Understanding non-adoption of "optimal" algorithmic tips for problem-solving: views and use of tips, barriers encountered, and relationship to personality

DAVID LEE*, UC Santa Cruz

WICHINPONG PARK SINCHAISRI*, UC Berkeley

Designing effective human-AI systems requires developing a deeper understanding of how humans use or choose not to use AI. In this paper, we explore how humans use (or don't use) "optimal" algorithmic tips in multi-step problem solving contexts, specifically when managing a virtual kitchen after disruptions necessitated a change of strategy. A qualitative analysis found that participants viewed or used tips in four ways: as rules, directional principles, options to try, and highlights. Even when workers rejected tips outright, tips could still be useful through creating focal points for worker sense-making. We also found that the challenge of operationalizing tips can lead to diverse barriers to adoption not just related to trust (counterintuitive tip or bad outcomes), but also to tip usability (tip clarity, difficulty to implement, and difficulty to track implementation) and environmental factors (misaligned incentives) that challenge a simple definition of what it means for a tip to be optimal. In a follow-up quantitative study, we validate these findings, identify three barriers that correlate significantly with intent to use tips, show that negative initial views of tips make one more likely to be impacted by experiencing bad outcomes, and show significant correlations between the "Orientation to Change" dimension of problem solving style and people's views of tips and barriers experienced.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: human-ai collaboration, algorithmic tips, barriers to adoption of tips, personality

ACM Reference Format:

David Lee and Wichinpong Park Sinchaisri. 2018. Understanding non-adoption of "optimal" algorithmic tips for problem-solving: views and use of tips, barriers encountered, and relationship to personality. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 19 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Human-AI interfaces are increasingly used to aid humans in the real world, including in consequential domains from healthcare [21] to legal decision-making [1]. However, these interfaces have not been fully adopted for many reasons such as the black-box nature of underlying algorithms resulting in a lack of transparency, accountability, or interpretability. The lack of human understanding of how algorithms work could pose serious problems to the society, from prisoners incorrectly denied parole to polluted air mistakenly identified as safe [35], and lead to aversion to the machine-generated recommendation among humans [6, 10]. For example, Dietvorst, Simmons, and Massey [10] showed that, in a forecasting task, humans prefer to follow the suggestion made by another human forecaster rather than by an algorithm and that their *confidence* in the algorithm declines at a faster rate when a flawed suggestion was made. Castelo, Bos, and Lehmann [6] further demonstrated that this is particularly true for subjective tasks where humans *incorrectly assume* that algorithms were only able to perform objective tasks. On the other hand, in tasks in which a

*Both authors contributed equally to this research.

ground truth exists, humans can make the opposite mistake of automation bias, where they trust algorithms when they should not [30]. A long thread of work has sought to better understand and address algorithmic aversion and automation bias, but much of this has focused on prediction, where the primary dynamic is whether or not a human should accept or reject an AI-generated decision. Less work has explored problem solving settings, where tips are not complete solutions that a human simply accepts or not, but hints augmenting a worker's problem solving process.

A recent paper coauthored by one of the authors (anonymized) introduced an algorithm for generating "optimal" tips, i.e. tips that bridge the largest gap between human strategies and optimal strategies. They studied their algorithm in a problem solving context in which humans need to make a sequence of interdependent decisions for managing a virtual kitchen with the aid of tips that provide users with guidance (e.g. "server should cook twice"), but not complete solutions (in which humans would only need to decide whether one should "adopt" or "not adopt"). They found that despite the algorithmic tip performing much better than alternatives in helping participants improve their score on average, people were less likely to adopt the tip compare to a more intuitive human-provided (but non-optimal) tip.

In this paper, we take a mixed-methods approach to understanding human non-adoption of machine-generated "optimal" tips in problem solving contexts. We first analyzed the qualitative responses from the prior large-scale behavioral experiment, finding that tips were viewed or used by participants in four different ways: as rules, directional principles, options to try, and highlights. Even in cases when workers rejected tips outright, tips could still be useful through creating focal points in the solution space for worker sense-making. We also found that in problem solving contexts, the challenge of operationalizing tips can lead to diverse barriers. Some related to participants not trusting the tip (due to it being counterintuitive or resulting in bad outcomes), but others related to tip usability (lacking clarity, being difficult to implement, or being difficult to track whether they were implementing) and to broader environment factors (misaligned incentives) that also challenge a simple definition of what it means for a tip to be optimal.

Building on these findings, and on hints of personality influences in some of the responses, we conducted a follow-up study which added new survey questions eliciting: (1) worker intents on following tips, (2) their views of tips and experiences of barriers, and (3) their problem-solving style based on a simplified version of the VIEW problem-solving style assessment [39]. Beyond allowing us to validate and quantify our identified barriers, the results also highlight three barriers (counterintuitive tips, bad outcomes, and difficulty to implement) as having significant correlations to future intention to use tips and personality (specifically, problem solving styles) as having several significant correlations with use of tips and experience of barriers. For example, those with a stronger "Developer" type in their "Orientation to Change" (see **Figure 2**) were less likely to express counterintuitive or misaligned incentives as barriers to adopting tips, and more likely to be successful in implementing a tip in early attempts to do so, among other results.

In what follows, after describing related work (**Section 2**) and the experimental setting (**Section 3**), we describe our qualitative study and its findings on views of tips and barriers to adoption (**Section 4**). This is followed by a description of our quantitative study and its findings that validate our observations and highlight significant correlations between intent to use tips, barriers experienced, and personality (**Section 5**). We end by discussing how this richer view of human interactions with tips suggest implications for design, raise questions about what it means for a tip to be "optimal", and point to directions on the use of diverse personalities in collaborative AI-assisted problem solving. Our hope is that a richer understanding of human adoption of algorithmic tips will ultimately help with the design of more effective human-AI interfaces that can better support humans while avoiding unintentional negative consequences.

2 RELATED WORK

Machine-learning algorithms have been increasingly employed as a tool to help guide human decision-making in various contexts, from healthcare [20, 21, 23] to criminal justice [1, 27]. Despite fast-growing advances in computing power and predictive capabilities outperforming humans in certain settings, these algorithmic tools also have important limitations and thus, have not been fully adopted by humans. In recent years, researchers have been documenting the phenomena of *automation bias*, where humans follow faulty algorithmic advice, and *algorithm aversion*, where humans reject correct advice from an algorithm [10] (see [26] for a comprehensive, systematic review of recent empirical studies on algorithm aversion). Various factors influence such aversion. Humans’ view of the machine-generated advice is greatly influenced by the characteristics of the algorithm, including the “black box" nature or the lack of transparency into how algorithms work [12, 33], complexity [13, 40], accuracy [3, 10, 45], and interface design [15, 25, 40]. When algorithms make a mistake, humans lose trust on their capabilities at a faster rate than when the mistakes were made by humans [3, 10]. The level of human trust on AI greatly depends not only on the stated accuracy but also the actual accuracy observed by humans [45]. Algorithms are also relied on less for certain types of tasks, such as subjective tasks that humans believe should be done by humans [6, 44], moral decisions [2, 31], or that they could compare with their own decisions [19, 41].

To address these challenges, researchers have sought to understand more deeply when humans are or are not able to detect and override errors [9]. For example, displaying confidence scores can help calibrate trust in AI models, but showing how much different features contribute to a prediction does not [8, 22, 36, 46]. More transparent models can actually make participants less able to detect mistakes due to information overload [34], unless cognitive forcing functions are used to force engagement with AI explanations [4]. Much of this work has focused on augmenting human decision-making in one-off tasks such as prediction and forecasting, in which the primary problem is whether to accept or not accept the AI-recommended decision.

In this work, we focus on human interactions with AI in sequential decision-making contexts where current decisions affect future states and outcomes for solving a larger problem or achieving a larger goal. This setting is challenging for both humans and AI models as humans need to make trade-offs between short- and long-term benefits of a decision while AI models need to account for decision-making strategies employed by humans. While there is a growing set of papers studying AI for supporting these complex problem-solving, e.g. for retail pricing [5], chess [28], and workforce management (anonymized), we are not aware of papers investigating algorithmic aversion or adoption of tips in such sequential decision-making contexts. This paper builds on a previous study analyzing a new algorithm for generating optimal tips. That paper observed that humans receiving advice from the algorithm were significantly less likely to adopt the advice compared to those receiving a human-generated advice, despite the algorithmic advice being more helpful for improving human performance. Our work examines this phenomena to develop a deeper understanding of non-adoption of algorithmic tips.

Finally, we contribute to a stream of research that investigates the impact of personality and problem-solving style on human interactions with AI. The majority of work in this area has been experimental and focused on humans’ attitudes about themselves [18]. For example, greater self-esteem and a bias towards one’s own decisions decreases human adoption of algorithmic advice [14, 24], especially among more experienced decision-makers [**? ?**]. Humans with higher self-efficacy are less likely to follow algorithmic advice due to a belief that they should have greater self-reliance [12, 19]. [11] finds that humans are more likely to adopt algorithmic advice when they can deviate to some degree from the advice. In a sequential decision-making setting, the focus of this paper, another important aspect to understand is the influence of humans' problem-solving styles, e.g. as formalized in Selby, Treffinger, and Isaksen's VIEW problem solving style assessment [39]. We are not aware of research that examines the relationship between problem solving style and algorithmic aversion. However, AI researchers have found that human problem-solving strategies can sometimes be recovered from trace data of their decisions [29], pointing towards a possible relationship. Our work includes a brief exploratory study of how personality and problem-solving style affect humans' views of and barriers to adopting machine-generated advice in a sequential decision-making contexts.

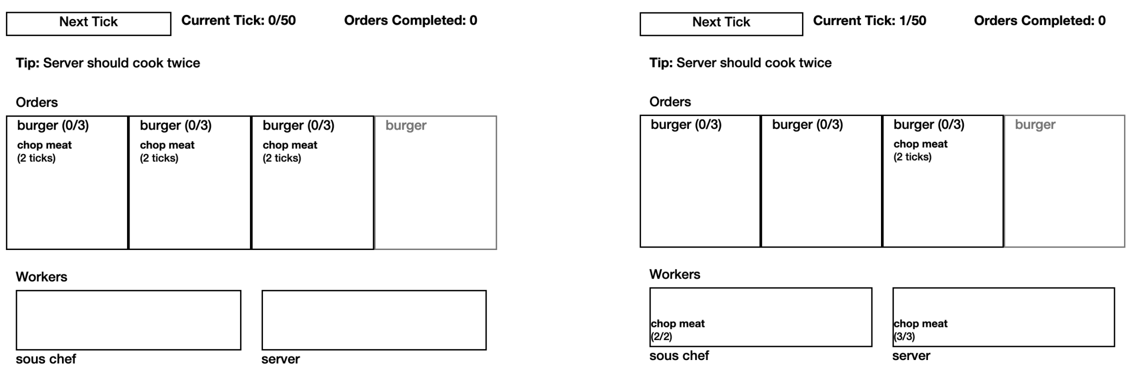
3 BACKGROUND AND SETTING: ALGORITHMIC TIPS FOR MANAGING A VIRTUAL KITCHEN

We consider human adoption of tips in a setting where humans receive advice while performing a sequential decision-making task. In the experimental setting, human participants act as managers for a virtual kitchen in which they must assign various cooking tasks (e.g., chopping, cooking, plating) to virtual kitchen workers (e.g., chef, sous-chef, server) with varying capabilities in a way that minimizes the completion time of all food orders (e.g., four burgers). Participants play the game for two rounds in which they operate a full-capacity kitchen (e.g., with all three virtual workers) and then for four additional rounds where the most capable virtual worker, the chef, is no longer available and the same food orders have to be made by the remaining two virtual workers, the sous-chef and the server. In these four rounds, workers try to reach a known optimal completion time and are provided with different tips (depending on their experimental condition). The key challenges of this game are that (i) human players need to uncover the actual skill level of each virtual worker and make short- and long-term trade-offs when allocating tasks to avoid bottlenecks, and (ii) the disruption caused by the departure of the chef requires the players to adapt to the new environment where a new decision-making strategy is needed.

The setting of the virtual kitchen management game can be considered as a general task allocation and scheduling problem that managers face in practice as it captures the key challenges of managing bottlenecks, learning workers' (initially) unknown skill levels, and adapting to a disruption. In practice, rather than knowing the true optimal value based on a fully-solved optimal policy, managers are often informed of performance benchmarks within the organization and across the industry. For example, according to the Bureau of Transportation Statistics, an American regional airline Endeavor Air has the highest on-time arrival performance of 89.16% compared to the industry's average of 84.63% [32]. Operations managers are constantly seeking to scheduling tasks to minimize delays and meet industry benchmarks.

In the prior paper we are building on, an extensive behavioral study was conducted on Amazon Mechanical Turk in two phases. In the first phase ($N = 172$), participants played the game without receiving a tip, and their sequences of decisions were recorded. This data was then fed into a novel machine learning algorithm that extracts the decision-making strategies employed by the participants and generates interpretable advice (the "tip") by comparing the actions taken by the humans with the optimal policy. Due to the sequential nature of the decision-making task, the optimal policy is a complex sequence of state-action pairs. For interpretability, the algorithm then chooses only a snapshot of the optimal policy, or one state-action pair to provide to the participants, that bridges the largest gap between human and optimal strategies. Therefore, even though the tip is based on the optimal policy, it is not guaranteed that the participants would be able to recover the remaining state-action pairs in the optimal policy. In the second phase ($N = 1,011$), a new set of participants were randomly assigned into one of the four conditions: *control* (e.g., not receiving any tip), *algorithm* (e.g., receiving a tip chosen by the novel algorithm), *human* (e.g., receiving a tip that received the most votes from participants in the first phase as "best tip to help future players"), and *baseline* (e.g., receiving a tip from a naive, frequency-based algorithm). Then, they proceeded to play the same virtual kitchen-management game where the tip (if any) was shown during the game in every round. **Figure ??** shows the example screenshots of the game interface for a participant randomly assigned into the *algorithm* condition. Performance was measured in terms of the completion time in each round of the game and the fraction of participants who achieved the optimal solution. The tips shown in this phase are:

- *Algorithm:* Server should cook twice (out of four burgers). This tip is part of the optimal policy.
- *Human:* Server should cook once (out of four burgers). This tip is not part of the optimal policy, but it is closer to the optimal policy under the full-capacity scenario so is more intuitive than the algorithmic tip (e.g., since the server takes the longest time to cook, cooking is typically assigned primarily to the chef)
- *Baseline:* Sous-chef should plate twice (out of four burgers). This tip is also part of the optimal policy, but chosen in a naive way that does not optimally bridge the gap between human and optimal strategies.



(a) The initial state where users observe available subtasks, median times to completion, and two idle virtual workers. The interface also shows the current time index, time limit, current progress, and potential tip.

(b) The next state after two previously available subtasks were assigned to the virtual workers and the true completion times were realized, revealing different levels of virtual workers’ skills

Fig. 1. Example screenshots from the virtual kitchen management game

The experimental results demonstrated that the algorithmic tip enabled human participants to substantially improve their performance compared to counterparts that were not shown the tip or were shown alternative tips. Specifically, participants in the *Algorithm* condition completed the final round of the game in 37.1 steps, outperforming those in other conditions: 37.9 (*Control*, $t(243) = -4.361$, $p = 0.00000806$), 37.5 (*Human*, $t(246) = -2.52$, $p = 0.00605$), and 38.4 (*Baseline*, $t(246) = -7.348$, $p < 10^{-12}$). Furthermore, 19% of participants in the *Algorithm* condition achieved optimal performance (34 steps) in the final round, compared to less than 1% in all other conditions. Thus, the algorithmic tip can be considered as the “optimal” tip. However, the authors noted that the adoption rate among participants in the *Algorithm* treatment was much lower (24–48%) than the adoption rate among those in the *Human* treatment (83–88%). Although the algorithmic tip was found to already be effective at improving performance (even with low adoption), increasing adoption of the optimal tip could potentially lead to a greater performance improvement. After the participants played all six rounds of the game, they responded to a survey about their experience, gameplay, and thoughts about the tip. This paper starts from an analysis of these survey responses to gain a deeper insight into human adoption and non-adoption of tips.

4 STUDY 1: DEVELOPING A RICH VIEW OF HOW WORKERS VIEW TIPS AND BARRIERS TO ADOPTION

4.1 Method: Qualitative Analysis

This paper takes a mixed-methods approach to developing a deeper insight into human adoption and non-adoption of tips. We started with a qualitative analysis of worker responses to one of the open-ended questions in the post-study survey, *"What did you think about the tip for these last four rounds and how did you incorporate it in your strategy?"*, specifically for those in the *Algorithm* condition ($N = 247$), i.e. those who received the optimal tip *"server cooks twice"*.

One author carried out an inductive coding process that drew on elements of grounded theory [7], in which they open coded anything relevant to characterizing the diverse ways in which workers engaged with provided tips. These open codes were then compared and grouped together to look for common themes, which led us to narrow down our coding to the following two research questions in subsequent rounds:

> **RQ1**: How did workers view or conceptualize the tips provided to them?
>
> **RQ2**: What barriers kept people from using tips, either initially or in later rounds?

Comparing and grouping the resulting codes led us to a preliminary codebook which was then used to code the responses in a final pass while checking for responses not fitting those codes. The final output was four different ways in which workers viewed and used tips (*rules*, *directional principles*, *options to try*, and *highlights*) and six different barriers to using tips (*counterintuitive*, *hard to implement*, *bad outcomes*, *lack of clarity*, *hard to track*, and *misaligned incentives*).

In the following subsections, we discuss our findings for each of our two research questions. To help convey the context underlying the participant quotes illustrating our findings, our blockquotes are in the form (PID, $c_3|c_4|c_5|c_6$, $d_3|d_4|d_5|d_6$), where PID is the participant, a_i refers to the number of times they had the server cook in rounds 3 to 6 (the disrupted rounds) and d_i refers to the duration that they were able to achieve in rounds 3 to 6 (with 34 being the minimum duration possible). For example, *(P90, 2|2|2|2, 36|37|36|37)* refers to a quote from participant 90 who had the server cook twice in each of the four disrupted rounds and achieved a duration of 36, 37, 36, and 37 in those rounds.

4.2 How workers view and used tips: rules, directional principles, options to try, and highlights

Of the workers who started off with an optimistic view of the tip, a few viewed the tips as *rules* that they *"had to figure out how to incorporate" (P90)*. For these workers, tips constrained the solution space they had to explore when sensemaking. For example, workers said:

> *"I knew that the server took longer to cook but HAD to cook twice so I had to figure out how to incorporate it" (P90, 2|2|2|2, 36|37|36|37)*
>
> *"I thought of it as a rule and not a tip, even though it didn't say it was a rule. So, I followed the tip..." (P108, 2|2|2|2, 34|36|34|34)*

Another group of workers also had an optimistic view of the tips, but described them as *directional principles* to focus or be more cognizant of. For these workers, they did not feel like they needed to follow it exactly, but the tip guided them in becoming more aware about using the server to cook. Workers said:

> *"I didn't try to have the serve cook twice, but I was cognizant and more aggressive with having them cook in general- I just didn't track the exact number of times." (P183, 1|2|1|2, 39|39|38|40)*
>
> *"It was very helpful. It made me focus on making sure the server cooked more even if that was not his obvious strength." (P43, 1|2|2|2, 38|34|34|34)*

Unlike the optimistic view of the previous two groups, others viewed it as an *option to try*. Some viewed it neutrally as something to test out before evaluating while others viewed it skeptically, but felt they were willing to give it a shot. They said:

> *“I tested it out the first round and found out that it worked, so I repeated it during later rounds.” (P214, 2|2|2|2, 36|34|34|34)*

> *“I tried it the first time, but I don’t think it was a good tip, so I ignored it the next times.” (P128, 2|1|1|1, 41|38|38|38)*

> *“To me It didn’t make sense. It basically went against everything I was taught, but I tried it.” (P86, 1|2|1|2, 38|35|36|36)*

Finally, there were many who were skeptical of the tip and chose not to follow it initially. However, for some of these workers, tips still played important roles as highlights making those options more salient for the workers’ own sensemaking and testing processes, or simply as something to try when nothing else worked. Like Schelling points [37], they provided focal points that made those options more prominent compared with other options in the environment. Workers said:

> *“I thought that it was kind of suspicious at first but as I was figuring the game out myself, I thought that it was correct.” (P195, 2|2|2|2, 35|40|34|34)*

> *“I did not listen to the tip the first two times since he takes more ticks but noticed when I incorporated it, I was more efficient” (P52, 1|1|2|2, 38|38|36|34)*

> *“At first I didn’t follow it because it seemed counter intuitive since they’re slow. But then I had trouble, so I tried it and came out ahead.” (P5, 1|1|2|2, 38|38|34|34)*

4.3 Barriers to adoption: counterintuitive, hard to implement, bad outcomes, lack of clarity, hard to track, and misaligned incentives

Workers described many barriers that kept them from using the tip or to using it successfully. One of the most common barriers expressed was that the tip was *counterintuitive* and did not make sense logically, which caused many to not follow the tip initially (and as will be seen later, the counterintuitive nature of the tip also compounded some of the later barriers):

> *“The first round, I ignored it because I knew the sous chef would do it quicker.” (P229, 1|2|2|2, 36|34|34|34)*

> *“At first I didn’t follow it because it seemed counter intuitive since they’re slow.” (P5, 1|1|2|2, 38|38|34|34)*

> *“I thought it didn’t make sense. The server took 12 ticks to cook, so I had them only cook once because the sous-Chef could finish in 8.” (P156, 2|1|1|1, 38|38|38|38)*

Other workers talked about how it was *hard to implement*, which seemed to relate to it being counter intuitive. Because tips need to be implemented within the context of a broader strategy, workers had to develop some sense of how it worked to apply it:

> *“I had a difficult time incorporating it and using it to my advantage. It always felt like the server took longer than needed when I could have had them doing other tasks.” (P167, 1|2|2|1, 38|35|40|38)*

> *“I tried to incorporate it into my strategy but somewhere along the way I got lost.” (P96, 2|1|1|2, 37|38|38|40)*

> *"It wasn't as useful as the tip in the first three rounds. I didn't really know how to implement it into my own strategy, or what it really implied." (P132, 1|2|1|1, 38|40|36|38)*

A third challenge was that incorporating the tip (by having the server cook twice) could result in worse outcomes. These *bad outcomes* caused people to abandon the tip:

> *"I tried it the first time, but I don't think it was a good tip, so I ignored it the next times." (P128, 2|1|1|1, 41|38|38|38)*

> *"The time I tried to incorporate the tip, I used more ticks than when I ignored it." (P42, 1|2|1|1, 38|40|36|38)*

> *"I let the server cook twice in the last couple of rounds and it didn't work well. If the game had continued I would have let the server only cook once." (P101, 1|1|2|2, 36|36|38|39)*

The previously described 3 barriers were the most commonly expressed, but there were also other barriers revealed in our analysis. For example, a few people felt there was a *lack of clarity* regarding what the tip actually meant concretely. They said:

> *"I wasn't sure what it meant. Does chopping count as cooking?" (P133, 0|1|1|1, 42|36|38|41)*

> *"I thought it was a little too broad, but maybe I'm just stupid because I could not figure out how to finish in less than 40 ticks." (P27, 1|2|2|2, 40|41|40|41)*

> *"I was really confused about this tip, I wasn't sure what it meant by let the server cook twice. I did this and it did not really help me, but maybe I misinterpreted the tip." (P135, 1|1|2|1, 38|39|36|39)*

Another barrier was that it was sometimes *hard to track* how many times the server had cooked so they did not know whether they had implemented the tip or not. This is a variant of the 'hard to implement' barrier, but unlike those quotes where participants focused on the challenge of getting it to work logically, this participant encountered more of a logistical challenge:

> *"It was confusing, I couldn't keep track of if he cooked or not" (P88, 1|1|1|1, 38|36|36|35)*

Finally, one participant touched on *misaligned incentives* in that trying to figure out how to implement the tip could result in lower short-term compensation:

> *"i didn't like it because i believe it took me longer to finish and i didnt receive any bonuses in those weeks" (P225, 1|2|1|2, 38|48|39|39)*

5 STUDY 2: QUANTIFYING AND RELATING INTENT, VIEWS, BARRIERS, AND PERSONALITY

5.1 Method: Quantitative Study Design and Analysis

Our qualitative study was naturally limited due to its reliance on short text responses to a single broad survey question. This meant that many responses were not relevant to our research questions or were not rich enough to add in interesting ways. Because of this, we conducted a follow-up study in which we sought to validate and quantify some of the observations elicited in our qualitative study and to surface relationships between tip views, barriers, and intent to use tips. Additionally, because we saw hints of the influence of personality on strategies and experiences (e.g. "logicians" emphasizing reasoning about why strategies work versus "experimenters" emphasizing trial and error), we added an exploration of personality into our follow-up study.

There were three new sets of questions that we added to the survey. First, we asked workers about their intent to use tips before each of the four disrupted rounds (Rounds 3-6): *"Do you intend to follow the tip in the coming round?"*[1]. Second, we asked three single-select matrix questions eliciting agreement to various views and experiences on a 5-point Likert-scale from "Strongly disagree" to "Strongly agree": (1) *"To what extent do you agree or disagree with the following statements about how you viewed the tip when you first received the tip going into Round 3?"*[2], (2) *"To what extent do you agree or disagree with the following statements about your thoughts about the tip evolved over the course of Rounds 3-6?"*[3], and (3) *"To what extent do you agree or disagree with the following statements about barriers that kept you from adopting the tip at any point within Rounds 3-6?"*[4]. Finally, we added three questions to elicit problem-solving style, based on a simplification of the VIEW problem-solving style assessment [39] in which we displayed a summarizing figure for each of the problem-solving dimensions (**Figure 2**) and asked workers to rate themselves along a scale between the two opposing styles for that dimension, e.g. *"Choose a number between 1 and 7 that indicates how well the description of 'Explorers' or 'Developers' in the figure above describes your orientation to change. For example, a "7" would indicate that you are described very well by 'Developers', while a "5" would indicate that you are better described by 'Developers' than 'Explorers', but still somewhat described by 'Explorers'."*

The VIEW problem-solving style assessment is a validated instrument that measures three dimensions integrating learning style, cognitive style, and psychological type stemming from psychology of the person and creative problem solving. The three dimensions are Orientation to Change, Manner of Processing, and Ways of Deciding, each of which have two opposing types (see **Figure 2**) that each correlate to different combinations of personality [17]. VIEW has been applied in particular to studying creative problem solving [16, 43] and to understanding and using individual differences in experiential learning [42, 43]. In consulting practices, VIEW has been used as a tool to support diverse needs including improving problem solving, communicating effectively, enhancing personal productivity, providing and receiving feedback, facilitating groups, managing change, developing leadership, designing instruction, building teams, and improving coaching and mentoring [38]. The instrument consists of 18 questions for Orientation to Change, 8 for Manner of Processing and 8 for Ways of Deciding, for a total of 34 items. As described in our Limitations section (**Section 6**), we created the simplified version due to practical constraints of our experiment already being too lengthy due to numerous rounds of gameplay.

We deployed the full experiment with these newly added survey questions to an additional 60 participants, again in the optimal *Algorithm* condition. We then investigate correlations between participants' views on the provided tips and their intent to follow the tips, success in following the tip, and performance using Pearson correlational analysis. In what follows, we highlight the significant findings along with corresponding figures for correlation matrices (with numbers shown for $p < 0.05$) and tables detailing p-values for significant entries.

5.2 Quantifying views on tips and barriers to adoption and relating them to tip intent and performance

5.2.1 Prevalence of tip views and barriers experienced. **Figures 3–4** depict how participants responded to questions regarding their views on tips and the barriers they experienced, respectively. The fraction of participants who responded

[1] Options were *"Yes, I intend to follow the tip this round"*, *"Maybe, I'll keep it in mind and see"*, *"No, I don't plan to follow the tip this round"*, and *"Other"*.

[2] Rows were: *"I viewed the tip positively as a rule I had to figure out how to follow"*, *"I viewed the tip positively as a hint in the right direction, but not required"*, *"I viewed the tip neutrally as an option to consider trying at some point"*, *"I viewed the tip negatively as likely flawed, but still planned to try it"*, and *"I viewed the tip negatively as likely flawed, so did not intend to follow it"*.

[3] Rows were: *"My view of the tip became more negative by the end"*, *"My view of the tip became more positive by the end"*, *"I went back and forth on whether to use the tip"*, and *"The tip highlighted new ideas I would not have thought of otherwise"*.

[4] Rows were: *"The tip felt counterintuitive"*, *"It was difficult to figure out how to implement the tip"*, *"Trying to follow the tip resulted in bad outcomes"*, *"I wasn't sure what the tip actually meant"*, *"I lost track of how many times the server cooked"*, and *"I was worried that exploring the tip would impact my payment in initial rounds"*

Preferences for Orientation to Change

Explorers

- Break away from the system, and redefine the problem
- View structure as limiting, confining
- May challenge authority, "bend" rules
- Emphasize originality and uniqueness—"ideas that stretch us"
- Press for extensive change and commitment to action
- Know newest trends and possibilities
- Ingenious and Unconventional
- Spontaneous and free-flowing
- Emphasize starting new tasks and the "big picture," often resist closure
- Produce ideas that others may not understand easily
- Individualistic, trust own judgment
- May emphasize generating

Developers

- Stay within existing paradigm or system, follow rules and procedures as given
- Find benefits and support in structure
- Emphasize improvement and usefulness
- Focus on gradual, incremental change
- Emphasize finding "just enough" new ideas
- Resourceful
- Dependable and Consistent
- Precise, Thorough, Efficient
- Good (Early) Planning and Organizing
- Emphasize thorough completion of tasks and attention to details, seek closure
- Know how to get their ideas accepted by others
- Look to authorities for guidance
- May emphasize focusing

Choose a number between 1 and 7 that indicates how well the description of 'Explorers' or 'Developers' in the figure above describes your orientation to change.

For example, a "7" would indicate that you are described very well by 'Developers', while a "5" would indicate that you are better described by 'Developers' than 'Explorers', but still somewhat described by 'Explorers'.

Preferences for Manner of Processing

When processing options or possibilities…

A person who prefers to process *Externally…*

- Starts talking about options right away
- Puts ideas out tentatively, ready to revise and reformulate along the way
- Derives energy from interaction with others (finds reflection challenging)
- Urges immediate action— "we can tune it up later if we need to"
- Shares ideas freely with a broad range of other people
- Seeks a great deal of input from others before reaching closure
- May often seem to be "bubbling over" with energy, rushing into things; impatient if asked to think too long about things

A person who prefers to process *Internally…*

- Wants time to think about options before discussing them
- Shares ideas with others after time to polish them– "when they're ready"
- Draws energy from reflection and consideration (challenged by feeling pressured to talk about ideas quickly)
- Seeks action after giving it careful consideration
- Shares ideas with others after establishing trust and confidence
- Builds ideas personally, then seeks feedback selectively
- Seems quiet, might be perceived as pensive or withdrawn; impatient if pressed to share too soon

Choose a number between 1 and 7 that indicates how well the description of 'Processing externally' or 'Processing internally' in the figure above describes how you process options or possibilities.

Preferences for Ways of Deciding

When making decisions about options or possibilities…

***People* preference involves giving one's primary attention to:**

- What will promote harmony and positive interpersonal relationships
- Criteria that are personal, sensitive to people's feelings, more subjective
- What's good, attractive, or pleasing about an option (which may not mean they really like it!)
- Considering the personal or inter-personal impact or consequences of a decision
- Seeking a solution or decision that all concerned can buy into

***Task* preference involves giving one's primary attention to:**

- What's logical or rational
- Criteria they consider objective, authoritative, and verifiable
- What's wrong with an option, what it lacks, or what it needs (which may not mean they really dislike it!)
- Considerations of standards, rigor, or quality
- "Letting the chips fall where they may…"
- Seeking the best solution or response, and being able to defend or justify the choice or decision

Choose a number between 1 and 7 that indicates how well the description of 'People preference' or 'Task preference' in the figure above describes how you make decisions about options or possibilities.

Fig. 2. Our simplfied elicitation of problem-solving styles using summaries of the three dimensions drawn from [39]

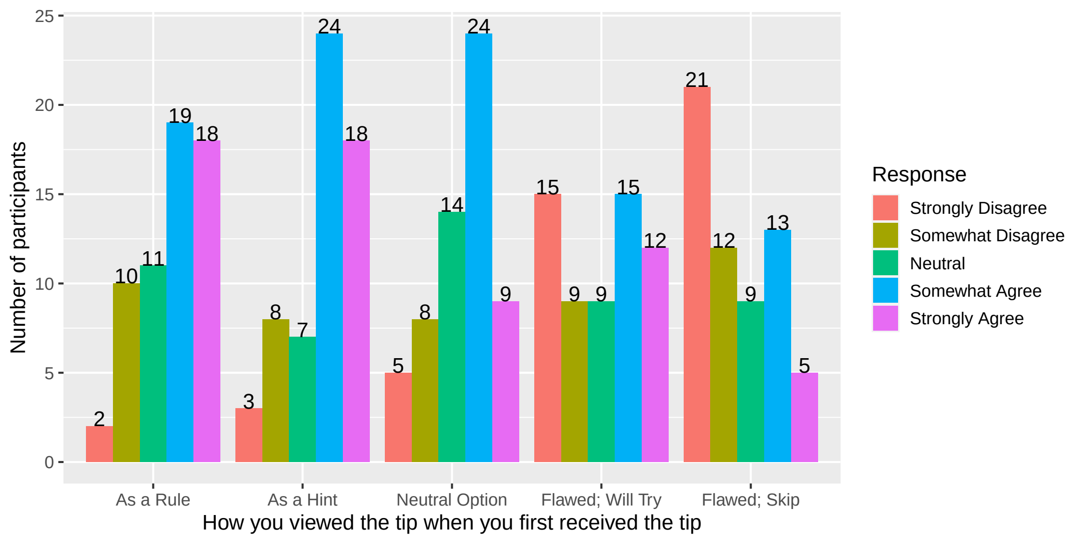


Fig. 3. Distributions of participants' responses to questions regarding their views on the provided tip when they first received it.

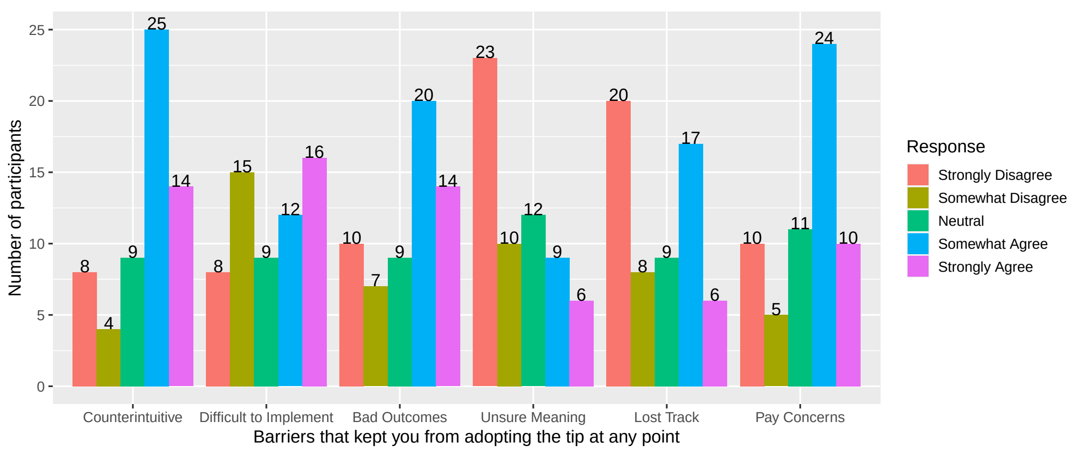


Fig. 4. Distributions of participants' responses to questions regarding the barriers that kept them from adopting the tip at any point.

As a Rule	As a Hint	Neutral Option	Flawed; Will Try	Flawed; Skip
70%	61.67%	55%	45%	30%

Table 1. The fraction of participants who responded "Somewhat agree" or "Strongly agree" for each of the questions regarding their views on the provided tip when they first received it.

Counterintuitive	Difficult to Implement	Bad Outcomes	Unsure Meaning	Lost Track	Pay Concerns
65%	46.67%	56.67%	25%	38.33%	56.67%

Table 2. The fraction of participants who responded "Somewhat agree" or "Strongly agree" for each of the questions regarding regarding the barriers that kept them from adopting the tip at any point.

"Somewhat agree" or "Strongly agree" for each of these questions is reported in Tables 1–2. As can be seen, all five views of types were common across participants, with a greater tendency towards initial views that were positive or neutral. For those who had negative reactions to tips (responded "Somewhat Agree" or "Strongly Agree" to *"I viewed the tip negatively as likely flawed, but still planned to try it"*, and *"I viewed the tip negatively as likely flawed, so did not intend to follow it"*), 57% and 47% respectively expressed that they "Somewhat Agree" or "Strongly Agree" with *"The tip highlighted new ideas I would not have thought of otherwise"*, validating the observation that tips can provide benefits to workers even when they are initially rejected, e.g. through highlighting areas of the solution space that can inform worker sense-making as seen in the qualitative results. All the identified barriers were experienced by a non-trivial portion of participants, though the first three barriers *counterintuitive*, *difficult to implement*, and *bad outcomes*, and the last one *pay concerns* were particularly prevalent.

5.2.2 The challenge of operationalizing tips. The challenge of operationalizing tips can also be seen in several analyses. For example, an intention to follow the tip only significantly correlated with actual compliance to the tip (successfully having the server cook twice, though not necessarily in an optimal way) in Rounds 5 ($r(58) = 0.29, p = 0.0267$) and 6 ($r(58) = 0.32, p = 0.0126$). When looking at performance, one can see that the completion time is spread out over a broad range of values when looking at those expressing an intent to follow the tip the 1st, 2nd, 3rd, and even the 4th time (see **Figure 5a**). And even when one succeeds in complying with the tip (having the server cook twice), one can see that there is still a significant range of performance values that can result (see **Figure 5b**). Despite "server cooks twice" being the "optimal" tip in terms of being the best for bringing human strategies closer to the optimal strategy, the tip can also result in people taking many non-performing paths.

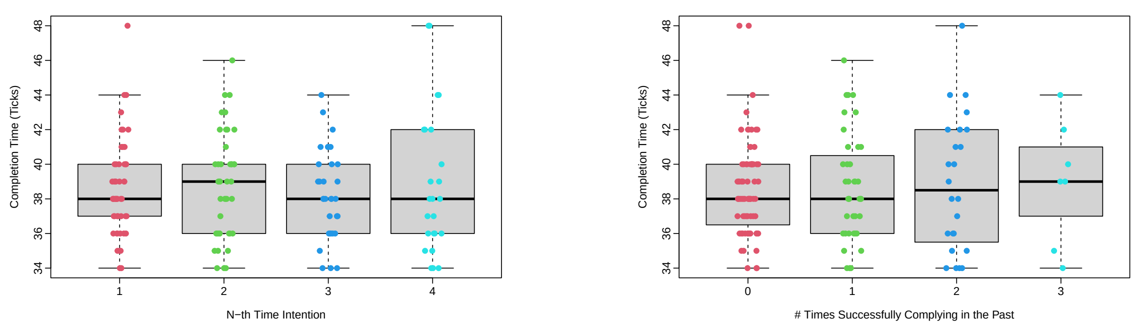


Fig. 5. Performance (completion time) across participants by how many times they (a) expressed an intent to follow the tip, or (b) succeeded in complying with the tip

5.2.3 What barriers impact tip adoption the most. Examining correlations between barriers and tip adoption show that the three barriers expressed as most prevalent (*counterintuitive*, *difficult to implement*, and *bad outcomes*) were also the only ones with significant (negative) correlations to worker intent to follow tips in Rounds 4-6 (all but the initial round) all at $p < 0.05$ (see **Figure 6**). Table 3 reports their correlations and p-values. Experiencing bad outcomes in particular had a strong impact in turning workers away from the tip.

5.2.4 How view of tips impacts barriers experienced. Examining correlations between view of tips and experienced barriers shows that those who have more negative views of the tips going into Round 3 (perhaps unsurprisingly) are

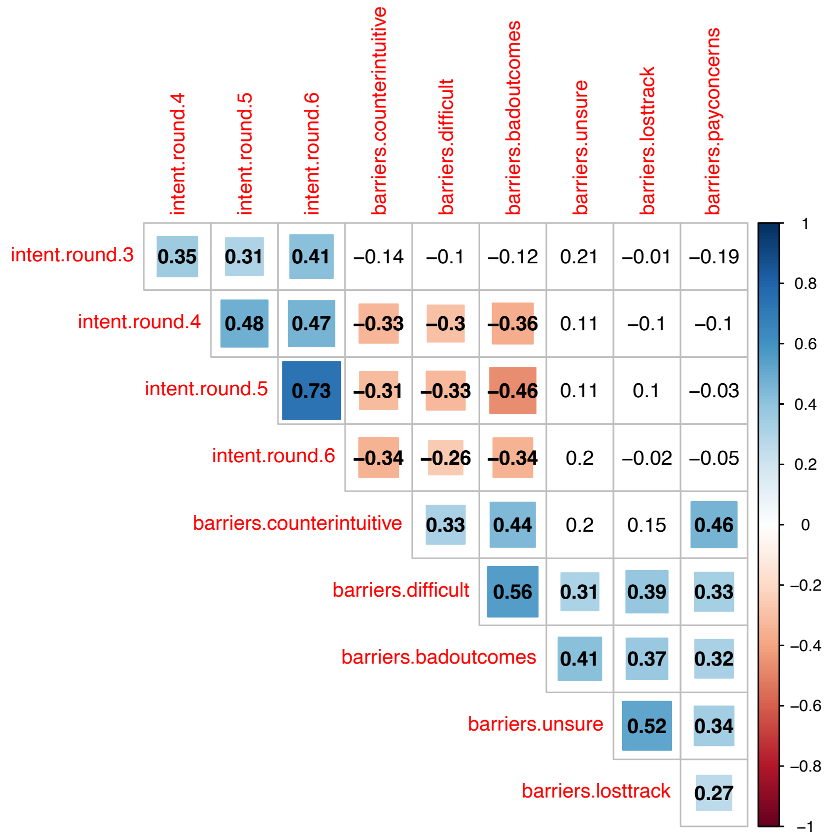


Fig. 6. Correlations among worker intent to follow the tip in Rounds 3 through 6 and barriers. Only statistically significant correlations at $p = 0.05$ are reported.

	Intent (Round 3)	Intent (Round 4)	Intent (Round 5)	Intent (Round 6)
Counterintuitive	-0.14 ($p = 0.3357$)	-0.33 ($p = 0.0094$)	-0.31 ($p = 0.0143$)	-0.34 ($p = 0.0071$)
Difficult to Implement	-0.10 ($p = 0.5201$)	-0.3 ($p = 0.0215$)	-0.33 ($p = 0.0102$)	-0.26 ($p = 0.0488$)
Bad Outcomes	-0.12 ($p = 0.4336$)	-0.36 ($p = 0.0042$)	-0.46 ($p = 0.0002$)	-0.34 ($p = 0.0074$)

Table 3. Correlations and p-values among worker intent to follow the tip in Rounds 3 through 6 and the three barriers (*counterintuitive*, *difficult to implement*, and *bad outcomes*).

more likely to express diverse things being barriers to them adopting tips. For example, those who said "I viewed the tip negatively as likely flawed, so did not intend to follow it" were more impacted by bad outcomes when they occurred ($r(58) = 0.32, p = 0.0135$) and also more likely to express "I wasn't sure what the tip actually meant" as a barrier ($r(58) = 0.37, p = 0.0035$). Those who said "I viewed the tip negatively as likely flawed, but still planned to try it" were similarly impacted by bad outcomes when they occurred ($r(58) = 0.32, p = 0.0128$) but not having the unsure meaning of the tip expressed as a barrier.

5.3 The influence of problem-solving style on tip views, barriers, and tip adoption

Figure 7 illustrates the correlations among problem-solving style on tip views and our measured outcomes. Our results also indicate that there are indeed relationships between personality (specifically, problem-solving style) and the ways people view and experience tips. In particular, "Developer" types in the "Orientation to Change" dimension had many significant correlations to other measured outcomes. They were more likely to express that "I viewed the tip negatively as likely flawed, but still planned to try it" going into Round 3 (the first disrupted round) ($r(58) = 0.26, p = 0.0436$), more likely to comply with the tip in the first disrupted round (though complying does not mean doing so in an optimal or high-performing way) ($r(58) = 0.3, p = 0.0219$), less likely to find the counterintuitive nature of tips ($r(58) = -0.28, p = 0.0320$) or the short-term loss to their payment ($r(58) = -0.38, p = 0.0031$) to be a barrier to use, and more likely to have a positive sentiment of tips at the end ($r(58) = -0.38, p = 0.0031$). This seems consistent with the description of Developers as people who "follow rules and procedures as given".

We also saw that those with a "Task Preference" in the "Ways of Deciding" dimension were more likely to view the tip as a rule ($r(58) = 0.39, p = 0.0022$) and not a flawed strategy ($r(58) = -0.38, p = 0.0165$). This also seems consistent with the description of those with Task Preferences as "logical or rational".

We note that these results are only a preliminary investigation (since we used a simplified version of the VIEW assessment tool) that point to the value of further investigations into the role of personality in human use of tips, as we will discuss further in the Discussion (**Section 7**).

6 LIMITATIONS

Our analysis provides a richer picture of the diverse ways in which workers view and use tips and the barriers the prevent the tips from being useful for workers. We note, however, that some of our findings may be specific to our unique type of problem-solving task. Specifically, our study considered a setting in which workers sought to achieve a *known* minimal duration for a sequence of actions (a *search problem*) with tips taking the form of *constraints* that an action sequence should satisfy (rather than, for example, specific actions recommended for specific points in time). Tips also came from *unknown sources* (workers were not told whether tips were from humans or algorithms). These caveats are important to note, but we also believe that these attributes are similar to many problem solving contexts in industry. As described in the Introduction, managers are often engaged in optimizing a complex sets of decisions towards known industry benchmarks.

Another limitation of our paper is that our study of correlations to problem solving style only used a very simplified version of VIEW in which participants self-report where they think they fall on each given dimension based on descriptions of the two opposing types. This was a practical constraint due to the length of our experiment (in which participants engage in training rounds, then 2 rounds of gameplay with all 3 workers, then 4 rounds of gameplay in the disrupted 2-worker setting) which made it impractical to add the 34 survey items making up the full VIEW assessment.

7 DISCUSSION AND FUTURE DIRECTIONS

With the limitations and caveats described in **Section 6** in mind, we discuss in this section implications of our findings and directions for future study.

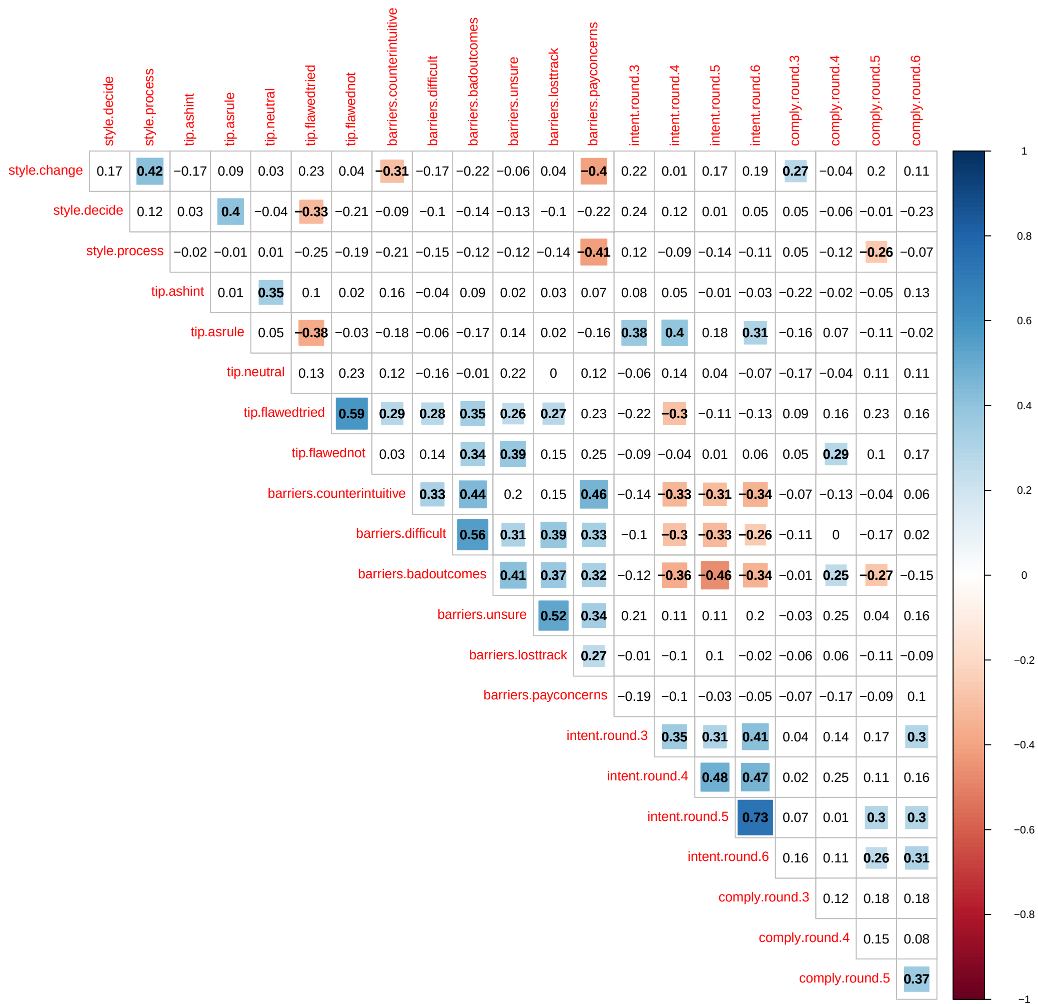


Fig. 7. Correlations among problem-solving style, tip views, barriers, and tip adoption. Only statistically significant correlations at $p = 0.05$ are reported.

7.1 Diverse reasons for lack of trust

First, we saw that there were diverse reasons why workers did not trust the algorithmic tip. Workers felt that the tip was counterintuitive or led them to bad outcomes. These suggest different solutions to supporting workers worth investigating in follow-up studies. For example, one might look for ways to provide more explainable tips that are less counterintuitive. But one could also try to increase worker confidence of the tip’s value by citing statistics of others who vouch for it or by showing that others also encountered bad outcomes on the path to implementing the tip successfully. Those with more optimistic views of tips going in seem to be less likely to express obstacles encountered as barriers.

We also found it interesting that people who ignored or discounted tips could still benefit from them and that tips could provide value through being highlights for making certain directions more salient. So if one is not able to provide workers with more optimistic views of perhaps, then perhaps one can think about new types of tips that aren't necessarily trying to convince people of a specific action to take, but are simply trying to support by shining a light on information or strategies that may be subtle or counterintuitive.

The fact that the tip could lead people down worse paths also raises questions about what it means for a tip to be "optimal". For example, while the server needed to cook twice to achieve the optimal duration of "34", it was also possible to get to a duration of "35" with the server cooking once. If the former case had a much higher likelihood of resulting in long durations than the latter case and may take a much longer time to figure out, then it could be true that the average payout for "server cooks once" could actually be higher, especially in the short-term. It may be worth investigating algorithms that generate tips that are not only good at bringing the human strategy closer to the optimal strategy, but also less likely to lead people down wrong paths.

7.2 Trust was not the only obstacle to benefiting from tips

We also saw that lack of trust was not the only obstacle preventing people from benefiting from tips. Several of the barriers we observed related to the "usability" of the tip with workers finding the tip lacked clarity, was hard to implement, or was just hard to even track whether or not they had used it. This partially related to the fact that the tip was a constraint on a complex and interdependent sequence of actions which workers were not always able to foresee in advance. While these were not the biggest factors impacting tip adoption, they were still experienced by non-trivial numbers of participants. These suggest solutions such as creating feedback or communication mechanisms for raising questions about clarity, creating more actionable tips guiding implementation, and creating tools for workers to track relevant statistics or map complex spaces.

We also saw that broader environmental factors could shape people's use of tips such as the incentives that workers are working within and whether those incentives encourage and support learning and exploration. This is an important reminder that it is not enough to design at the level of the human-ai interactions themselves. Agent interactions cannot be divorced from the broader organizational, community, or societal contexts in which the interactions are situated.

7.3 Personalities and collaboration in AI-assisted problem solving

Our qualitative findings seemed to reflect different approaches that people take to problem solving. Some people are logicians emphasizing reasoning about why strategies work. Others are experimenters emphasizing trial and error. These different approaches may affect how open people are to trying out counterintuitive tips, the barriers that affect a tips usefulness to them, and the design interventions that would be effective for encouraging tip use.

Our quantitative studies indicated that "Orientation to Change" may be particularly related to views of tips and the experience of barriers. It would be interesting to study this further. For example, how might teams of people with different combinations of these personality types use tips? What implications might this have on forming teams or using collaborative interactions to facilitate more effective use of algorithmic tips? We see this as a particularly interesting direction because collaboration might be an effective way to help workers reason about counterintuitive tips or to figure out what a tip means and how to implement it.

8 CONCLUSION

In this paper, we undertook a mixed methods approach to develop a deeper understanding of why people choose not to adopt “optimal” algorithmic tips in problem solving contexts. We found that even when people choose to not follow tips initially, they can still benefit from them over time. We found that barriers to adoption included trust, but also extend to tip usability and broader environmental factors. We identified three barriers most correlated to not following tips (*counterintuitive*, *hard to implement*, and *bad outcomes*), found that those with more negative initial views of tips were more likely to be impacted by bad outcomes, and that the “Orientation to Change” dimension of problem-solving styles had several correlations to people’s views of tips and barriers experienced. Our analysis add new perspectives to AI-assisted decision-making by extending beyond forecasting to study sequential decision-making and problem-solving contexts. It also raises a number of questions that we view as promising directions for future research. How might one design algorithms that optimize the usability or implementability of tips? How might the influence of personality on tip adoption be used to design collaborative interactions that facilitate better use of algorithmic tips? With the rapidly increasing use of AI across society, it is important to continue developing a richer view of human-AI interaction that can inform more human-centric development of AI algorithms.

REFERENCES

[1] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2022. Machine Bias. In *Ethics of Data and Analytics*. Auerbach Publications, Boca Raton, 254–264.

[2] Yochanan E Bigman and Kurt Gray. 2018. People are averse to machines making moral decisions. *Cognition* 181 (2018), 21–34.

[3] Eric Bogert, Aaron Schecter, and Richard T Watson. 2021. Humans rely more on algorithms than social influence as a task becomes more difficult. *Scientific reports* 11, 1 (2021), 1–9.

[4] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1 (April 2021), 1–21.

[5] Felipe Caro and Anna Sáez de Tejada Cuenca. 2018. *Believing in analytics: Managers adherence to price recommendations from a DSS.* Technical Report. Working Paper.

[6] Noah Castelo, Maarten W Bos, and Donald R Lehmann. 2019. Task-dependent algorithm aversion. *Journal of Marketing Research* 56, 5 (2019), 809–825.

[7] Kathy Charmaz and J Smith. 2003. Grounded theory. *Qualitative psychology: A practical guide to research methods* 2 (2003), 81–110.

[8] Hao-Fei Cheng, Ruotong Wang, Zheng Zhang, Fiona O’Connell, Terrance Gray, F Maxwell Harper, and Haiyi Zhu. 2019. Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI ’19, Paper 559)*. Association for Computing Machinery, New York, NY, USA, 1–12.

[9] Maria De-Arteaga, Riccardo Fogliato, and Alexandra Chouldechova. 2020. A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12.

[10] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. 2015. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General* 144, 1 (2015), 114.

[11] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. 2018. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science* 64, 3 (2018), 1155–1170.

[12] Mary T Dzindolet, Linda G Pierce, Hall P Beck, and Lloyd A Dawe. 2002. The perceived utility of human and automated aids in a visual detection task. *Human factors* 44, 1 (2002), 79–94.

[13] Emir Efendić, Philippe PFM Van de Calseyde, and Anthony M Evans. 2020. Slow response times undermine trust in algorithmic (but not human) predictions. *Organizational Behavior and Human Decision Processes* 157 (2020), 103–114.

[14] Gavan J Fitzsimons and Donald R Lehmann. 2004. Reactance to recommendations: When unsolicited advice yields contrary responses. *Marketing Science* 23, 1 (2004), 82–94.

[15] M Sinan Gönül, Dilek Önkal, and Michael Lawrence. 2006. The effects of structural characteristics of explanations on use of a DSS. *Decision support systems* 42, 3 (2006), 1481–1493.

[16] John C; Selby Houtz. 2009. Problem Solving Style, Creative Thinking, and Problem Solving Confidence. *Educational Research Quarterly; West Monroe* 33, 1 (Sept. 2009), 18–30.

[17] Scott G Isaksen, Astrid H Kaufmann, and Bjørn T Bakken. 2016. An examination of the personality constructs underlying dimensions of creative problem-solving style. *J. Creat. Behav.* 50, 4 (Dec. 2016), 268–281.

[18] Timothy A Judge, Edwin A Locke, Cathy C Durham, and Avraham N Kluger. 1998. Dispositional effects on job and life satisfaction: the role of core evaluations. *Journal of applied psychology* 83, 1 (1998), 17.

[19] Kohei Kawaguchi. 2021. When will workers follow an algorithm? A field experiment with a retail business. *Management Science* 67, 3 (2021), 1670–1695.

[20] Matthieu Komorowski, Leo A Celi, Omar Badawi, Anthony C Gordon, and A Aldo Faisal. 2018. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nature medicine* 24, 11 (2018), 1716–1720.

[21] I Kononenko. 2001. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif. Intell. Med.* 23, 1 (Aug. 2001), 89–109.

[22] Johannes Kunkel, Tim Donkers, Lisa Michael, Catalin-Mihai Barbu, and Jürgen Ziegler. 2019. Let Me Explain: Impact of Personal and Impersonal Explanations on Trust in Recommender Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19, Paper 487)*. Association for Computing Machinery, New York, NY, USA, 1–12.

[23] Min Hun Lee, Daniel P Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. 2021. A human-ai collaborative approach for clinical decision making on rehabilitation assessment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[24] Jennifer M Logg, Julia A Minson, and Don A Moore. 2019. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes* 151 (2019), 90–103.

[25] Poornima Madhavan and Douglas A Wiegmann. 2007. Effects of information source, pedigree, and reliability on operator interaction with decision support systems. *Human factors* 49, 5 (2007), 773–785.

[26] Hasan Mahmud, AKM Najmul Islam, Syed Ishtiaque Ahmed, and Kari Smolander. 2022. What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change* 175 (2022), 121390.

[27] Keri Mallari, Kori Inkpen, Paul Johns, Sarah Tan, Divya Ramesh, and Ece Kamar. 2020. Do i look like a criminal? examining how race presentation impacts human judgement of recidivism. In *Proceedings of the 2020 Chi conference on human factors in computing systems*. 1–13.

[28] Reid McIlroy-Young, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. 2020. Aligning superhuman ai with human behavior: Chess as a model system. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1677–1687.

[29] Reid McIlroy-Young, Yu Wang, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. 2021. Detecting Individual Decision-Making Style: Exploring Behavioral Stylometry in Chess. *Advances in Neural Information Processing Systems* 34 (2021), 24482–24497.

[30] Kathleen L Mosier, Linda J Skitka, Susan Heers, and Mark Burdick. 2017. Automation bias: Decision making and performance in high-tech cockpits. In *Decision Making in Aviation*. Routledge, 271–288.

[31] Paweł Niszczota and Dániel Kaszás. 2020. Robo-investment aversion. *Plos one* 15, 9 (2020), e0239277.

[32] Bureau of Transportation Statistics. 2021. Airline on-time performance and causes of flight delays. https://www.bts.gov/explore-topics-and-geography/topics/airline-time-performance-and-causes-flight-delays

[33] Dilek Önkal, Paul Goodwin, Mary Thomson, Sinan Gönül, and Andrew Pollock. 2009. The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making* 22, 4 (2009), 390–409.

[34] Forough Poursabzi-Sangdeh, Daniel G Goldstein, Jake M Hofman, Jennifer Wortman Wortman Vaughan, and Hanna Wallach. 2021. Manipulating and Measuring Model Interpretability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21, Article 237)*. Association for Computing Machinery, New York, NY, USA, 1–52.

[35] Cynthia Rudin. 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence* 1, 5 (2019), 206–215.

[36] James Schaffer, John O'Donovan, James Michaelis, Adrienne Raglin, and Tobias Höllerer. 2019. I can do better than your AI: expertise and explanations. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 240–251.

[37] Thomas C Schelling. 1980. *The Strategy of Conflict: With a New Preface by the Author.* Harvard University Press.

[38] Edwin C Selby, Donald J Treffinger, and Scott G Isaksen. 2021. Applying View.

[39] Edwin C Selby, Donald J Treffinger, Scott G Isaksen, and Kenneth J Lauer. 2004. Defining and assessing problem-solving style: Design and development of a new tool. *J. Creat. Behav.* 38, 4 (Dec. 2004), 221–243.

[40] Jan-Philipp Stein, Markus Appel, Alexandra Jost, and Peter Ohler. 2020. Matter over mind? How the acceptance of digital entities depends on their appearance, mental prowess, and the interaction between both. *International Journal of Human-Computer Studies* 142 (2020), 102463.

[41] Sharifa Sultana, Md Mobaydul Haque Mozumder, and Syed Ishtiaque Ahmed. 2021. Chasing Luck: Data-driven Prediction, Faith, Hunch, and Cultural Norms in Rural Betting Practices. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–17.

[42] Treffinger and Selby. 2004. Problem solving style: A new approach to understanding and using individual differences. *Korean Journal of Thinking and Problem Solving* (2004).

[43] Donald J Treffinger, Edwin C Selby, and Scott G Isaksen. 2008. Understanding individual problem-solving style: A key to learning and applying creative problem solving. *Learn. Individ. Differ.* 18, 4 (Oct. 2008), 390–401.

[44] Michael Yeomans, Anuj Shah, Sendhil Mullainathan, and Jon Kleinberg. 2019. Making sense of recommendations. *Journal of Behavioral Decision Making* 32, 4 (2019), 403–414.

[45] Ming Yin, Jennifer Wortman Vaughan, and Hanna Wallach. 2019. Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–12.

[46] Yunfeng Zhang, Q Vera Liao, and Rachel K E Bellamy. 2020. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 295–305.