CESifo WORKING PAPERS

10796
2023

Original Version: November 2023
This Version: March 2024

# The Virtues of Lab Experiments

*Gary Charness, James Cox, Catherine Eckel, Charles Holt, Brian Jabarian*

CESifo

## Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10796*

# The Virtues of Lab Experiments

## Abstract

Physical lab experiments have played an instrumental role in sculpting the history of experimental economics, facilitating controlled information conditions, efficient monetary inducements, and exclusive advantages via immediate human interaction and engaging experiences. These unique benefits render in-person lab experiments essential for the future of experimental economics, complementing the growth of online experiments and the emerging AI revolution. We characterize the environments in which it seems particularly important to conduct lab-in-the-lab experiments. Overall, the lab benefits culminate in a comprehensive research procedure that produces precise and enlightening outcomes, ultimately enriching the domain of experimental economics, and potentially extending benefits to the broader realm of social science.

*Gary Charness*
*University of California*
*Santa-Barbara / CA / USA*
*charness@econ.ucsb.edu*

*James Cox*
*Andrew Young School of Policy Studies*
*Georgia State University, Atlanta / GA / USA*
*jccox@gsu.edu*

*Catherine Eckel*
*Department of Economics, Texas A&M University, College Station / TX / USA*
*ceckel@tamu.edu*

*Charles Holt*
*Department of Economics, University of Virginia, Charlottesville / VA / USA*
*cah2k@virginia.edu*

*Brian Jabarian**
*Booth Business School*
*University of Chicago / IL / USA*
*jabarian@uchicago.edu*

*corresponding author

March 7, 2024
We thank Colin Camerer, Tim Cason, Ingar Haaland, John List, Andreas Ortmann, Marie Claire Villeval, Daniel Zizzo for useful comments. We thank Rishane Dassanayake for excellent research assistance. Brian Jabarian acknowledges funding from Effective Ventures US.

# 1 Introduction

In experimental economics, the research question ideally drives the choice of the setting — lab, field, online, or a mixture. Recent years, however, have seen a marked shift toward online experiments at the expense of traditional lab settings, as shown by Fréchette et al. (2022). We suspect that many of these research questions might have been more effectively addressed in a lab environment. These might involve complex settings and strategic and social interactions. Nevertheless, this growing preference for online settings has several valid underpinnings. At least three main factors contribute. First, research funding constraints and power calculation requirements lead to a compelling draw toward collecting data at a reduced cost. Second, online experiments can reduce logistical burdens. They minimize the time, material, human resources, and financial commitments associated with creating a sample pool, managing a lab, and conducting the experiment. Third, online platforms provide researchers access to more heterogeneous and representative data, which can partly mitigate concerns about external validity (Snowberg and Yariv, 2021).

In light of the accelerating digital economy and the generative AI revolution (Bommasani et al., 2022; Acemoglu and Johnson, 2023), compounded by the shift to remote work in the wake of Covid-19 (Barrero et al., 2021; Hansen et al., 2023), this trend toward online experimentation is expected to intensify. Although the potential of AI in all stages of experimentation is explored in the community (Charness et al., 2023b), we argue that lab experiments retain unique, irreplaceable benefits. Rather than being rendered obsolete by generative AI and other experimental settings, these advantages complement them. One obvious advantage of lab experiments is that the people subjects see are obviously humans, and invisible manipulation would violate no-deception traditions. Calling on the rich history of lab experiments, we champion the continued use of physical lab experiments when they are best suited to the research question. This article advocates for recognizing and preserving the uniqueness of physical lab experiments as a valuable research tool for current and future generations of economists. Calling on the rich history of lab experiments, we champion the continued use of physical lab experiments when they are best suited to the research question and as a valuable research tool for current and future generations of economists.

The rest of the paper is organized as follows. Section 2 addresses the unique advantages of participant engagement and comprehension facilitated by laboratory settings. Section 3 considers the crucial role of maintaining strict information conditions in laboratory experiments. Section 4 discusses the pivotal role of physical procedures and

credibility in the context of laboratory experiments. Aside from discussing these benefits, Section 5 presents several specific case studies that illustrate the advantages of in-person lab experiments. Section 6 discusses and attempts to identify the environments and circumstances where lab experiments seem to be the best tool available and concludes.

## 2 Participants' Engagement and Comprehension

Establishing mutual understanding among participants is crucial as a basis for making choices in economics experiments; this is especially true for decisions in interactive environments, where outcomes depend in part on the choices of others. Providing precise instructions and ensuring informed choices helps to ensure a controlled experiment that yields reliable results. In this context, in-person laboratory experiments have a distinct edge as they foster social engagement. One might expect subjects who are physically present to be inherently more engaged and committed. We mention three factors. First, being physically present in a lab may induce a greater sense of participating in something serious and valuable. In this regard, laboratory experiments also tend to have lower dropout rates, as highlighted by Dandurand et al. (2008), Zhou and Fishbach (2016), and Jun et al. (2017). Second, the learning environment reinforces participants' engagement. Technical issues, which could disrupt group processes in online settings, can be mitigated in lab environments, where researchers can closely monitor the experiment in real time. Third, the synchronicity of the experiment in terms of place and time can also strengthen participants' attention commitment.

Furthermore, trust and transparency are integral to laboratory experiments, fostering a mutually beneficial relationship between researchers and participants. In-person laboratory settings offer an environment conducive to real-time problem-solving on the part of the researchers, enhancing the efficiency and effectiveness of the experiment. This immediate troubleshooting ranges from sorting out technical difficulties to mitigating participant distress. Still, despite the benefits of heightened engagement, it is important to recognize that the level of engagement observed in lab settings may not accurately reflect behavior in some real-world scenarios where engagement is more limited and distraction is prevalent.

Experimental best practices include offering avenues for enhancing (or ensuring) participant comprehension. For example, many challenges to participant comprehension can be addressed during pilot experiments, thereby refining the instructions, and facilitating the smooth execution of the primary research sessions. While pilot exper-

iments can of course be conducted online, an important feature of pilot experiments is receiving questions and feedback from the participants (this is akin to having in-person focus groups with long sessions of feedback). De-briefing, especially after a pilot experiment, is a valuable practice that is easier in person. While we are unaware of studies on this point, we suspect that nearly everyone who has taught both remotely and in person would attest to questions being more frequent when people are physically present; such queries can alleviate confusion and build confidence. This is particularly important when experiments involve single decisions, such as one-shot games, as comprehension of the strategic layout is imperative.

Best practices also now include carefully providing examples for the participants. While this could potentially influence behavior due to an anchoring effect, as noted by Li et al. (2021), they can substantially enhance comprehension by lowering the cognitive load in cases where decision-making is complex or difficult. In such cases, using examples that cover a broad spectrum of possible choices and outcomes is advisable. Furthermore, reading the instructions aloud is another best practice, since this ensures that the participants have common information. This common information is only possible in the lab, since one doesn't know if all other participants have received the same information. This issue can affect behavior in interactive decisions. This is the best one can do, given that common knowledge is impossible (no one really knows what information other people have successfully processed).

Another effective method for enhancing comprehension is active learning, such as inviting participants to answer questions about the material. When managed thoughtfully, this can significantly boost engagement and understanding while fostering a positive environment for social interactions and suggestions. For example, the six rows in the instructions for Dufwenberg and Muren (2006) experiment covered all six possible cases and stated the outcome in each case. The experimenter read the first case and asked a participant for the outcome if this case occurred. The experiment then proceeded to the next case. Now everyone in the room realized that they might get called upon for an answer and one suspects this increased their comprehension and level of engagement.

Another approach involves allowing participants to create their own examples and then having research assistants on-site to verify and to correct comprehension errors. Goeree and Holt (2001) used this technique to ensure understanding before subjects played a sequence of single-shot games with no rematching, where it was essential to prevent earnings confusion in advance. In the "travelers' dilemma" game, for example, the subjects were asked to provide an example where their numerical "claim" on

a wide interval was higher than the other person's, another example where the ranks were reversed, along with a third case where the claims were equal. Then the subjects' earnings calculations were privately checked to ensure comprehension without affecting other subjects' beliefs. This approach is easier to supervise with physical presence. Finally, while this point applies to both online and lab experiments, it is worth noting that experiments often primarily utilize college students as subjects. Although this practice has faced criticism, college students offer several advantages. They are typically easier to instruct, intelligent, and willing to ask questions if they do not understand something. Additionally, they can be motivated by relatively small monetary incentives. Snowberg and Yariv (2021) found that student subjects exhibited the lowest noise levels compared to samples from the general population. Furthermore, supporting the value of laboratory experiments, Rigotti et al. (2023) contended that despite the cost-effectiveness of online observations, the advantages are counterbalanced by increased noise levels. They also observed that subjects in a lab setting were more responsive to treatments, a major consideration.

# 3 Information Conditions

Laboratory experiments in economics and other social sciences require strict control over the information environment. While the first section highlighted how laboratory experiments enhance participant comprehension and engagement, leading to more reliable results, this section focuses on the equally significant role of implementing stringent information conditions. Specifically, this involves ensuring that all participants are privy to identical information and meticulously minimizing the influence of external variables. Such control is vital, since it minimizes or eliminates uncertainties that could otherwise skew participants' behavior, thereby bolstering the validity and reliability of the experimental results. This control shelters the factors under study from external noise and fosters a more refined understanding of economic and social phenomena. This distinct focus illustrates another dimension of the strengths that laboratory settings bring to empirical research in economics and social sciences.

First, uniformity in experimental setup facilitates optimal information conditions. For instance, employing identical computer devices for all participants eliminates any variation in results attributable to hardware or software disparities, allowing researchers to identify and rectify problems on-site. For example, a researcher who is reading the instructions along with subjects is able to promptly identify and correct a setup error pertaining to random matchings or value incentives, This close review is essential,

since errors of this nature might otherwise persist in a series of loosely monitored sessions.

Second, laboratory experiments possess a significant advantage in controlling participant access to external information. This control is inherently more straightforward in hand-conducted experiments. Still, it can also be accomplished in computerized lab experiments by limiting internet access and restricting participants to only the software necessary for the experiment. Further steps to manage information access include requiring participants to remain silent and storing personal electronic devices in a designated area. Monitoring participants in a controlled environment eliminates the potential for external information-seeking by subjects. In view of these factors, physical laboratory settings also mitigate the likelihood of cheating or collaboration (a persistent challenge in online experiments). In comparison, online experiments present unique challenges in controlling information access, as participants have greater autonomy.

Finally, timing can subtly influence experimental outcomes, and laboratory experiments offer the unique advantage of allowing researchers to have granular control over this crucial aspect. For instance, Uliana Popova's pioneering "stripped down poker" experiment revealed that "second mover" subjects used the decision delay of the "first mover" with whom they were matched to infer whether that person's card draw was high (Ace) or low (King). During a post-session debriefing, some first movers even admitted to intentional delay after drawing a strong hand (Ace) in order to trick the second mover into not folding. The researchers had not considered this timing issue, since the experiment was computerized. The solution was to delay releasing any of the first movers' raise or fold decisions until after the decisions for all matched pairs had been submitted and confirmed. The capacity to exert control over timing is indispensable in ensuring the reliability of experimental results. In the case of Popova's experiment, the controlled environment of the laboratory was instrumental in identifying a timing bias where second movers gained unintended information; this facilitated subsequent procedural adjustments. Such nimble optimization of timing conditions is not easily achieved in field experiments or online surveys due to constraints in controlling responses or interactions. Through precise control over the timing in lab experiments, researchers can isolate variables of interest and minimize confounding timing cues, enhancing the reliability and internal validity of the results.

# 4 Phyisical Procedures and Credibility

For economic experiments to produce reliable and valid results, researchers need to have credibility with the participants. Laboratory experiments allow researchers to establish credibility via payment systems and the enforcement of fair and transparent procedures. It is generally accepted as best practice in the experimental field to provide performance incentives in the form of cash, directly tying participants' outcomes to tangible monetary outcomes.

There is evidence that this mechanism may elicit more authentic responses from participants (Camerer and Hogarth, 1999; Gneezy and Rustichini, 2000). This is because incorporating real monetary consequences augments the salience of decision-making contexts and prompts participants to deliberate with a level of seriousness akin to real-world economic environments (Harrison and List, 2004). This alignment between participants' interests and the experiment's objectives facilitates a more genuine representation of decision-making behavior, which is integral to ecological validity (Belot et al., 2015).

Of course, online experiments may also have performance-based incentives. But expecting an online payment is not quite the same thing as cash. While it is true that students in a lab session are perfectly happy to be paid through Venmo, there tends to be a degree of uncertainty about payments made remotely; now is not the same as later. For example, a co-author recently received many anxious inquiries about being paid for a Prolific task when no payment had been received after 12 hours. Thus, even though online experiments also typically employ monetary incentives, the laboratory's immersive and controlled environment may well have a more pronounced effect on the seriousness with which participants approach tasks (Charness et al., 2013) and on the belief that they will be paid as stated. Thus, we feel that conducting in-person experiments is the gold standard for establishing credibility, and enhancing the reliability and validity of empirical findings through transparent and immediate payment systems.

For payments, instilling confidence that earnings will be adequately compensated can be achieved by paying show-up fees: a pre-announced amount, such as $5 or $10, for punctuality. This guarantee can be communicated through the recruitment email and repeated in the consent form, further reinforced by having the cash visible on each participant's desk at the beginning of the session. It is also possible to distribute the cash show-up fee to the participants at the beginning of a session, making it clear that cash is involved. Payments at the session's conclusion should remain private. Currently, private payments through platforms like Paypal or Venmo are widely accepted. Inves-

tigating if these payment methods impact behavior in sensitive contexts, such as the dictator game, would be valuable (Umer, 2023).

The enhancement of credibility also hinges on establishing transparent and fair procedures. For example, as discussed in Section 2, providing consistent and clear instructions to all participants can mitigate potential confounding factors from misunderstandings or misinterpretations, and facilitating participant queries during the experiment can clarify any ambiguities, ensuring a comprehensive understanding of the experimental process. Furthermore, laboratory experiments in economics frequently incorporate safeguards to bolster credibility by convincingly demonstrating non-deception. The assurance that the information presented to participants is authentic and forthright is integral to the validity and reliability of the data gathered (Jamison et al., 2008).

When participants are confident that the experimenter is not employing deceptive tactics, they are likelier to engage in the task earnestly and provide responses reflective of their genuine preferences and behavior (Ortmann and Hertwig, 2002). One effective approach includes projecting participant instructions and experiment procedures on a large screen for all to view and understand. The integration of public randomization devices, such as bingo cages or dice, also aids in improving participants' grasp of probabilities and dispelling misconceptions. As a result, participants could be considerably more likely to accept the randomness of outcomes when they observe the physical manifestation of randomization processes.

# 5 Case Studies

This section encompasses several case studies that illustrate some of the unique advantages of laboratory experiments.

## 5.1 Belief in Real-Time Impact on Other Participants

A key characteristic in laboratory experiments lies in the knowledge of participants interacting with authentic individuals. In experiments centered around social preferences or social dilemmas, it is paramount that participants believe their decisions impact other participants' rewards. Physical laboratory experiment procedures can effectively foster such beliefs. For example, a transparent and credible protocol is to initially assemble all of the participants in the same room, for visual observation of all the others, before dividing them into separate rooms for decisions. This setup helps participants believe that their decisions have a real-time impact on other participants'

rewards, as they can see the other subjects, and also observe one of the subjects being randomly selected and publicly given the responsibility for ensuring that the stated procedures are accurately followed.

Cox (2004)'s investment-game experiment highlights the benefit of laboratory experiments by gathering all participants in the lab and using a physical device to randomly assign them as first or second movers. Similarly, the moonlighting game experiment by Cox et al. (2008) further emphasizes the unique benefit of laboratory experiments by having participants draw a folded piece of paper marked with either a  or # from a container and directing them to rooms labeled accordingly. This process of categorizing participants as first or second movers in separate rooms strengthens the belief in the real-time impact of their decisions on others.

In other papers, Frohlich et al. (2001, 2004), evaluated whether participants in the dictator game harbored skepticism regarding other participants' actual presence or the experimental procedures' veracity. In this context, "social distance" refers to the perceived closeness or separation between the decision-maker and the recipient in the experiment. This can include aspects such as the extent to which participants believe they are making decisions that have real consequences for actual individuals—by systematically manipulating this social distance – ranging from conditions where the connection between the participants' choices and their impacts on others was made very explicit to conditions where this connection was more abstract or ambiguous. Frohlich and Oppenheimer observed that increased doubt, corresponding to greater social distance, influenced the allocational patterns. This underscores the significance of cultivating a belief among participants in the immediate and tangible effects of their decisions on others in laboratory experiments.

Moreover, Eckel and Grossman (1996) were concerned about participant belief in the authenticity of their donations in their charity experiment. To address this, they selected a participant to monitor the collection and writing of the charity check and walked with the monitor to deposit the check into a mailbox. This procedure solidified the participants' beliefs in the veracity of the donations and persuaded the accounting office to reimburse the payments, showcasing the unique benefit of laboratory experiments in fostering a plausible belief in real-time impact on other participants.

## 5.2 Physical Randomization Devices and Experimental Protocols

Using physical devices to generate random variable realizations can refine research questions and enhance confidence in data interpretation. In the experiments on the

representativeness heuristic reported by Kahneman and Tversky (1973) and Tversky and Kahneman (1974), two groups of subjects were given the same thumbnail descriptions of five individuals. One group was told that the five descriptions had been randomly selected from a set for 30 engineers and 70 lawyers. The other group was told they had randomly selected from a set for 30 lawyers and 70 engineers. Subjects were asked to report the probability that each thumbnail description described an engineer. Despite the stated difference in prior odds, participants reported almost the same posterior probabilities, suggesting that people disregard prior odds when making decisions and favor the representativeness heuristic over Bayes' rule.

Grether (1980) explored this empirical question further using physical randomization devices. In his experiment, a monitor elected by participants validated the experiment procedures. Quantitative prior and posterior odds were established through bingo cages, which were used to realize random variables. The monitor verified (but did not report to subjects) the initial draw from the cage that determined which of two other bingo cages (with different known contents) was subsequently used to draw a sample. Subjects could observe the sample but not the cage used. Participants were paid based on their choice's alignment with Bayes' rule. Grether's data did not support the conclusion that subjects ignore prior odds information. His data did support the conclusion that experienced subjects give more weight to the observed sample proportions and less to the known prior odds than Bayes' rule prescribes.

In physical lab experiments, devices such as coins, dice, or bingo cages can generate outcomes in front of participants, thus supporting belief in the truthful execution of the experimental procedures as per the subject instructions. Evidence for transparent and credible procedures is especially important when the payoffs involve hundreds of dollars for a single decision task. This is the reason that Holt and Laury (2002) used actual dice thrown at each subject's desk to determine the outcomes in the 90x payoff scale used in a risk preference experiment. Researchers can investigate subjects' decisions under risk when they apply either their subjective notions of probabilities or when physical devices are used to inform classical relative frequency notions. The application of physical randomization devices has a rich history in paper and pencil response protocols and in computerized experiments (Cox and Grether, 1996; Cox et al., 2015).

### 5.3 Privacy, Social Distance, and Reputation Effects: Probing Internalized and Socially-Enforced Norms

In the context of privacy, social distance, and reputation effects, physical laboratory experiments facilitate the credible exploration of privacy features in payoff protocols, revealing how regard for reputation influences decision-making. Distinct payment protocols in laboratory experiments offer valuable insights into the dichotomy of internalized norms dictating behavior and socially-enforced norms where the desire to maintain reputation could alter decisions. While the double-anonymous (or double-blind) payment protocol provides an environment in which a subject's decisions are not linked to their identity, this may come at some cost. For example, Loewenstein (1999) suggests that by emphasizing that choices are anonymous Hoffman et al. (1994)'s experiments might have implied to participants that selfishness was expected. In contrast, the single-blind protocol involves a more public disclosure of decisions, where subjects' responses are personally identified, and they individually approach the experimenter to claim their earnings. Notably, research conducted by Hoffman et al. (1994) and Cox and Deck (2005) divulges significant variation in decisions within certain fairness games under the single-blind versus double-blind protocols.

Physical laboratory experiments, thus, facilitate the credible exploration of privacy features in payoff protocols, revealing how regard for reputation influences decision-making. The study by Hoffman et al. (1996) shows this, demonstrating how each facet of their double-blind procedure affects generosity. Further, distinct payoff protocols can be considered; an example is provided by Dufwenberg and Muren (2006), where a randomly selected student had to publicly announce their dictator choices in an auditorium, observed by the other subjects and experimenters. Although this "stage" protocol introduces certain confounds, this issue can be accurately explored and varied within the physical laboratory setting—a potentially challenging or unfeasible process in other settings, for example, online.

## 6 Physical Experience and Knowledge Acquisition

An experimental paper by Charness et al. (2023a) illustrates the value of in-person laboratory experiments in discerning whether individuals appear to learn their risk preferences. The study alleviates the artificial nature of unfamiliar decisions, such as selecting a gamble from a table or determining investment in a risky asset, by familiarizing participants with various choice mechanisms and potential outcomes.

Participants initially chose from six gambles in a modified Eckel-Grossman risk-elicitation design, with a 50% chance that this initial choice would count towards a monetary payoff (Dave et al. (2010)). Twenty-four practice rounds were then undertaken, half involving a pre-specified gamble, the other half allowing a free choice of a gamble. Participants physically rolled dice to determine the potential outcome and noted it on a record sheet. A final gamble choice followed, with the remaining 50% chance of being chosen for payoff implementation. The study found that 55% of participants altered their choice of gambles, with distinct patterns and statistically significant changes oriented towards higher expected payoffs and diminished risk aversion.

Finally, experiments like Gneezy et al. (2014) work underline another advantage of the laboratory setup. They mimicked an overpayment scenario in person and observed the subjects' responses to it. This simulated real-world experience brought the field into the lab. While it might be possible to conduct an overpayment scenario online, it is difficult to believe (but has not yet been tested) that remote participants would make the effort to report overpayment and somehow refund it through the online platform.

## 6.1 Experiments on Macroeconomics

Ricciuti (2008) explores the potential of laboratory experiments in macroeconomic research. The paper revolves around the applicability of laboratory experiments in macroeconomic settings, showing how experiments can contribute to our understanding of macroeconomic phenomena. It establishes two general classes of laboratory experiments in macroeconomics - experiments concerned with general equilibrium and those concerned with testing specific macroeconomic theories. The paper argues that laboratory experiments can provide valuable insights into issues, such as monetary policy, fiscal policy, and market dynamics, by allowing researchers to control the experimental environment and isolate specific factors that influence economic behavior. Specifically, the paper demonstrates the benefit of real-time observation in laboratory experiments. Ricciuti's paper highlights the importance of observing participants' behavior and decision-making processes in real-time to better understand the impact of various economic policies and individual decision-making processes in macroeconomic research.

Laboratory experiments in macroeconomics serve as valuable tools for resolving ambiguities associated with multiple equilibria predicted by theoretical models. In cases where multiple equilibria are theoretically plausible, empirical identification of the most probable outcomes can be challenging. Laboratory experiments, however,

can provide controlled environments that allow for a detailed exploration of such situations, as illustrated by the work of Heinemann et al. (2004).

In their study, Heinemann et al. (2004) applied global games to examine equilibrium multiplicity in the context of a speculative currency-attack model. The global game approach reduces equilibrium multiplicity by introducing uncertainty about economic fundamentals, leading individuals to adopt a unique threshold strategy based on their perception of these fundamentals. The researchers manipulated the information set, creating both complete and incomplete information treatments, where participants made decisions based on different knowledge levels about economic fundamentals. Their findings showed that subjects in both treatments adopted threshold strategies, suggesting that the interplay between available information and perceived economic fundamentals drives individuals' choices. This observation aids in disentangling the possible equilibria in a theoretical model with multiple equilibria. Furthermore, the variance in decision-making was lower, and there was more significant coordination when public information was available, supporting the notion that public information could guide participants' beliefs and coordinate their choices around a common equilibrium.

Therefore, by enabling researchers to manipulate and control various elements, such as information availability and perceived economic fundamentals, laboratory experiments provide unique insights into which equilibria are most likely to occur in practice. This attribute enhances their value in macroeconomic research, particularly when multiple theoretical equilibria exist.

## 6.2 Comparative Analysis of Individual and Small Group Behavior

Small entities, such as families, committees, and boards, frequently make decisions. Laboratory experiments excel at investigating the distinctions and similarities between individual and small-group behavior. Such groups can either be "natural", meeting in person to discuss and reach decisions, or "virtual", communicating via digital means. If the lab setup includes separate rooms, each group can hold private discussions, and the decisions can be communicated on paper or via computer systems linking the rooms. Additional insights into decision-making processes can be acquired by capturing audio and video recordings of discussions if the lab is suitably equipped.

Cox (2002) compared the actions of individuals and three-person groups within the investment game framework, inquiring whether groups sent or returned more or fewer per capita resources than individuals. Following individual decisions in a 40-

person lab, participants were randomly formed into three-person groups and directed to separate (breakout) rooms to make decisions within the same game framework. The groups had the liberty to reach decisions in any manner they chose. Individual decisions were privately recorded with ID numbers, also included in the group decision data. This facilitated the analysis of whether group transfers corresponded to the low, medium, or high amounts transferred by group members when they made decisions as individuals.

Similarly, Cox and Hayne (2006) examined the disparities between individual and group bidding in first-price common value auctions to determine if small groups exhibited more or less rationality than individuals. In this context, a bid is rational if it does not exhibit the winner's curse: the bid is not higher than the expected value of the auctioned item conditional on the bidder's value estimate being the highest of $N$ value estimates (where $N$ is the number of individual or group bidders). Discussion within groups can lead them to underbid their unconditional item values by enough to escape the winner's curse that plagues individual bidders. Or, alternatively, group discussion could promote higher bids. Physical labs with small group breakout rooms can support experiments that compare rationality of bidding by natural groups with that of individuals. If a lab has audio and video recording equipment, it can support research on whether group members articulate reasoning over multiple bidding rounds that leads them to underbid their unconditional signals by enough to avoid the winner's curse.

Furthermore, laboratory experiments possess unique advantages in studying macroeconomic phenomena due to their ability to capture, isolate, and analyze the behavior of individuals and small groups under tightly controlled conditions. Duffy (2016) argues that lab experiments help overcome limitations encountered in empirical field data, which often grapple with identification problems and difficulties in controlling for all crucial variables. For example, in money demand, lab experiments allow the creation of economic settings where only the transaction motive exists, a feat challenging to accomplish with field data. This facilitates a more accurate, focused analysis of specific economic theories or principles.

Additionally, Ricciuti (2008) points to the issues posed by the endogeneity of policy in real-world economies, which complicates the analysis and correct inference of data regarding changes and policy reactions. Lab experiments, in contrast, provide complete control over parameters and the flexibility to conduct "what if" scenarios, thereby offering a deep understanding of the impact and direction of change in these parameters.

Consequently, lab experiments can provide rich, detailed insights into individual and small group behaviors under various conditions, which can be used to better understand macroeconomic phenomena when scaled up. This flexibility, control, and replicability make laboratory experiments a powerful macroeconomic research tool.

## 6.3 Unveiling Subject Characteristics

It is common for experiments to explore scenarios where subjects' behavior might vary based on with whom they are paired. Often, the focus is on potential differences stemming from gender or ethnicity. A challenge for researchers is revealing these characteristics without alerting subjects to the experiment's purpose and triggering experimenter demand, especially in discrimination studies. This issue becomes more complex in online experiments.

Creative solutions can be found in the laboratory setting, as shown by two studies in the April 2001 issue of Economic Inquiry that explored gender differences in the ultimatum game. Both sought to understand how gender influenced behavior and how one's counterpart's gender affected this behavior, albeit using different methods. (Eckel and Grossman, 2001) grouped subjects across the room, creating all-male, all-female, and mixed-gender groups facing each other in all possible configurations. Subjects were aware of the gender mix they faced but not the specific identities of their counterparts. Conversely, (Solnick, 2001) used first names to signal gender, visible on subjects' forms. Interestingly, the results from the two studies differed. The reasons for this discrepancy and which method better mirrors real-world behavior remain unexplored.

Niederle and Vesterlund (2007) sought to inform subjects of their group composition (two people who appeared to be female and two people who appeared to be male) without emphasizing gender. They arranged subjects in four rows. In each, there were two who appeared to be female and two who appeared to be male. Sujects were told that the row comprised their group. Eckel and Füllbrunn (2015) explored the gender composition of financial markets by recruiting the desired composition and allowing subjects to infer the gender composition by observing the waiting room. Eckel and Füllbrunn (2017) and Holt et al. (2017) instead recruited a sizeable mixed-gender group, and privately gave them session cards for one of two different asset market logins. In both cases, gender sorting was achieved: in the first case the gender mix was openly revealed, while in the second, single-gender groups were achieved covertly while minimizing the risk of experimenter demand effects.

(Cox and Deck, 2006) adopted trust and dictator games in a computerized experiment with varying money payoff levels and social distance protocols to address the question: "When are women more generous than men?" In the high social-distance protocol, subjects privately collected payoffs using numbered keys, with the number being private information. On the other hand, the low social-distance protocol had subjects enter their names into computers and individually sign for their payoffs, facing the experimenter at the experiment's end. This approach to recording gender did not imply that the experiment focused on gender effects on behavior. Replicating this lab experiment's controls online, where gender is self-identified and payoff protocols have high social distance, could prove challenging.

# 7 Conclusion

This article underlines the importance of laboratory experiments, noting the distinct benefits they afford. These include controlling environmental variables, enabling real-time human interactions, managing procedures effectively, and precisely assessing participant engagement and comprehension. Laboratory environments contribute to the credibility of social impact research and facilitate real-time observations of the effects of decisions, enriching the authenticity of the experimental experience. Lab settings also offer enhanced privacy protections, an essential facet given the increasing importance of safeguarding participant data.

We would like to outline those environments where laboratory experiments seem particularly useful. Charness (2015) points out that there can be "many ladders to the same roof", and that one should use the tool that is best for the task at hand. Regarding the relevant characteristics, a first consideration is the greater degree of control possible in the lab; for example, it is difficult to control informational access remotely and monitor as effectively as when people are physically present. This control seems especially important when participants can acquire additional information that is relevant to their decisions or when the task is interactive with other participants, as in games.

A second, and vitally important, issue is that of comprehension. If people do not understand the task or the underlying rules, the experimental decisions are random and not useful. Some experiments are easy to understand, while some are very complex and require detailed instructions. If the task is complex or highly nuanced, we feel that one should best use a physical laboratory when this is feasible. A closely-related concern is the flow of information, since information control is critical. People can access other sources of information when responding at home. In addition, it is far easier

(if indeed it is possible) to implement common information in the lab than online.

Third, tasks that require more engagement work best when others are physically present. This also applies to social engagement. For example, telling people online that what they give to a paired person in a dictator game seems likely to not only generate smaller allocations to the recipient than otherwise but also to result in more random outcomes. It is our belief (supported by Bohnet and Frey (1999)) that people will be more pro-social when they physically see others present. Field environments that feature considerable face-to-face interaction are thus best simulated in a laboratory environment. The importance of conducting experiments in a synchronous fashion is closely related. On the other hand, remote and asynchronous field environments might best be conducted online.

Finally, credibility is a key aspect when conducting experiments asking participants to trust either other participants or the experimenter. If people have doubts about the veracity of experimenter claims (as considered by Frohlich et al. (2001)), behavior may well be affected. Our experience is that it is more difficult to lie to one's face than remotely, so credibility is likely to be higher when an experiment is conducted in person. This aspect is not as likely to be a factor in individual decision-making experiments.

The physicality of laboratory environments extends beyond the realm of standard experimental benefits. It paves the way for a rich, experiential learning platform for participants and researchers. The immediacy of observing experiments unfold offers valuable insights, particularly useful in pilot studies, enabling researchers to fine-tune methodologies based on direct feedback. In research on group behavior, a topic of increasing interest (see Charness and Chen (2020)), lab experiments offer an irreplaceable tool for exploring the dynamics and interactions that underpin collective action.

In light of rising new technologies and removing work life after Covid-19, we wish to reaffirm the continued importance of laboratory experiments in economics and social sciences research. We especially encourage junior researchers, who may not have significant exposure to this methodology, to explore its potential and consider the broad array of insights it can provide in complement to other experimental settings, online and field, and imagine how generative AI can help to nature with the unique benefits of physical experiments.

# References


Acemoglu, Daron and Simon Johnson (2023), *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*. Hachette UK.

Bardsley, Nicholas (2000), "Control without deception: Individual behaviour in free-riding experiments revisited." *Experimental Economics*, 3, 215–240.

Bargh, J.A., M. Chen, and L. Burrows (1996), "Automaticity of social behavior: direct effects of trait construct and stereotype-activation on action." *Journal of Personality and Social Psychology*, 71, 230–244.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis (2021), "Why working from home will stick." Working Paper w28731, National Bureau of Economic Research.

Belot, Michèle, Raymond Duch, and Luis Miller (2015), "A comprehensive comparison of students and non-students in classic experimental games." *Journal of Economic Behavior & Organization*, 113, 26–33.

Bohnet, Iris and Bruno S. Frey (1999), "Social distance and other-regarding behavior in dictator games: Comment." *American Economic Review*, 89, 335–339.

Bommasani, Rishi et al. (2022), "On the opportunities and risks of foundation models." *arXiv preprint arXiv:2108.07258*.

Camerer, Colin F. and Robin M. Hogarth (1999), "The effects of financial incentives in experiments: A review and capital-labor-production framework." *Journal of Risk and Uncertainty*, 19, 7–42.

Charness, Gary (2015), "The hammer and the screwdriver." In *Handbook of Experimental Methodology* (Fréchette and Schotter, eds.), 197–199, Academic Press.

Charness, Gary, Nir Chemaya, and Dario Trujano-Ochoa (2023a), "Learning your own risk preferences." *Journal of Risk and Uncertainty*. Forthcoming.

Charness, Gary and Yan Chen (2020), "Social identity, group behavior, and teams." *Annual Review of Economics*.

Charness, Gary, Uri Gneezy, and Michael A. Kuhn (2013), "Experimental methods: Between-subject and within-subject design." *Journal of Economic Behavior & Organization*, 81, 1–8.

Charness, Gary, Brian Jabarian, and John A List (2023b), "Generation next: Experimentation with ai." Technical report, National Bureau of Economic Research.

Cox, James C. (2002), "Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females." In *Advances in Experimental Business Research*, Kluwer Academic Publishers.

Cox, James C. (2004), "How to identify trust and reciprocity." *Games and Economic Behavior*, 46, 260–281.

Cox, James C. and Cary Deck (2005), "On the nature of reciprocal motives." *Economic Inquiry*, 43, 623–635.

Cox, James C. and Cary Deck (2006), "When are women more generous than men?" *Economic Inquiry*, 44, 587–598.

Cox, James C. and David M. Grether (1996), "The preference reversal phenomenon: Response mode, markets, and incentives." *Economic Theory*, 7, 381–405.

Cox, James C. and Stephen C. Hayne (2006), "Barking up the right tree: Are small groups rational agents?" *Experimental Economics*, 9, 209–222.

Cox, James C., Klarita Sadiraj, and Vjollca Sadiraj (2008), "Implications of trust, fear, and reciprocity for modeling economic behavior." *Experimental Economics*, 11, 1–24.

Cox, James C., Vjollca Sadiraj, and Ulrich Schmidt (2015), "Paradoxes and mechanisms for choice under risk." *Experimental Economics*, 18, 215–250.

Dandurand, F., T.R. Shultz, and K.H. Onishi (2008), "Comparing online and lab methods in a problem-solving experiment." *Behavior Research Methods*, 40, 428–434.

Dave, Chetan, Catherine C. Eckel, Cathleen A. Johnson, and Christian Rojas (2010), "Eliciting risk preferences: When is simple better?" *Journal of Risk and Uncertainty*, 41, 219–243.

Doyen, Stephane, Olivier Klein, Cora-Lise Pichon, and Axel Cleeremans (2012), "Behavioral priming: It's all in the mind, but whose mind?" *PLOS One*.

Duffy, J. (2016), *Macroeconomics: A Survey of Laboratory Research*. URL https://escholarship.org/uc/item/7267v17q.

Dufwenberg, Martin and Astri Muren (2006), "Generosity, anonymity, gender." *Journal of Economic Behavior & Organization*, 61, 42–49.

Eckel, Catherine C. and Sascha C. Füllbrunn (2015), "Thar she blows? gender, competition, and bubbles in experimental asset markets." *American Economic Review*, 105, 906–920.

Eckel, Catherine C. and Sascha C. Füllbrunn (2017), "Hidden vs. known gender effects in experimental asset markets." *Economics Letters*, 156, 7–9.

Eckel, Catherine C. and Philip J. Grossman (1996), "Altruism in anonymous dictator games." *Games and Economic Behavior*, 16, 181–191.

Eckel, Catherine C. and Philip J. Grossman (2001), "Chivalry and solidarity in ultimatum games." *Economic Inquiry*, 39, 171–188.

Frohlich, N., J. Oppenheimer, and A. Kurki (2004), "Modeling other-regarding preferences and an experimental test." *Public Choice*, 119, 91–117.

Frohlich, Norman, Joe Oppenheimer, and J. Bernard Moore (2001), "Some doubts about measuring self-interest using dictator experiments: the costs of anonymity." *Journal of Economic Behavior & Organization*, 46, 271–290.

Fréchette, Guillaume R., Kim Sarnoff, and Leeat Yariv (2022), "Experimental economics: Past and future." *Annual Review of Economics*, 14, 777–794.

Gneezy, Uri, Alex Imas, and Kristóf Madarász (2014), "Conscience accounting: Emotion dynamics and social behavior." *Management Science*, 60, 2645–2658.

Gneezy, Uri and Aldo Rustichini (2000), "Pay enough or don't pay at all." *Quarterly Journal of Economics*, 115, 791–810.

Goeree, Jacob K. and Charles A. Holt (2001), "Ten little treasures of game theory and ten intuitive contradictions." *American Economic Review*, 91, 1402–1422.

Goeree, Jacob K. and Charles A. Holt (2016), *Quantal Response Equilibrium: A Stochastic Theory of Games*. Princeton University Press.

Grether, David M. (1980), "Bayes rule as a descriptive model: The representativeness heuristic." *Quarterly Journal of Economics*, 95, 537–557.

Hansen, Stephen, Peter John Lambert, Nicholas Bloom, Steven J. Davis, Raffaella Sadun, and Bledi Taska (2023), "Remote work across jobs, companies, and space." Working Paper w31007, National Bureau of Economic Research.

Harrison, Glenn W. and John A. List (2004), "Field experiments." *Journal of Economic Literature*, 42, 1009–1055.

Heinemann, Frank, Rosemarie Nagel, and Peter Ockenfels (2004), "The theory of global games on test: Experimental analysis of coordination games with public and private information." *Econometrica*, 72, 1583–1599.

Hoffman, Elizabeth, Kevin McCabe, Keith Shachat, and Vernon Smith (1994), "Preferences, property rights, and anonymity in bargaining games." *Games and Economic Behavior*, 7, 346–380.

Hoffman, Elizabeth, Kevin McCabe, and Vernon Smith (1996), "Social distance and other-regarding behavior in dictator games." *American Economic Review*, 86, 653–660.

Holt, C.A., M. Porzio, and M.Y. Song (2017), "Price bubbles, gender, and expectations in experimental asset markets." *European Economic Review*, 100, 72–94.

Holt, Charles A. (2019), *Markets, Games, and Strategic Behavior: An Introduction to Experimental Economics*. Princeton University Press.

Holt, Charles A. and Susan K. Laury (2002), "Risk aversion and incentive effects." *American Economic Review*, 92, 1644–1655.

Holt, Charles A., William Shobe, Dallas Burtraw, Karen L. Palmer, and Jacob K. Goeree (2007), "Auction design for selling co2 emission allowances under the regional greenhouse gas initiative." URL https://ssrn.com/abstract=2459462.

Horton, John, David Rand, and Richard Zeckhauser (2011), "The online laboratory: Conducting experiments in a real labor market." *Experimental Economics*, 14, 399–425.

Jamison, Julian, Dean Karlan, and Laura Schechter (2008), "To deceive or not to deceive: The effect of deception on behavior in future laboratory experiments." *Journal of Economic Behavior & Organization*, 68, 477–488.

Jun, E., G. Hsieh, and K. Reinecke (2017), "Types of motivation affect study selection, attention, and dropouts in online experiments." *Proceedings of the ACM on Human-Computer Interaction*, 1, 1–15.

Kahneman, D. and A. Tversky (1973), "On the psychology of prediction." *Psychological Review*, LXXX, 237–251.

Li, Lunzheng, Zacharias Maniadis, and Constantine Sedikides (2021), "Anchoring in economics: a meta-analysis of studies on willingness-to-pay and willingness-to-accept." *Journal of Behavioral and Experimental Economics*, 90.

Loewenstein, George (1999), "Experimental economics from the vantage-point of behavioural economics." *The Economic Journal*, 109, 25–34.

Niederle, Muriel and Lise Vesterlund (2007), "Do women shy away from competition? do men compete too much?" *The Quarterly Journal of Economics*, 122, 1067–1101.

Ortmann, Andreas and Ralph Hertwig (2002), "The costs of deception: Evidence from psychology." *Experimental Economics*, 5, 111–131.

Ricciuti, Roberto (2008), "Bringing macroeconomics into the lab." *Journal of Macroeconomics*, 30, 216–237.

Rigotti, Luca, Alistair Wilson, and Neeraja Gupta (2023), "The experimenters' dilemma: Inferential preferences over populations."

Snowberg, Erik and Leeat Yariv (2021), "Testing the waters: Behavior across participant pools." *American Economic Review*, 111, 687–719.

Solnick, Sarah (2001), "Gender differences in the ultimatum game." *Economic Inquiry*, 39, 189–200.

Tversky, Amos and Daniel Kahneman (1974), "Judgment under uncertainty: Heuristics and biases." *Science*, CLXXXV, 1124–1131.

Umer, Hamza (2023), "Effectiveness of random payment in experiments: A meta-analysis of dictator games." *Journal of Economic Psychology*, 96.

Zhou, H. and A. Fishbach (2016), "The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions." *Journal of Personality and Social Psychology*, 111, 493–504.