the “Instrument” section).

We used block-randomization to present scenarios, ensuring that conditions within each scenario were divided into subgroups and randomized in sequence. In total, three blocks with four scenarios each and one with five scenarios were randomized. Scenarios within a block were presented in a fixed order. This approach minimized the consecutive presentation of conditions from the same scenario and was applied across all scenarios with multiple conditions (ranging from 2 to 4). We then assessed participants’ financial literacy for exploratory purposes. Before collecting demographic data, we included a self-reported data quality question. At the end of the survey, we debriefed participants by explaining the study’s purpose, the data quality question, and the instruments used. Participants also had the opportunity to withdraw their consent for data usage. We deleted the data of those who withdrew. A copy of the entire survey, an example of the block randomization (survey flow), and the full analysis plan are available on our OSF page (<https://osf.io/kgesy>).

Economic and Social Country Factors

After data collection, we retrieved the GDP per capita index and Gini coefficient, Democracy Index, and currency denomination for each country for exploratory purposes. We obtained GDP per capita and Gini coefficient data from the World Bank website (<https://data.worldbank.org/>) (World Bank n.d.), where GDP per capita represents a country’s total economic output divided by its midyear population (see section 4.3 in the web appendix) and Gini coefficient represents income inequality within a country (values ranging from 0 = perfect equality to 100 = perfect inequality).

The Democracy Index quantifies the state of democracy across countries, categorizing them into four regime types based on their scores. It is calculated on a 0–10 scale using multiple qualitative and quantitative indicators. The four regime types are full democracy (score > 8), flawed democracy ($8 > \text{score} > 6.01$), hybrid regime ($6 > \text{score} > 4.01$), and authoritarian regime (score < 4; [Economist Intelligence Unit 2024](#)).

To measure currency denomination, we used its exchange rate against US\$, indicating how many units of the local currency equaled US\$. We extracted exchange rates from October 1, 2022, aligning with the data collection period.

RESULTS

Our results offer three key contributions ([table 2](#)). First, we confirmed that mental accounting is broadly replicable across various countries, although the effect sizes were generally smaller than those in the original scenarios. Second, we demonstrated that several variables did not significantly account for variations in mental accounting. These included individual differences such as age, education, income (measured as ordinal brackets relative to each country's median, reflecting participants' relative position within their own country's income distribution), financial literacy, and country-level characteristics (i.e., currency denomination, democracy, and inequality levels). Conversely, exploratory analyses showed that a higher GDP per capita was associated with stronger mental accounting effects, up to approximately \$50,000.

Third, further exploratory analyses suggest that mental accounting manifests differently depending on the nature of the decision context. Specifically, scenarios seem to vary consistently along three characteristics: whether the decision involves others or is made individually (social context), whether individuals are making choices for themselves or evaluating outcomes for others (decision perspective), and whether they are actively setting prices or responding to predetermined ones (role in price determination). This heterogeneity in how mental accounting operates across financial decisions offers new insight into the driving psychological mechanisms.

Mental Accounting Widely Replicates

The primary pre-registered aim of this work was to test the replicability of seven foundational scenarios of mental accounting using a large and diverse sample. We consistently replicated all the included scenarios, with only slightly lower effect sizes compared to the original studies, demonstrating that mental accounting is a robust phenomenon that generalizes across time and cultures. We consider as replication any effect that

aligns with the direction reported in the original scenario and with a 95% statistical credibility (i.e., there is 95% statistical confidence that the effect exists and goes in the same direction as originally reported).

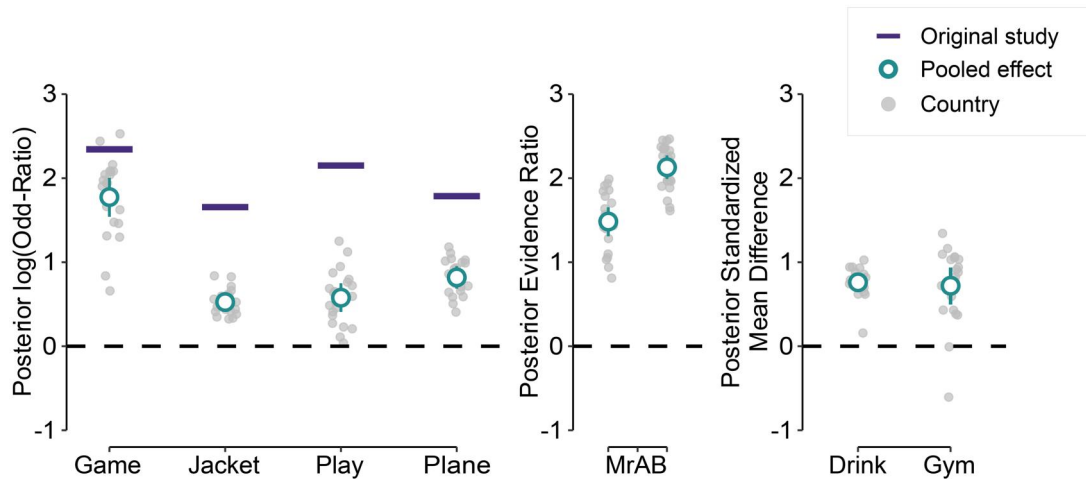
We found an overall replication of 90.5% (133 out of 147 effects), where MrAB and Game achieved a 100% (21/21) replication rate, Drink 95% (20/21), Jacket 81% (17/21), Play 71% (15/21), Gym 90% (19/21), and Plane 95% (20/21). To produce these findings, we calculated effect sizes for each scenario using a fixed-effect Bayesian regression approach ([figure 1](#) and [table 2](#) in the [web appendix](#)), setting weakly informative priors for all regression coefficients (normal distribution with mean = 0 and SD = 2.5). This prior choice reflects our assumption that very large effects are unlikely, while remaining sufficiently wide to allow the data to dominate the inference.

For the MrAB scenario, we analyzed participants' responses to scenarios involving gains (MrAB1 and MrAB3) and losses (MrAB2 and MrAB4) separately. According to mental accounting theory, when experiencing multiple monetary events, consumers should prefer experiencing gains separately (e.g., winning \$50 and \$25 in two different lotteries rather than winning \$75 in one lottery), but prefer experiencing losses together (e.g., receiving one fine of \$150 rather than two separate fines of \$100 and \$50). They should also prefer a smaller gain to a larger one that is reduced by a separate loss (e.g., receiving \$20 rather than receiving \$100 and losing \$80) but prefer a loss that is reduced by a separate gain than a smaller loss (e.g., spending \$200 but receiving \$25 rather than spending \$175). To quantify these effects, we coded each participant's pair of responses within each condition group as a single binary variable: 1 if the participant gave the combination predicted by mental accounting theory (e.g., for gains: choosing MrA in MrAB1 and MrB in MrAB3), and 0 otherwise. Since each sub-scenario has three possible responses (MrA, MrB, or no difference), there are nine possible combinations per condition group, and the probability of the predicted combination under chance responding is 1/9.¹ We then estimated the proportion of participants showing the predicted pattern for each country using a Bayesian binomial regression model and computed an evidence ratio comparing the observed proportion to this chance baseline (see [web appendix](#), section 4.2 for the full formulation). Gains and losses are

¹ The evidence ratio tests only whether the predicted combination exceeds its 1/9 chance rate, not whether it is the most frequent of the nine combinations; in principle, it could beat chance while another combination dominates. [Web appendix figure 11](#) (section 6) shows that in every MrAB scenario the predicted combination is in fact the modal choice, both pooled and within countries, confirming a genuine preference rather than a minority pattern that merely exceeds chance.

FIGURE 1

HIERARCHICAL BAYESIAN META-ANALYSIS.



NOTE.—Comparisons between the effects reported by the original scenarios (horizontal bars) and the current sample (open circles). We could not calculate the effect size of the original MrAB, Drink, and Gym scenarios as the information provided in the original papers was limited. Each small dot represents one country (random effect), while the open circles and their error bars indicate the mean pooled effect and its 95% CIs.

analyzed separately because they test distinct theoretical predictions: the segregation of gains versus the integration of losses.

The evidence ratio is a log odds ratio. A value of 0 indicates that the predicted pattern occurs exactly at the chance rate of 1/9, and positive values indicate it occurs more often, with exponentiation giving the factor by which the odds of the predicted pattern exceed chance (e.g., an evidence ratio of about 1.5 corresponds to odds roughly 4.5 times higher than chance). Expressing the effect this way places it on the same log-odds scale as the other binary scenarios, allowing direct comparison in figure 1 and table 3.

For the Drink and Gym scenarios, which involved continuous responses rather than categorical choices, we used Bayesian linear regression models to quantify the mental accounting effects. In the Drink scenario, mental accounting theory predicts that people will find it more acceptable to pay higher prices in upscale contexts compared to modest ones. To quantify it, we used a Bayesian linear regression model predicting the logarithm of prices from the store type (fancy resort hotel vs. run-down grocery store), while accounting for country-specific variations in both baseline prices and the effect of store type. Similarly, for the Gym scenario, the theory predicts that people perceive wasting money differently depending on how payments are framed. To test this, we predicted participants' ratings about wasting money from the payment frame (per-session vs. yearly

TABLE 3

HIERARCHICAL BAYESIAN META-ANALYSIS

	Estimate	95% CIs	$\log_{10}(\text{BF})$
MrAB1	1.48	1.31–1.65	13.24
MrAB2	2.13	1.99–2.27	31.67
Game	1.78	1.54–2.01	9.44
Drink	0.76	.66–.86	10.63
Jacket	0.53	.42–.63	7.96
Play	0.58	.41–.75	4.54
Gym	0.72	.50–.94	4.56
Plane	0.82	.69–.95	11.63

NOTE.—The estimate corresponds to the log odd-ratio for the scenarios Game, Jacket, Play, and Plane, all of which have a binary dependent variable. For the MrAB scenario, we estimated the Evidence Ratio, and for the Drink and Gym scenarios, we estimated the standardized mean difference (Cohen's d). This table reports the estimate, CI, and $\log_{10}(\text{BF})$ values corresponding to the green dots shown in figure 1.

membership), while accounting for country-specific variations in both baseline ratings and the effect of the payment frame.

For the remaining scenarios (Game, Jacket, Play, and Plane), which involved binary choices, we calculated effect sizes using Bayesian binomial regression models. In the Game scenario, mental accounting theory predicts that people set different prices for a ticket based on their relationship with the buyer (friend vs. stranger) and the ticket's acquisition context. To quantify this, we modeled the probability that participants would set a

price equal to or above market value as a function of buyer type while accounting for country-specific variations. In the Jacket scenario, the theory predicts that people's willingness to make an extra trip to save money depends on the relative rather than the absolute price difference. To test this, we modeled the probability of making the trip to another store as a function of the price magnitude (high vs. low), while accounting for country-specific variations. In the Play scenario, the theory predicts that people make different decisions about attending an event depending on whether they previously lost the ticket or an equivalent amount of cash. To examine this, we modeled the probability of paying \$40 as a function of the loss type (cash vs. ticket), while accounting for country-specific variations. Finally, for the Plane scenario, the theory predicts that people's willingness to pay for an upgrade coupon depends on how the coupon was acquired. To test this, we modeled the probability that participants would choose to pay their friend \$35 as a function of the coupon type (purchased vs. free), while accounting for country-specific variations.

To complement unpooled analyses, we performed a hierarchical Bayesian meta-analysis to pool data across countries while accounting for country-specific variations. This approach allows us to draw more generalizable conclusions about mental accounting effects while still preserving information about country-level differences. For each scenario, we fitted the Bayesian model to the observed effect size and standard error of each country and found a 100% (7/7 scenarios) replication rate (see figure 1 and table 3; see section 4 in the web appendix for the analysis plan). Again, we used weakly informative priors for all regression coefficients (normal distribution with mean = 0 and SD = 2.5). For all the scenarios where we could compare the effect sizes, we found a smaller effect than in the original scenarios. This finding aligns with other replication efforts (Open Science Collaboration 2015) and suggests that the use of smaller sample sizes and student samples in the original scenarios might have resulted in inflated estimates of effect size. In addition, the smaller effect sizes may reflect the fact that we included lower GDP-per-capita countries in the sample.

While our survey utilized a within-subjects design for all scenarios, some of the original scenarios (Jacket, Play, and Plane) employed a between-subjects design. To test whether this design choice affected the results, we leveraged the randomized order in which the four scenario blocks were presented. For these scenarios, the two conditions always appeared in different blocks, so we re-estimated each effect using only the condition each participant encountered first, before any exposure to the other condition. Because block order was randomized, this constitutes a between-subjects test with

random assignment. All three effects replicated (log odds ratio: Jacket = 0.65, 95% CI [.52, .78]; Play = 0.93, 95% CI [.74, 1.13]; Plane = 0.87, 95% CI [.71, 1.03]). The between-subjects estimates were not credibly different from the within-subjects estimates for Jacket (95% CI of the difference [−.29, .04]) and Plane ([−0.26, 0.16]), while for Play the between-subjects estimate was credibly larger ([−.61, −.10]), indicating that, if anything, the within-subjects design was conservative (see figure 8 in the web appendix).

Factors Related to Inter-Country Variability

Having demonstrated the robustness of mental accounting effects across countries (pre-registered analyses), we next turn to exploratory analyses investigating what factors might explain the variations in effect sizes between countries, as seen in figure 1. Results demonstrated that individual differences did not explain such variations in our sample, while economic country characteristics measured as GDP per capita did. Furthermore, the association between GDP per capita and mental accounting was significant also when controlling for countries' currency denomination, democracy, and inequality levels. However, this GDP association should be interpreted with caution. It would be an ecological fallacy to conclude that wealth directly drives mental accounting, as individual income does not consistently predict mental accounting within countries. We therefore do not claim a causal link between GDP per capita and mental accounting, but rather report a robust statistical association. The potential mechanisms underlying this association are considered further in the Discussion.

We first examined whether individual differences such as Age, Education, Income, and Financial Literacy could account for these variations. We converted the log odds ratios from the Game, Jacket, Play, and Plane scenarios into the same metric used for the Drink and Gym scenarios (standardized mean difference, also known as Cohen's *d*), to combine the mental accounting effects of different scenarios (Chinn 2000). MrAB was excluded from these analyses since its measure of effect size could not be mapped onto a Cohen's *d*. We then ran a hierarchical Bayesian analysis in which Financial Literacy, Education, Age, and Income were modeled as fixed-effects and random slopes, and the variable scenario was modeled as a random intercept. We set weakly informative priors for all regression coefficients (normal distribution with mean = 0 and SD = 2.5). All four predictors were z-scored (standardized to mean = 0 and SD = 1) to ensure comparability across variables measured on different scales. We did not find any credible effects for any of the four predictors (table 4).

TABLE 4

ASSOCIATIONS OF FINANCIAL LITERACY AND DEMOGRAPHIC VARIABLES WITH MENTAL ACCOUNTING

	Standardized Estimate	95% CIs	log ₁₀ (BF)
Financial Literacy	-0.03	-.10, .05	-1.31
Education	0.00	-.09, .09	-1.42
Age	-0.07	-.15, .02	-0.70
Income	0.01	-.05, .08	-1.46

NOTE.—Standardized estimates are based on z-scored country-level means for Financial Literacy, Education, Age, and Income.

To exclude the possibility that the lack of significant effects was due to the inclusion of all demographic variables in a single analysis, we additionally modeled each predictor separately. We fitted individual models for Financial Literacy, Education, Age, and Income, each including both fixed effects and random slopes by scenario. These separate analyses confirmed our initial findings, showing no credible effects for any of the individual difference variables when examined independently. These results suggest that mental accounting does not depend on the individual difference factors that we considered, even when potential correlations between predictors are eliminated from the analysis (table 4 in the web appendix). Finally, to rule out the possibility that the relationship between individual differences and mental accounting might be non-linear rather than absent, we fitted a non-linear (logistic growth) model to each demographic variable. In no case did the model identify a credible non-linear pattern (web appendix table 7).

To further validate this conclusion, we conducted more granular analyses where we considered each participant, country, and scenario separately. Also in this case, we fitted individual models for Financial Literacy, Education, Age, and Income, each including both fixed effects and random slopes by country, as well as a model with all the predictors. Both approaches confirmed the absence of a robust association between the four demographic variables and mental accounting (figure 2 in the web appendix).

Next, we explored whether the broader economic context of a country might influence the extent of mental accounting by testing the relationship between GDP per capita, used here as a proxy for national economic prosperity, and mental accounting. We found a positive association between GDP per capita and mental accounting up to approximately \$50,000 (slope = 2.5, 95% CI [.69, 4.48]) and a plateau effect for higher values (figure 2B). To obtain this result, we tested for the presence of both a linear and a non-linear relationship, fitting a Bayesian hierarchical linear regression to test the first hypothesis and a Bayesian hierarchical logistic growth model to test the second. We then compared the predictive

performance of the two models using leave-one-out cross-validation, and found that the non-linear model outperformed the linear model. The difference in expected log pointwise predictive density (ELPD) was -12.7 (SE = 4.5), indicating a substantially higher predictive accuracy for the non-linear model. These findings suggest that, in our sample, stronger effects for mental accounting were observed in countries with higher GDP per capita, but this effect levels off at higher economic thresholds.

Because GDP per capita is related to other country-level differences, the probability of introducing omitted variable bias is a concern. Therefore, in an effort to isolate the GDP per capita association from related country-level metrics like political climate and economic inequality, we tested whether other related constructs—currency denomination, the Democracy Index, and the Gini coefficient—could also explain the effect. We first applied the same linear and logistic growth model comparison to each control variable individually. In no case did the non-linear model provide a meaningful improvement over the linear model (all ELPD differences within one standard error; web appendix table 7).

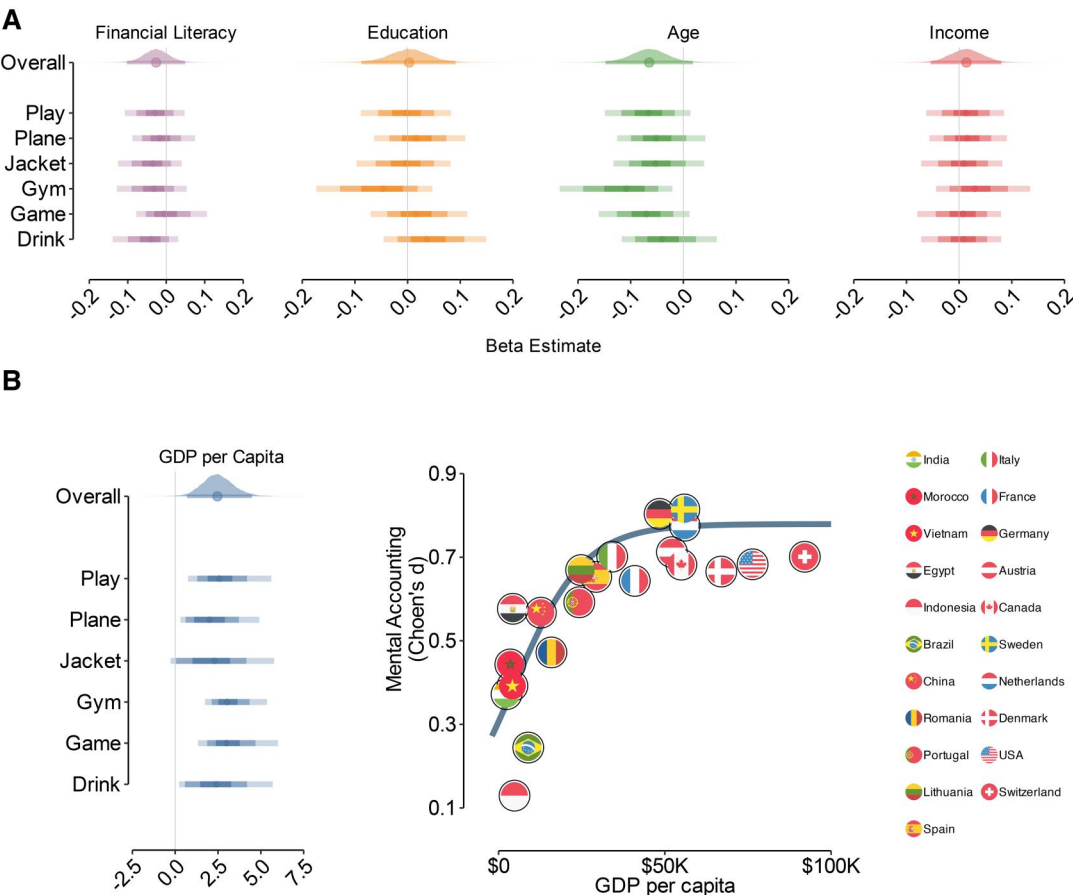
Having established that these control variables do not exhibit non-linear patterns, we included them as linear covariates alongside a non-linear GDP term. We reasoned that controlling for these variables that might have an impact on mental accounting effects, could capture other country-level differences that are correlated with GDP per capita. By doing so we could rule out at least some of the potential alternative explanations for the relationship between GDP per capita and mental accounting across countries. We found no evidence of an association of currency denomination ($\beta = -0.07$, 95% CI [-0.20, .05]), Democracy Index ($\beta = -0.06$, 95% CI [-0.15, .04]), or Gini coefficient ($\beta = -0.07$, 95% CI [-0.18, .04]) with mental accounting, but we still found an association between GDP per capita (growth rate = 1.90, 95% CI [.11, 4.50]) and mental accounting.

It is important to note that the absence of an individual income effect reported above does not contradict the GDP finding. Income in our study was measured as ordinal brackets relative to each country's median, reflecting participants' relative position within their own country's income distribution, whereas GDP per capita is an absolute measure of a country's economic output. The two results are therefore not contradictory, provided that the GDP association is not interpreted as evidence that wealth directly causes stronger mental accounting.

Heterogeneity in Scenarios with Multidimensional Scaling (MDS) Analysis

The preceding analyses examined what predicts the strength of mental accounting effects. We found that

FIGURE 2
RELATIONSHIP BETWEEN INDIVIDUAL DIFFERENCES VARIABLES, GDP PER CAPITA, AND MENTAL ACCOUNTING



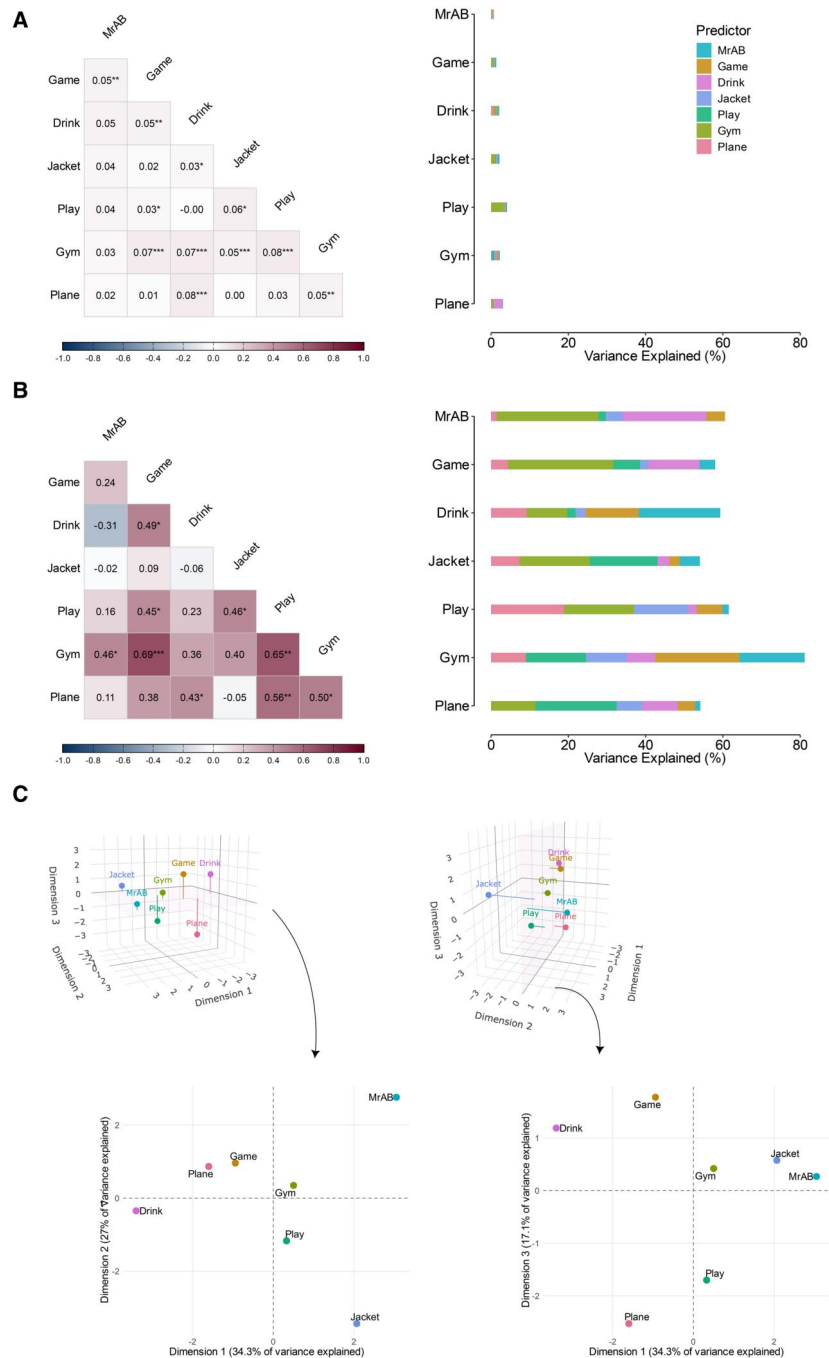
NOTE.—(A) Study-level posterior estimates of Financial Literacy, Education, Age, and Income as predictors of mental accounting. For each predictor, horizontal bars represent study-specific estimates (fixed + random effects) at three credible interval levels (thin: 95%; medium: 80%; thick: 50%). The density plots at the top of each column show the overall fixed effect. The vertical gray line marks zero. (B) GDP per Capita analysis. Left: study-level estimates of the logistic growth rate parameter, displayed in the same format as Panel A. Right: scatter plot of GDP per capita against mental accounting (Cohen's *d*), with each country represented by its national flag and the fitted logistic curve overlaid.

individual demographics do not explain who shows stronger effects, while there is an association with country-level wealth (i.e., GDP). Throughout these analyses, we adopted a theory-guided approach and treated the different scenarios as interchangeable instances of the same mental accounting construct, since they were all developed to tap into it. Next, we sought to leverage the unique structure of our dataset, with multiple scenarios administered to the same participants across 21 countries, to directly test this assumption. Specifically, we asked whether the various scenarios actually tap into the same underlying process. If they do, individuals who show strong effects on one scenario should also show strong effects on others.

To test this, we computed individual-level correlations across all scenario pairs, treating individual participants as the unit of analysis within each country, using appropriate methods for binary and continuous outcomes, and meta-analyzed the results across countries. Because these analyses ask how similar each scenario is to the others, and MrAB is the only scenario measured twice (gains and losses), we represented it with a single index by averaging the two. Their high mutual similarity reflects the shared paradigm rather than a relationship between distinct scenarios. Individual-level correlations were uniformly weak (mean $|r| = 0.04$, ranging from -0.005 to 0.078 ; figure 3A, left), indicating that an individual's mental accounting response on one scenario

FIGURE 3

THREE-DIMENSIONAL STRUCTURE OF MENTAL ACCOUNTING SCENARIOS



NOTE.—(A) Individual-level analysis. Left: meta-analyzed individual-level correlation matrix across the seven mental accounting scenarios (combining MrAB gain and loss). Significance: * $p < .05$, ** $p < .01$, *** $p < .001$. Right: partial R^2 decomposition showing each predictor scenario's contribution to the total variance explained when predicting each target from all others (mixed-effects models with country random intercepts; drop-one ΔR^2 approach, normalized). (B) Country-level analysis. Left: Pearson correlation matrix of country-level effect sizes. Right: partial R^2 decomposition (LMG method). (C) Multidimensional scaling (MDS) based on Euclidean distances between standardized country-level effect sizes. The three-dimensional solution explained 78.3% of the total variance. Left: Dimension 1 (social context; 34.3%) $\times$ Dimension 2 (decision perspective; 26.9%). Right: Dimension 1 $\times$ Dimension 3 (role in price determination; 17.1%). Insets show the full three-dimensional configuration. Numerical coordinates are reported in table 8 of the web appendix.

provides almost no information about their response on another. To validate this result, we used mixed-effects models to predict each scenario's mental accounting score from all other scenarios, with random intercepts for the country. This approach more directly addresses the question of how well one can predict an individual's mental accounting response on a target scenario given their responses on all other scenarios, naturally accounts for the hierarchical structure of the data, and accommodates the different nature of outcomes through appropriate link functions. Although all full models were significantly better than an intercept-only baseline with country random effects (all $ps < .001$), the variance explained remained low: other scenarios collectively explain only 4.3% of variance (figure 3A, right). These findings suggest that the various scenarios, despite all being designed to capture mental accounting, do not tap into a single underlying process at the individual level.

While the scenarios do not appear to measure a common process at the individual level, we observed substantial variation in how they replicated across countries (ranging from 71% to 100%), and country-level GDP was associated with effect size variation. This suggests that scenarios may differ in their sensitivity to country-level factors. We therefore assessed whether a structure emerges at the aggregate level: if scenarios share sensitivity to country-level factors, countries showing strong effects on one scenario should also show strong effects on others.

Country-level correlations were indeed substantially stronger (mean $|r| = 0.34$, ranging from -0.31 to 0.69 ; figure 3B, left). Using the same predictive modeling approach at the country level, we predicted each scenario's effect size from all other scenarios across the 21 countries. On average, other scenarios collectively explained 45% of variance in a held-out scenario (mean adjusted R^2 across the seven scenarios, accounting for the number of predictors; figure 3B, right), confirming that a structure emerges at the aggregate level. We note that the individual-level and country-level estimates are not directly comparable: because country means average over many participants, aggregation removes the individual-level noise that attenuates the within-person correlations, so a stronger association at the country level is expected and reflects shared sensitivity to country-level factors (such as GDP or culture) rather than a common process within individuals (figure 3A vs. 3B).

Having established that scenarios covary at the country level, we next sought to characterize this structure. We conducted a multidimensional scaling (MDS) analysis, which projects scenarios into a lower-dimensional space based on the similarity of their between-country effect size patterns, revealing which scenarios cluster together and along which dimensions. We constructed a

matrix with 21 countries as rows and seven mental accounting scenarios as columns, using the mean effect size for each scenario within each country. For scenarios with binomial outcomes (e.g., *MrAB*, *Game*, *Jacket*, *Play*, *Plane*), effect sizes were converted to standardized mean differences using the logit approximation. We then standardized the effect sizes across countries and computed Euclidean distances between scenarios. This distance matrix served as input to the MDS, from which we extracted three principal coordinate dimensions that together explained 78.3% of the variance in inter-scenario relationships.

By visualizing the scenarios within this three-dimensional space (figure 3C), the MDS analysis clarifies how mental accounting effects vary along interpretable scenario features associated with different patterns of between-country variability. We propose the following interpretations for these three dimensions, while recognizing that alternative explanations may be possible. The first dimension (34.3% of variance) appears to reflect differences in social context involvement. That is, whether the decision is made in interaction with others (e.g., *Drink* scenario: "a friend offering to buy beer") or individually (e.g., *Jacket* scenario: "Would you make the trip to the other store?"). The second dimension (26.9% of variance) seems to capture the decision perspective: whether participants are evaluating outcomes for others (e.g., *MrAB* scenario: "Who is happier?") or making decisions for themselves (e.g., *Drink* scenario: "What price do you tell them?"; figure 3C). The third dimension (17.1% of variance) distinguishes the role in price determination, ranging from actively setting a price (e.g., *Game* scenario: "What price do you ask for the ticket?"), to evaluating or choosing among fixed options (e.g., *Gym* scenario: "The membership to the gym allows you to pay. . ."), to responding to a predetermined price (e.g., *Play* scenario: "Would you pay \$40 for another ticket?"; figure 3C, table 5).

DISCUSSION

Mental accounting is a foundational framework in the behavioral and economic sciences, which has influenced hundreds of studies and driven numerous real-world applications across businesses, nonprofits, and governmental agencies. Though widely accepted, it has now been decades since much of the work was initially produced. At present, there is widespread and growing recognition of the need to determine the extent to which such insights apply more globally across diverse populations (Arnett 2008; Henrich, Ensminger *et al.* 2010; Henrich, Heine *et al.* 2010; Ioannidis 2005; Krishna 2021; Loken and Gelman 2017; Munafò *et al.* 2017; Open Science Collaboration 2015; Simmons *et al.* 2011;

TABLE 5
MDS DIMENSIONS OF MENTAL ACCOUNTING SCENARIOS

Dimension	Explanation	Associated scenarios
1. Social context	Differentiates whether the decision is made in isolation or involves a social interaction/transaction with another person.	Individual: Jacket Social/Interactive: Drink, Plane, Game
2. Decision perspective	Distinguishes between making a choice for oneself versus evaluating for someone else (third-person perspective).	Self-focused: Drink, Jacket, Play, Gym, Plane Other-focused: MrAB
3. Role in price determination	Reflects the participant's agency in the financial transaction, ranging from setting a price to reacting to a fixed price.	Active (setting price): Game, Drink Evaluative (choosing/evaluating): Gym, Plane Passive (responding to fixed price): Play, Jacket

Simonsohn *et al.* 2014). Applying a rigorous standard for sampling and methods will support insights to this end.

Responding to such a need, our study makes three key contributions. First, we replicated seven core scenarios related to mental accounting (Kahneman and Tversky 1982; Shafir and Thaler 2006; Thaler 1985; Tversky and Kahneman 1981), while extending its findings to a broader set of countries. We demonstrated that mental accounting effects are robust and, crucially, that they were observed in all countries, including those rarely studied in marketing, psychology, and other behavioral sciences. Our replication involved over 5,500 participants from 21 countries, thereby providing strong evidence in support of this model, which has played a central role in shaping our understanding of how consumers manage money and make economic decisions. Adapting the original scenarios to enhance inclusivity and comprehensibility allowed our replication to show that mental accounting is pervasive and replicates over time and across cultures. Indeed, results revealed remarkable similarities across countries and an attenuation of the effect sizes. In planning our replication, we selected scenarios that represented the heterogeneity of effects associated with mental accounting (e.g., transaction utility, framing of outcomes, sunk costs) while maintaining consistency and focusing mainly on how different economic situations are presented to consumers.

Second, we contribute to the mental accounting literature by systematically assessing other factors that potentially influence this effect. Beyond the importance of replicating the original findings, the lack of a significant moderating effect of financial literacy on mental accounting suggests that this model applies broadly. Similarly, demographic factors such as age, education, and household income were not associated with mental accounting, indicating that its effects are consistent across diverse consumer groups. To account for potential correlations among these factors, we tested them both independently

and concurrently, yet their association with mental accounting remained negligible. We observed similar results in a Causal Forest analysis performed to further explore heterogeneous treatment effects (see section 5 in the web appendix). This again reinforces the conclusion that these demographic variables do not play a meaningful role in explaining our results.

We did observe some variability at the country level, with mental accounting effects relating positively and non-linearly to GDP per capita. Specifically, we found the association between GDP per capita and mental accounting only up to a level of \$50,000. While this is an exploratory finding and should be understood with appropriate caution, it is compelling. We considered, for example, that GDP per capita could be a proxy for other country-level differences, such as a country's democracy, inequality levels, or the specific currency denominations (Antonides and Ranyard 2017; Raghuram 2006; Raghuram and Srivastava 2009; Reutskaja *et al.* 2020). We included these factors in the analyses but GDP per capita remained the only significant predictor of mental accounting.

We caution against presuming that the GDP per capita finding is evidence that wealth alone amplifies mental accounting; we also found that individual income does not consistently predict mental accounting within countries. Therefore, GDP per capita should not be interpreted as a causal factor, but may be better understood as a proxy for some other country-level characteristic, essentially the broader economic environment rather than singular indicators of wealth.

Despite our controls, it is possible that other omitted variables covary with GDP. For instance, differences in institutional trust or market integration may drive mental accounting principles in general or influence how these specific scenarios were understood. Moreover, methodological factors, such as the scaling of monetary amounts and the translation of scenarios, may not have maintained full psychological equivalence across

different economic contexts. Finally, the stronger mental accounting effects in higher-income countries might reflect that these psychological tests were initially developed and validated in high-income per capita contexts, potentially making them more attuned to how people make financial decisions in more economically developed environments.

Third, we extend the mental accounting model by showing that the seven scenarios vary meaningfully in their between-country effect size patterns, clustering along three interpretable dimensions. While all scenarios measure effects related to mental accounting, they differ systematically in how financial decisions are framed. Specifically, the scenarios vary in terms of social context involvement (individual vs. interactive decisions), decision perspective (self vs. other), and role in price determination (setting vs. evaluating prices). This multidimensional structure suggests that mental accounting effects are not uniform across scenario types, and that the decision context shapes how these effects manifest across countries. Recognizing these dimensions helps explain why some mental accounting effects are more robust than others and may inform the design of more targeted interventions for different types of financial decision-making. Future research should examine whether other mental accounting scenarios align with this three-dimensional framework or whether additional dimensions are needed.

Notably, these dimensions connect to established research traditions in consumer psychology. For instance, the third dimension (the role of price determination) aligns with the well-documented violations of procedural invariance (e.g., differences in people's preferences across willingness-to-pay and choice tasks; [Tversky, Slovic, and Kahneman 1990](#)). Research on preference reversals has shown that pricing and choice elicitation methods engage different psychological processes and can yield systematically different preferences ([Lichtenstein and Slovic 1971](#)). Scenarios where participants must generate a price (e.g., *Drink, Game*) versus evaluate fixed options (e.g., *Jacket, Play*) may tap into these different valuation processes, and the variation in effect sizes along this dimension is consistent with this possibility. That mental accounting scenarios naturally vary along the WTP-choice continuum suggests that the response mode may moderate the strength of mental accounting effects, a hypothesis that future research could directly test by manipulating the elicitation format within a single scenario. The first and second dimensions similarly connect to active research areas: social context in consumer decision-making ([Argo, Dahl, and Manchanda 2005](#)) and self-other differences in judgment and choice ([Polman 2012a, 2012b](#)).

Limitations and Future Directions

Our study does not encompass all mental accounting effects, and we cannot assume that untested findings will replicate as consistently as those we examined. Future research should extend these efforts by replicating additional scenarios, particularly those exploring how mental accounting helps consumers maintain self-control over their spending ([Koch and Nafziger 2016](#)).

Another potential limitation of this work is the use of a within-subjects design, whereas mental accounting research has traditionally employed a between-subjects design, given that each condition presents a different version of the same scenario. As such, our results may be conservative estimates of the between-subjects effects. To address this, we carefully randomized the presentation of scenarios to ensure that participants did not see different versions of the same scenario consecutively. This approach follows a similar within-subjects replication of prospect theory effects, which were originally conducted as between-subjects experiments ([Ruggeri et al. 2020](#)). Additionally, analyzing only the first condition each participant encountered (a between-subjects test enabled by the randomized block order) replicated all mental accounting effects, with estimates equal to or larger than the within-subjects ones. Our MDS analysis was conducted at the country level, using mean effect sizes per scenario within each country. This approach identifies which scenarios show similar patterns of between-country variability, that is, scenarios that are similarly affected by country-level factors such as GDP, cultural norms, or translation. Importantly, this does not necessarily imply that the same dimensional structure operates at the individual level. Aggregation to country means removes within-country individual variation, which constitutes the majority of the total variance. The weak individual-level correlations across scenarios are consistent with this conclusion: scenarios that covary between countries do not necessarily covary across individuals within countries.

This is critical for real-world applications: while behavioral sciences largely promote individual choice autonomy, concepts such as mental accounting, loss aversion, and temporal discounting are typically better observed in the aggregate (i.e., full populations). Attempts to corral, leverage, or prevent harms of such choice patterns may thus be limited in value and effect when applied only at the level of an individual; aggregate effects of such policy may be valuable but also overstate individual benefits to consumers ([Consumer Financial Protection Bureau 2013](#)). Broader system-level thinking will be necessary to promote wider value to individuals ([Thomas, Kay, and Laurin 2023](#)).

Additionally, because the MDS dimensions were derived from the effect sizes themselves, incorporating them as moderators of the same data would be circular. The alignment between the extracted dimensions and interpretable scenario features (particularly the WTP-choice distinction) generates testable predictions that future research could examine through experimental manipulation of these features within scenarios.

We also acknowledge that our approach to minimizing bias through extensive randomization and broad data collection resulted in demographic samples with varying representativeness. Due to budget constraints on participant compensation, we employed convenience sampling to gather a large and diverse sample. To reduce self-selection bias, however, we implemented the Demić-Večkalov method, which has been used in previous large-scale studies (Ruggeri *et al.* 2020, 2022).

Our results offer reassurance for the mental-accounting-based policies already implemented while providing a reliable base to create new ones, for instance, to help citizens better deal with debt management or fraudulent financial practices. Product designers and policymakers could use this insight to design interventions that encourage better credit management, such as structuring repayment platforms to promote principal reduction through highlighting the benefits of small-but-meaningful payment increases or bundling smaller debts to reduce perceived burden. Similarly, policies could require home banking apps to factor in and leverage mental accounting to ensure customers understand how to manage their savings and enable avoidance of expenses that impede optimal savings.

Although we replicated mental accounting in all countries included in this work, we did find some differences, especially among countries with lower income levels. As a result, and in line with the nudging approach to policy-making (Thaler and Sunstein 2021), we caution that any real-world applications should be properly tested before large-scale interventions are put in place in a given context, regardless of their successful adoption elsewhere.

CONCLUSION

In this study, we provide robust evidence that mental accounting is a global and pervasive behavioral phenomenon, replicating the original findings and extending them to a larger, more diverse participant sample. The results revealed an attenuation of the effect sizes and remarkable between-country similarities. We also uncovered a three-dimensional structure of the mental accounting effects, offering new nuance and insight to the consumer behavior literature. Finally, we also explored the role of nation-level wealth, showing that mental accounting effects are similarly strong in countries with

higher-income levels, but weaker as income levels decrease in our broad sample of 21 countries.

DATA COLLECTION STATEMENT

Data collection was conducted between September and October 2022, involving collaborators from 21 countries: Austria, Brazil, Canada, China, Denmark, Egypt, France, Germany, India, Indonesia, Italy, Lithuania, Morocco, Netherlands, Portugal, Romania, Spain, Sweden, Switzerland, the United States, and Vietnam. All 102 authors collected data in countries where they had connections, either because they were originally from or currently working there. Data were collected exclusively online and coordinated through a single institutional Qualtrics account owned by the JDM Lab at the University of Padova (the main coordinator of the project). An advertising message was disseminated via social media platforms and online forums using the Demić-Večkalov method as outlined in Ruggeri *et al.* (2022). When this method was not sufficiently effective, student samples and collaborators' networks were utilized. The data analysis followed a pre-registered plan and was executed entirely by the fourth author. There are no exemptions to the data sharing and archiving requirements for this study. All data are de-identified and comply with the relevant ethical guidelines. Currently, collected data are stored in the JDM Lab's Qualtrics account, accessible only to current lab members. Materials, datasets, and code are also available in our Open Science Framework project folder at <https://osf.io/kgesy>.

continued from page 1

vacondio@unitn.it) is a Junior Assistant Professor at the Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Corso Bettini, 841, 38068 Rovereto (TN), Italy, and at the Motivational Digital Systems (MODIS) Unit, Fondazione Bruno Kessler (FBK), Via Santa Croce 77, 38122 Trento, Italy. Simone D'Ambrogio (simone.dambrogio@psy.ox.ac.uk) is a Postdoctoral Researcher at the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3EL, United Kingdom. Marta Caserotti (marta.caserotti@unipd.it) is an Assistant Professor at the Department of Developmental Psychology and Socialisation, University of Padova, Via Venezia, 8, 35131 Padova, Italy. Beatrice Conte (conte.beatrice.97@gmail.com) is an Independent Researcher. Prisca De Roni (prisca.deroni@hsancamillo.it) is a Postdoctoral Researcher at the Healthcare Innovation Technology Lab, IRCCS San Camillo Hospital, Via Alberoni 70, 30126 Venezia, Italy. Hilda Du Plooy (hilda.duplooy@univ.it) is a Doctoral Researcher at the Department of Human Sciences, University of Verona, Via S. Francesco, 22, 37129 Verona, Italy. Vivian D. Grillo (vivian.darlene.grillo@unitn.it) is a Doctoral Researcher at the Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Corso Bettini, 841, 38068 Rovereto (TN), Italy. Libera Y. Mastromatteo (libera.mastromatteo@unipd.it) is a Postdoctoral Researcher at the Department of Developmental Psychology and Socialisation, University of Padova, Via Venezia, 8, 35131 Padova, Italy. Elisa Tedaldi (elisa.tedaldi@uzh.ch) is a Postdoctoral Researcher at the Department of Communication and Media Research, University of Zurich, Andreasstrasse 15, 8050 Zürich,

Switzerland. Filippo Toscano (filippo.toscano@uzh.ch) is a Doctoral Researcher at the Department of Psychology (Chair of Social and Economic Psychology), University of Zurich, Binzmühlestrasse 14, 8050 Zurich, Switzerland. Jesús Aguilar-Armijo (jesusaguilararmijo@gmail.com) is an Independent Researcher. Parisa Ahmadi Ghomroudi (p.ahmadighomroudi@unitn.it) is a Postdoctoral Researcher at the Department of Psychology and Cognitive Sciences (DiPSCo), University of Trento, Corso Bettini, 84I, 38068 Rovereto (TN), Italy. Lucian Alexa (lucian.alex@ieconomics.ro) is a Deputy Managing Director at Intuitive Economics, Calea Vacaresti, 391, District 4, Bucharest, Romania. Mathias H. Andersen (mathias@drmr.dk) is a Research Collaborator at the Department of Radiology and Nuclear Medicine, Copenhagen University Hospital—Amager and Hvidovre, Danish Research Centre for Magnetic Resonance, Kettegaard Allé 36, 2650 Hvidovre, Denmark. Per A. Andersson (per.a.andersson@liu.se) is an Assistant Professor at the Department of Behavioral Sciences and Learning, Division of Psychology, Jedi-Lab, Linköping University, 581 83 Linköping, Sweden. Karine Aoun Barakat (kaoun@inseec.com) is an Associate Professor at the Department of Marketing, INSEEC Grande Ecole, 37 Quai de Grenelle, 75015 Paris, France. Carolina Barros (carolinaabarro.psi@gmail.com) is a Doctoral Researcher at the Department of Cognition in Context (CO2), Faculdade de Psicologia da Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal. Ruggero Basanisi (ruggero.basanisi@imtlucca.it) is a Research Collaborator at the MoMiLab Research Unit, IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100 Lucca, Italy. Tara Beilner (tara.beilner@psy.lmu.de) is a Doctoral Researcher at the Department of Psychology, Ludwig-Maximilians-Universität (LMU) München, Leopoldstraße 13, 80802 München, Germany. Sergiu Burlacu (sburlacu@irvapp.it) is a Research Fellow at the Research Institute for the Evaluation of Public Policies (IRVAPP), Fondazione Bruno Kessler (FBK), Via Santa Croce 77, 38122 Trento, Italy. Thai Cao (thai.caoquoc@hust.edu.vn) is a Researcher at the Department of Education, Hanoi University of Science and Technology and Karlsruhe Institute of Technology, No. 1 Dai Co Viet Street, Hanoi, Vietnam. Alessandra Carella (alessandra.carella@studenti.unipd.it) is a Doctoral Researcher at the Department of Developmental Psychology and Socialisation, University of Padova, Via Venezia, 8, 35131 Padova, Italy. Arianna Chiappi (arianna.chiappi@unibocconi.it) is a Senior Research Assistant at the Bocconi Experimental Laboratory for the Social Sciences, Bocconi University, Via Röntgen n. 1, 20136 Milan, Italy. Zafer Çiftçi (zaferc48@gmail.com) is an Independent Researcher. Claudia Civi (claudia.civi@citystgeorges.ac.uk) is a Senior Lecturer at the Department of Psychology and Neuroscience, City St George's, University of London, Northampton Square, London, EC1V 0HB, United Kingdom. Alana Daly (alanadaly@correo.ugr.es) is a Doctoral Researcher at the Department of Psychology, University of Granada, Campus Universitario de Cartuja s/n, 18011 Granada, Spain. Valdonė Darškuvienė (valdone.darskuviene@ism.lt) is a Professor at the Department of Economics, ISM University of Management and Economics, Gedimino ave. 7, LT-01103 Vilnius, Lithuania. Marta De Pedis (marta.depedis@gmail.com) is an Assistant Lecturer at the Department of Human Sciences and Education, Public University of Navarre (UPNA/NUP), Los Magnolios building, Arrosadia Campus, 31600 Pamplona-Iruña, Spain. Earle J. Du Plooy (earle.duplooy@uct.ac.za) is a Director of Faculty Operations at the Faculty of Health Sciences, University of Cape Town, Anzio Road, Observatory, Cape Town 7925, South Africa. Mohammed El-Mir (mohammed.elmir1@usmba.ac.ma) is a Full Professor at the Department of Psychology, Sidi Mohamed Ben Abdellah University, 3 Lots Leila, APT 2, Hay Azhar, Fes, Morocco. Christian T. Elbak (chel@mgmt.au.dk) is an Associate Professor at the Department of Management, Aarhus University, Universitetsbyen 61, 8000 Aarhus C, Denmark. Sondas Elkot (elkotsondos@gmail.com) is a Senior Research Executive at the Market Strategy and Understanding Unit (MSU), Ipsos Egypt for Consulting Services, Cairo 27003, Egypt. Valeria Fanghella (valeria.fanghella@edhec.edu) is an Assistant Professor at EDHEC Business School, France. Eman Farahat (emanyasser98@hotmail.com) is a Behavioral Scientist at the Behavioral Science Group (BSG), UAE

Government, Al Nahyan, Abu Dhabi, UAE. Amy G. Fehl (amy.fehl@neoma-bs.fr) is an Associate Professor at the Department of Marketing, NEOMA Business School, 59, rue Pierre Taittinger, 51100 Reims, France. Ama P. Fenny (apfenny@ug.edu.gh) is an Associate Professor at the Institute of Statistical Social and Economic Research, University of Ghana, 1 Botanical Gardens, Legon, Accra, Ghana. Paul A. G. Forbes (paul.forbes@hhu.de) is a Postdoctoral Researcher at Comparative Psychology, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany. Gemma Garbi (gemmaarbi@hotmail.it) is an Independent Researcher. André Gonçalves (andregoncalves1@campus.ul.pt) is a Doctoral Researcher at the Department of Cognition in Context (CO2), Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal. Sevilas Guvuriro (guvuriros@ufs.ac.za) is an Associate Professor at the Department of Economics & Finance, University of the Free State, 205 Nelson Mandela Drive, Bloemfontein 9300, South Africa. Ali Hajian (ahajian@umass.edu) is a Doctoral Researcher at the Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Hicks Way 135, Amherst 01003, USA. Steve Heinke (steve.heinke@unifr.ch) is a Senior Researcher at the Mobiliar Center for Resilience, University of Fribourg, Avenue de l'Europe 20, CH-1700 Fribourg, Switzerland. Austin W. Howard (austinwilliamhoward@gmail.com) is a Doctoral Researcher at the Department of Psychology, Durham University, Upper Mountjoy, South Rd, Durham, DH1 3LE, United Kingdom. Sudharsana Jagatheesh Jayanand (sudharsana.jagatheesh@gmail.com) is a Postgraduate Student at the Department of Government, London School of Economics, Centre Building, Houghton St, London, WC2A 2AZ, United Kingdom. Peiran Jiao (P.Jiao@maastrichtuniversity.nl) is an Associate Professor at the Department of Finance, Maastricht University, Tongersestraat 53, 6211 LM Maastricht, The Netherlands. Gabriela M. Jiga-Boy (g.jiga@swansea.ac.uk) is a Senior Lecturer at the School of Psychology, Swansea University, Singleton Park, Swansea, Wales, SA2 8PP, United Kingdom. Alejandra Jordano De Castro (ajdc22@bath.ac.uk) is a Trainee Clinical Psychologist (DCLinPsy) at the Department of Psychology, University of Bath, Claverton Down, BA2 7AY, United Kingdom. Tobias Kalenscher (Tobias.Kalenscher@hhu.de) is a Professor at the Department of Comparative Psychology, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany. Austėja Kazemekaitytė (austėja.kazemekaityte@unibz.it) is a Postdoctoral Research Fellow at the Faculty of Economics and Management, Free University of Bozen-Bolzano, Piazzetta dell'Università, 1, 39031 Bolzano, Italy. Afreen S. Khalid (afreen.khalid@ru.nl) is a Doctoral Researcher at the Behavioural Science Institute, Radboud University, Postbus 9104, 6500HE Nijmegen, The Netherlands. Kiana Kothe (kianakothe@gmail.com) is a Doctoral Researcher at the School of Health, University of the Sunshine Coast Thompson Institute, 12 Innovation Parkway, Birtinya, QLD 4575, Australia. Philip Krüger (philipkr@edu.aau.at) is affiliated with the Department of Psychology, Alpen-Adria-Universität Klagenfurt, Hauptgebäude, Nordtrakt, Ebene 2, Universitätsstraße 65-67, 9020 Klagenfurt am Wörthersee, Austria. Ngán Lê Thị Kiêu (kieungan.le@gmail.com) is an Independent Researcher. Gintarė Leckė (gintare.lecke@vdu.lt) is an Assistant Professor at the Faculty of Economics and Management, Vytautas Magnus University, K. Donelaicio g. 58, LT-44248 Kaunas, Lithuania. Yanina Ledovaya (y.ledovaya@gmail.com) is a DevEx and AI Evangelist in the Strategy and Communications Department at JetBrains GmbH, Christoph-Rapparin-Bogen 23, 80639 Munich, Germany. Mengyu Lim (mengyu001@e.ntu.edu.sg) is affiliated with Psychology in the School of Social Sciences, Nanyang Technological University, 48 Nanyang Avenue, Singapore 639818, Singapore. Luca M. Lüpken (luca.luepken@hhu.de) is a Postdoctoral Researcher at the Department of Comparative Psychology, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany. Huong Mai Thi Xuan (xuong.mai@gmail.com) is a Market Researcher in the Client Service Department at Kantar Vietnam, Vo Van Tan, District 3, Ho Chi Minh City, Vietnam. Laura Mangold (laura.mangold@unibas.ch) is a Doctoral Researcher at the Department of Psychology, University of Basel, Missionsstrasse 62a, 4055 Basel, Switzerland.

Alfarisi Maulana (alfarisimaulana@outlook.com) is a Government Assistant Trainer at The Regional Centre for Teachers and Education Personnel of North Sumatra, Ministry of Primary and Secondary Education of the Republic of Indonesia, Jl. Kenanga Raya No. 64, Tj. Sari, Kec. Medan Selayang, Kota Medan, Sumatera Utara 20122, Indonesia. Maya Maze (mayamaze21@gmail.com) is an Independent Researcher. Hajdi Moche (hajdi.moche@liu.se) is a Postdoctoral Researcher at the Department of Behavioural Sciences and Learning, Division of Psychology, Jedi-Lab, Linköping University, 581 83 Linköping, Sweden. Zahra Moradi (z.moradi@ucl.ac.uk) is a Senior Lecturer at the Institute of Education, Department of Psychology and Human Development, University College London, Gower Street, London, WC1E 6BT, United Kingdom. Adel Moumin (moumin@hotmail.it) is an Independent Researcher. Valeria Nava (Valeria.nava@live.it) is a Retail Banking Advisor at Banca Monte dei Paschi di Siena, S.p.A., Piazza Salimbeni 3, 53100 Siena, Italy. Michelle J. Y. Neoh (michelle008@e.ntu.edu.sg) is affiliated with Psychology in the School of Social Sciences, Nanyang Technological University, 48 Nanyang Avenue, Singapore 639818, Singapore. Leonardo Nicolao (leonardo.nicolao@ufrgs.br) is an Assistant Professor at the Department of Management, Universidade Federal do Rio Grande do Sul—UFRGS, Rua Washington Luiz, 855, Porto Alegre, RS 90010-460, Brazil. Hamza O. K. El Hallaoui (hamza.oueld@proton.me) is a Doctoral Researcher at the Institut de Neurosciences de la Timone, Aix-Marseille University, Faculté de Médecine, 27 Bd Jean Moulin, 13005 Marseille, France. Sebastian Olschewski (sebastian.olschewski@unibas.ch) is a Postdoctoral Researcher at the Department of Psychology, University of Basel and Warwick Business School, University of Warwick, Missionsstrasse 62a, 4055 Basel, Switzerland. Adobea Y. Owusu (ayowusu@ug.edu.gh) is a Full Professor at the Institute of Statistical, Social and Economic Research, University of Ghana, E.N. Omaboe Building, Opposite Maths Dept, University of Ghana, Legon, Ghana. Ahmet F. Ozates (ahmetfaikozates@gmail.com) is a Category Manager in the Indirect Procurement Department at C&A GMBH. Sofia Pelica (sofiapelica@gmail.com) is a Doctoral Researcher at the Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Sonja Perkovic (sp@mgmt.au.dk) is an Associate Professor at the Department of Management, Aarhus University, Universitetsbyen 61, 8000 Aarhus C, Denmark. Ananda W. M. Puteri is an Independent Researcher. Hagai Rabinovitch (h.rabinovitch@uva.nl) is a Postdoctoral Researcher at the School of Economics, University of Amsterdam, Roetersstraat 11, WB Amsterdam, 1018, Netherlands. Guilherme A. Ramos (guilhermear7@insper.edu.br) is an Assistant Professor at Insper Institute of Education and Research and Pensi Institute, Rua Quatá, 300, São Paulo, Brazil. Nicole Robitaille (nicole.robitaille@queensu.ca) is an Associate Professor at the Smith School of Business, Queen's University, 143 Union Street, Kingston, ON K7L 3N6, Canada. Caroline Roux (caroline.roux@concordia.ca) is a Professor at the Department of Marketing, Concordia University, John Molson School of Business, 1455 Boul. De Maisonneuve Ouest, Montreal, QC H3G 1M8, Canada. Benjamin Scheibehenne (scheibehenne@kit.edu) is a Professor at the Faculty of Economics and Management, Karlsruhe Institute of Technology, Rüppurrer Str. 1a, 76137 Karlsruhe, Germany. Martin Schoemann (martin.schoemann@tu-dresden.de) is a Research Associate at the Department of Psychology, TUD Dresden University of Technology, Zellescher Weg 17, 01069 Dresden, Germany. Mohammad Seidisarouei (dsarouei@gmail.com) is a Neuropsychologist in the Neurology Department at Evangelische Klinikum Niederrhein-EVKLN, Fahrner Straße 133, 47169 Duisburg, Germany. Sanjay Singh (sanjay.singh@iimsirmaur.ac.in) is an Assistant Professor at the Department of Organizational Behaviour and Human Resource Management, Indian Institute of Management Sirmaur, Dhaula Kuan, Himachal Pradesh 173031, India. Mustafa Z. Söyük (zeyd.soyuk@uni-marburg.de) is a Doctoral Researcher at the Faculty of Psychology, Philipps-Universität Marburg, Gutenbergstraße 18, 35037 Marburg, Germany. Liza Steiner (liza.steiner77@gmail.com) is affiliated with the Department of Psychology, Alpen-Adria-Universität Klagenfurt,

Hauptgebäude, Nordtrakt, Ebene 2, Universitätsstraße 65-67, 9020 Klagenfurt am Wörthersee, Austria. Amira TarekAl Rai (amiratarek.at@gmail.com) is a Senior Market Research Analyst in the Research and Business Development Department at Cairo 3A, 5th Settlement, Cairo, Egypt. Berto Usman (berto_usman@unib.ac.id) is an Associate Professor at the Department of Management, University of Bengkulu, Jl. W.R Supratman, Kandang Limun, Bengkulu 38122, Indonesia. Hannah Van Alebeek (hannah.vanalebeek@plus.ac.at) is a Senior Scientist at the Department of Psychology, University of Salzburg, Hellbrunner Straße 34, 5020 Salzburg, Austria. Mohammad H. Vazirian (m.vazirian@modares.ac.ir) is affiliated with the Department of Psychology, Tarbiat Modares University, Tehran, Iran. Evgeniya Vedernikova (e.vedernikova@tilburguniversity.edu) is a Doctoral Researcher at the Tilburg School of Social and Behavioral Sciences, Tilburg University, Prof. Cobbenhagenlaan 225, 5037 DB Tilburg, The Netherlands. Janet L. Wijaya (janetlouis696@gmail.com) is a Teacher at the Department of Social Studies, CHIS School Bali, Jl. Pulau Moyo Gg. Umasari No. 8, Pedungan, Denpasar Selatan, Kota Denpasar, Bali 80223, Indonesia. Xinxin Zhu (xinxinzhu@dufe.edu.cn) is an Assistant Professor at the School of Finance, Dongbei University of Finance and Economics, 217 Jianshan Street, Shahekou District, Dalian, Liaoning 116023, China. Jichuan Zong (zong@dufe.edu.cn) is a Professor at the School of Finance, Dongbei University of Finance and Economics, 217 Jianshan Street, Shahekou District, Dalian, Liaoning 116023, China. Leaf Van Boven (vanboven@colorado.edu) is a Professor at the Department of Psychology and Neuroscience, University of Colorado Boulder, UCB 345, Boulder, CO 80309, USA. Stephan Dickert (stephan.dickert@hnu.de) is a Full Professor at the Faculty of Business and Economics, University of Applied Sciences Neu-Ulm, Wileysstraße 1, 89231 Neu-Ulm, Germany. Lorella Lotto (lorella.lotto@unipd.it) is a Full Professor at the Department of Developmental Psychology and Socialisation, University of Padova, Via Venezia, 8, 35131 Padova, Italy. David J. Hardisty (david.hardisty@sauder.ubc.ca) is an Associate Professor at the UBC Sauder School of Business, The University of British Columbia, 2053 Main Mall, Vancouver, BC V6T 1Z2, Canada. Justin Pomerance (Justin.Pomerance@unh.edu) is an Assistant Professor at the Marketing Department, University of New Hampshire, Peter T. Paul College of Business and Economics, 10 Garrison Ave, Durham, NH 03824, USA. Beatriz Pereira (beatriz.pereira@monash.edu) is a Senior Lecturer at the Department of Marketing, Monash University, 26 Sir John Monash Drive, Caulfield East, VIC 3145, Australia. Kai Ruggeri (kai.ruggeri@columbia.edu) is a Professor at the Department of Health Policy and Management, Columbia University, 722 W 168th Street, New York, NY 10032, USA; and a Senior Visiting Fellow at the Policy Research Group, Centre for Business Research, Judge Business School, University of Cambridge, Trumpington Street, Cambridge, United Kingdom. Enrico Rubaltelli (enrico.rubaltelli@unipd.it) is an Associate Professor at the Department of Developmental Psychology and Socialisation, University of Padova, Via Venezia, 8, 35131 Padova, Italy. Address all correspondence to Enrico Rubaltelli. This article reports a large-scale international replication project involving more than 100 authors (see the preprint at <https://osf.io/apc26/>). Shared first authorship is attributed to Giulia Priolo, Federica Stabulum, and Martina Vacondio, with Enrico Rubaltelli supervising the project and Simone D'Ambrogio conducting the statistical analyses. David J. Hardisty, Justin Pomerance, and Beatriz Pereira served as senior supervisors during manuscript preparation and provided relevant contributions during the review and resubmission process. The authors gratefully acknowledge the JDM Lab at the University of Padova for providing the supportive environment in which this project was conceived and developed. Special thanks are extended to the members of the lab whose contributions were indispensable to the completion of this research. Additionally, the authors would like to express their sincere gratitude to Leaf Van Boven, Stephan Dickert, Lorella Lotto, and Kai Ruggeri for their invaluable comments and assistance at every stage of the project and during the preparation of the early manuscript draft. **Supplementary materials** are included in the **web appendix** accompanying the online version of this article.

REFERENCES

- Abeler, Johannes and Felix Marklein (2017), "Fungibility, Labels, and Consumption," *Journal of the European Economic Association*, 15 (1), 99–127.
- Altmejd, Adam, Anna Dreber, Eskil Forsell, *et al.* (2019), "Predicting the Replicability of Social Science Lab Experiments," *PLoS One*, 14 (12), e0225826.
- Antonides, Gerrit and Rob Ranyard (2017), "Mental Accounting and Economic Behaviour," in *Economic Psychology*, Chichester, West Sussex, UK: John Wiley & Sons, Ltd, 123–38.
- Argo, Jennifer J., Darren W. Dahl, and Rajesh V. Manchanda (2005), "The Influence of a Mere Social Presence in a Retail Context," *Journal of Consumer Research*, 32 (2), 207–12.
- Arnett, Jeffrey J. (2008), "The Neglected 95%: Why American Psychology Needs to Become Less American," *The American Psychologist*, 63 (7), 602–14.
- Atlas, Sara A. and Daniel M. Bartels (2018), "Periodic Pricing and Perceived Contract Benefits," *Journal of Consumer Research*, 45 (2), 350–64.
- Bai, R. (2023), "Impact of Financial Literacy, Mental Budgeting and Self Control on Financial Wellbeing: Mediating Impact of Investment Decision Making," *PloS One*, 18 (11), e0294466.
- Barberis, Nicholas and Ming Huang (2001), "Mental Accounting, Loss Aversion, and Individual Stock Returns," *The Journal of Finance*, 56 (4), 1247–92.
- Baucells, Manel and Woonam Hwang (2017), "A Model of Mental Accounting and Reference Price Adaptation," *Management Science*, 63 (12), 4201–18.
- Cheng, Alex and Carey K. Cryder (2018), "Double Mental Discounting: When a Single Price Promotion Feels Twice as Nice," *Journal of Marketing Research*, 55 (2), 226–38.
- Chinn, Susan (2000), "A Simple Method for Converting an Odds Ratio to Effect Size for Use in Meta-Analysis," *Statistics in Medicine*, 19 (22), 3127–31.
- Choe, Young and Carol M. Kan (2021), "Budget Depreciation: When Budgeting Early Increases Spending," *Journal of Consumer Research*, 47 (6), 937–58.
- Choi, James J., David Laibson, and Brigitte C. Madrian (2009), "Mental Accounting in Portfolio Choice: Evidence from a Flypaper Effect," *The American Economic Review*, 99 (5), 2085–95.
- Comelli, Martino (2021), "The Impact of Welfare on Household Debt," *Sociological Spectrum*, 41 (2), 154–76.
- Consumer Financial Protection Bureau (2013), "A Review of the Impact of the CARD Act on the Consumer Credit Card Market," October 1, https://files.consumerfinance.gov/f/201309_cfpb_card-act-report.pdf.
- Deuflhard, Florian, Dimitris Georgarakos, and Roman Inderst (2019), "Financial Literacy and Savings Account Returns," *Journal of the European Economic Association*, 17 (1), 131–64.
- Economist Intelligence Unit (2024), "Democracy Index 2023: Age of Conflict," <https://policycommons.net/artifacts/11334740/democracy-index-2023-final-report/12223667/>.
- Erat, Sanjog and S. Raghunath Bhaskaran (2012), "Consumer Mental Accounts and Implications to Selling Base Products and Add-Ons," *Marketing Science*, 31 (5), 801–18.
- Evers, Ellen R., Alex Imas, and Chulwoo Kang (2022), "On the Role of Similarity in Mental Accounting and Hedonic Editing," *Psychological Review*, 129 (4), 777–89.
- Fei, Lin, Daniel M. Bartels, and Wenwen Zhang (2026), "Consumers' Mental Representation of Expenditures: Implications for Spending and Savings Decisions," *Journal of Consumer Research*, 52 (6), 1076–101.
- Ghai, Sangeet (2021), "It's Time to Reimagine Sample Diversity and Retire the WEIRD Dichotomy," *Nature Human Behaviour*, 5 (8), 971–2.
- Gourville, John T. (1998), "Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation," *Journal of Consumer Research*, 24 (4), 395–403.
- Hahnel, Ulf J. J., Gilles Chatelain, Beatrice Conte, Valentino Piana, and Tobias Brosch (2020), "Mental Accounting Mechanisms in Energy Decision-Making and Behaviour," *Nature Energy*, 5 (12), 952–8.
- Hastings, Justine S. and Jesse M. Shapiro (2013), "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks," *The Quarterly Journal of Economics*, 128 (4), 1449–98.
- Heath, Chip and Jack B. Soll (1996), "Mental Budgeting and Consumer Decisions," *Journal of Consumer Research*, 23 (1), 40–52.
- Henrich, Joseph, Jean Ensminger, Richard McElreath, *et al.* (2010), "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment," *Science (New York, N.Y.)*, 327 (5972), 1480–4.
- Henrich, Joseph, Steven J. Heine, and Ara Norenzayan (2010), "The Weirdest People in the World?," *The Behavioral and Brain Sciences*, 33 (2–3), 61–83.
- Hossain, Mohammad Tazul (2018), "How Cognitive Style Influences the Mental Accounting System: Role of Analytic versus Holistic Thinking," *Journal of Consumer Research*, 45 (3), 615–32.
- Ioannidis, John P. A. (2005), "Why Most Published Research Findings Are False," *PLoS Medicine*, 2 (8), e124.
- Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47 (2), 263–91.
- Kahneman, Daniel and Amos Tversky (1982), "The Psychology of Preferences," *Scientific American*, 246 (1), 160–73.
- Kahneman, Daniel and Amos Tversky (1984), "Choices, Values, and Frames," *American Psychologist*, 39 (4), 341–50.
- Kelting, Kelsey, Scott Robinson, and Richard J. Lutz (2019), "Would You Like to Round Up and Donate the Difference? Roundup Requests Reduce the Perceived Pain of Donating," *Journal of Consumer Psychology*, 29 (1), 70–8.
- Klein, Richard A., Kate A. Ratliff, Michelangelo Vianello, *et al.* (2014), "Investigating Variation in Replicability," *Social Psychology*, 45 (3), 142–52.
- Koch, Alexander K. and Julia Nafziger (2016), "Goals and Bracketing under Mental Accounting," *Journal of Economic Theory*, 162, 305–51.
- Kőszegi, Botond and Filip Matějka (2020), "Choice Simplification: A Theory of Mental Budgeting and Naive Diversification," *The Quarterly Journal of Economics*, 135 (2), 1153–207.
- Krishna, Aradhna (2021), "The Need for Synergy in Academic Policies: An Introduction to the Dialogue on Pre-

- Registration," *Journal of Consumer Psychology*, 31 (1), 146–50.
- Lee, Choong Yeol and Carey K. Morewedge (2023), "Mental Accounting of Product Returns," *Journal of Consumer Psychology*, 33 (3), 583–90.
- Lian, Chen (2021), "A Theory of Narrow Thinking," *The Review of Economic Studies*, 88 (5), 2344–74.
- Lichtenstein, Sarah and Paul Slovic (1971), "Reversals of Preference between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, 89 (1), 46–55.
- Liu, Hsin-Hsien and Hsuan-Yi Chou (2018), "Promotional Formats and Inaction Inertia," *Journal of Economic Psychology*, 66, 22–32.
- Liu, Yunxin and Siegfried Dewitte (2021), "A Replication Study of the Credit Card Effect on Spending Behavior and an Extension to Mobile Payments," *Journal of Retailing and Consumer Services*, 60, 102472.
- Loken, Eric and Andrew Gelman (2017), "Measurement Error and the Replication Crisis," *Science (New York, N.Y.)*, 355 (6325), 584–5.
- Lusardi, Annamaria and Olivia S. Mitchell (2014), "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52 (1), 5–44.
- Lusardi, Annamaria and Olivia S. Mitchell (2017), "How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness," *Quarterly Journal of Finance*, 07 (03), 1750008.
- Mack, Ashley (2023), "The Sneaky Sticker Shock of Subscription Culture," *New York Times*, November 10.
- Muehlbacher, Stephan and Erich Kirchler (2019), "Individual Differences in Mental Accounting," *Frontiers in Psychology*, 10, 2866.
- Munafò, Marcus R., Brian A. Nosek, Dorothy V. M. Bishop, *et al.* (2017), "A Manifesto for Reproducible Science," *Nature Human Behaviour*, 1 (1), 0021.
- Open Science Collaboration (2015), "Estimating the Reproducibility of Psychological Science," *Science*, 349 (6251), aac4716.
- Prelec, Drazen and George Loewenstein (1998), "The Red and the Black: Mental Accounting of Savings and Debt," *Marketing Science*, 17 (1), 4–28.
- Polman, Evan (2012a), "Effects of Self–Other Decision Making on Regulatory Focus and Choice Overload," *Journal of Personality and Social Psychology*, 102 (5), 980–93.
- Polman, Evan (2012b), "Self–Other Decision Making and Loss Aversion," *Organizational Behavior and Human Decision Processes*, 119 (2), 141–50.
- Raghubir, Priya (2006), "An Information Processing Review of the Subjective Value of Money and Prices," *Journal of Business Research*, 59 (10–11), 1053–62.
- Raghubir, Priya and Joydeep Srivastava (2009), "The Denomination Effect," *Journal of Consumer Research*, 36 (4), 701–13.
- Reinholtz, Nicholas, Daniel M. Bartels, and Jeffrey R. Parker (2015), "On the Mental Accounting of Restricted-Use Funds: How Gift Cards Change What People Purchase," *Journal of Consumer Research*, 42 (4), 596–614.
- Reutskaja, Elena, Jeremiah Iyamabo, Priya Raghubir, and Inigo Gallo (2020), "Price-Denomination Effect: Choosing to Pay With Denominations That Are the Same as the Product Prices," *Frontiers in Psychology*, 11, 552888.
- Ruggeri, Kai, Amma Panin, Milica Vdovic, *et al.* (2022), "The Globalizability of Temporal Discounting," *Nature Human Behaviour*, 6 (10), 1386–97.
- Ruggeri, Kai, Sonia Ali, Mari Louise Berge, *et al.* (2020), "Replicating Patterns of Prospect Theory for Decision under Risk," *Nature Human Behaviour*, 4 (6), 622–33.
- Setterfield, Mark and Yun K. Kim (2024), "How Financially Fragile Can Households Become? Household Borrowing, the Welfare State, and Macroeconomic Resilience," *Review of Evolutionary Political Economy*, 5 (1), 121–51.
- Shafir, Eldar and Richard H. Thaler (2006), "Invest Now, Drink Later, Spend Never: On the Mental Accounting of Delayed Consumption," *Journal of Economic Psychology*, 27 (5), 694–712.
- Shah, Muhammad Umer Draz, Imran Ullah Khan, and Nasir Uddin Khan (2025), "The Role of Digital Payments in Overspending Behavior: A Mental Accounting Perspective," *International Journal of Emerging Markets*, 20 (10), 4031–53.
- Shefrin, Hersch M. and Meir Statman (2000), "Behavioral Portfolio Theory," *The Journal of Financial and Quantitative Analysis*, 35 (2), 127–51.
- Shefrin, Hersch M. and Richard H. Thaler (1988), "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 26 (4), 609–43.
- Shiller, Robert (2017), "Richard Thaler Is a Controversial Nobel Prize Winner—but a Deserving One," *The Guardian*, October 11. <https://www.theguardian.com/world/2017/oct/11/richard-thaler-nobel-prize-winner-behavioural-economics>.
- Shy, Oz (2023), "Cash Is Alive: How Economists Explain Holding and Use of Cash," *Journal of Economic Literature*, 61 (4), 1465–520.
- Silva, Emmanuel Marques, Rafael de Lacerda Moreira, and Patricia Maria Bortolon (2023), "Mental Accounting and Decision Making: A Systematic Literature Review," *Journal of Behavioral and Experimental Economics*, 107, 102092.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn (2011), "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant," *Psychological Science*, 22 (11), 1359–66.
- Simonsohn, Uri, Leif D. Nelson, and Joseph P. Simmons (2014), "P-Curve: A Key to the File-Drawer," *Journal of Experimental Psychology. General*, 143 (2), 534–47.
- Skagerlund, Kenny, Thérèse Lind, Camilla Strömbäck, Gustav Tinghög, and Daniel Västfjäll (2018), "Financial Literacy and the Role of Numeracy—How Individuals' Attitude and Affinity with Numbers Influence Financial Literacy," *Journal of Behavioral and Experimental Economics*, 74, 18–25.
- Skwara, Florian (2023), "Effects of Mental Accounting on Purchase Decision Processes: A Systematic Review and Research Agenda," *Journal of Consumer Behaviour*, 22 (5), 1265–81.
- Soman, Dilip and Amar Cheema (2011), "Earmarking and Partitioning: Increasing Saving by Low-Income Households," *Journal of Marketing Research*, 48 (SPL), S14–22.
- Soman, Dilip and John T. Gourville (2001), "Transaction Decoupling: How Price Bundling Affects the Decision to Consume," *Journal of Marketing Research*, 38 (1), 30–44.

- Soman, Dilip and Vivian M. W. Lam (2002), "The Effects of Prior Spending on Future Spending Decisions: The Role of Acquisition Liabilities and Payments," *Marketing Letters*, 13 (4), 359–72.
- Sussman, Abigail B. and Adam L. Alter (2012), "The Exception Is the Rule: Underestimating and Overspending on Exceptional Expenses," *Journal of Consumer Research*, 39 (4), 800–14.
- Sussman, Abigail B., Eesha Sharma, and Adam L. Alter (2015), "Framing Charitable Donations as Exceptional Expenses Increases Giving," *Journal of Experimental Psychology: Applied*, 21 (2), 130–9.
- Sussman, Abigail B., Yusu Wang, and Anastasiya Apalkova (2023), "Consumer Financial Decision Making," in *Cambridge Handbook of Consumer Psychology*, ed. D. Rucker, C. Lamberton, and S. Spiller, New York: Cambridge University Press.
- Tarlin, Steve (2021), *The Future of Cash* (Federal Reserve Bank of Philadelphia Discussion Paper, 20-03).
- Thaler, Richard H. (1985), "Mental Accounting and Consumer Choice," *Marketing Science*, 4 (3), 199–214.
- Thaler, Richard H. (1990), "Anomalies: Saving, Fungibility, and Mental Accounts," *Journal of Economic Perspectives*, 4 (1), 193–205.
- Thaler, Richard H. (1999), "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12 (3), 183–206.
- Thaler, Richard H. and Cass R. Sunstein (2021), *Nudge: The Final Edition*, New Haven, CT: Yale University Press.
- Thomas, Ian A., Neil R. Kay, and Kristin Laurin (2023), "The I-Frame and the S-Frame: How Focusing on Individual-Level Solutions Has Led Behavioral Public Policy Astray," *Behavioral and Brain Sciences*, 46, 1–25.
- Thomas, Manoj, Kalpesh Kaushik Desai, and Satheeshkumar Seenivasan (2011), "How Credit Card Payments Increase Unhealthy Food Purchases: Visceral Regulation of Vices," *Journal of Consumer Research*, 38 (1), 126–39.
- Tully, Stephanie M. and Eesha Sharma (2018), "Context-Dependent Drivers of Discretionary Debt Decisions: Explaining Willingness to Borrow for Experiential Purchases," *Journal of Consumer Research*, 44 (5), 960–73.
- Tversky, A. and D. Kahneman (1981), "The Framing of Decisions and the Psychology of Choice," *Science (New York, N.Y.)*, 211 (4481), 453–8.
- Tversky, Amos, Paul Slovic, and Daniel Kahneman (1990), "The Causes of Preference Reversal," *American Economic Review*, 80 (1), 204–17.
- van Rooij, Maarten C. J., Annamaria Lusardi, and Rob J. M. Alessie (2012), "Financial Literacy, Retirement Planning and Household Wealth," *The Economic Journal*, 122 (560), 449–78.
- Wang, Mei, Marc Oliver Rieger, and Thorsten Hens (2016), "How Time Preferences Differ: Evidence from 53 Countries," *Journal of Economic Psychology*, 52, 115–35.
- Wood, Stacy (2024), "The Future of Consumer Research Methods: Lessons of a Prospective Retrospective," *Journal of Consumer Research*, 51 (1), 151–6.
- World Bank (n.d.), "World Bank Open Data," <https://data.worldbank.org>.