Behavioral Economics for Human-in-the-Loop Control Systems Design

OVERCONFIDENCE AND THE HOT HAND FALLACY

G3 WERBUNG, PADERBORN

MARIUS PROTTE, RENÉ FAHR, and
DANIEL E. QUEVEDO

This century brought interesting challenges and opportunities that derive from the way digital technology is shaping the lives of individuals and society as a whole. A key feature of many engineered systems is that they interact with humans. Rather than solely affecting humans, people often make decisions that affect the engineered system. As an example, when driving cars, people often decide to take a route that differs from that suggested by the navigation system. This information is fed back to the service provider and henceforth used when making route suggestions to other users. The analysis and design of such *cyberphysical human systems* (CPHSs) would benefit from an understanding of how humans behave. However, given their immense complexity, it is unclear how to formulate appropriate models for human decision makers, especially when operating in closed-loop systems (see "Summary" for an overview of this article).

Digital Object Identifier 10.1109/MCS.2020.3019723
Date of current version: 13 November 2020

A key feature of many engineered systems is that they interact with humans.

CPHSs fit into the general context of human–machine systems and pose multidisciplinary challenges that have been addressed in a number of domains. For example, interesting surveys on signal processing approaches include [1] and [2], whereas [3] adopts a computer science viewpoint. More focused on specific applications, the articles in the special issue [4] and [5] study the interaction between humans and vehicles. Reference [6] surveys systems comprised of humans and robot swarms, and [7] reviews telemanipulation by small unmanned systems. Within the control systems community, significant advances have been made [8]–[11]. It is common to model the effect of humans as a limited actuation resource or unknown deterministic dynamics to be identified. This opens the door to using various robust control and game-theoretic methods to design closed-loop systems. However, most paradigms adopted in the systems control literature to date hide the decision capability of humans and limitations to their cognitive and computational capabilities. Notable exceptions to the literature include [12] and [13], where human decision making is characterized by a partially observed Markov decision process (MDP) [14], with states representing basic physical human aspects, such as operator fatigue or proficiency. Specifically, [13] investigates a stochastic two-player game resulting from the interaction between a human operator and an unmanned aerial vehicle (UAV). Here, the UAV is allowed to react to the decisions made by the operator to compensate for potential noncooperative behavior. For a CPHS with multiple human decision makers, [15] models limitations in their decision-making capabilities using "level-k reasoning" and associated game theory [16]: each agent optimizes a cost function using only reduced knowledge of the other agents' policies.

The present work investigates human decision making from a behavioral economics perspective. The question of how behavioral biases affect decision making is posed, since behavioral influences are expected to potentially hamper technological optimization efforts. More generally, the aim is to direct attention to the human as a poorly observed uncertainty source in CPHSs. Thus, humans are considered as being "bounded rational," that is, they are strategic thinkers who want to maximize their benefits but, at the same time, have only limited cognitive and computational capabilities. We distinguish between closed-loop scenarios (where the decision maker receives feedback about the success of its actions) and open-loop situations (where no such feedback information is available). Special attention is paid to the role played by feedback on behavioral biases in the form of overconfidence [17] and the so-called hot hand fallacy [18]. The latter effect may arise when a decision maker becomes aware of past success and overestimates future success probabilities. Some prominent findings from behavioral economics are introduced next that can be regarded as influential for CPHSs. These will be subsequently applied to an experimental human-in-the-loop framework inspired by the UAV piloting scenario of [13]. This article highlights the value of giving greater consideration to human behavior and behavioral biases in control engineering. The central outcome of the study is that frequent feedback to human decision makers may lead to suboptimal results. This stands in contrast to situations wherein computers carry out optimizations and more information is always beneficial [14], [19], [20]. As will be shown, the situation is more ambiguous when humans act as decision makers.

Summary

The successful design of human-in-the-loop control systems requires appropriate models for human decision makers. While most paradigms adopted in the control systems literature hide the (limited) decision capability of humans, in behavioral economics, individual decision making and optimization processes are well known to be affected by perceptual and behavioral biases. Our goal is to enrich control engineering with insights from behavioral economics research through exposing such biases in control-relevant settings.

This article addresses the following two key questions:

1) How do behavioral biases affect decision making?
2) What is the role played by feedback in human-in-the-loop control systems?

The experimental framework shows how individuals behave when faced with the task of piloting an unmanned aerial vehicle under risk and uncertainty, paralleling a real-world decision-making scenario. The findings support the notion of humans in the underlying behavioral biases of cyberphysical systems, regardless of—or even because of—receiving immediate outcome feedback. Substantial shares of drone controllers act inefficiently by either flying excessively (overconfident) or overly conservatively (underconfident). Furthermore, human controllers self-servingly misinterpret random sequences by being subject to a "hot hand fallacy." Control engineers are advised to mind the human component to not compromise technological accomplishments through human issues.

Humans are considered as being "bounded rational," that is, they are strategic thinkers who want to maximize their benefits but, at the same time, have only limited cognitive and computational capabilities.

CONTRIBUTIONS FROM BEHAVIORAL ECONOMICS

Behavioral economics is an emerging research field concerned with the overarching question of how humans behave in economic decisions, in contrast to how they are prescribed to behave by economic theory [21]. Therefore, traditional economic assumptions and models are revised and enriched by insights from psychology. This serves to improve the understanding and predictability of human behavior, especially regarding recurring errors and biases in decision making [22]. The perception of humans thereby shifts from the assumption of a rationally optimizing Markov decision maker toward regarding humans as rather bounded rational agents. Humans constantly use cognitive mechanisms of simplification, namely, decision heuristics, to process information and make decisions under uncertainty [21], [23]. Such heuristics are used subconsciously due to individuals not managing to adequately process the complexity of a decision problem to account for all relevant information [22]. In fact, when making judgments or estimations of events, frequencies, or probabilities under uncertainty, individuals do not always obey Bayesian rules and statistical logic (as they are meant to do in models assuming perfectly rational agents and Markov decision makers). Such heuristic simplifications sometimes yield reasonable judgments but may also lead to severe and systematic errors [24]–[26]. Even experienced researchers and professionals often have the same judgment heuristics and biases as laypersons [25].

Overconfidence

Overconfidence is well known to be one of multiple behavioral-stylized facts influencing and thereby biasing human decision making. Overconfidence can be defined as a general miscalibration in beliefs [27], more specifically, the discrepancy between confidence and accuracy [17]. In being overconfident, people overestimate their own capabilities, knowledge, or the general favorability of future prospects [28]. The authors of [29] go so far as labeling overconfidence the "most robust finding in the psychology of judgment."

Early contributions to research on overconfidence found that people systematically tended to be unrealistically optimistic about their future, judging themselves to be more likely to experience a variety of positive events and less likely to experience negative events than others. This pattern has been traced back to the degree of desirability influencing the perceived probability of such events [30]. In the original study [31], more than 80% of the subjects regarded themselves as more skillful and less risky car drivers than the average driver. Moreover, half of the subjects estimated themselves to be among the top 20% of the sample, vividly illustrating overconfidence as the discrepancy between confidence and accuracy (belief and reality) in doing so. Meanwhile, evidence on overconfidence does not depend on whether or not experimental subjects are familiar with the given tasks or actions [32]. The task's difficulty, on the other hand, is known to have an impact on overconfidence, as more difficult tasks were shown to facilitate overconfidence than easy tasks [27], [33]. Overconfidence can express itself (and consequently has been studied) in multiple ways, such as "better-than-average" beliefs [32] or overprecision [34]. The facet of overconfidence most fitting for this article appears to be overestimation, described as self-servingly overestimating the likelihood of desirable outcomes, supposedly fueled by wishful thinking for such desirable outcomes to occur (overconfidence is therefore regarded as closely related to self-deception [35]) [36], [37]. An experimental approach for studying overconfidence was requested in [32], as most research in this field merely relies on verbal statements or subjective estimations rather than monitoring human decision behavior. The current analysis makes a contribution to this call. Overconfidence is connected to a variety of other behavioral phenomena, which will be revised in future work.

Underestimation of Systematic Risks

Analogous to overestimating favorable outcomes, overconfidence also implies underestimating risks (or, rather, the variance of risky processes), indicating a too-narrow distribution in one's subjective probability beliefs [38]. Often, humans assess probabilities incorrectly or rather draw incorrect conclusions from them through not adhering to Bayesian rules or neglecting base rates [24], [39]. In doing so, smaller probabilities are usually overestimated, while larger probabilities are underestimated [40], [41]. In contrast, decision makers are also found to underestimate the overall likelihood of the occurrence of risks with small single probabilities of occurrence in some instances. Risks of disqualification [42] through low-probability events generate less concern than their probability warrants, on average [43]. A cognitive process behind this underestimation of especially low probability risks (for example, being involved in car crashes or natural disasters) can be characterized as

Humans constantly use cognitive mechanisms of simplification, namely, decision heuristics, to process information and make decisions under uncertainty.

subconsciously approximating low probabilities with zero to not have to mentally address them [44].

Consequences of this issue can be observed in various areas of economic and social decision making. Overconfident individuals (who underestimate health, financial, or driving risks) are more likely to have insufficiently low insurance coverage against such risks [45]. Furthermore, employees are observed to underestimate the risk of their own company's stock and be overly optimistic about its future performance. They tend to include such stocks too heavily in their retirement plans (despite the respective stock being riskier than the overall market) as a consequence of excessive extrapolation of positive past performance [46]. Money being at stake neither changes overconfident behavior nor leads to better estimation results concerning abilities, probabilities, or risk. Experimental evidence was obtained on overconfidence, leading to overly rushed market entries that often precede business failures due to entrepreneurs overestimating their relative chances to succeed [47]. Overconfident top managers were further shown to underestimate the volatility of cash flows of S&P 500 companies, resulting in erroneous investment and financing decisions, in a large-scale survey over 10 years [38]. Managers competing for leadership positions display overconfidence by taking on riskier projects due to underestimating their risks [48]. Overconfident CEOs also overestimate the positive impact of their leadership and ability to successfully complete a merger to generate future company value. These CEOs are found to execute value-destroying mergers, thereby exposing the affected companies to high financial or even existential risks [49]. Multiple experimental studies found that overconfident investors trade more excessively than others, overestimate the impact of the little information they have [50], [51], and do not adjust their trading volume to new negative information [52]. These findings, which are inconsistent with the Markov assumption and Bayesian updating, contradict theoretical expectations of rationally optimizing economic agents and reemphasize the notion of individuals as bounded rational decision makers [21].

Attribution Theory

Another mechanism contributing to overconfident decision making is a biased perception of causality in a self-serving way. Accordingly, individuals express the tendency to attribute positive past outcomes to themselves and their abilities while blaming negative outcomes on external circumstances [53]. Successes are internalized, and failures are externalized [54]. In an early experiment, teachers attributed the learning success of their pupils to their teaching skills while blaming bad learning performances on the pupils themselves [55]. This self-serving attribution bias was identified to be the major driver of the aforementioned CEO overconfidence leading to risky mergers [56]. In [57], overconfidence evolves when traders get feedback about their ability through experience in a multiperiod market model and overweigh the role of their ability in prior success. Overconfident investors subsequently tend to become even more overconfident in their future investments. Meanwhile, financial losses are attributed to environmental circumstances, such as unfavorable macroeconomic developments or simply bad luck, with the investor's degree of overconfidence remaining constant [58], [59]. Professionals were also found to be as likely as laypersons to express overconfidence in making economic decisions as well as reevaluating the quality of their own previous decisions in hindsight [60].

People overestimate their own capabilities as well as their control over future events that are actually determined by external factors, especially chance. This was illustrated in an experimental study in which participants were either given a lottery ticket (control group) or allowed to select a lottery ticket themselves (treatment group). When asked to name a price they would sell it for, subjects who chose their tickets were, on average, demanding four times more money than those who were simply given their tickets. Those who chose the ticket themselves mistakenly assumed they would be more likely to win, although the winners would be determined entirely by chance. This mechanism has been labeled the "illusion of control" [54], although critics of this concept have argued that it identifies a pattern of people overestimating their ability to predict future outcomes rather than people overestimating their ability to control future outcomes [61]. Both interpretations are regarded as sufficient for the purpose of the current analysis.

The Role of Feedback in Overconfidence

Providing feedback is commonly regarded as an efficient learning mechanism in MDPs to compute optimal behavior [14], [19], [62] and offers great value for deriving logical inference rules in accordance with Bayesian reasoning in machine learning [63], [64]. While these points may hold true for cyberphysical systems without human decision makers, humans are often observed to react to feedback differently. Contradicting Markovian assumptions, behavioral economic research constantly observes errors in updating the information set [65], [66]. This indicates that providing more feedback information to humans does not necessarily lead to better estimations and decisions. The importance of outcome feedback in counteracting certain behavioral biases (through continuous "monitoring of progress through a judgment-action-outcome loop" [67] and offering corrective adjustments) has been presented before [68] yet may not apply equally to every type of bias.

While research studying the connection of high information supply to overconfidence is comparably scarce, in various instances, a larger amount of feedback was not found to improve behavioral rationality with regard to overconfidence, as would be expected by theory. Both overconfidence and underconfidence were shown to persist in employees' choices of incentive schemes, although they received clear feedback revealing the optimal alternative for them [69]. Experiments on sports outcome estimation showed that overconfidence did not decrease but rather increased with additional information, as the subjects' confidence increases more than their accuracy does [70]. Similarly, CEO overconfidence in forecasts was found to persist against corrective feedback [71], and venture capitalists were shown to be more overconfident when having access to more information [72]. In general knowledge tasks that are related to each other, corrective outcome feedback has been shown to mitigate underconfidence but not overconfidence [73]. Feedback appears to affect behavioral biases rather ambiguously, with overconfidence tending to largely persist or grow in response to additional information.

Misperceptions of Random Sequences

With regard to human sequential decision processes, another potential threat of high information supply leading to biased decision making is the danger of falling victim to the hot hand fallacy. The study of this error in the assessment of past outcomes originates in observations from basketball. If players repeatedly scored on their past shots, they are commonly judged to have a "hot hand" by teammates and the audience. Therefore, the player's likelihood of scoring on future shots is judged to increase by the audience, even though future shots are independent of past shots. In doing so, the player's objective probability of scoring is overestimated based on past successes [18]. This hot hand fallacy results from misjudging statistically independent favorable events to be connected to one another, implying a positive autocorrelation between them. Repeated past positive outcomes are therefore erroneously expected to occur again in the near future, thereby overestimating their objective probability of occurrence. The hot hand fallacy is a vivid illustration of biased probability judgments through presumed representativeness of recent information. Misinterpretations of random sequences through such extrapolation of recent outcomes into the future were empirically found to apply to several contexts. In basketball, the findings in [18] were supported by studies on basketball betting market odds, although only small effects were found [74], [75]. In a laboratory experiment simulating a blackjack game, gamblers were observed to bet more money after a series of wins than after a series of losses [76]. In more sensitive contexts, individual decisions also show signs of the hot hand fallacy. For instance, financial investors with negative prior experiences with low stock returns exhibit a decreased willingness to take financial risks in future investments [77]. Attribution theory (especially the illusion of control mentioned earlier) may factor into the hot hand assumption as well, since subjects are more likely to attribute recent sequences with low alternation to human skilled performance, while sequences with high alternation are perceived as chance processes [78].

Before proceeding, note that the hot hand fallacy's opposing bias, the so-called *gambler's fallacy*, presents the erroneous expectation of systematic reversals in stochastic processes. These are grounded in the belief that a small sample should be representative for the underlying population, that is, a Bernoulli random process should balance out, even across a few rounds [79]. Both phenomena have recently been strongly disputed in their status as biases [80], [81].

SURVEILLANCE DRONE PILOTING FRAMEWORK AND HYPOTHESES

Our experimental design for analyzing behavioral biases in CPHSs builds upon a basic setup involving UAVs that interact with human operators, as discussed in [13] and graphically illustrated in Figure 1. Partly autonomous UAVs are used for road network surveillance inside a network that connects multiple traffic junctions. The drone pilot's objective is to gain the maximum information about the traffic at all junctions using a single drone that follows a predesigned path (see Figure 1). To increase the picture quality (and thereby the information obtained), the operator may choose to fly up to eight rounds over each junction j, taking an additional photo each round. The individual photos are combined into a higher-resolution image whose total value will be denoted I_j. This total value depends on the number of flights f_j and the random additional information content of each photo taken. The situation is modeled via the random process $\sigma_j(i)$, which quantifies the information content of the combined photo of junction j at round i. The process is initiated at $\sigma_j(1) = 25$ and, through iteration of

$$\sigma_j(i+1) = \sigma_j(i) + s_j(i)\rho(\sigma_j(i)), \quad i = 1, 2, \ldots, f_j - 1, \tag{1}$$

leads to $I_j = \sigma_j(f_j)$.

Here, the additional information content of the photo taken at round i is quantified by the concave function $\rho(\cdot)$, which describes a decreasing marginal yield

$$\rho(25) = 25, \rho(50) = 20, \rho(70) = 10, \rho(80) = \rho(85) \\ = \rho(90) = \rho(95) = 5. \quad (2)$$

This reflects the fact that each additional picture often comes with less information value added, since a rough image of the traffic flow is already gained by the earlier pictures and subsequent pictures only help to further sharpen the image. In (1), $s_j(i) \in \{0, 1\}$ is an independently identically distributed Bernoulli random process with probability $P[s_j(i) = 1] = p$. This models instances wherein the new picture offers no information value added over the previous one due to, for instance, bad weather causing poor visibility, strong wind, or lacking flow of traffic.

While taking more photos will often lead to combined images of higher quality and with more information content, there exists after each flight a (small) probability r of the drone crashing, leading to a loss of D (the value of the UAV) and the inability to continue flying again over the current or later junctions.

Given these findings, the value of the image obtained at each junction j belongs to the finite set {0, 25, 50, 70, 80, 85, 90, 95, 100}. The current combined value of the drone and images after flying i rounds at junction j satisfies

$$V_j(i) = Dc_j(i) + \sigma_j(i) + \sum_{\ell=1}^{j-1} I_\ell, \quad (3)$$

where

$$c_j(i) = \begin{cases} 1, & \text{if the drone is still intact after flying } i \text{ rounds at junction } j, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

The total value gained by the operator at the end of the mission is

$$V = 400c_{10}(f_{10}) + \sum_{j=1}^{10} I_j. \quad (5)$$

Decision Problem

Given the possibility of the drone crashing, rather than simply intending to fly the maximum number of eight rounds over each junction, the operator is faced with the decision of how many rounds to fly over each junction to maximize its profit V. This amounts to a sequential stopping rule problem defined over a finite horizon. To make decisions of when to fly (or switch) to the next junction (in addition to knowing the system model and its probabilities), the operator receives feedback from the drone (see Figure 2). Every time the UAV passes over a junction, it sends back $\sigma_j(i)$, the information gain obtained so far over the current junction. Further, the operator is informed whether the drone has crashed. Using this information, the

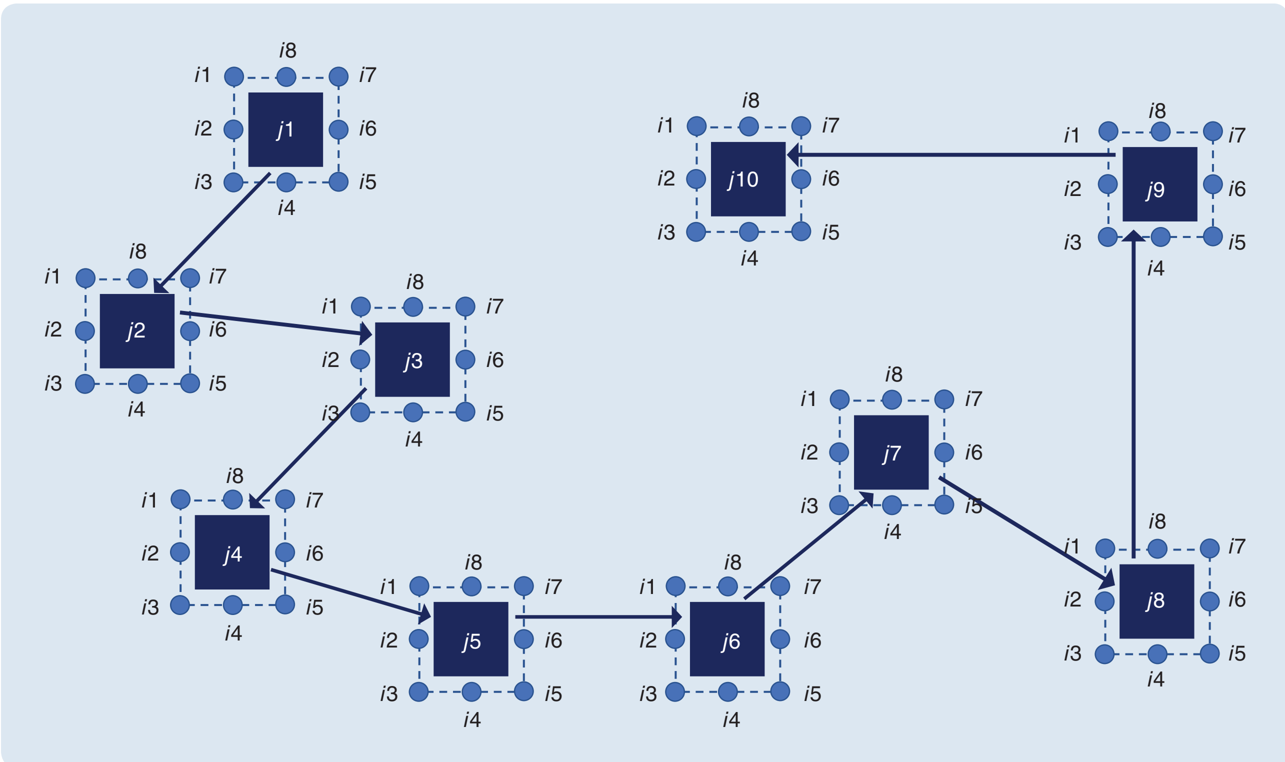

FIGURE 1 An illustration of a road network for unmanned aerial vehicle missions, adapted from [13]. A single drone follows a given path overflying 10 junctions. At each junction $j \in \{1, 2, \ldots 10\}$, a maximum of $N = 8$ rounds can be flown.

human operator is faced with the task of designing a closed-loop flight policy.

The exact stopping rule can, in principle, be derived using dynamic programming [20], [82]. However, the value function depends not only on the value $\sigma_j(i)$ but also the current round i and junction j (with later junctions having less value). Thus, instead of pursuing an optimal strategy, a one-stage look-ahead rule within the current junction becomes a reasonable alternative. Using such a myopic policy, the operator chooses to fly another round over the current junction j if and only if the drone is intact, $(c_j(i) = 1)$, $i \leq 8$, and the marginal gain of flying one more round is positive:

$$g(\sigma) \triangleq E[V_j(i+1) - V_j(i) \,|\, \sigma_j(i) = \sigma] = p\rho(\sigma) - Dr > 0, \quad (6)$$

where (3) is used (see also Figure 3).

The parameters will be fixed to $D = 400$, $p = 0.5$, and $r = 0.02$, so that (2) provides

$$\begin{aligned} &g(25) = 4.5, g(50) = 2, g(70) = -3, \\ &g(80) = g(85) = g(90) = g(95) = -5.5. \end{aligned} \quad (7)$$

This makes flying until an information value $\sigma_j(i) = 70$ is accumulated in the current junction the optimal choice. From the fourth node onward, the marginal gain from

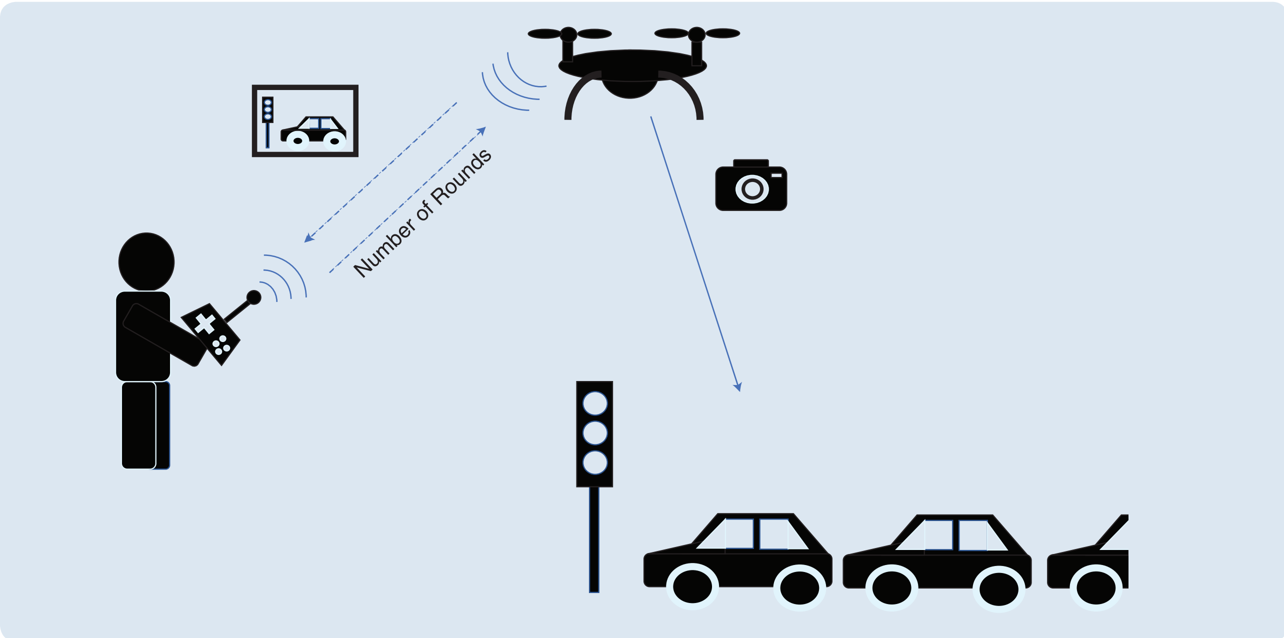

FIGURE 2 The human-in-the-loop control system. The operator decides how many times the unmanned aerial vehicle should fly over each junction to gain accurate information about traffic conditions in the city. To assist in the decision making, the drone may send feedback information about the picture quality, that is, information gained.

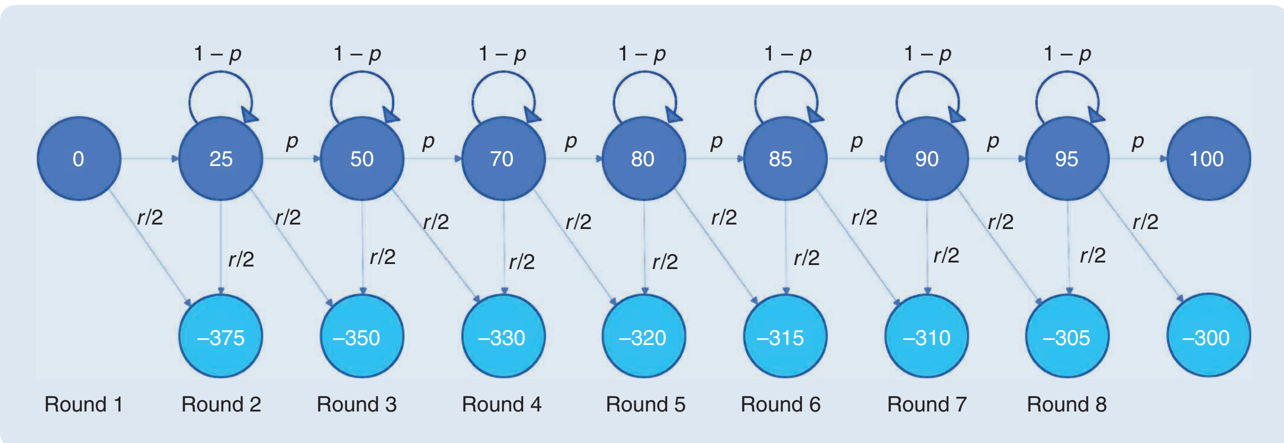

FIGURE 3 The marginal gains within each junction, for $D = 400$. Diagonal arrows describe an increase in information value with a subsequent crash, and vertical arrows represent a nonincrease with a subsequent crash. Note the maximum information value gain attainable at each traffic junction, namely, 100, can only be reached if a subject decides to fly all eight rounds possible while obtaining an increase in each round.

flying another round is lower than the marginal loss. Therefore, a rational (and risk-neutral) agent should refrain from flying further rounds.

In addition to this full feedback case, a situation is considered wherein the drone does not transmit the feedback information $\sigma_j(i)$ to the operator. This requires the operator to design an open-loop policy [19], as detailed later.

Experimental Drone Framework

In our behavioral economic engineering project (see "Behavioral Economics Engineering"), the framework is abstracted

Behavioral Economics Engineering

The deviation between theoretically prescribed behavior and actual human decision making is often labeled "behavioral messiness" [S1] and poses a huge challenge for organizational policy makers. Experiments thereby represent a tool for bridging economic theory and real-world institutional design to generate practical value from applied economic science [S1]. The laboratory basically acts as a "wind tunnel" for practice to test behavioral reactions to institutional interventions in a controlled environment before implementing them into the real world. Projects following this approach are summed under the label of "behavioral economic engineering" [S1]. Prominent examples include the implementation of optimized retirement savings plans [S2] and the redesign of matching algorithms for American physicians [S3].

REFERENCES

[S1] G. E. Bolton and A. Ockenfels, "Behavioral economic engineering," *J. Econ. Psychol.*, vol. 33, no. 3, pp. 665–676, June 2012. doi: 10.1016/j.joep.2011.09.003.

[S2] R. H. Thaler and S. Benartzi, "Save more tomorrow: Using behavioral economics to increase employee saving," *J. Polit. Econ.*, vol. 112, no. S1, pp. 164–187, 2004. doi: 10.1086/380085.

[S3] A. E. Roth and E. Peranson, "The redesign of the matching market for american physicians: Some engineering aspects of economic design," *Amer. Econ. Rev.*, vol. 89, no. 4, pp. 748–780, 1999. doi: 10.1257/aer.89.4.748.

Experimental Research Method and Induced Value Theory

Behavioral economic findings are often based on experimental research, especially controlled laboratory experiments, as "experimental control is exceptionally helpful for distinguishing behavioral explanations from standard ones" [22]. Laboratory experiments, which were considered unfeasible for economic disciplines and the privilege of natural sciences up until the late 20th century [S4], aim to parallel real-world situations in laboratory settings while abstracting from environmental factors. This means experiments isolate certain variables of interest from more complex, real-world contexts while simultaneously controlling for conditions of the subjects' economic and social environments [S5]. Participants, commonly referred to as "subjects," work on computerized or analog tasks, such as solving math problems [S6] or assembling LEGO figures [S7], while anonymously making decisions that are observed by the experimenter. Economic theories or concepts that the experimenter aims to test in the experiment are often in relation to a given task. A constituting factor of experiments is being incentivized, that is, the participants are paid for solving tasks and making decisions by the experimenter, with the concrete amount of payments depending on the respective experimental design [S5]. Subjects know about the payments they will receive beforehand, as they receive instructions containing the experimental procedure and the payoff function at the beginning of every experiment. Herein lies the biggest advantage of experimental economic research: actual human behavior can be observed, with the subject's actions having actual monetary consequences for their payoffs, such that they must "put their money where their mouth is." This is not the case for other research methods used in economic or social science, like surveys or scenario studies, in which participants only state how they would behave in certain situations, while no information is gained whether they would actually behave the way they stated when faced with the decision problem in reality. The problems of intention–behavior gaps [S8] and giving socially desirable answers [S9] are well known.

As a basic principle of economic experiments, there is no deception by the experimenter. Subjects will be asked to do exactly what is stated as their task in the instructions they receive before the experiment, and they will be paid exactly according to the payoff function given in the instructions [S10]. Experiments can be used for various purposes, such as testing economic theories and theoretical equilibria; testing policies and environments; establishing phenomena, stylized facts, and new theories; or deriving political recommendations [S5]. Furthermore, experimental evidence can be replicated and reevaluated by other experimenters using the same experimental setup and instructions, which are usually published alongside the results [S11].

To experimentally test certain theories or economic interventions, subjects are randomly assigned into one control group and (at least) one treatment group to avoid (self-)selection biases, with a treatment being the intervention that will be tested. Differences between the control and treatment groups must thereby be reduced to exactly one factor, namely the so-called treatment variable, that is, the intervention that the experimenter wants to evaluate [S10]. Herein lies another major into an economic experiment (see "Experimental Research Method and Induced Value Theory") to gain insights on actual individual behavior in human-in-the-loop control. The experiment translates the CPHS concepts of open- versus closed-loop control into a sequential decision-making game, framed in a context in which the subjects are owners of a UAV with a photo function. They are told that they were hired by the local city administration to support the city's traffic surveillance by taking photos of 10 important traffic junctions. As outlined previously, the subjects' task is to decide how many rounds the UAV should fly over each of the 10 traffic junctions (up to a maximum of eight times) while the drone autonomously flies and takes pictures. Subjects are incentivized, as they are paid according to the amount of information value gained through the pictures their drones take in the fictitious currency "taler," with one taler equaling one unit of information value gathered. Talers are exchanged into euros after all parts of the experiment are completed. However, during each round in which a UAV flies, it faces a constant crash risk of $r = 2\%$. At the end of this job (that is, after all 10 traffic junctions), subjects are able to sell their drone for an additional fixed payment of $D = 400$ taler, as long as it is still intact [see (5)].

Each traffic junction represents one stage of the experiment. For simplicity, subjects do not have to bear any costs from flying the drone. Further, the drone's value does not diminish from usage, and subjects are told that batteries for the drones as well as (ground) transport between the traffic junctions are provided free of charge by the city administration. One battery charging allows the drone to fly up to eight rounds before it must land to recharge. Therefore, subjects are able to decide for a maximum of eight rounds per traffic junction and consequently take up to eight pictures. This corresponds to a maximum of eight nodes that can be reached per junction. In the first round flown over each traffic junction, an information value of $\sigma_j(1) = 25$ is gained with certainty, as any picture taken represents an increase in information value over no picture taken at all. From the second round onward, the information value potentially gained per round decreases [see (2)]. For simplification, a crash can only happen after a picture is taken. As per (5), the information value gained at the junction where the crash occurs up to the point of it still qualifies for the payment, as subjects were told that the drone's memory chip would survive the crash.

advantage of controlled economic experiments, as they allow causal inferences. Due to the experimental environment being held constant and environmental factors being controlled for all groups, they cancel out through comparative statics between groups. Therefore, the outcome in a treatment group compared to a control group can only be caused by an exogenous change, that is, by the treatment intervention itself, since it represents the only factor that differs between groups [22], [S10], [S12].

As noted, a constituent factor of economic experiments consists of the provision of decision-dependent monetary incentives. Economic experiments underlie the general assumption that any kind of utility an individual experiences can be expressed by an equivalent monetary incentive. Through these incentives, the experimenter is able to induce a certain utility function $U(x)$ onto the subjects to neutralize the subject's inherent preferences for the duration of the experiment. This "induced value theory" was originally introduced by Nobel Memorial Prize in Economic Sciences Laureate Vernon Smith [S13]. According to him, subjects should derive all utility from the monetary incentives provided in the experiment itself during their participation.

While the internal validity of experiments is usually recognized to be very strong, the external validity, that is, the transferability of their results to the world outside the laboratory (the external validity), is commonly discussed and criticized by opponents of this methodology [44], [S10]. For instance, in the context of perceptional and judgmental biases, laboratory settings are accused of changing the environment in which a human makes efficient decisions to an artificial one. However, it can be argued oppositely that laboratory experiments show people at their best by providing all information necessary and eliminating distractions [44]. If people fall victim to biases in this isolated, safe environment, they will likely do so outside of it as well.

REFERENCES

[S4] P. A. Samuelson and W. Nordhaus, *Principles of Economics*, 12th ed. New York: McGraw-Hill, 1985.

[S5] A. E. Roth, "Laboratory experimentation in economics: A methodological overview," *Econ. J.*, vol. 98, no. 393, pp. 974–1031, Dec. 1988. doi: 10.2307/2233717.

[S6] N. Mazar, O. Amir, and D. Ariely, "The dishonesty of honest people: A theory of self-concept maintenance," *J. Marketing Res.*, vol. 45, no. 6, pp. 633–644, 2008. doi: 10.1509/jmkr.45.6.633.

[S7] D. Ariely, E. Kamenica, and D. Prelec, "Man's search for meaning: The case of legos," *J. Econ. Behav. Org.*, vol. 67, nos. 3–4, pp. 671–677, Sept. 2008. doi: 10.1016/j.jebo.2008.01.004.

[S8] M. J. Carrington, B. A. Neville, and G. J. Whitwell, "Why ethical consumers don't walk their talk: Toward a framework for understanding the gap between the ethical purchase intentions and actual buying behaviour of ethically minded consumers," *J. Bus. Ethics*, vol. 97, no. 1, pp. 139–158, Nov. 2010. doi: 10.1007/s10551-010-0501-6.

[S9] A. J. Nederhof, "Methods of coping with social desirability bias: A review," *Euro. J. Soc. Psychol.*, vol. 15, no. 3, pp. 263–280, 1985. doi: 10.1002/ejsp.2420150303.

[S10] J. Weimann and J. Brosig-Koch, *Einführung in die exerimentelle Wirtschaftsforschung*. Berlin: Springer-Verlag, 2019.

[S11] G. Charness, "Laboratory experiments: Challenges and promises: A review of "theory and experiment: What are the questions?" by Vernon Smith," *J. Econ. Behav. Org.*, vol. 73, no. 1, pp. 21–23, 2010. doi: 10.1016/j.jebo.2008.11.005.

[S12] J. Pearl, "Causality: Models, reasoning, and inference," *Econ. Theory*, vol. 19, no. 4, pp. 675–682, Aug. 2003. doi: 10.1017/S0266466603004110.

[S13] V. L. Smith, "Experimental economics: Induced value theory," *Amer. Econ. Rev.*, vol. 66, pp. 274–279, Feb. 1976.

The operators must decide how many rounds a UAV should fly over each traffic junction to maximize the total gain V in (5) and consequent payoff. To accomplish this, some receive feedback on the results of the previous rounds—depending on the treatment—which they can base their decision on. As previously noted, the resulting optimization problem belongs to the class of general stopping problems. Its solution (both with and without feedback information) can, in principle, be derived but requires careful computations. Since individuals are well known to have limited cognitive and computational capabilities, they will not be able to compute a globally optimal strategy. Instead, optimizing subjects will use the decision heuristic in (6), that is, the fourth node is identified as the one to reach.

These findings lead to two decision heuristics, depending on whether feedback information is available or not. In the former (*closed-loop heuristic*), the UAV should aim to fly as many rounds as needed at every junction until reaching an information value of $\sigma_j(i) = 70$. To be more precise, introduce $a_j(i) \in \{0, 1\}$, where $a_j(i) = 1$ corresponds to the decision to fly another round over the current junction. The closed-loop heuristic can be stated as

$$a_j(i) = 1, \quad \text{if and only if } \sigma_j(i) < 70,\ i \leq 8, \text{ and the drone is intact.} \tag{8}$$

Note that in this case, the total number of rounds flown over a junction j, namely, $f_j = \Sigma a_j(i)$, depends on the sequence of increases and nonincreases the subject experiences.

In the absence of feedback from the drone, a suitable *open-loop heuristic* amounts to attempting to fly $f_j = 5$ times at every junction:

$$a_j(i) = 1, \quad \text{if and only if } i \leq 4 \text{ and the drone is intact.} \tag{9}$$

This yields the expected result of three increases in information value, with the first being granted with certainty in the first round. These rules of thumb present reasonable heuristic approximations to the task's optimal solution when implemented for all junctions. A rational and risk-neutral decision maker, who acts according to expected utility theory [83], would follow these heuristics. Deviations from these decision rules upward or downward represent indicators of overconfidence or underconfidence, respectively, depending on the direction.

The experiment ends either after all junctions are completed and the drone is sold or once the drone has crashed. A crash could be caused by external factors, like weather conditions, the drone hitting some object or animal, or technical failure. This risk of a total loss of the drone's value due to a crash imposes the prospect of high one-time costs onto the participants and, thus, prevents them from simply choosing the maximum number of rounds possible at every traffic junction. As subjects are paid according to the information value gained, striving for the optimal information value (using the just-described heuristic) presents the rational strategy. Always flying the maximum number of eight rounds is not efficient, as gaining information value has to be weighed against maintaining the chance of being able to sell the drone at the end of the experiment, as indicated in (5).

Hypotheses

The variety of behavioral phenomena and stylized facts regarding decision biases and errors in relation to overconfidence presented in the preceding section suggests inefficient human decision making. In fact, individuals are commonly observed to act overconfidently. In doing so, they tend to underestimate systematic risks, insufficiently process joint and conditional probabilities, self-servingly misattribute causal relationships, and erroneously expect favorable outcome sequences to continue in the future. These aspects should be particularly relevant in our example, since the probability of the drone crashing may appear small for one specific round (2%), while the overall risk of the UAV crashing at any time over the course of the experiment is substantially higher. Modeling the crash probability as a Bernoulli process, the chance of a crash when attempting to fly the maximum number of rounds at each of the 10 junctions would be $1 - 0.98^{80} > 0.8$. Therefore, a narrow focus on the low single-decision crash probability [84] could lead subjects to underestimate the overall crash probability and consequently take higher risk. Providing feedback is commonly assumed to diminish these behavioral biases but known to not entirely resolve them, even aggravating them in some instances. Subjects may rather attribute positive feedback to their own decision performance instead of realizing that it is caused by chance. All of these aspects point in the direction of expecting overconfident drone piloting. Manager overconfidence in economic decision making can be summarized to originate in either overestimating expected cash flows or underestimating risk [85]. Translated to our design, the drone pilots' overconfidence could either result from overestimating the likelihood of information value improvements or underestimating the risk of the UAV crashing. While the task is framed in a quite understandable context, it is nontrivial in its solution. Most subjects are expected to approach the task rather intuitively, that is, not solving it systematically through comparing marginal gain and marginal costs, and they consequently act overconfidently through one of its facets described previously, if not through both. According to the general consensus of literature on overconfidence in human decision making, the following hypothesis is posed.

Hypothesis 1

Drone pilots will generally act overconfidently in flying more than the optimal number of rounds, regardless of feedback.

As is also known from the literature, individuals are prone to unreflectively repeat decisions that previously yielded them positive outcomes, even though such past successes depended entirely on chance and have no impact on future outcomes. This decision heuristic, the hot hand fallacy previously

Sequences of immediate positive feedback will lead to drone pilots falling victim to the hot hand fallacy.

discussed, presents a major misinterpretation of random sequences. While the hot hand fallacy in the context of the original study on basketball shots [18] is subject to criticism today (with the data being reanalyzed and the conclusion being reversed [80]), the general idea of individuals misjudging the meaning of random outcome sequences remains to be tested in the context of CPHSs. For sequential decision problems featuring feedback from the system, decision makers appear especially vulnerable to the hot hand fallacy. Translated to our drone framework, it is therefore expected that those drone pilots who immediately reach the optimal information value of a junction through experiencing gains in repeated rounds in the closed-loop system (as the possibility is only given there) perceive themselves to be "on a roll," that is, as having the equivalent of a hot hand. Our experiment defined a hot hand as three consequent increases from the start of a junction. With the junctions being distinct from each other and the first success being certain anyway, there is no possibility of pattern overlays [81]. Thus, including the outcome of a fourth flight as unimportant for judging the decision inefficient, we do not consider the results subject to "streak selection bias" [80]. Since decision makers are tempted to expect such apparent trends to continue, they will likely fly beyond the optimal number of rounds, overconfidently making their operation inefficient with regards to risk and reward. In this way, subjects fall victim to the hot hand fallacy. The literature observes a shift from hot hand beliefs for shorter outcome streaks toward a gambler's fallacy pattern for increasing streak length [79], [86]. Drone controllers who experience multiple increases in a row would therefore begin to overestimate the probability of a nonincrease at some point and stop flying to prevent a crash. Compared to the gambler's fallacy, a hot hand is considered more suitable to portray a drone controller's behavior in an attempt to gather information value, with humans more prone to expect favorable outcomes to repeat under their apparent control [78]. A hot hand fallacy can also be considered the more problematic error than the gambler's fallacy since, in the case of the latter, subjects would stop earlier and not risk a drone crash through flying in periods that yield marginal losses (although sacrificing marginal gain).

Hypothesis 2

Sequences of immediate positive feedback will lead to drone pilots falling victim to the hot hand fallacy.

These two hypotheses were tested in an experiment involving students at Paderborn University in September 2019. Details are given next.

EXPERIMENTAL DESIGN

To compare behavioral effects in closed- and open-loop operation, a between-subject design experiment is conducted, meaning each subject can only participate in exactly one treatment, with the groups being compared afterward.

Treatments

To evaluate the effect of feedback in human-in-the-loop control, two treatments are presented: closed loop and open loop. The groups differ only in the fact that subjects in the former receive immediate outcome feedback after every decision made, while subjects in the latter do not receive any feedback until the end of the experiment.

In the *closed-loop treatment*, subjects receive feedback directly after each round of each junction on whether an increase in information value was gained in this specific round (see Figure 3). They were also informed about the current total information value gained for the respective junction and whether the drone was still intact. Subsequently, they were asked whether they wanted to fly another round. The framing in the instructions said that the drone would transmit the pictures taken to the drone pilot's laptops immediately, so feedback was given just in time for the next round's decision.

In the *open-loop treatment*, subjects are provided no feedback. Subjects make decisions on how many rounds the drone should fly over each junction in advance. Afterward, they are informed about the total information value they obtained and whether the drone is still intact and sold at the end of the experiment. The framing in the instructions states that the UAV must be deconstructed in a complex procedure to extract and read out the memory chip to review the pictures taken. Therefore, no feedback on the pictures would be provided until either all junctions were completed or the drone crashed.

Procedure

The experiment was computerized using the experiment software oTree [87] and hosted centrally on a university server, allowing subjects to remotely access the experiment and not have to physically leave the laboratory. A total of 500 subjects who previously enrolled voluntarily into the BaER-Lab student participant pool were chosen randomly, with 250 being randomly assigned to the two treatments in advance. Note that BaER-Lab stands for "Business and Economic Research Laboratory." For further information, visit www.baer-lab.org. These subjects were contacted via the online recruitment system ORSEE [88] and invited to participate within the following five days.

The invitation e-mail included a hyperlink that directed the subjects to their respective treatment, where they received detailed instructions for the experiment. The instructions of both treatments only differed regarding the provision of feedback, as explained earlier. To progress to the drone flying task, subjects had to correctly solve four control questions. The questions revolved around the central parameters of the experimental design to assure that the subjects read and understood the critical parts of the instructions. Subsequently, subjects advanced to the drone flying task, in which they had to make decisions by clicking single-choice buttons. Subjects would receive one unit of taler for each unit of information value gained as pay for flying the drone. During the experiment, one subject's payoffs did not depend on the decisions of any others. After completing the drone flying task, a results screen presented the total information value accrued over all traffic junctions, the state of the drone (that is, whether it was still intact and consequently sold to earn 400 additional talers or not), and the corresponding payoffs the subject generated in Euros. Payoffs were translated from taler to euro at an exchange rate of €1 per 120 talers. Subjects were then asked to answer a standardized questionnaire, which included questions on demographics, perceived task difficulty, and the subject's reasons for choosing the number of rounds completed. Afterward, subjects were presented with the multiple price list of [89] (see "Multiple Price List Format") to measure their risk preferences.

Completing the price list was incentivized, and for every 15th subject, one out of the list's 20 rows was randomly selected, and the payoffs that the subject's chosen alternative row yielded was added to their total payoffs. If Alternative A was selected by the subject, the fixed amount stated in the respective row was simply added to its payoffs. If Alternative B was selected, the lottery was automatically completed and either €0 or €30 was added to the subject's payoffs. After completing the price list, subjects were informed of whether the list came into effect for their payoffs, and, if so, which row was selected. On a final screen, subjects were informed of their total payoffs in euros and thanked for their participation.

A subject's total payoff function can consequently be formalized as

$$\Delta m = \frac{1}{120} V + d(\ell)_{i=15x}, \tag{10}$$

where V is the total value, as in (5). The term $d(\ell)$ represents the additional payoff from the multiple price list, conditional on the selected row ℓ of the list that a subject was paid for being the 15th participant, that is, a subjects participant's identifier (ID) was a multiple of 15.

Subjects were able to collect their payoffs in cash at the chair's secretariat following their participation by stating a unique eight-figure ID code they created at the beginning of the experiment. This allowed the correct payoffs to be dispersed to each subject while maintaining anonymity about the subject's decisions.

EXPERIMENTAL RESULTS

Out of the 105 subjects who completed the online experiment, 57 participated in the closed-loop and 48 in the open-loop treatment, respectively. The average age of the subjects was 23.87 years and varied by approximately 1.5 years between treatments. Of the total number of subjects, 31 (29.52%) were female, while one subject classified themselves as "diverse." On average, subjects earned €4.89 from the experiment, plus the amount they would earn from completing the multiple price list (in the event they were paid due to being a 15th participant).

Subjects in the closed-loop treatment flew, on average, 5.89 rounds over each traffic junction, thereby exceeding the average number of rounds from the open-loop treatment by 0.79 rounds per junction. Over all rounds flown, subjects in the closed-loop treatment gained a cumulative information value of 651.84, on average. This represents a 27.83% surplus compared to the average information value gained in the open-loop treatment. Similarly, a higher drone crash rate of nearly 70% was observed in the open-loop group. These two dimensions generally interact (since the information value is fixed) once a drone crashes. Averages of rounds flown and accumulated information value as well as average earnings and crash rates for each treatment are displayed in Table 1.

As expected from control theoretic results [19], the closed-loop system is clearly observed to be more effective than the open-loop piloting system with regard to accumulating information about traffic conditions in our framework. Considerably more information value was aggregated with a closed loop of immediate outcome feedback being implemented, which also translates to a higher monetary payoff for the subjects. However, subjects displayed behavioral biases, as will be shown. Indeed, the average number of rounds in the closed-loop treatment being nearly six already hints that several subjects tended to fly beyond the heuristically optimal number of rounds. In doing so, they would have taken inefficiently high risks, since marginal losses exceeded marginal gains in later rounds of each junction. For instance, instead of aiming to add five or 10 more units of information value when already sitting at 70, subjects should stop and continue with the next traffic junction to obtain 25 units of information value with certainty at the same crash risk.

Overconfidence and Underconfidence

To determine whether a subject acts overconfidently, optimally, or even underconfidently at a certain traffic junction, the number of rounds flown is compared to the heuristics discussed before [see (6)]. If the subject flew beyond the heuristically optimal number of rounds, the decision at this junction is noted as overconfident. If the subject decides to fly fewer rounds than optimal, the junction's decision

Multiple Price List Format

Multiple price list formats offer a simple but effective tool to reveal subjects' actual risk preferences. The basic concept was designed with the goal of testing the standard economic assumption of risk neutrality as well as the behavioral assumption of constant relative risk aversion. The list consisted of a table with two columns that represented two lotteries the subjects had to choose between for 10 rows, featuring varying probabilities of winning certain amounts of money that remained constant for all rows [S14]. As the original price list contained monetary values with two decimals and varying probabilities to be weighed, it can be assumed that it was difficult for subjects to compare the different options. The subsequent work [89] designed a more clear-cut version of the multiple price list (see Table S1), in which subjects had to choose between a safe payoff that increased by €1 with each row and a fair lottery that remained constant for all rows.

Subjects are asked to indicate for each row whether they prefer the offered safe payment (Option A) or the offered fair lottery (Option B). The safe payment equals the number of the respective row minus one (in euros), starting at €0 in the first row and increasing up to €19 in the 20th row. The lottery is the same in each row, containing a 50% chance to win €30 and a 50% chance of winning nothing. A subject's risk attitude can be determined by observing the row in which the choice changes from B to A. When continuously choosing B until row 16 and continuously choosing A afterward, a subject is classified as risk neutral, as both options yield the exact same expected payoffs of €15 in row 16. At this point, a risk-neutral subject is indifferent to receiving the safe payment and playing the lottery. If the change from B to A occurs before the 16th row, then a subject is classified as risk averse, as he or she willing to sacrificing the chance to play a lottery that yields a higher expected payoff in favor of a safe payment of a lower amount. If the change occurs after the 16th row, a subject is regarded as risk seeking. Contrary to risk-averse subjects, risk-seeking subjects are willing to sacrifice a safe payment to play a lottery that yields an expected payoff that is lower than the safe payment but yields a higher maximum payoff.

After making the choices for every row, a fraction of the participant group (for example, every seventh [89] or 15th [S15] subject) is chosen at random and paid for one randomly chosen row of the table, according to his or her respective choice made in this row. This procedure incentivizes subjects to choose in accordance with his or her true preferences in each row, since they would forfeit expected utility if he or she did not, thereby revealing them to the experimenter [89]. Therefore, the multiple price list format presents an incentive-compatible instrument to measure individual risk preferences.

TABLE S1 The multiple price list from [89].

	Option A	Option B
1)	€0 safe	€30 with a probability of 50%, €0 with a probability of 50%
2)	€1 safe	€30 with a probability of 50%, €0 with a probability of 50%
3)	€2 safe	€30 with a probability of 50%, €0 with a probability of 50%
4)	€3 safe	€30 with a probability of 50%, €0 with a probability of 50%
⋮	⋮	⋮
19)	€18 safe	€30 with a probability of 50%, €0 with a probability of 50%
20)	€19 safe	€30 with a probability of 50%, €0 with a probability of 50%

REFERENCES

[S14] C. A. Holt and S. K. Laury, "Risk aversion and incentive effects," *Amer. Economic Rev.*, vol. 92, no. 5, pp. 1644–1655, Dec. 2002. doi: 10.1257/000282802762024700

[S15] B. M. Djawadi and R. Fahr, "The impact of risk perception and risk attitudes on corrupt behavior: Evidence from a petty corruption experiment," IZA—Institute of Labor Economics, Bonn, Germany, IZA Discussion Paper No. 7383, 2013.

indicates underconfidence. If the heuristic is met, the subject acts as optimizing.

As noted before, flying until the fourth node is reached (associated with an information value of 70) is a suitable heuristic for a risk-neutral decision maker. This induces different strategies for the treatments. In the closed-loop system, it is optimal for the subjects to fly as many rounds as necessary to reach the fourth node. The actual number of rounds needed depends on the sequence of increases and nonincreases experienced, as illustrated in Figure 3. In the open-loop system, flying five rounds is a suitable heuristic. As subjects must decide on the number of rounds to fly for all traffic junctions in advance, this strategy applies to all of them, as explained before.

TABLE 1 The average values of rounds flown, gained information value V, earnings Δm, and crash rates, by treatment. Standard deviations are presented in parentheses.

	Closed Loop	Open Loop
Average rounds flown	5.89 (2.08)	5.10 (1.82)
Average information value gained in total	651.84 (397.56)	509.90 (380.13)
Average earnings, in Euros	5.43 (3.31)	4.30 (3.17)
Crash rate	52.63%	68.75%

Overconfidence characterizes the predominating behavioral tendency in the closed-loop system.

An indicator of overconfidence at the subject level is created by calculating the quotient of the number of overconfident junctions and the total number of junctions per subject (since the number of traffic junctions played differed between subjects due to some drones crashing prematurely). This relative frequency of overconfident junctions by a subject will be labeled "overconfidence degree." Degrees of underconfidence and optimizing behavior are computed analogously, such that all three degrees sum to one. The average degrees of each of these behavioral tendencies are displayed in Table 2.

The average overconfidence degree of 0.44 in the closed-loop system means that subjects, on average, decided overconfidently for 44% of the traffic junctions. Comparatively, overconfidence degrees barely differ between the two treatments, with subjects deciding overconfidently for more than 42% of the junctions. The average underconfidence degree was higher in the open-loop treatment. Degrees of overconfidence and underconfidence are relatively close to each other for the open loop, while overconfidence characterizes the predominating behavioral tendency in the closed-loop system. The optimization degree drops off in both treatments, with subjects only deciding optimally at one-fourth and one-fifth of the junctions, respectively. This gap also becomes apparent in the graphic illustration of the average degrees of overconfidence, underconfidence, and optimization displayed in Figure 4. Degrees of overconfidence did not differ significantly by gender in either treatment.

TABLE 2 **The average shares of overconfident, underconfident, and optimal decisions. Standard deviations are presented in parentheses.**

	Closed Loop	Open Loop
Overconfidence degree	0.4410 (.3628)	0.4167 (.3497)
Underconfidence degree	0.3091 (.3099)	0.3833 (.3309)
Optimizing degree	0.2521 (.2581)	0.200 (.2388)

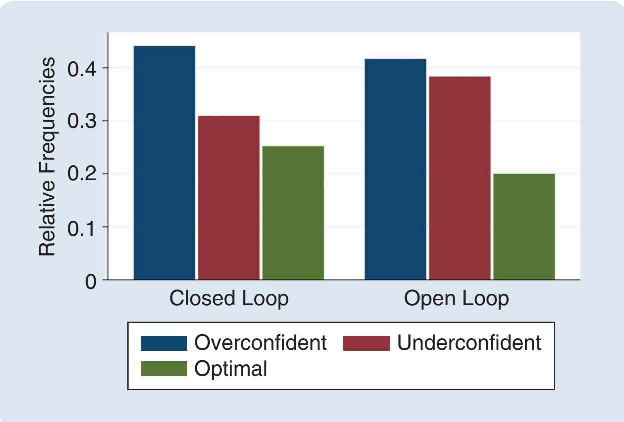

FIGURE 4 The average shares of overconfident, underconfident, and optimal decisions.

Seven subjects in the closed-loop group and eight subjects in the open-loop group behaved overconfidently in every phase they flew in, resulting in the maximum overconfidence degree of one. Alternately, only three subjects in the closed-loop system and two in the open-loop system decided optimally for every junction, while three subjects in each treatment acted underconfidently at every junction. Among those junctions in which subjects act overconfidently, they flew, on average, 2.88 [standard deviation (SD) = 1.36] more rounds after reaching the optimal node in the closed-loop treatment. In the open-loop group, subjects flew, on average, 1.88 (SD = 0.85) rounds beyond the optimum. It appears that those subjects who decide to fly more rounds than optimal do it noticeably, that is, usually by more than one round. Also, subjects exceed the optimal number of rounds by, on average, one more round in the closed-loop system compared to the open-loop system once they fly beyond the optimum. The difference is statistically significant in a Mann–Whitney U-test ($Z = 7.393$, $p = 0.000$) and match the observation of subjects in the closed-loop treatment, flying more rounds overall and accruing more information value in total.

To gain a more precise impression about the distribution of subjects' deviations from optimal behavior within each treatment, we classify all subjects into broad categories. Subjects that act in accordance with the strategy of flying until reaching the fourth node at all traffic junctions are classified as *optimizing*. A subject whose overconfidence degree exceeds one-third up to 0.5 is classified as *rather overconfident*. Once a subject flies beyond the optimal number of rounds in more than half of the junctions (that is, displaying an overconfidence degree of more than 0.5), this subject is considered *strongly overconfident*. Analogously, the categories *rather underconfident* and *strongly underconfident* are defined using the same thresholds in terms of underconfidence degrees as the ones for overconfidence. The sixth category, *mixed*, defines scenarios in which subjects show overconfident or underconfident behavior in at least one junction but in fewer than one-third of all junctions they played (equaling overconfidence or underconfidence degrees between 0 and 0.333). Since these subjects act optimizing in some junctions (but not in all), their behavior still contradicts the notion of a strictly optimizing agent, although no clear dominant decision pattern can be assigned to this behavior. The distribution of subjects into these categories by treatment is displayed in Table 3. Two subjects from the closed-loop treatment were excluded because they crashed in the first round at the first traffic junction, wherefore no statement on their decision-making behavior can be made.

According to Table 3, only a few subjects in each treatment act optimally throughout. Rather, individual behavior deviates from the optimal number of rounds in both directions. Approximately half of the subjects in both groups (closed loop: 51.79%, open loop: 46.81%) can be considered rather overconfident or strongly overconfident. Additionally, large shares of rather underconfident or strongly underconfident individuals (closed loop: 37.50%, open loop: 44.68%) are observed as well. From the assumptions of MDPs and standard economic theory follow notions of individuals as rationally optimizing decision makers that leave no room for overconfident or underconfident behavior. Consequently, the empirical results do not meet these theoretical propositions. In fact, according to a binomial test of the data for the closed-loop treatment, the probability of observing 29 overconfident (rather or strongly overconfident) subjects out of a sample of 56 is virtually zero ($p < 0.000001$), given the assumption of optimizing behavior. The same result is obtained for overconfidence in the open-loop systems. Consequently, we can fully support Hypothesis 1, indicating a strong general tendency of individuals to act overconfidently in CPHSs, regardless of whether receiving immediate outcome feedback or not. While not part of the hypothesis, we also observe the equivalent results for underconfident behavior in both treatments.

Hot Hand Fallacy

In contrast to an open loop, only a closed-loop system provides the threat of sequences of immediate outcome feedback luring subjects into a hot hand fallacy. In the experiment, a subject in the closed-loop treatment is defined to have a "hot hand" once it reaches the optimal node through obtaining three increases in information value in a row, even though the first increase was certain. This corresponds to the common perception of three repeated outcomes as a streak [90]. In this case, an optimizing subject would stop flying when following the one-stage optimal strategy, since three increases are necessary to reach the optimum and flying beyond the optimum yields a negative marginal gain. The outcome of prior rounds should have no relevance for the subject's decision, since the rounds' outcomes are independent. If a subject decides to continue flying, it violates the heuristically optimal decision rule, falling victim to the hot hand fallacy. At 84 junctions, the respective subject experienced three continuous increases in information. In 70 of these 84 cases (83.33%), subjects decided to fly at least one more round, thereby falling victim to the hot hand fallacy. This share of hot hand fallacies among situations that pose the threat of falling victim to it is highly statistically significant: under the assumption of subjects being optimizing Markov decision makers, the probability of such a high proportion disobeying the optimal strategy is practically zero [binomial test: Pr(all subjects act optimally | 83.33% of the hot hand situations result in the hot hand fallacy): $p < 0.0000001$]. Comparing nonoverconfident and overconfident decision making between junctions with and without a hot hand in Table 4, those subjects who experience a hot hand situation (70 out of 84, 83.33%) have a significantly higher chance to act overconfidently at that junction than those subjects who do not reach the optimum through three consecutive gains in information value (123 out of 260, 47.31%). Running a chi-squared test to test the hypothesis statistically, a highly significant relationship is obtained between hot hand situations and subsequent overconfident behavior, that is, the hot hand fallacy, with an error probability p of virtually zero $[\chi^2(df = 1){:}33.46, p = 0.000]$. Consequently, Hypothesis 2 is strongly supported.

Risk Preferences

Subjects' individual risk preferences are broadly classified to be risk neutral, risk averse, or risk seeking (see [89] and "Risk Preferences"). Given that the presented strategy is based on the assumption of a risk-neutral decision

TABLE 3 **The categorization of decision behavior, by treatment. Relative shares referring to each treatment are presented in parentheses.**

	Closed Loop	Open Loop	Total
Optimal	3 (5.36%)	2 (4.26%)	5 (4.85%)
Rather overconfident	5 (8.93%)	6 (12.77%)	10 (10.68%)
Strongly overconfident	24 (42.86%)	16 (34.04%)	40 (38.83%)
Rather underconfident	8 (14.29%)	5 (10.64%)	13 (12.62%)
Strongly underconfident	13 (23.21%)	16 (34.04%)	29 (28.16%)
Mixed	3 (5.36%)	2 (4.26%)	5 (4.85%)
Total	56 (100%)	47 (100%)	103 (100%)

TABLE 4 **The hot hand fallacy–chi-squared table. The test captures every instance in which a subject started flying at any traffic junction in the closed-loop treatment.**

	Not Overconfident	Overconfident	Total
No hot hand	137	123	260
Hot hand	14	70	84
Total	151	193	344

maker, the observed distribution of risk preferences in each treatment is displayed in Table 5. Risk preferences of 15 subjects in total could not be determined, since their choices switched between Option A and Option B three times or more. Consistent with the literature, the clear majority of approximately two-thirds of the subjects in both treatments display risk aversion. Overall, the distribution of risk preferences appears very similar between the treatments, with the relative shares of each risk attitude varying only slightly between the groups. This impression is supported by a two-sided Kolmogorov–Smirnov test, in which the null hypothesis that the risk preference distribution in the two treatments is not statistically different from each other and cannot be rejected ($D = 0.030$, $p > 0.1$). In addition, degrees of overconfidence and underconfidence did not differ significantly in between-treatment comparisons for each risk attitude using Mann–Whitney U-tests.

Further, risk-seeking subjects were not observed displaying significantly more overconfident behavior (open loop: $\varnothing = 0.5257$, closed loop: $\varnothing = 0.175$) than risk-neutral (open loop: $\varnothing = 0.4333$, closed loop: $\varnothing = 0.700$) or risk-averse (open loop: $\varnothing = 0.4140$, closed loop: $\varnothing = 0.3818$) individuals in either treatment (Kruskal–Wallis equality of populations tests, closed: $H = 0.450$, $p = 0.7984$, open: $H = 5.437$, $p = 0.0660$). Note that statistical tests on risk preferences must be interpreted with caution, since the sample size of risk-neutral and risk-seeking individuals is extremely low. Alternately, in the closed-loop system, risk-averse subjects were significantly more underconfident than the risk-neutral and risk-seeking subjects ($\varnothing$ underconfidence degree: 0.3480 (SD = 0.3018), Kruskal–Wallis equality of populations test: $H = 8.891$, $p = 0.0117$), while no effect was found in the open-loop system.

Besides overconfidence, risk-seeking preferences could (in theory) present an alternative explanation for drone pilots flying beyond the optimal number of rounds, with subjects primarily aiming to achieve a large information value while hoping their drones don't crash. However, substantial differences were not found in overconfidence, conditional on risk attitude between treatments, and overall mainly observe individuals who can be classified as risk averse. This would theoretically predict that subjects decide rather conservatively through settling for a lower number of rounds to remain in the game and protect the drone from crashing, even at the cost of missing higher payments

Risk Preferences

Individual risk preferences represent a factor commonly influencing human behavior that should necessarily be considered to explain individual decision making under risk and uncertainty. Human risk preferences are generally differentiated into risk-averse, risk-neutral, and risk-seeking behavior. They are usually determined by the individual's subjective evaluation of a probabilistic payment from a lottery compared to a certainty equivalent, that is, a safe fix payment [89], [S14]. If an individual subjectively judges the utility from the expected value E of a payment X as higher than the expected utility of such payment, namely, $u(X)$, then this individual is classified as risk averse [S16]:

$$u(E[X]) > E[u(X)]$$

As a consequence of Jensen's inequality, a risk-averse individual's personal utility function is concavely shaped. This means that the individual is willing to sacrifice the chance of higher, but uncertain, payments to receive a smaller, but certain, payment (the certainty equivalent), thereby paying the so-called insurance premium [S16]. In an example by [S17] displayed in Figure S1, individuals with €10 at their disposal must decide whether or not to participate in a lottery in which €5 are gained with a probability of 50% and €5 are lost with a probability of 50%. For a risk-averse individual, the utility of the expected value $u(10)$ is greater than the expected utility of wealth $0.5u(15) + 0.5u(5)$.

In contrast, if an individual derives less utility from the expected value of X than X's expected utility,

$$u(E[X]) < E[u(X)],$$

then this individual is considered risk seeking. Its subjective utility function is convex (see the analogous example in Figure S2). Risk-seeking individuals are willing to give up the certainty equivalent to get the chance to play the lottery, through which they may obtain a higher payment, that is, they favor gambling. The payment obtainable beyond the certainty equivalent through playing the lottery is called a risk premium. If an individual is indifferent between both alternatives, the individual can be considered risk neutral, with risk and insurance premium canceling out [S16].

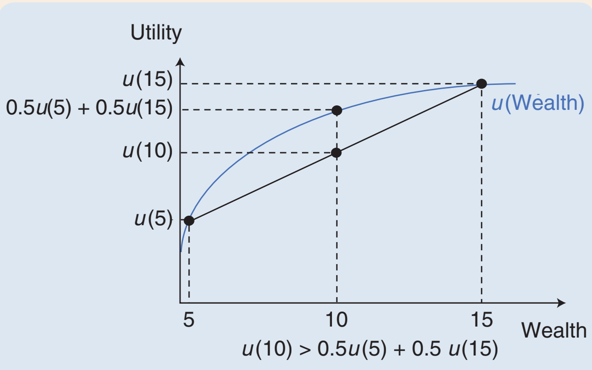

FIGURE S1 An example for a risk-averse individual's utility function (adapted from [106]).

Risk preferences do not explain overconfident behavior, although it appears to explain the observed degrees of underconfidence.

from information value gains. Since the experimental results display a large share of subjects flying beyond the optimum, the case for individual overconfidence as the dominant explanation for inefficient drone piloting beyond the optimum is strengthened. Consequently, we conclude risk preferences do not explain overconfident behavior, although it appears to explain the observed degrees of underconfidence.

DISCUSSION AND CONCLUSION

The experimental drone framework demonstrates how individuals behave when faced with the task of piloting a UAV under risk and uncertainty, paralleling a real-world decision problem. Even though a closed-feedback loop was identified to be the more successful system than open-loop operation, inefficient drone piloting was still observed from the vast majority of subjects. Individuals expressed both overconfident and underconfident behavioral tendencies, regardless of receiving immediate outcome feedback. Specifically, results indicate that immediate outcome feedback, originally intended to support optimal decision making, can be rather counterproductive in

TABLE 5 The distribution of risk categories, by treatment. Percentage values in parentheses display relative frequencies of risk preferences in the respective treatment.

	Closed Loop	Open Loop	Total
Risk averse	36 (63.16%)	33 (68.75%)	69 (65.71%)
Risk neutral	6 (10.53%)	5 (10.42%)	11 (10.48%)
Risk seeking	6 (10.53%)	4 (8.33%)	10 (9.52%)
Not identifiable	9 (15.79%)	6 (12.50%)	15 (14.29%)
Total	57 (100%)	48 (100%)	105 (100%)

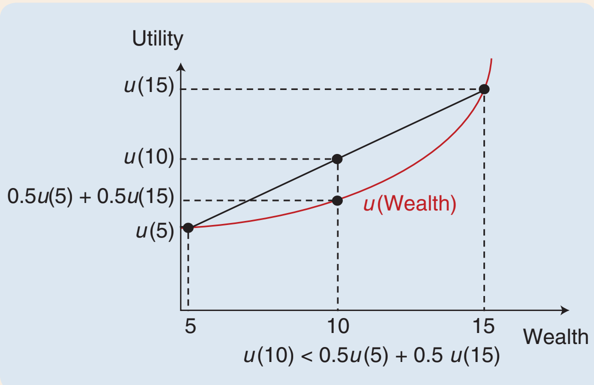

FIGURE S2 An example for a risk-seeking individual's utility function (adapted from [106]).

Standard economic theories, in the tradition of the *homo economicus*, are usually based upon the explicit or implicit assumption of individual risk neutrality, reflecting the notion of individuals as rationally optimizing ("Markovian") decision makers [S18]. On the contrary, behavioral economic research usually observes risk-averse decision makers in the laboratory [S19], [S20] as well as in the field [S21].

The degree of human risk aversion was found to increase with the size of incentives at stake [S14], [S22], contradicting theories of constant relative risk aversion [S14]. However, individuals are observed to act risk averse, even with only low incentives at stake, wherefore models that assume risk-neutral human behavior do not adequately represent reality and may lead to biased inference [S14]. Regardless of this clear evidence, only sporadic effort exists to incorporate risk preferences other than risk neutrality into Markov decision processes [S23], [S24].

REFERENCES

[S16] J. W. Pratt, "Risk aversion in the small and in the large," *Econometrica*, vol. 32, no. 1/2, pp. 122–136, 1964. doi: 10.2307/1913738.

[S17] H. R. Varian, Intermediate Microeconomics: A Modern Approach. 8th ed., New York: Norton, 2006.

[S18] S. C. Jaquette, "A utility criterion for Markov decision processes," *Manage. Sci.*, vol. 23, no. 1, pp. 43–49, Sept. 1976. doi: 10.1287/mnsc.23.1.43.

[S19] G. W. Harrison, J. A. List, and C. Towe, "Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion," *Econometrica*, vol. 75, no. 2, pp. 433–458, Mar. 2007. doi: 10.1111/j.1468-0262.2006.00753.x.

[S20] C. C. Eckel and P. J. Grossman, "Men, women and risk aversion: Experimental evidence," in *Handbook of Experimental Economic Results*, vol. 1, C. R. Plott and V. L. Smith, Eds. Amsterdam, The Netherlands: Elsevier., 2008, ch. 113, pp. 1061–1073.

[S21] G. W. Harrison, M. I. Lau, and E. E. Rutström, "Estimating risk attitudes in Denmark: A field experiment," *Scand. J. Econ.*, vol. 109, no. 2, pp. 341–368, 2007. doi: 10.1111/j.1467-9442.2007.00496.x.

[S22] S. J. Kachelmeier and M. Shehata, "Examining risk preferences under high monetary incentives: Experimental evidence from the People's Republic of China," *Amer. Econ. Rev.*, vol. 82, pp. 1120–1141, Dec. 1992.

[S23] R. A. Howard and J. E. Matheson, "Risk-sensitive Markov decision processes," *Manage. Sci.*, vol. 18, no. 7, pp. 365–369, Mar. 1972. doi: 10.1287/mnsc.18.7.356.

[S24] A. Majumdar, S. Singh, A. Mandlekar, and M. Pavone, "Risk-sensitive inverse reinforcement learning via coherent risk models," in *Proc. Robotics: Science and Systems*, 2017. doi: 10.15607/RSS.2017.XIII.069.

Researchers and practitioners should carefully account for the behavioral component in the control of cyberphysical systems.

this regard. Overconfident decisions and consequently inefficient drone piloting can be facilitated by the hot hand fallacy as a misinterpretation of random sequences of immediate feedback on positive outcomes, since subjects fail to realize such sequences are caused by chance and therefore are history independent. In fact, a handful of subjects in the closed-loop treatment stated in the questionnaire that their decision strategy was to fly as long as they achieved steady increases in information value, thereby trying to exploit an apparent hot hand. The fact that the possibility for this fallacy is only given in a closed-loop system presents an obvious weakness that should be considered in designing such feedback policies.

In general, the current work exposes the human as an underobserved source of errors in human-in-the-loop control systems. Researchers and practitioners should carefully account for the behavioral component in the control of cyberphysical systems and the potential problems that arise from it, aside from only mathematical model optimization. Specifically, our findings illustrate the impact of behavioral biases regarding the effects of immediate feedback and the (mis)understanding of history independence in chance processes.

While more information is commonly needed to result in better decisions in cyberphysical systems, human susceptibility for perceptual biases in response to high information supply must be considered. Therefore, identifying an optimal quantity and frequency of feedback remains a goal for future research. We expect a carefully crafted intermittent feedback to be better suited for this purpose and stress the need for an intelligent feedback design that adapts to an individual's rationality to provide suitable amounts of information.

Considering the effect of humans in control loops more seriously presents an important issue for research and practice. Overall, humans were mostly shown to not act optimizing in the given decision problem in our experimental framework, which strongly places the Markov decision maker as an adequate characterization of human decision making in question. Models of human decision processes should be revisited to account for limits of cognitive capacities and behavioral biases that result from them to not jeopardize technological accomplishments through erroneous human decisions. Otherwise, individuals in human-in-the-loop control might take unnecessarily high risks and render thoughtfully designed policies inefficient, as seen for highly frequent feedback in the case of the hot hand fallacy.

Lastly, our study further provides a methodological contribution to research on CPHSs, making a first attempt to incorporate insights from behavioral economics into control engineering. Further, it introduces incentivized economic experiments as a viable option to reveal how individuals actually behave, in contrast to how they are theoretically prescribed to behave. An experimental UAV framework featuring a sequential decision problem is presented, with a focus on behavioral biases in relation to feedback policies. Future experimental research in this area may intensify efforts of incorporating various other behavioral phenomena and stylized facts into control engineering by building upon this framework to design or test behavioral interventions that are able to proactively counteract them.

AUTHOR INFORMATION

Marius Protte (marius.protte@uni-paderborn.de) is a Ph.D. student at the Chair of Corporate Governance at Paderborn University's Faculty of Business Administration and Economics. He received the M.Sc. degree in business administration in 2019 from Paderborn University. His research interests are in behavioral economics, business ethics, and human–machine interaction.

René Fahr received the diploma (M.Sc.) degree in economics in 1998 from the University of Bonn. In 2003, he received the Ph.D. degree jointly from the University of Bonn and the London School of Economics in quantitative economics. Since 2008, he has served as the head of the Chair of Corporate Governance at Paderborn University and research director of the BaER-Lab at the Faculty of Economics and Business Administration. Since 2019, he has been vice president for technology transfer and a member of the university's executive board and, since 2020, head of the Behavioral Economic Engineering and Responsible Management research group at the Heinz Nixdorf Institute. Since 1998, he has been affiliated with the Institute of Labor Economics in Bonn as a resident research affiliate. From 2003 to 2008, he was a postdoctoral researcher in personnel economics at the University of Cologne. His research interests are in behavioral economics, behavioral business ethics, and human–machine interaction.

Daniel E. Quevedo is a professor of cyberphysical systems in the School of Electrical Engineering and Robotics, Queensland University of Technology (QUT), Brisbane, Australia. Before joining QUT, he was the chair of automatic control at Paderborn University, Germany. In 2000, he received the ingeniero civil electrónico and M.Sc. degrees from Universidad Técnica Federico Santa María, Valparaíso,

Chile, and the Ph.D. degree from the University of Newcastle, Australia in 2005. He received the IEEE Conference on Decision and Control Best Student Paper Award in 2003 and was a finalist in 2002. In 2009, he was awarded a five-year research fellowship from the Australian Research Council. He is corecipient of the 2018 IEEE Transactions on Automatic Control George S. Axelby Outstanding Paper Award. He currently serves as associate editor for *IEEE Control Systems* and is a member of the editorial board of *International Journal of Robust and Nonlinear Control*. From 2015 to 2018, he was chair of the IEEE Control Systems Society Technical Committee on Networks and Communication Systems. His research interests are in control of networked systems and power converters. He is a Senior Member of IEEE.

REFERENCES

[1] A. Vempaty, B. Kailkhura, and P. K. Varshney, "Human-machine inference networks for smart decision making: Opportunities and challenges," in *Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing (ICASSP)*, 2018, pp. 6961–6965.

[2] S. Narayanan and P. G. Georgiou, "Behavioral signal processing: Deriving human behavioral informatics from speech and language," *Proc. IEEE*, vol. 101, no. 5, pp. 1203–1233, 2013. doi: 10.1109/JPROC.2012.2236291.

[3] G. Schirner, D. Erdogmus, K. Chowdhury, and T. Padir, "The future of human-in-the-loop cyber-physical systems," *Computer*, vol. 46, no. 1, pp. 36–45, 2013. doi: 10.1109/MC.2013.31.

[4] M. Tanelli, R. Toledo-Moreo, and L. M. Stanley, "Guest editorial: Holistic approaches for human–vehicle systems: Combining models, interactions, and control," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 5, pp. 609–613, 2017. doi: 10.1109/THMS.2017.2749939.

[5] F. Dressler, "Cyber physical social systems: Towards deeply integrated hybridized systems," in *Proc. Int. Conf. Computing, Networking and Communications*, 2018, pp. 420–424. doi: 10.1109/ICCNC.2018.8390404.

[6] A. Kolling, P. Walker, N. Chakraborty, K. Sycara, and M. Lewis, "Human interaction with robot swarms: A survey," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 1, pp. 9–26, 2016. doi: 10.1109/THMS.2015.2480801.

[7] S. N. Young and J. M. Peschel, "Review of human–machine interfaces for small unmanned systems with robotic manipulators," *IEEE Trans. Human-Mach. Syst.*, vol. 50, no. 2, pp. 131–143, 2020. doi: 10.1109/THMS.2020.2969380.

[8] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006. doi: 10.1016/j.automatica.2006.06.027.

[9] P. J. van Overloop, J. M. Maestre, A. D. Sadowska, E. F. Camacho, and B. de Schutter, "Human-in-the-loop model predictive control of an irrigation canal," *IEEE Control Syst. Mag.*, vol. 35, no. 4, pp. 19–29, Aug. 2015. doi: 10.1109/MCS.2015.2427040.

[10] Z. Ercan, A. Carvalho, M. Gokasan, and F. Borelli, "Modeling, identification, and predictive control of a driver steering assistance system," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 5, pp. 700–710, 2017. doi: 10.1109/THMS.2017.2717881.

[11] M. Inoue and V. Gupta, "Weak control for human-in-the-loop systems," *IEEE Control Syst. Lett.*, vol. 3, no. 2, pp. 440–445, 2019. doi: 10.1109/LCSYS.2019.2891489.

[12] C.-P. Lam and S. S. Sastry, "A POMDP framework for human-in-the-loop system," in *Proc. IEEE Conf. Decision and Control*, 2014, pp. 6031–6036. doi: 10.1109/CDC.2014.7040333.

[13] L. Feng, C. Wiltsche, L. Humphrey, and U. Topcu, "Synthesis of human-in-the-loop control protocols for autonomous systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 450–462, Apr. 2016. doi: 10.1109/TASE.2016.2530623.

[14] M. L. Puterman, *Markov Decision Processes*. Hoboken, NJ: Wiley-Interscience, 1994.

[15] B. M. Albaba and Y. Yildiz, "Modeling cyber-physical human systems via an interplay between reinforcement learning and game theory," *Annu. Rev. Control*, vol. 48, pp. 1–19, Oct. 2019. doi: 10.1016/j.arcontrol.2019.10.002.

[16] C. F. Camerer, T.-H. Ho, and J.-K. Chong, "A cognitive hierarchy model of games," *Quart. J. Econ.*, vol. 119, no. 3, pp. 861–898, 2004. doi: 10.1162/0033553041502225.

[17] P. S. Schaefer, C. C. Williams, A. S. Goodie, and W. K. Campbell, "Overconfidence and the big five," *J. Econ. Res. Pers.*, vol. 38, no. 5, pp. 473–480, 2004. doi: 10.1016/j.jrp.2003.09.010.

[18] T. Gilovich, R. Vallone, and A. Tversky, "The hot hand in basketball: On the misperception of random sequences," *Cogn. Psychol.*, vol. 17, no. 3, pp. 295–314, July 1985. doi: 10.1016/0010-0285(85)90010-6.

[19] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Trans. Autom. Control*, vol. 19, no. 5, pp. 494–500, Oct. 1974. doi: 10.1109/TAC.1974.1100635.

[20] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Belmont, MA: Athena Scientific, 2005.

[21] D. Kahneman, "Maps of bounded rationality: Psychology for behavioral economics," *Amer. Econ. Rev.*, vol. 93, no. 5, pp. 1449–1475, Dec. 2003. doi: 10.1257/000282803322655392.

[22] C. F. Camerer and C. Loewenstein, "Behavioral economics: Past, present, future," in *Advances in Behavioral Economics* C. F. Camerer, G. Loewenstein, and M. Rabin, Eds. Princeton, NJ: Princeton Univ. Press, 2003, pp. 3–52.

[23] D. Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York: Harper Collins, 2009.

[24] D. Kahneman and A. Tversky, "On the psychology of prediction," *Psychol. Rev.*, vol. 80, no. 4, pp. 237–251, 1973. doi: 10.1037/h0034747.

[25] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *Science*, vol. 185, no. 4157, pp. 1124–1131, Sept. 1974. doi: 10.1126/science.185.4157.1124.

[26] P. M. Todd and G. Gigerenzer, "Bounding rationality to the world," *J. Econ. Psychol.*, vol. 24, no. 2, pp. 143–165, Apr. 2003. doi: 10.1016/S0167-4870(02)00200-3.

[27] S. Lichtenstein and B. Fischhoff, "Do those who know more also know more about how much they know?" *Org. Behav. Human Perform.*, vol. 20, no. 2, pp. 159–183, Dec. 1977. doi: 10.1016/0030-5073(77)90001-0.

[28] B. M. Barber and T. Odean, "Boys will be boys: Gender, overconfidence, and common stock investment," *Quart. J. Econ.*, vol. 116, no. 1, pp. 261–292, 2001. doi: 10.1162/003355301556400.

[29] W. F. M. DeBondt and H. Thaler, R, "Financial decision-making in markets and firms: A behavioural perspective," in *Handbook of Operations Research and Management Science*, vol. 9, A. R. Jarrow, V. Maksimovic, and W. T. Ziemba, Eds. Amsterdam, The Netherlands: Elsevier, 1995, pp. 385–410.

[30] N. D. Weinstein, "Unrealistic optimism about future life events," *J. Pers. Soc. Psychol.*, vol. 39, no. 5, pp. 806–820, Nov. 1980. doi: 10.1037/0022-3514.39.5.806.

[31] O. Svenson, "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychol.*, vol. 47, no. 2, pp. 143–148, Feb. 1981. doi: 10.1016/0001-6918(81)90005-6.

[32] E. Hoelzl and A. Rustichini, "Overconfident: Do you put your money on it?" *Econ. J.*, vol. 115, no. 503, pp. 305–318, 2005. doi: 10.1111/j.1468-0297.2005.00990.x.

[33] D. A. Moore and P. J. Healy, "The trouble with overconfidece," *Psychol. Rev.*, vol. 115, no. 2, pp. 502–517, 2008. doi: 10.1037/0033-295X.115.2.502.

[34] J. R. Radzevick and D. A. Moore, "Competing to be certain (but wrong): Market dynamics and excessive confidence in judgment," *Manage. Sci.*, vol. 57, no. 1, pp. 93–106, Jan. 2011. doi: 10.1287/mnsc.1100.1255.

[35] W. van Hippel and R. Trivers, "The evolution and psychology of self-deception," *Behav. Brain Sci.*, vol. 34, no. 1, pp. 1– 56, 2011. doi: 10.1017/S0140525X10003018.

[36] D. A. Moore and D. Schatz, "The three faces of overconfidence," *Soc. Pers. Psychol. Compass*, vol. 11, no. 8, pp. 1–12, 2017. doi: 10.1111/spc3.12331.

[37] G. Mayraz, "Wishful thinking," Centre for Economic Performance (CEP), London, CEP Discussion Papers, 2011. [Online]. Available: https://ssrn.com/abstract=1955644

[38] I. Ben-David, J. R. Graham, and C. R. Harvey, "Managerial miscalibration," *Quart. J. Econ.*, vol. 128, no. 4, pp. 1547–1584, 2013. doi: 10.1093/qje/qjt023.

[39] P. Sedlmeier and G. Gigerenzer, "Teaching Bayesian reasoning in less than two hours," *J. Exp. Psychol.*, vol. 130, no. 3, pp. 380–400, 2001. doi: 10.1037/0096-3445.130.3.380.

[40] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, 1979. doi: 10.2307/1914185.

[41] P. Slovic, B. Fischhoff, and S. Lichtenstein, "Why study risk perception?" *Risk Anal.*, vol. 2, no. 2, pp. 83–93, 1982. doi: 10.1111/j.1539-6924.1982.tb01369.x.

[42] K. Abbink, B. Irlenbusch, and E. Renner, "An experimental bribery game," *J. Law, Econ. Org.*, vol. 18, no. 2, pp. 428–454, 2002. doi: 10.1093/jleo/18.2.428.

[43] E. U. Weber, "Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet)," *Climate Change*, vol. 77, nos. 1–2, pp. 103–120, 2006. doi: 10.1007/s10584-006-9060-3.

[44] P. Slovic, B. Fischhoff, and S. Lichtenstein, "Accident probabilities and seat belt usage: A psychological perspective," *Accid. Anal. Prev.*, vol. 10, no. 4, pp. 281–285, 1978. doi: 10.1016/0001-4575(78)90030-1.
[45] A. Sandroni and F. Squintani, "Overconfidence and asymmetric information: The case of insurance," *J. Econ. Behav. Org.*, vol. 93, pp. 149–165, Sept. 2013. doi: 10.1016/j.jebo.2012.10.015.
[46] S. Benartzi, "Excessive extrapolation and the allocation of 401(k) accounts to company stock," *J. Finance*, vol. 56, no. 5, pp. 1747–1764, Oct. 2001. doi: 10.1111/0022-1082.00388.
[47] C. Camerer and D. Lovallo, "Overconfidence and excess entry: An experimental approach," *Amer. Econ. Rev.*, vol. 89, no. 1, pp. 306–318, Mar. 1999. doi: 10.1257/aer.89.1.306.
[48] A. M. Goel and A. V. Thakor, "Overconfidence, CEO selection, and corporate governance," *J. Finance*, vol. 63, no. 6, pp. 2737–2784, 2008. doi: 10.1111/j.1540-6261.2008.01412.x.
[49] U. Malmendier and G. Tate, "CEO overconfidence and corporate investment," *J. Finance*, vol. 60, no. 6, pp. 2661–2700, Dec. 2005. doi: 10.1111/j.1540-6261.2005.00813.x.
[50] T. Odean, "Do investors trade too much?" *Amer. Econ. Rev.*, vol. 89, no. 5, pp. 1279–1298, Dec. 1999. doi: 10.1257/aer.89.5.1279.
[51] B. M. Barber and T. Odean, "Trading is hazardous to your wealth: The common stock investment performance of individual investors," *J. Finance*, vol. 55, no. 2, pp. 773–806, Apr. 2000. doi: 10.1111/0022-1082.00226.
[52] I. Trinugroho and R. Sembel, "Overconfidence and excessive trading behavior: An experimental study," *Int. J. Bus. Manage.*, vol. 6, no. 7, pp. 147–152, July 2011. doi: 10.5539/ijbm.v6n7p147.
[53] I. Frieze and B. Weiner, "Cue utilization and attributional judgments for success and failure," *J. Pers.*, vol. 39, no. 4, pp. 591–605, 1971. doi: 10.1111/j.1467-6494.1971.tb00065.x.
[54] E. J. Langer, "The illusion of control," *J. Pers. Soc. Psychol.*, vol. 32, no. 2, pp. 311–328, 1975. doi: 10.1037/0022-3514.32.2.311.
[55] T. J. Johnson, R. Feigenbaum, and M. Weiby, "Some determinants and consequences of the teacher's perception of causation," *J. Educ. Psychol.*, vol. 55, no. 5, pp. 237–246, 1964. doi: 10.1037/h0043389.
[56] M. T. Billett and Y. Qian, "Are overconfident CEOs born or made? Evidence of self-attribution bias from frequent acquirers," *Manage. Sci.*, vol. 54, no. 6, pp. 1037–1051, Apr. 2008. doi: 10.1287/mnsc.1070.0830.
[57] S. Gervais and T. Odean, "Learning to be overconfident," *Rev. Financial Stud.*, vol. 14, no. 1, pp. 1–27, 2001. doi: 10.1093/rfs/14.1.1.
[58] G. Hilary and L. Menzly, "Does past success lead analysts to become overconfident?" *Manage. Sci.*, vol. 52, no. 4, pp. 489–500, Apr. 2006. doi: 10.1287/mnsc.1050.0485.
[59] K. Daniel and D. Hirshleifer, "Overconfident and investors, predictable returns, and excessive trading," *J. Econ. Perspect.*, vol. 29, no. 4, pp. 61–88, 2015. doi: 10.1257/jep.29.4.61.
[60] G. Törngren and H. Montgomery, "Worse than chance? performance and confidence among professionals and laypeople in the stock market," *J. Behav. Finance*, vol. 5, no. 3, pp. 148–153, 2004. doi: 10.1207/s15427579jpfm0503_3.
[61] P. K. Presson and V. A. Benassi, "Illusion of control: A meta-analytic review," *J. Social Behavior Personality*, vol. 11, pp. 493–510, Jan. 1996.
[62] M. van Otterlo and M. Wiering, "Reinforcement learning and markov decision processes," in *Reinforcement Learning*, vol. 12, M. Wiering and M. van Otterlo, Eds. Berlin: Springer-Verlag, 2012, ch. 1, pp. 3–42.
[63] G. Kasneci, J. Van Gael, R. Herbrich, and T. Graepel, "Bayesian knowledge corroboration with logical rules and user feedback," in *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2010*. J. L. Balcázar, F. Bonchi., A. Gionis., and M. Sebag., Eds. Berlin: Springer-Verlag, 2010, pp. 1–18.
[64] D. P. Bertsekas, *Reinforcement Learning and Optimal Control*. Belmont, MA: Athena Scientific, 2019.
[65] C. Camerer, "Behavioral economics: Reunifying psychology and economics," *Proc. Nat. Acad. Sci.*, vol. 96, no. 19, pp. 10,575–10,577, Sept. 1999. doi: 10.1073/pnas.96.19.10575.
[66] G. Charness and D. Levin, "When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity and affect," *Amer. Econ. Rev.*, vol. 95, no. 4, pp. 1300–1309, Sept. 2005. doi: 10.1257/0002828054825583.
[67] D. N. Kleinmuntz, "Cognitive heuristics and feedback in a dynamic decision environment," *Manage. Sci.*, vol. 31, no. 6, pp. 680–702, June 1985. doi: 10.1287/mnsc.31.6.680.
[68] R. M. Hogarth, "Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics," *Psychol. Bull.*, vol. 90, no. 2, pp. 197–217, 1981. doi: 10.1037/0033-2909.90.2.197.
[69] I. Larkin and S. Leider, "Does commitment or feedback influence myopic loss aversion? An experimental analysis," *Amer. Econ. J., Microecon.*, vol. 4, no. 2, pp. 184–214, 2012. doi: 10.1257/mic.4.2.184.
[70] C. I. Tsai, J. Klayman, and R. Hastie, "Effects of amount of information on judgment accuracy and confidence," *Org. Behav. Human Decision Process.*, vol. 107, no. 2, pp. 97–105, Nov. 2008. doi: 10.1016/j.obhdp.2008.01.005.
[71] G. Chen, C. Crossland, and S. Luo, "Making the same mistake all over again: CEO overconfidence and corporate resistance to corrective feedback," *Strateg. Manage. J.*, vol. 36, no. 10, pp. 1513–1535, Oct. 2015. doi: 10.1002/smj.2291.
[72] A. L. Zacharakis and D. A. Shepherd, "The nature of information and overconfidence on venture capitalists' decision making," *J. Bus. Ventur.*, vol. 16, no. 4, pp. 311–332, July 2001. doi: 10.1016/S0883-9026(99)00052-X.
[73] V. Subbotin, "Outcome feedback effects on under- and overconfident judgments (general knowledge tasks," *Org. Behav. Human Decision Process.*, vol. 66, no. 3, pp. 268–276, June 1996. doi: 10.1006/obhd.1996.0055.
[74] C. F. Camerer, "Does the basketball market believe in the 'hot hand'?" *Amer. Econom. Rev.*, vol. 79, pp. 1257–1261, Dec. 1989.
[75] W. O. Brown and R. D. Sauer, "Does the basketball market believe in the hot hand? Comment," *Amer. Econ. Rev.*, vol. 83, pp. 1377–1386, Dec. 1993.
[76] A. W. Chau and J. G. Phillips, "Effects of perceived control upon wagering and attributions in computer blackjack," *J. Gen. Psychol.*, vol. 122, no. 3, pp. 253–269, 1995. doi: 10.1080/00221309.1995.9921237.
[77] U. Malmendier and S. Nagel, "Depression babies: Do macroeconomic experiences affect risk taking?" *Quart. J. Econ.*, vol. 126, no. 1, pp. 373–416, Feb. 2011. doi: 10.1093/qje/qjq004.
[78] P. Ayton and I. Fischer, "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?" *Memory Cogn.*, vol. 32, no. 8, pp. 1369–1378, 2004. doi: 10.3758/BF03206327.
[79] M. Rabin and D. Vayanos, "The gambler's and hot-hand fallacies: Theory and applications," *Rev. Econ. Stud.*, vol. 77, no. 2, pp. 730–778, 2009. doi: 10.1111/j.1467-937X.2009.00582.x.
[80] J. B. Miller and A. Sanjurjo, "Surprised by the hot hand fallacy? A truth in the law of small numbers," *Econometrica*, vol. 86, no. 6, pp. 2019–2047, 2018. doi: 10.3982/ECTA14943.
[81] J. B. Miller and A. Sanjurjo, "How experience confirms the gambler's fallacy when sample size is neglected," OSF Preprints, Center for Open Science, Charlottesville, VA, Oct. 2018. [Online]. Available: https://doi.org/10.31219/osf.io/m5xsk
[82] T. S. Ferguson, "Optimal stopping and applications," Mathematics Department, Univ. California, Los Angeles, 2008. [Online]. Available: https://www.math.ucla.edu/people/ladder/tom
[83] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, 2nd ed., Princeton, NJ: Princeton Univ. Press, 1944.
[84] D. Read, G. Loewenstein, and M. Rabin, "Choice bracketing," *J. Risk Uncertainty*, vol. 19, no. 1/3, pp. 171–197, 1999. doi: 10.1023/A:1007879411489.
[85] D. Hirshleifer, A. Low, and S. H. Teoh, "Are overconfident CEOs better innovators?" *J. Finance*, vol. 67, no. 4, pp. 1457–1498, Aug. 2012. doi: 10.1111/j.1540-6261.2012.01753.x.
[86] M. E. Jarvik, "Probability learning and a negative recency effect in the serial anticipation of alternative symbols," *J. Exp. Psychol.*, vol. 41, no. 4, pp. 291–297, 1951. doi: 10.1037/h0056878.
[87] D. L. Chen, M. Schonger, and C. Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *J. Behav. Exp. Finance*, vol. 9, pp. 88–97, Mar. 2016. doi: 10.1016/j.jbef.2015.12.001.
[88] B. Greiner, "Subject pool recruitment procedures: Organizing experiments with ORSEE," *J. Econ. Sci. Assoc.*, vol. 1, no. 1, pp. 114–125, May 2015. doi: 10.1007/s40881-015-0004-4.
[89] T. Dohmen, A. Falk, D. Huffman, and U. Sunde, "Are risk aversion and impatience related to cognitive ability," *Amer. Econ. Rev.*, vol. 100, no. 3, pp. 1238–1260, June 2010. doi: 10.1257/aer.100.3.1238.
[90] K. A. Carlson and S. B. Shu, "The rule of three: How the third event signals the emergence of a strak," *Org. Behav. Human Decision Process.*, vol. 104, no. 1, pp. 113–121, 2007. doi: 10.1016/j.obhdp.2007.03.004.