

1 **Quadratic Survey Interface**

2
3 **ANONYMOUS AUTHOR(S)***

4
5 Quadratic Surveys (QS) elicit more accurate individual preferences than traditional surveys, such as Likert-scale surveys. However, the
6 cognitive load associated with QS has hindered its adoption in digital surveys for collective decision-making. We introduce a two-phase
7 “organize-then-vote” QS interface based on decision-making and preference construction theories designed to lessen the cognitive load.
8 Since interface design significantly impacts survey results and accuracy, our design scaffolds survey takers’ decision-making while
9 managing the cognitive load imposed by QS. In a 2x2 between-subject in-lab study on public resource allotment, we compared our
10 interface with a traditional text interface across QS with 6 (short) and 24 (long) options. Our interface reduced satisficing behaviors
11 arising from cognitive overload in long QS conditions. Participants using our interface in the long QSs shifted their cognitive effort
12 from mechanical operations to constructing more comprehensive preferences. This research clarifies how human-centered design
13 improves preference elicitation tools for collective decision-making.

14
15 CCS Concepts: • **Human-centered computing → Collaborative and social computing systems and tools; Collaborative and**
16 **social computing design and evaluation methods; User studies; HCI design and evaluation methods; Interactive systems**
17 **and tools; Empirical studies in interaction design.**

18
19 Additional Key Words and Phrases: Quadratic Voting; Survey Response Format; Collective decision-making; Preference Construction

20
21 **ACM Reference Format:**

22 Anonymous Author(s). 2024. Quadratic Survey Interface. *Proc. ACM Hum.-Comput. Interact.* 1, 1 (September 2024), 41 pages. <https://doi.org/10.1145/nmnnnnnn.nmnnnnnn>

23
24 **1 Introduction**

25
26 Designing effective survey interfaces is crucial for accurately capturing respondents’ preferences, which directly impact
27 the quality and reliability of the data collected. For instance, voice assistants have been used to gather user feedback [1],
28 and a recent Human-Computer Interaction (HCI) study highlights that certain survey response formats can increase
29 errors [2]. In this paper, we introduce **Quadratic Surveys (QS)**, a survey tool designed to elicit preferences more
30 accurately than traditional methods [3]. Despite the promise of QS, there has been no research on designing interfaces
31 to support its unique quadratic mechanisms [4], where participants must rank and rate items — a task that poses
32 significant cognitive challenges. To popularize QS and ensure high-quality data, this paper addresses the question: *How*
33 *can we design interfaces to support participants in completing Quadratic Surveys (QS) more effectively?*

34
35 We envision an effective interface that navigates participants through the complex mechanism and preference
36 construction process, tailored to the unique challenges of QS. QS improves accuracy in individual preference elicitation
37 compared to traditional methods like Likert scales by requiring participants to make trade-offs using a fixed budget
38 of credits, where purchasing k votes for an option in QS costs k^2 credits [5, 3]. This quadratic cost structure forces
39 respondents to carefully evaluate their preferences, balancing the strength of their support or opposition against the
40 limited budget. While this cost structure forces them to make thoughtful trade-offs, this complexity also increases

41
42
43
44
45 Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not
46 made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components
47 of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on
48 servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

49 © 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

50 Manuscript submitted to ACM

51
52 Manuscript submitted to ACM

53 cognitive load, making it mentally taxing to weigh costs, evaluate options, and construct rankings [6]. Moreover, QS,
54 often referred to as Quadratic Voting (QV) in voting scenarios, can involve hundreds of options [7, 8], increasing the
55 risk of cognitive overload and satisficing [9, 10, 11]. We propose that an effective interface tailored to support users
56 through these complex trade-offs can reduce cognitive load and facilitate better preference construction.
57

58 To date, existing quadratic mechanism-powered applications simply present options, allow vote adjustments and
59 automatically calculate votes, costs, and budget usage. These designs focused heavily on the mechanism, rather than
60 supporting possible challenges these application users faced. Survey interface literature, while addressing decision-
61 making and usability, most focus on traditional surveys that do not share the unique option-to-option trade-offs that QS
62 introduces [12, 13, 14, 15, 16, 2]. The challenges, in particular, cognitive load [17, 18, 15, 19] and scaffolding challenging
63 tasks [20, 21, 22] like preference construction, are opportunities where user interfaces have shown their promises. As
64 a result, it remains unclear how different interface designs might support QS in reducing cognitive load and aiding
65 preference construction.
66

67 After multiple design iterations, we propose a novel interactive two-phase “organize-then-vote” QS interface (two-
68 phase interface, Figure 1) that integrates three key elements. First, the interface scaffolds the preference construction
69 process by having participants initially categorize the survey options into “Positive,” “Neutral,” or “Negative.” This
70 serves as a cognitive warm-up, easing participants into the more complex QS voting task. Second, the interface arranges
71 the options according to these categorizations, providing a structured visual layout. Third, participants can refine the
72 positions of these options using drag-and-drop functionality, giving them greater control and agency in the preference-
73 construction process. These design features are aligned with preference construction theory and build upon prior
74 research in interface design to reduce cognitive load and enhance user engagement.
75

76 To explore how these interface elements mitigate the cognitive load and support preference construction in Quadratic
77 Surveys, we pose the following research questions:
78

- 79 • RQ1. How does the number of options in Quadratic Surveys impact respondents’ cognitive load?
- 80 • RQ2a. How does the two-phase interface impact respondents’ cognitive load compared to a text interface?
- 81 • RQ2b. What are the similarities and differences in sources of cognitive load across the two interfaces?
- 82 • RQ3. What are the differences in Quadratic Survey respondents’ behaviors when coping with long lists of
83 options across the two-phase interface and the text interface?

84 We invited 41 participants to a lab study comparing our two-phase interface with a baseline to understand how
85 different interface designs and option lengths (6 options or 24 options) impact cognitive load. Qualitative findings,
86 measured using the NASA Task Load Index (NASA-TLX) and semi-structured interviews, revealed that participants
87 using the two-phase interface experienced cognitive demand more from strategic, holistic thinking compared to personal
88 relevance and operational tasks, particularly in longer surveys. Quantitative results showed that, although participants
89 spent more time per option, they made faster decisions during the voting phase, suggesting a more efficient distribution
90 of cognitive effort. We concluded that the two-phase interface prevented cognitive overload in long QS surveys and
91 shifted mental load toward more strategic thinking, reducing reliance on mental shortcuts like satisficing [9].
92

93 *Contributions.* We contribute to the HCI community by proposing the first interface specifically designed for QS
94 and QV-like applications, aimed at reducing cognitive challenges and scaffolding preference construction through a
95 two-phase interface with direct manipulation. Before our work, no research had explored QS interfaces, particularly for
96 long surveys prone to cognitive overload. Few studies in HCI address interfaces for surveys and questionnaires. Our
97 study demonstrated how user interfaces can facilitate preference construction in situ and reduce satisficing behaviors
98

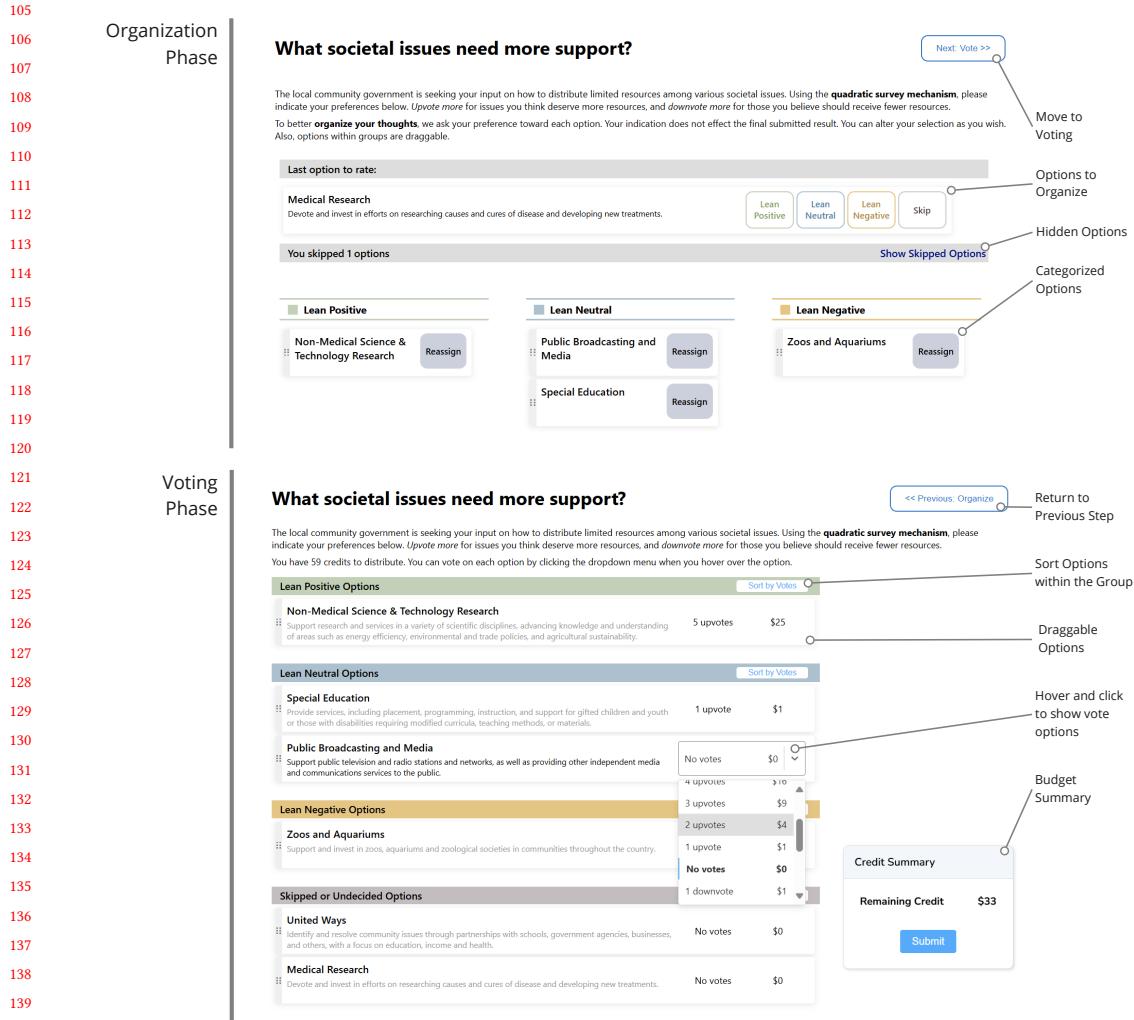


Fig. 1. The Two-Phase Interface: The interface consists of two phases. Survey respondents can navigate between phases using the top right button. In the organization phase, the interface presented one option at a time to the respondents where they choose among four choices: “Lean Positive”, “Lean Neutral”, “Lean Negative”, or “Skip”. Skipped options are hidden and can be evaluated later. The chosen options will be listed below. Items can be dragged and dropped across categories or returned to the stack. In the voting phase, options are listed in the order of the four categories. When hovering over each option, respondents can select a vote for that option using the dropdown. Each dropdown contains the cost associated with the vote. A sort button allows ascending sorting within each category. A summary box tracks the remaining credit balance.

by promoting incremental updates and deeper engagement with survey items through interface elements. Additionally, our interview is also the first in-depth qualitative analysis of user experiences with Quadratic Mechanism applications, identifying key factors contributing to cognitive load. The impact of our contribution extends beyond QS, offering design implications for other preference-elicitation tools in complex scenarios. By making QS easier to use and more accurate, our design also encourages wider adoption among researchers and practitioners. Finally, our work lays the

157 groundwork for future studies on qualitative insights and future interface designs of quadratic mechanisms, supporting
 158 decision-makers in eliciting accurate respondents' preferences.
 159

160 2 Related Work

162 This research lies at the intersection of three core areas: quadratic surveys, survey and voting interface design, and
 163 choice overload along with cognitive challenges. In this section, we review the related works in each of these areas.
 164

166 2.1 Quadratic Survey and the Quadratic Mechanism

167 We introduce the term **Quadratic Survey (QS)** to describe surveys that utilize the quadratic mechanism to collect
 168 individual attitudes. The **quadratic mechanism** is a theoretical framework designed to encourage the truthful revelation
 169 of individual preferences through a quadratic cost function [4]. This framework gained popularity through **Quadratic**
 170 **Voting (QV)**, also known as plural voting, which uses a quadratic cost function in a voting framework to facilitate
 171 collective decision-making [23].
 172

173 To illustrate how QS works, we formally define the mechanism: each survey respondent is allocated a fixed budget,
 174 denoted by B , to distribute among various options. Participants can cast n votes for or against option k . The cost c_k for
 175 each option k is derived as:
 176

$$178 \quad c_k = n_k^2 \quad \text{where} \quad n_k \in \mathbb{Z}$$

180 The total cost of all votes must not exceed the participant's budget:
 181

$$182 \quad \sum_k c_k \leq B$$

185 Survey results are determined by summing the total votes for each option:
 186

$$187 \quad \text{Total Votes for Option } k = \sum_{i=1}^S n_{i,k}$$

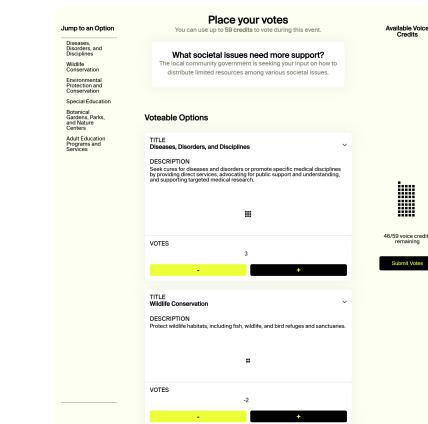
189 where S represents the total number of participants, and $n_{i,k}$ is the number of votes cast by participant i for option k .
 190 Each additional vote for each option increases the marginal cost linearly, encouraging participants to vote proportionally
 191 to their level of concern for an issue [24].
 192

193 QS adapts these strengths of the quadratic mechanism in *voting* to encourage truthful expression of preferences in
 194 *surveys*. Unlike traditional surveys that elicit either rankings *or* ratings, QS allows for *both*, enabling participants to cast
 195 multiple votes for or against options, incurring a quadratic cost. Cheng et al. [3] showed that this mechanism aligns
 196 individual preferences with behaviors more accurately than Likert Scale surveys, particularly in resource-constrained
 197 scenarios like prioritizing user feedback on user experiences.
 198

199 In recent years, empirical studies on QV have expanded into various domains [25, 26]. Applications based on the
 200 quadratic mechanism have also grown, including Quadratic Funding, which redistributes funds based on outcomes
 201 from consensus made using the quadratic mechanism [27, 28]. Recent work by South et al. [29] applies the quadratic
 202 mechanism to networked authority management, later used in Gov4git [30]. Despite the increasing breadth and depth
 203 of applications utilizing the quadratic mechanism, little attention has been paid to user experience and interface design,
 204 which support individuals in expressing their preference intensity. Our work aims to address this by designing interfaces
 205 supporting quadratic mechanisms.
 206

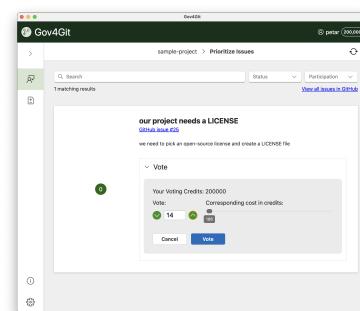


(a) Software by WeDesign, used in the first empirical QV research [5]. Little information is available about the software, except for an image from Posner and Weyl [24]. In the image, each prompt has thumbs up and down icons to update the vote in the center. The remaining budget appears as a progress bar at the top.

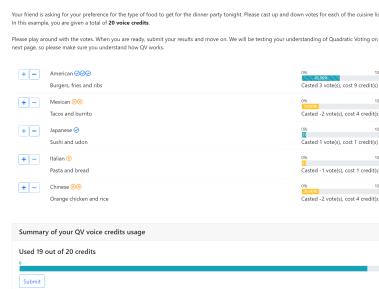


(b) An open-sourced QV interface [31] forked from GitCoin [32], used by the RadicalxChange community [33]. This interface presents total credits as small blocks. Votes are updated using plus and minus buttons, with numerical counts shown under each option and surface area as costs.

(c) An open-source QV interface [34] offers a publicly available service. Options show only the current number of votes, with credits displayed in the top right corner. This interface does not show the costs of votes but supports sorting options.



(d) The interface designed for gov4git [30] updates votes using arrows under each option, with the associated cost shown as a percentage bar to the right. A search bar exists for searching specific pull requests or issues.



(e) The interface used in the research by Cheng et al. [3] employs the most visual components. Icons depict the current number of votes, with progress bars signifying the current spending.

Fig. 2. Recent interface for applications using the quadratic mechanism.

Manuscript submitted to ACM

2.2 Design Implications from Surveys and Voting Interfaces

We began by examining existing QV applications, which share the same mechanism as QS, and identifying shared interface components. All QV interfaces generally include:

- Option list: A list of options for voting.
 - Vote controls: Buttons to increase or decrease votes for each option.
 - Individual vote tally: A display of the votes cast per option.
 - Summary: An auto-generated summary of costs and the remaining budget.

These components present options, calculate costs, and allow vote adjustments. However, these interfaces focus purely on mechanics without little understanding of voter's usability needs or offering cognitive support to help them complete the task effectively. In addition, the HCI community conducted few research [35, 36] on survey and questionnaire interfaces components, with more work focusing more on alternative input modalities like bots, voice, and virtual reality [37, 38, 39, 40]. Thus, we turn to marketing and research literature for insights into how digital survey interface elements can impact user behavior and usability.

Research in the marketing and research communities focusing on survey and questionnaire design, usability, and interactions examines the influence of presentation styles and ‘response format.’ Weijters et al. [41] demonstrated that horizontal distances between options are more influential than vertical distances, with the latter recommended for reduced bias. Slider bars, which operate on a drag-and-drop principle, show lower mean scores and higher nonresponse rates compared to buttons, indicating they are more prone to bias and difficult to use. In contrast, visual analog scales that operate on a point-and-click principle perform better [42]. These studies show how even small design changes can have a large impact on usability, highlighting the importance of designing interfaces that prioritize human-centered interaction rather than focusing solely on functionality.

Voting interfaces are a specialized type of survey interface that not only elicit individual choices but often have consequential impacts. For example, the butterfly ballot, an atypical design, may have influenced the outcome of the 2000 U.S. Presidential Election [43]. Research has shown that ballot interfaces can significantly influence democratic processes [12, 44, 45]. Several studies also highlighted how voting interface designs shift voter decisions [12], reduce usability errors [46, 47], or improve interaction [48, 49, 50, 51, 52]. We provide more details to these voting interfaces in the Appendix A.

From the QV implementations, response format literature, and voting interfaces, we identified how interfaces significantly influence respondent behavior, decision accuracy, and cognitive load. While these systems are functional, they lack the human-centered design needed to reduce cognitive load and make them truly usable, rather than simply operable. These burdens are especially problematic for complex systems like QS, where high cognitive demands may deter researchers and users alike. Developing effective, human-centered interfaces for QS could enhance usability, reduce cognitive overload, and increase adoption in both research and practical applications.

2.3 Cognitive Challenges and Choice Overload

The challenge of respondents making difficult decisions using quadratic mechanisms remains unexplored in the literature. Lichtenstein and Slovic [6] identified three key elements that make decisions difficult. These elements include making decisions in unfamiliar contexts, being forced to make tradeoffs due to conflicting choices, and quantifying the value of one's opinions. QS fits all three elements: participants may encounter unfamiliar options set by the decision

313 maker, are constrained by budgets that require tradeoffs, and cast final votes as numerical values. Thus, we believe QS
314 introduces high cognitive load.
315

316 Cognitive overload can adversely affect performance, leading individuals to rely on heuristics rather than deliberate,
317 logical decision-making [53]. When presented with excessive information, such as too many options, individuals
318 'satisfice', settling for a 'good enough' solution rather than an optimal one [9, 10, 11]. Subsequently, too many options
319 can overwhelm individuals, resulting in decision paralysis, demotivation, and dissatisfaction [54].
320

321 Additionally, Alwin and Krosnick [55] highlighted that the use of ranking techniques in surveys can be time-
322 consuming and potentially more costly to administer. These challenges are compounded when ranking numerous items,
323 requiring substantial cognitive sophistication and concentration from survey respondents [56].
324

325 Notable applications of Quadratic Voting include the 2019 Colorado House, which considered 107 bills [57], and
326 the 2019 Taiwan Presidential Hackathon, which featured 136 proposals [58]; both used a single QV question with
327 hundreds of options. Psychological and behavioral research highlights the importance of understanding how individuals
328 navigate and benefit from new interfaces under long-list QS conditions. These empirical applications of QV suggest
329 QS's potential to elicit individual preferences, emphasizing the need to study cognitive load and interface design.
330

331 3 Quadratic Survey Interface Design

332 In this section, we present the QS interface. Using components identified in current QV interfaces, together with
333 prior literature, we iterated over paper prototypes and three design iterations based on each pre-test result. During
334 our initial paper prototyping iterations, two challenges emerged: identifying relative preferences among options and
335 deciding the degree of trade-offs between options. In this study, we focus on the first challenge to inform our interface
336 design iterations. For coherence and brevity, we provide detailed descriptions of these iterations and additional design
337 considerations in Appendix B.
338

341 3.1 The Two-Phase Interface

342 3.1.1 *Justifying a two-phase approach.* The main objective for the two-phase interface is to facilitate preference
343 construction and reduce cognitive load. The two-phase interface, shown in Figure 1 consists of two steps: An organization
344 phase and a voting phase. Throughout both phases, survey respondents can drag-and-drop options across the presented
345 option list.
346

347 A *two-phase approach*. If preferences are constructed, by nature, they consist of a series of constructed decision-making
348 processes [6]. Two major decision-making theories informed this two-step interaction interface design: Montgomery
349 [59]'s Search for a Dominance Structure Theory (Dominance Theory) and Svenson [60]'s Differentiation and Consoli-
350 dation Theory (Diff-Con Theory). The former suggested that decision-makers prioritize creating dominant choices
351 to minimize cognitive effort by focusing on evidently superior options [59]. The latter described a two-phase process
352 where decisions are formed by initially *differentiating* among alternatives and then *consolidating* these distinctions to
353 form a stable preference [60]. Both theories supported the design decision to reduce the dimensions during the initial
354 decision process and help emphasize relatively important options to form decisions. Hence, the two-phase design –
355 organize then vote – aimed to facilitate this cognitive journey explicitly. The first phase focused on differentiating and
356 identifying dominant options, enabling survey respondents to preliminarily categorize and prioritize their choices. The
357 second phase presented these categorized options in a comparable manner, with drag-and-drop functionality, enhancing
358

365 one's ability to consolidate preferences. This structured approach aimed to construct a clear decision-making procedure
 366 that reduced cognitive load and enhanced clarity and confidence in the decisions made.
 367

368 *Phase 1: Organization Phase.* The goal of the organization phase was to support participants in identifying dominating
 369 options or partitioning options into differentiable groups. In this section, we first describe how the interaction worked,
 370 then we detail the reasons for the implemented design decisions.
 371

372 The organizing interface, depicted on the top half of Figure 1, sequentially presented each survey option. Participants
 373 selected a response among three ordinal categories – “lean positive”, “lean negative”, or “lean neutral”. Once selected, the
 374 system moved that option to the respective category. Participants could skip the option if they did not want to indicate
 375 a preference. Options within the groups were draggable and rearrangeable to other groups should the participants wish.
 376

377 Strack and Martin [61]’s research showed that upon understanding a survey question, respondents either recalled a
 378 prior judgment or constructed a new one when completing an attitude survey. In addition, revealing one option at a
 379 time gated the amount of information presented to the survey respondent and thereby reduced the extraneous load [62].
 380 This process allowed participants to form or express opinions on individual options incrementally. This design also
 381 mitigated the original concern from early prototypes where participants accidentally treated the organizing task as a
 382 ranking task.
 383

384 The three possible options, positive, neutral, and negative, aimed to scaffold participants in constructing their own
 385 choice architecture [63, 64], which strategically segmented options into diverse and alternative choice presentations
 386 while avoiding the biases from defaults. We believed that these three categories were sufficient for participants to
 387 segment the options. However, we chose not to limit the number of options one could place into a category to prevent
 388 restricting user agency, a core user interface design principle [18].
 389

390 Immediate feedback displaying the placement of options and allowing participants to rearrange them via drag-and-
 391 drop adhered to key interface design principles [18]. At the same time, it allowed finer grain control for individuals to
 392 surface dominating options and create differentiating groups of options.
 393

394 *Phase 2: Interactive Voting Phase.* The objective of the voting phase was to facilitate the consolidation of differentiated
 395 options through interactive elements while reinforcing the differentiation across options constructed by participants from
 396 the previous phase. This facilitation was achieved by retaining the drag-and-drop functionality for direct manipulation
 397 of position and enabling sorting within each category.
 398

399 Options were displayed as they were categorized within each category from the previous step and in the following
 400 section orders – lean positive, lean neutral, lean negative, and skipped or undecided as detailed on the bottom half
 401 of Figure 1. The Skipped or Undecided category contained options left in the organization queue, possibly because
 402 survey respondents had a pre-existing preference or chose not to organize their thoughts further. The original order
 403 within these categories was preserved to maintain and reinforce the differentiated options. This new ordering sequence
 404 mitigated the concern from prototype 3 where options without a category are left at the top of the voting interface.
 405 Respondents had the flexibility to return to the organization interface at any point during the survey to revise their
 406 choices.
 407

408 In the voting interface, options remained draggable, enabling participants to modify or reinforce their preference
 409 decisions as needed. Each category featured a sort-by-vote function that enabled reordering within the same category.
 410 Although these interactions did not influence the final voting outcome, they were designed to support consolidation
 411 and positional proximity in information organization. This design aimed to automate the grouping of similar options
 412 while providing an intuitive drag-and-drop mechanism, thereby facilitating decision-making by placing similar options
 413
 414 Manuscript submitted to ACM
 415
 416

417 near each other. This echoed the principles of the proximity compatibility principle, particularly emphasizing spatial
418 proximity and mental compatibility [65]. The interface design anticipated that participants would find it easier to
419 consolidate their choices when similar options were positioned close together, thereby reducing cognitive load.
420

421 While multiple interaction mechanisms exist, drag-and-drop has been extensively explored in rank-based surveys.
422 For instance, Krosnick et al. [66] demonstrated that replacing drag-and-drop with traditional number-filling rank-based
423 questions improved participants' satisfaction with little trade-off in their time. Similarly, Timbrook [67] found that
424 integrating drag-and-drop into the ranking process, despite potentially reducing outcome stability, was justified by the
425 increased satisfaction and ease of use reported by respondents. The trade-off was deemed worthwhile as QS did not use
426 the final position of options as part of the outcome if it significantly enhanced user satisfaction and usability [68].
427

428 Together, these design decisions led to our belief that a two-phase interface with direct interface manipulation could
429 reduce the cognitive load for survey respondents to form preference decisions when completing QS.
430

431 *3.1.2 Aesthetic Design Decisions.* We made three aesthetic design decisions. First, we decided to remove all visual
432 elements. Prior literature suggested that the use of emojis might influence the interpretations of surveys [69] and
433 decrease user satisfaction [15]. Prior literature also noted that not all data visualization elements reduce cognitive
434 demand [70]. Even though effective visualization can aid decision-making, it remains an open question that this study
435 does not aim to address. Thus, we also removed all visualization elements, such as blocks, progress bars, and percentage
436 indicators.
437

438 Second, the final interface has all options presented on the screen at the same time, intentionally. Unlike all the
439 prototypes and existing interfaces, prior literature emphasized the importance of placing all the options on the same
440 digital ballot screen to avoid losing votes. This echoes the proverb "out of sight, out of mind," where individuals might
441 be biased toward options that are shown to them, and additional effort is required for individuals to retrieve specific
442 information if options are hidden.
443

444 Last, we decided to use a dropdown positioned to the right of the option such that control of votes and the budget
445 summary are placed near one another. The layout of the votes and cost was inspired by online shopping cart checkout
446 interfaces where quantities are supplied next to the itemized costs followed by the total checkout amount. We chose
447 a dropdown after iterating with two alternative input methods (Figure 3): the original click-based buttons and a
448 wheel-based implementation. The former design requires survey respondents to click multiple times to reach their
449 desired vote values. Thus, we wanted to look for a solution to aid respondents in reaching their intended value faster. A
450 wheel-based approach allows intuitive control of the votes by using the wheel on the mouse and clicking to fine-tune
451 the values. However, in our early pilot studies, not all participants were familiar with wheel control. Thus, we opted for
452 a dropdown menu for vote selection.
453

454 **3.2 Baseline Interface: Text-based Interface**

455 To compare how the designed interactive components influenced participants' cognitive load and behavior, we designed
456 the text-based interface as our control condition. The text-based interface shares the other functionalities of the two-
457 phase interface without the two-phase interactive design and the drag-and-drop (Figure 4.) The interface contained
458 the question prompt at the top of the screen. The options were presented in a list underneath the prompt. Survey
459 respondents could update the votes by selecting from a dropdown that provided all possible voting options and costs
460 given the number of credits available. A small summary box to the right of the interface showed the current total cost
461 and the remaining credits for the respondent. The interface randomly presented options to avoid ordering bias [71, 72].
462

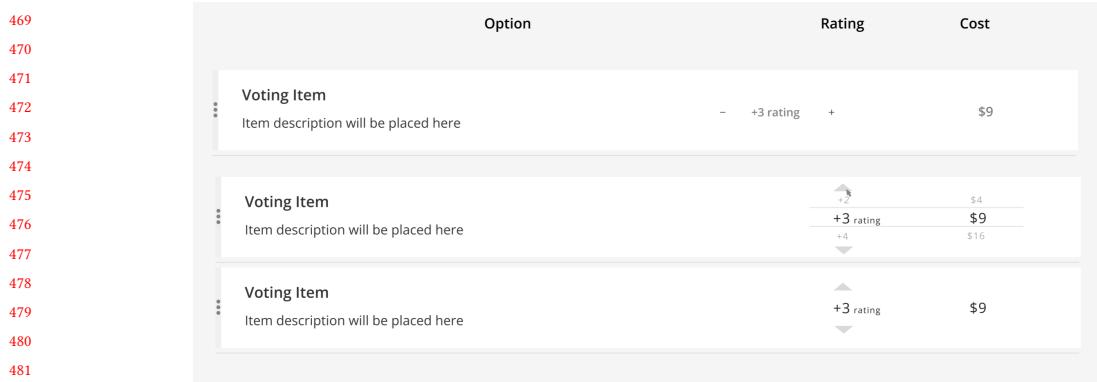


Fig. 3. Alternative vote control. The click-based design (upper) mirrors traditional vote control used in other QV interfaces, where each click controls one vote. The wheel-based design (the latter two) allows control through both clicks and mouse wheel rotation.

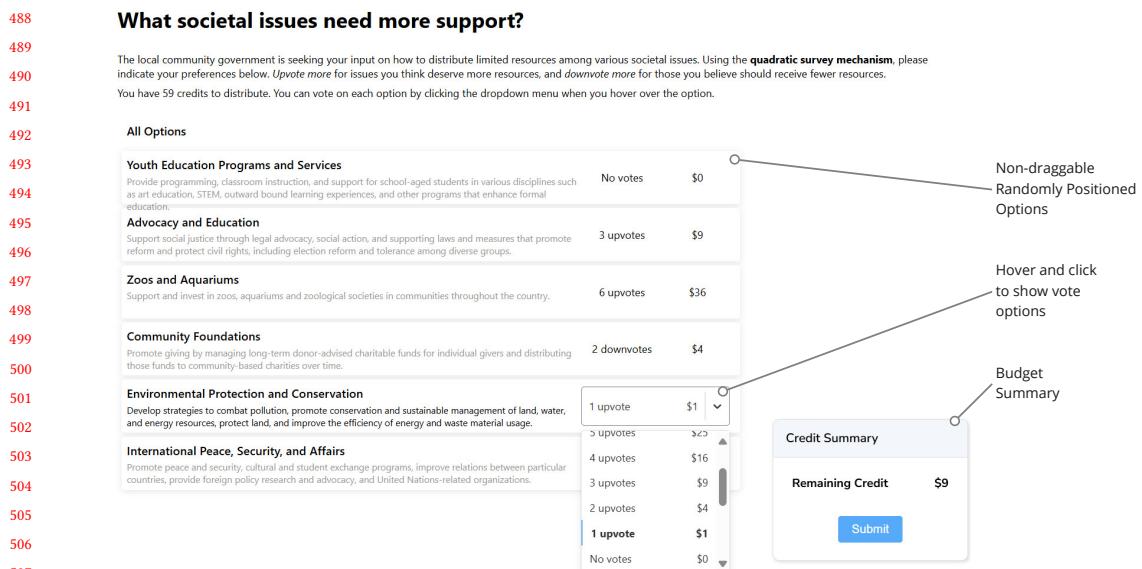


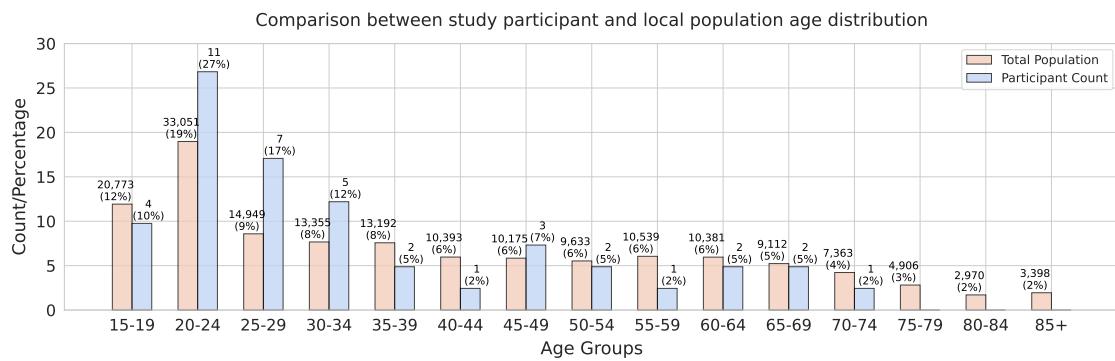
Fig. 4. The text-based interface: This interface is based on the interactive version but does not include the two-phase interactive support and lacks the drag-and-drop functionality. Options are randomly positioned.

Both experiment interfaces are developed with a React.js frontend and a Next.js backend powered on MongoDB. We open-sourced both interfaces.¹

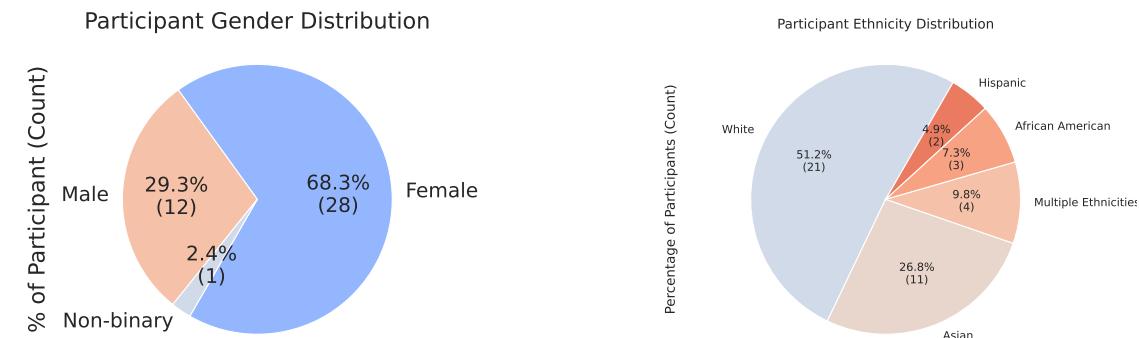
¹[link-to-github](#)

521 4 Experiment Design

522 We recruited 41 participants, with one excluded² due to data quality concerns, from a United States college town
 523 using online ads, digital bulletins, social media posts, online newsletters, and physical flyers in public spaces beyond
 524 campus. To ensure participant diversity, we prioritized non-students by selectively accepting them as we monitored
 525 demographics. Study participants' mean age was 34.63 years old, with an age distribution similar to the county's
 526 demographic profile (Figure 5a) albeit a slightly higher representation of younger adults. Gender and race demographics
 527 are aggregated in Figure 5b and 5c. The study was framed as focusing on societal attitudes to avoid response bias. The
 528 university's Institutional Review Board reviewed and approved this study.
 529



545 (a) Age distribution of the participants skewed slightly younger despite following similar trends as the county population.



(b) Gender distribution of our participants skewed towards female participants.

(c) Ethnicity distribution remains diverse with less Hispanic and African American participants.

Fig. 5. Demographic distributions: Age, Gender, and Ethnicity

Figure 6 visually represents the study protocol. Participants completed the study in the lab to control for external influences. Participants used a 32-inch vertical monitor displaying all options on the survey to prevent hidden information during decision-making. After consenting, participants watched a video explaining the quadratic mechanism without

²The participant reported not completing the survey seriously because they believed the experiment was fake.

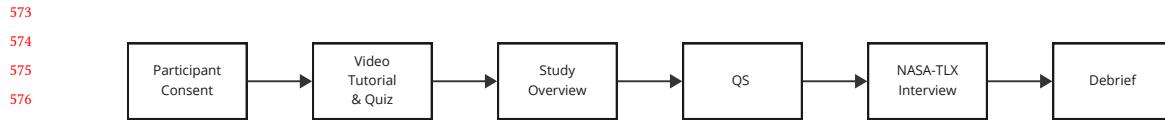


Fig. 6. Study protocol: Participants are asked to learn about the mechanism of QS after consenting to the study. The researcher explained the study overview and asked participants to complete the QS. A NASA-TLX survey followed by interviews to understand participants' cognitive load. We debriefed participants after the study.

hints of interface operation followed by a quiz to ensure understanding. Participants rewatched the video or consulted the researcher until they could select the correct answers. The participant's screen was captured throughout the study. The researcher primed the participant that the study aimed to help local community organizers understand preferences on societal issues to better allocate resources. Participants were randomly assigned to one of four groups:

- Short Text (ST): A text interface with 6 options. ($N = 10$)
- Short Interactive (SI): A two-phase interface 6 options. ($N = 10$)
- Long Text (LT): A text-based interface 24 options. ($N = 10$)
- Long Interactive (LI): A two-phase interface with 24 options. ($N = 10$)

Participants completed the survey independently, without the researcher's presence. They then contacted the researcher for the NASA-TLX survey, followed by a short audio-recorded semi-structured interview. The session concluded with a debriefing and a \$15 cash compensation, during which participants were informed of the study goal on cognitive load and interface design.

We made several experimental design choices. First, we selected a between-subject design to minimize study fatigue, considering the complexity of QS, and avoiding the learning effect that could influence how participants evaluated the options. Second, we chose the context of public resource allotment, where participants expressed their preferences regarding their preference across 6 or 24 societal issue options, following the methodology of Cheng et al. [3]. These issues are relevant to all citizens and effectively demonstrate the need to prioritize limited public resources. We curated 26 societal issues used by Charity Navigator [73] which evaluates over 20,000 charities in the United States. The interface randomly presents options from this list to participants. Appendix C contains the full list.

We relied on prior literature to test 6 and 24 options, representing short and long lists, as identifying the 'breaking point' for cognitive overload would require impractical time and resource commitments which we detail below. Constant sum surveys and the Analytic Hierarchy Process (AHP) recommend fewer than ten and seven options, respectively [74, 75, 76]. Miller [77]'s classic work on cognitive processing capacity and Saaty and Ozdemir [78]'s theoretical proof supported the use of 7 ± 2 items. A meta-analysis by Chernev et al. [79] identified 6 and 24 as common values for short and long lists in choice overload studies, rooted in the original experiment by Iyengar and Lepper [54].

Finally, we deployed self-report subjective surveys and analytical measures (i.e., time and clickstream data). We adopted the paper-based weighted NASA Task Load Index (NASA TLX), a widely used multidimensional tool that averages six subscale scores to represent overall workload after completing a task [80, 81, 82]. NASA-TLX is favored for its low cost and ease of administration [83], with less variability compared to one-dimensional workload scores [84], making it suitable for our study. While cognitive load can be measured through performance measures, psychophysiological

625 measures, subjective measures, and analytical measures [83], given the extended nature of QS we chose not to adopt
 626 them.
 627

628 5 Cognitive Load and Sources across Experiment Conditions

630 This section presents cognitive load across experiment groups and the sources contributing to each cognitive load
 631 dimension. Due to the smaller sample size, we focus on descriptive statistics and qualitative assessments. We present
 632 quantitative data from survey tasks and qualitative insights from post-survey interviews.
 633

634 To analyze the qualitative data, the first author conducted an inductive thematic analysis process [85] after transcribing
 635 the interview. They coded snippets from each transcript based on specific research questions and topics of interest for
 636 the qualitative analysis. Similar codes were merged within each research question or topic to form relevant themes.
 637 When differences were hypothesized, they applied a deductive coding process to text snippets related to a specific
 638 research question or topic of interest.
 639

640 The results for this section are organized as follows: we start with participant demographics and then provide an
 641 overview of our cognitive load findings. We then examine the six dimensions used in the NASA-TLX survey: mental
 642 demand, physical demand, temporal demand, performance, effort, and frustration.
 643

644 5.1 Overall Cognitive Load

645 To answer RQ1 and RQ2a, we derive the weighted NASA-TLX scores across the four experiment conditions. We show
 646 these results in Figure 7. Weighted NASA-TLX uses a continuous 0-100 score, with higher values indicating greater
 647 cognitive load. We use predefined mappings of NASA-TLX scores to cognitive levels: low, medium, somewhat high,
 648 high, and very high, as listed by Hart and Staveland [80]. We show value interpretations in Figure 7b.
 649

650 Surprisingly, the long text interface had lower mean ($\mu = 34.60$) and median ($\tilde{x} = 33.85$) cognitive load scores than
 651 both the short text ($\mu = 43.23$, $\tilde{x} = 39.00$) and long interactive interfaces ($\mu = 42.02$, $\tilde{x} = 42.70$). Additionally, the
 652 two-phase interface decreased cognitive load for the short survey but increased it for the long survey. Notably, the
 653 short text interface had the most participants ($N = 8$) reporting somewhat high or higher cognitive loads, whereas the
 654 other conditions had a more balanced distribution, with about half reporting medium to high loads.
 655

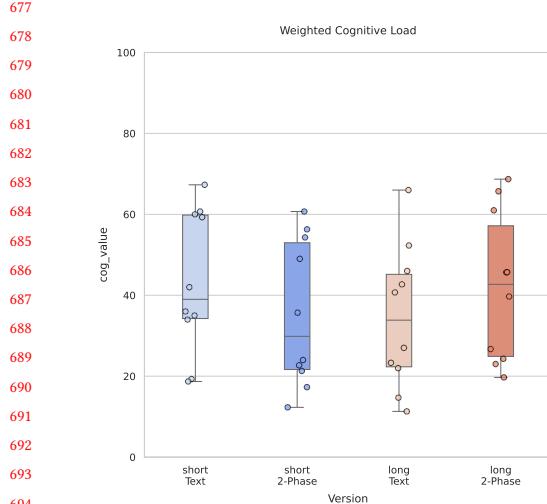
656 We acknowledge that these results may not fully reflect actual cognitive load due to noise from factors like small
 657 sample size, task nature, or participants' interpretation of the cognitive load scale. To explore these possibilities further,
 658 we turn to qualitative insights from post-task interviews.
 659

660 5.2 Sources of Mental Demand

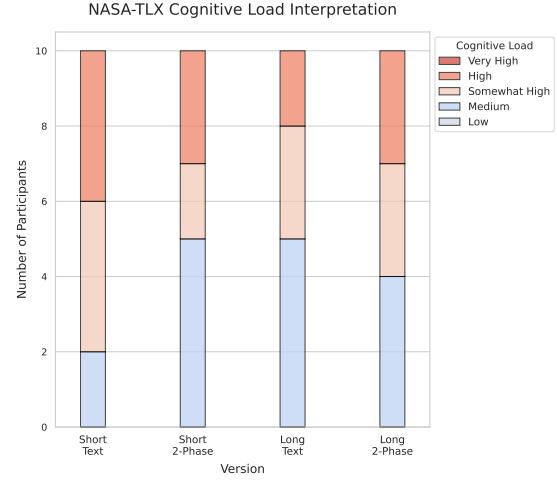
661 **Key Differences:** First, slightly more participants using the text interface reported mental demand from
 662 precisely determining the number of votes for options compared to the two-phase interface. Second, when it comes
 663 to long QS, participants using the long two-phase interface considered broader societal impacts and evaluated
 664 options holistically, while those in the long text interface focused on personal relevance and individual issues.
 665

666 Mental demand refers to the degree of mental and perceptual activity required to complete a task. Interview results
 667 showed primary drivers of participants' mental demand were *Budget management* and *Preference construction*³.
 668

669
 670 ³The full table is in Appendix D.1.
 671
 672
 673
 674
 675
 676



(a) NASA-TLX Weight Score: The Long Two-Phase Interface exhibits the highest weighted cognitive load with a median of 42.70, a mean of 42.02. This is higher than the Long Text Interface, which has a median cognitive load of 33.85, a mean of 34.60. However, the Short Text Interface demonstrates a higher cognitive load with a median of 39.00, a mean of 43.23, compared to the Short Two-Phase Interface which has a median of 29.85, a mean of 35.36. Standard deviation are similar across groups at around 18.



(b) NASA-TLX Cognitive Interpretation: More participants in the Short Text Interface, totaling 8, reported a somewhat high or above cognitive load, which is significantly higher compared to the 5 participants who reported similarly for the Short Two-Phase Interface. However, the Long Two-Phase Interface saw slightly more participants, 6 in total, reporting somewhat high or above cognitive load compared to the Long Text Interface.

Fig. 7. This figure shows the box plot results for weighted NASA-TLX scores across experiment groups and participant counts based on individual score interpretations. In 7a, we observe a downward trend in cognitive load for the short QS, while the long QS shows an upward trend. Interestingly, there is a counterintuitive downward trend between short and long text interfaces. In 7b, these trends are clearer when NASA-TLX scores are grouped into five tiers.

5.2.1 Mental Demand Source #1: Budget management. 14 participants expressed demand from budgeting within limited credit (*S032 Q [...] for certain societal issues you had to ... take away from other societal issues that you could support. , N = 5*), tracking remaining credits (*S006 Q [...] looking at the remaining credits, I'm trying to mentally divide that up before I start allocating , N = 10*), and maximizing credit use (*S032 Q [...] I used all the credit that I had available , N = 8*).

We categorized budget management-induced mental demand as operational (single interface-level action, e.g., using the last remaining credit) or strategic (higher-level goal, e.g., evenly distributing credits across options). Long survey participants more often report operational causes suggesting that more survey options induced short-term thinking.

Manuscript submitted to ACM

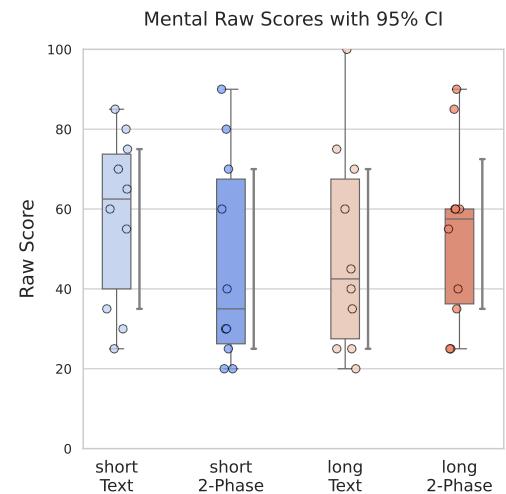


Fig. 8. Mental Demand Raw Score: Across all four experiment groups, participants' reported mental demand is spread across a wide range with many participants experiencing high mental demand.

729 5.2.2 *Mental Demand Source #2: Preference construction.* All
 730 but one participant reported increased mental demand due to
 731 preference construction. We further break it down into three
 732 sources: comparative preference evaluation (i.e., evaluating
 733 relative importance between options; S002 *Figuring out*
 734 ... how much I prioritize option 1 over option 2 , N = 16), resource-constraint prioritization (i.e., trading off between
 735 options due to resource constraints, S005 [...] very hard to take decisions ... because I felt that multiple options deserve
 736 equal amount of credit ... but you have given very limited amount of credit . , N = 17), and precise resource allocation (S023 [...] having to pick how many upvotes would go to each one , N = 30).

740 Almost all participants mentioned preference construction as a source of mental demand, supporting the theory that
 741 preference construction is a difficult and mentally demanding task. Notably, more participants using the text interface
 742 reported mental demand from precise resource allocation compared to the two-phase interface (18 vs. 12). We conjecture
 743 that the first pass on the survey items in the organization phase helped participants reduce their mental demands in
 744 this area once they got around to vote allocation.
 745

746 In addition, when categorizing preference construction-induced mental demand, participants (N = 8) in the long text
 747 interface tend to consider a smaller scope that focuses on personal relevance. Conversely, participants (N = 9) in the
 748 long two-phase interface considered the broader societal impact and evaluated options more comprehensively. Compare
 749 the following two quotes, where one focused on adjusting credits between two options and the other reflecting across
 750 broader societal values:
 751

752 *Trying to figure out what upvotes I should give [...] I kind of went back and forth between those two. [...] So it was very mentally*
 753 *tasking for me.* S015 (LT)

754 [...] really having to think, especially with so many different societal issues. How do I personally prioritize them? And to what extent
 755 do I prioritize them? S009 (LI)

756 Inspecting both causes, while we did not notice significant differences in mental demand raw values (Figure 8)
 757 across the four experiment groups, especially between the interfaces across long QS, they exhibited different sources
 758 contributing to their mental demand.
 759

760 5.3 Sources of Physical Demand

761 **Key Differences:** Two-phase interface experienced higher physical demand from increased mouse usage.

762 Physical demand refers to the physical effort required to complete a task, such as physical exertion or movement.
 763 Most participants reported minimal physical demand (N = 32), reflected in the low NASA-TLX physical demand
 764 scores (Figure 9). Notably, 11 out of 20 participants who used the two-phase interface mentioned physical demand
 765 from using the mouse, reflecting their increased interaction with the interface. This is further supported by the raw
 766 NASA-TLX physical demand scores (Figure 9), which show a significant visual difference between short and long
 767 two-phase interfaces as well as between text and two-phase interfaces in long surveys.
 768

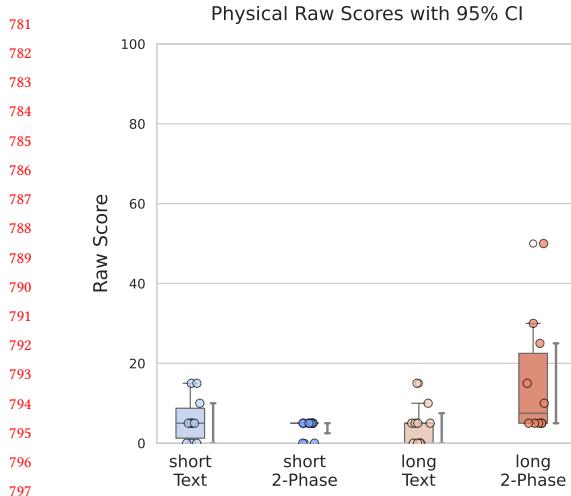


Fig. 9. Physical Demand Raw Score: Participants other than the long two-phase interface reported minimal physical demand. The long two-phase interface had the highest physical demand, likely due to increased mouse clicks and extended time spent looking at the vertical screen.

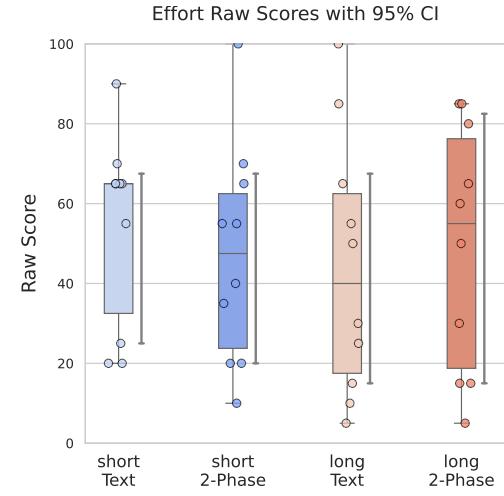


Fig. 10. Effort Raw Score: Effort scores shows indifference across groups.

Table 1. Effort Sources: Participants using the text interface focused more on operational tasks, while those using the two-phase interface focused more on strategic planning.

[Effort]	Total	Version				Experiment Conditions			
		ST	SI	LT	LI	Short	Long	Text	Inter
Operational	21	6	5	8	2	11	10	14	7
Strategic	28	6	8	5	9	14	14	11	17
Personal	22	4	7	5	6	11	11	9	13
Global	11	2	3	2	4	5	6	4	7
None/Little/a bit	9	2	1	3	3	3	6	5	4

5.4 Source of Effort

Key Differences: First, participants in the text interface associated effort with operational tasks more often than participants from the two-phase interface. Conversely, participants in the two-phase interface cited more sources from strategic planning than those in the text interface. We observed that participants experienced effort when considering a comprehensive view while using the two-phase interface.

Effort refers to how hard participants felt they worked to achieve the level of performance they did. Since effort includes both mental and physical resource intensity, refer to Section 5.2 and Section 5.3 for definitions.

833 While the raw NASA-TLX effort scores (Figure 10) showed a similar spread across experiment groups, the qualitative
 834 analysis showed more distinction that participants using the two-phase interface considered options more comprehensively
 835 and felt less effort on completing operational tasks, similar to what we found on mental demands (Section
 836 5.2).
 837

838
 839 *5.4.1 Effort Source #1: Operational Tasks.* 14 of the 20 participants using the text interface mentioned Operational
 840 Tasks as effort sources, compared to 7 using the two-phase interface, with the lowest mention by the long two-phase
 841 interface group ($N = 2$).
 842

843
 844 *5.4.2 Effort Source #2: Strategic Planning.* Different from Operational Tasks, 11 participants in the text interface
 845 compared to 17 participants described strategic planning as sources of effort, with almost all participants ($N = 9$) from
 846 the long two-phase interface. We further categorize strategic planning into *narrow* and *broad* scopes as we did for
 847 mental demand section 5.2. Participants using the two-phase interface ($N = 7$) had nearly mentioned double ($N = 4$)
 848 times regarding global strategies.
 849

850 851 5.5 Sources of Temporal Demand

852
 853 **Key Differences:** First, participants using the short text interface wanted to complete the task quickly and
 854 reported the highest temporal demand. The two-phase interface lowered the temporal demand for short QS. Second,
 855 participants using the long text interface showed the lowest temporal demand score. The two-phase interface
 856 increased the temporal demand for long QS.
 857

858
 859 Temporal demand measures the time pressure participants feel during a task. Lower demand indicates participants
 860 felt comfortable taking a more leisurely pace. We categorize the main sources of increased temporal demand as the time
 861 pressure on *Decision Making* (S024 *maybe I should just hurry up and make a decision.*, $N = 15$) and *Operational Tasks* (S032 *to be able to move through this quickly and efficiently*, $N = 16$) (Table 5). *Budget* also came up occasionally (S034 *as the money decreases I felt kind of rushed*, $N = 4$).
 862

863
 864 *5.5.1 Two-phase Interface Reduced Temporal Demand on Short QS.* The raw NASA-TLX values in Fig 11 show that
 865 participants in the Short Text Interface condition reported the highest temporal demand among all. They framed
 866 concerns as time spent on making decisions. Five participants felt they invested more time and effort than anticipated,
 867 prompting them to rush. The two-phase interface reduced this, with only one participant in the short survey group
 868 reporting similar concerns.
 869

870
 871 *5.5.2 Long QS on Text Interface Showed the Lowest Temporal Demand.* Based on the raw NASA-TLX values in Fig 11, partic-
 872 ipants in the long text interface exhibited the lowest temporal demand, explaining why this group had the lowest overall cog-
 873 nitive load as discussed in Section 5.1. This is counter-intuitive
 874 since participants in this condition made more decisions and
 875 operations compared to the short text group. There are two
 876 potential explanations. First, we noticed that more participants
 877

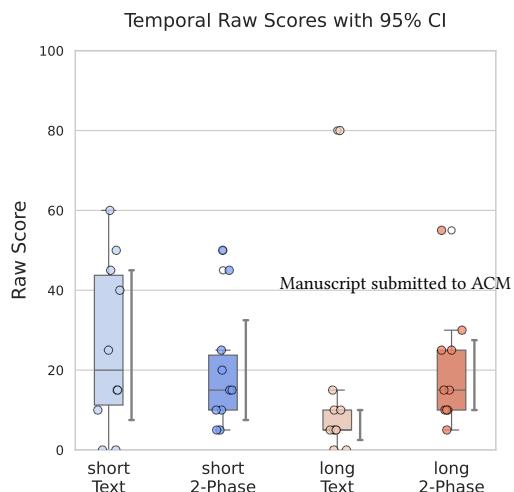


Fig. 11. Temporal Demand Raw Score: The short text interface results in the highest temporal demand, while the long text interface is the lowest. Two-phase interfaces, show moderate temporal demand, suggesting that interactive ele-

885 who experienced a short survey expressed a desire to com-
886 plete the task efficiently ($N = 7$) than those in the long survey
887 groups ($N = 1$). They often expressed things like:
888

889 I wanna get through things in an efficient manner [...] to move
890 through this quickly and efficiently. Q S032 (ST)
891

A second possible explanation for the unexpectedly low temporal demand in the Long Text condition was the participants' satisficing behaviors.

897 I didn't really do the math, so I was like \$2 is not that much left so
898 I tried my best to use up most of it.
899 Q S035 (LT)

901 Due to cognitive overload from the long list of options, par-
902 ticipants may have spent less time than they expected on the decisions. We will discuss this possibility in-depth in
903 Section 7.1.

5.5.3 *Two-phase Interface Increased Temporal Demand in Long QS.* Despite the unexpectedly low temporal demand in long QS with a text interface, completing the long QS with our two-phase interface raised the temporal demand to the level of short QS groups. All five participants who mentioned a feeling of time pressure on decision-making in the Long Two-phase group described the pressure affirmatively. This means their pressure stemmed from having too many remaining decisions to make (*S022 Q So it didn't take too much time. but obviously there was a lot of things to consider, so there was some temporal demand.*), not from the time they have already spent (i.e., framed negatively) as that in the Short Text group.

5.6 Source of Frustration

❑ **Key differences:** We observed evidence that participants in the long text interface showed the least amount of frustration from operational causes compared to other experiment conditions.

922 Frustration refers to the extent to which the participant is annoyed, irritated, or discouraged during the task.
923 Sources of frustration were generally related to either *Operational Actions* (e.g., credit management ($N = 6$) and
924 managing quadratic vote costs ($N = 5$)), or *Societal Concerns* (e.g., regretful tradeoffs ($N = 8$) or pessimism about other's
925 vote ($N = 6$)). We provide the full table in Appendix D.5.
926

In general, the frustration derived from societal concerns did not seem strongly affected by any of the experimental conditions. With respect to operational action-driven frustration, however, we saw some discrepancies. The long text interface condition had the fewest participants expressing operational frustration, with half expressing no frustration, mirroring the trends in the actual scores (Figure 12). Similar to the finding that the long text group has the lowest temporal demand, this is counter-intuitive as more options and dense text are known to lead to more frustration in interface design. Participants engaging in satisficing behaviors in the long text interface may explain this phenomenon – prior literature [86, 87] indicates that satisficers tend to be less frustrated and happier than maximizers.

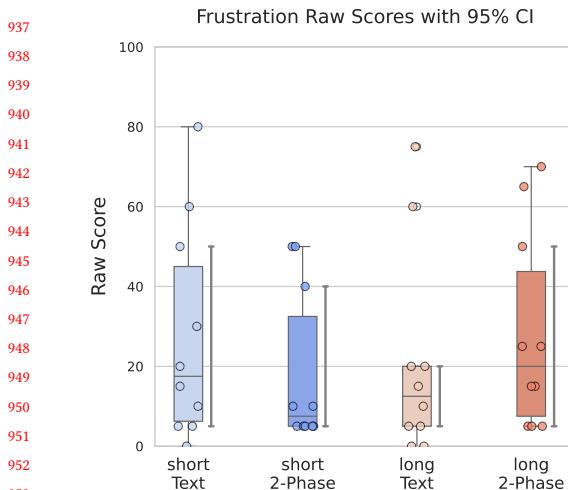


Fig. 12. Frustration Raw Score: Participants other than the long text interface highlighted several operational tasks that led to frustration. All groups share causes from strategic planning.

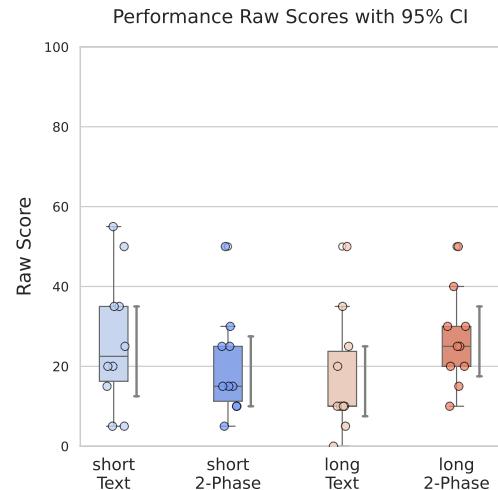


Fig. 13. Performance Demand Raw Score: Participants showed indifferent performance raw scores across experiment conditions, all trending toward satisfactory.

5.7 Source of Performance

Q Key Differences: Participants who used a two-phase interface were generally more positive about their final outcome – they were twice as likely to report "feeling good" about their final results

Performance refers to a person's perception of their success in completing a task. Lower values mean good perceived performance; higher values mean poor perceived performance. We found minimal qualitative differences between experiment groups regarding factors influencing perceived performance. Two influencing factors emerged: *Operational Actions* and *Social Responsibility*⁴. Despite most participants reporting positively on their performance, nuances exist in how different groups interpret their performance.

5.7.1 Operational Actions. Operational actions, like the theme presented in temporal demand, refer to specific, executable procedures participants perform in the survey. This could involve: pressure to spend all credits or stay within budget ($N = 6$), fears that final vote choices did not reflect true preferences ($N = 5$), or concerns that they had finished the task inefficiently ($N = 6$).

5.7.2 Social Responsibility. Social responsibility-based concerns around performance came up when participants reflected on how their final vote counts would be perceived by others (S041 *Q I don't want people to think that I just like don't care about <ethnicity> people at all*) or influence real-world decision-making (S027 *Q Some of these things might ... have outcomes that I didn't foresee*).

All groups cited social responsibility as source to evaluate effort. Raw NASA-TLX scores (Figure 13) show participants had indistinguishable performance scores. This aligns with the interview results where most participants felt positive about their final submission.

⁴The full performance table is at Appendix D.3

989 To dig deeper, we also analyzed participants' language when they described their performance. Expressions like
 990 "good enough" may be indicative of satisficing behaviors – our results suggest participants are satisfied at similar rates
 991 regardless of the interface. 1/4 of the participants in the text interface expressed "done their best," referring to exhausting
 992 their effort. Participants who used a two-phase interface were generally more positive about their final outcome – they
 993 were twice as likely to report "feeling good" about their final results ($N = 11$ v.s. $N = 6$).
 994

995 5.8 Summary across all cognitive load dimensions

996 In this subsection, we gathered differences across dimensions and synthesized them into a list.
 997

- 1000 • **Mental Demand:** Participants using the text interface reported higher mental demand from determining the
 1001 number of votes for options. Those using the long two-phase interface considered broader societal impacts and
 1002 evaluated options holistically.
- 1003 • **Physical Demand:** Physical demand was higher for participants using the two-phase interface due to increased
 1004 mouse usage.
- 1005 • **Effort:** Effort sources varied, with text interface participants focusing more on operational tasks, and two-phase
 1006 interface participants engaging more in strategic planning, reflecting deeper, more comprehensive consideration
 1007 of options.
- 1008 • **Temporal Demand:** Temporal demand was highest in the short text interface, where participants aimed to
 1009 complete tasks quickly, while the long text interface showed the lowest temporal demand.
- 1010 • **Frustration:** Frustration levels from operational causes were lowest in the long text interface.
- 1011 • **Performance:** Participants using the two-phase interface felt more positive about their performance, being
 1012 twice as likely to report "feeling good" about their results compared to the text interface users.

1016 Overall, participants using the two-phase interface tend to think more comprehensively and critically, while those
 1017 using the text interface focus more narrowly on operational tasks. In addition, we suspect that participants who
 1018 completed the long QS on a text interface engaged in satisficing behaviors based on the counter-intuitive results that
 1019 they had the lowest temporal demand and frustration level. We will interpret these results in the discussion section. To
 1020 better understand participants' behavior, we analyze click-stream data across experiment conditions in the next section.
 1021

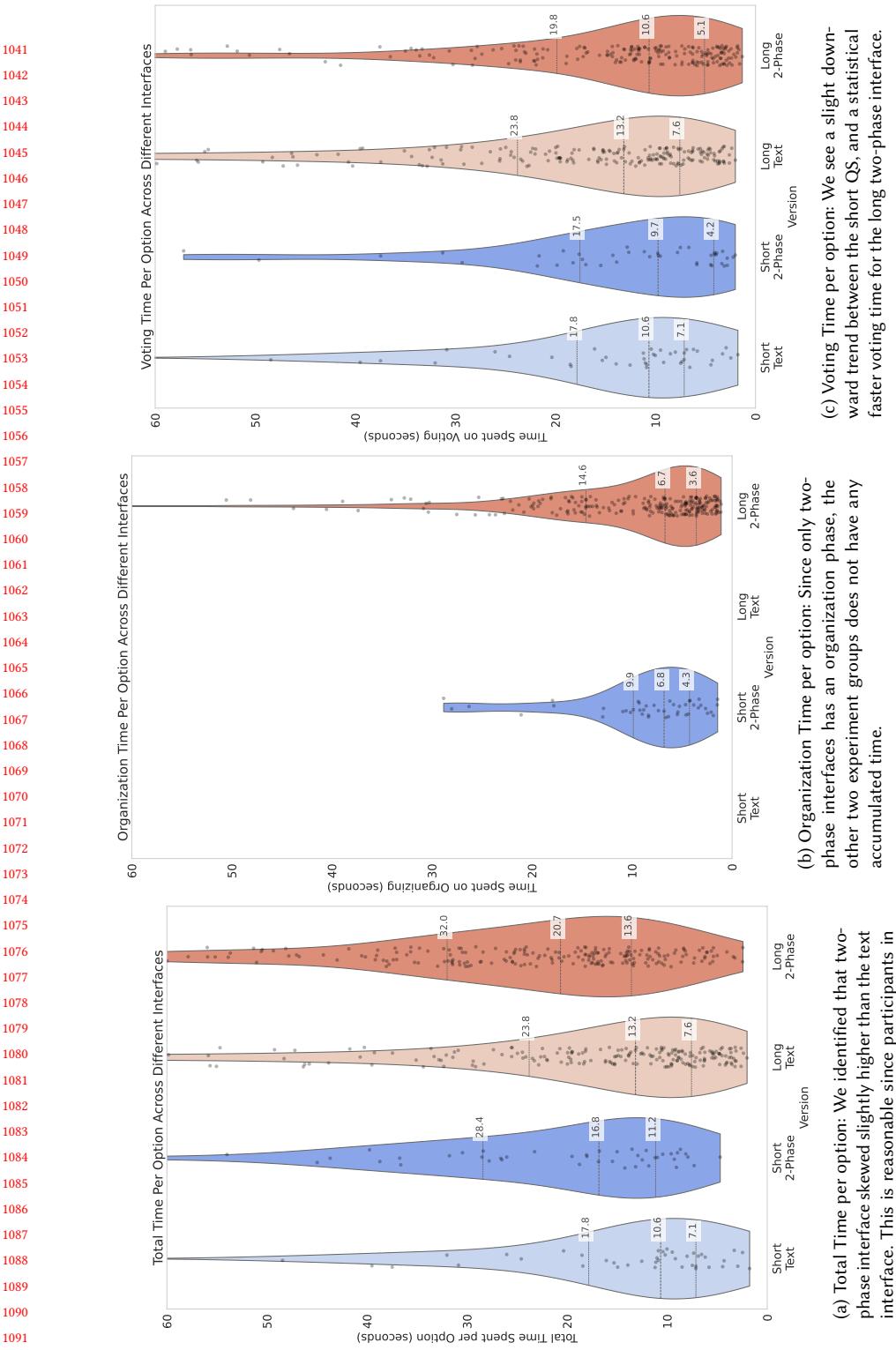
1022 6 Behavior Results

1025 To answer RQ3, we investigate time-to-action and remaining credit differences across experiment conditions. Time-to-
 1026 action is a widely used metric in decision sciences, where longer decision time often indicates more complex cognitive
 1027 processing [88]. Additionally, resource allocation strongly influences decision making. Cheng et al. [3] showed that the
 1028 number of given credits influences the validity of QV. Decision science studies like Shah et al. [89] and [90] showed
 1029 how scarcity influences decisions, increases risk aversion, and adds cognitive load. These measures serve as proxies for
 1030 participant behavior, and all analyzed data is publicly available⁵ for transparency and to facilitate further research.
 1031

1032 6.1 Time Spent per Option

1035 Our first analysis focuses on understanding how much time participants spent per option across different stages and
 1036 experiment conditions. Based on the QS system log, we can derive the following detailed logs of participant actions: *the*
 1037

1038 ⁵link-to-github



- (a) Total Time per option: We identified that two-phase interface skewed slightly higher than the text interface. This is reasonable since participants in these two conditions require additional time to complete the organizing steps.
- (b) Organization Time per option: Since only two-phase interfaces has an organization phase, the other two experiment groups does not have any accumulated time.
- (c) Voting Time per option: We see a slight downward trend between the short QS, and a statistical faster voting time for the long two-phase interface.

Fig. 14. Time per option across all experiment conditions. In each of these violin plots, a dot represents the total time a participant spent on that option. Fig 14a one dot represents the total time a person spent for one option. Figure 14b and Figure 14c are decomposition of the total time.

1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144

option involved in the interaction, *the type of interaction* (such as updating a certain number of votes), and *the time* between this interaction and the previous one.

We aggregate all the time spent on each option as the total time spent for that option. Organization time covers both placing options into categories and the drag-and-drop time during the organization phase. Voting time strictly refers to the time participants took to update vote values for each option. To minimize noise, we intentionally drop all the time participants spent on the first option in the organization phase or voting phase. The goal is to exclude time spent on reading the prompt, forming their preference, or understanding the interface.

Figure 14 each dot represents one option for one participant. Figure 14a shows total time, figure 14b shows organization time, and figure 14c shows voting time. The violin plot shows the distribution of the dots and the three horizontal lines represent the median, 25th percentile, and 75th percentile of the time spent for that interface. We limited the y-axis to 1 minute to improve visualization clarity.

Participants spent slightly more time per option on the two-phase interface than the text interface. A non-parametric Mann-Whitney U test showed a small effect size (long QS: $p < 0.0000001$, Rank-biserial: -0.304 , Cohen's d: -0.030 ; short QS: $p = 0.01$, Rank-biserial: -0.37 , Cohen's d: -0.082). This is expected as the two-phase interface has an additional step of organizing the options. We break down the total time spent into organization time and voting time in Figure 14b and Figure 14c.

We observed minimal difference in organization time per option (Figure 14b) between short and long surveys, as options are shown one at a time for drag-and-drop. In terms of the voting time (Figure 14c), participants spent significantly less time voting on the two-phase interface than on the text interface with a small effect size in the long QS ($U = 24053$, $p < 0.005$, Rank-biserial: 0.167 , Cohen's d: 0.017), but not in the short survey ($p > 0.4$, Power=0.051). This supports our hypothesis that the two-step design in the two-phase interface facilitates more efficient decision-making, especially in longer surveys.

6.2 Budget and Voting Behaviors

To further analyze participant behaviors, we break down the aggregated time from the previous analysis and examine fine-grain interactions. Specifically, we examine if there are differences among behavior across interfaces. As we outlined, credit scarcity might influence decision making. Figure 15 plots the time of voting actions over the remainder of the participant's budget across the text and two-phase interface across all four groups. Each bar shows the number of actions accumulated across participants at specific percentages of remaining credits. A kernel density estimate (KDE) plot is provided to visualize the trends. We did not follow Quarfoot et al. [5] in counting accumulated votes over time due to varying total times across individuals.

Comparing experiment groups, we see fewer differences in the short QS but different interaction distributions between the two interfaces in the long QS. Given the significant differences in voting time between the text and two-phase interface for the long QS, we focus on deciphering the voting action changes between these two conditions in this subsection.

In Figure 15, we see two distinct patterns between the short survey and the long survey in terms of participant behaviors. In long surveys, participants exhibited more actions both when the budget was abundant and when it began to run out. This pattern was more pronounced with the long two-phase interface. We further separated the behaviors where participants made large or small changes to the options, specifically for the long version. In Figure 16, we define an adjustment of four or more votes as large, which we plotted in the first row of the figure. Adjustments of two or fewer votes are considered small, which is 10% of the possible values one can choose among the maximum of 21 votes.

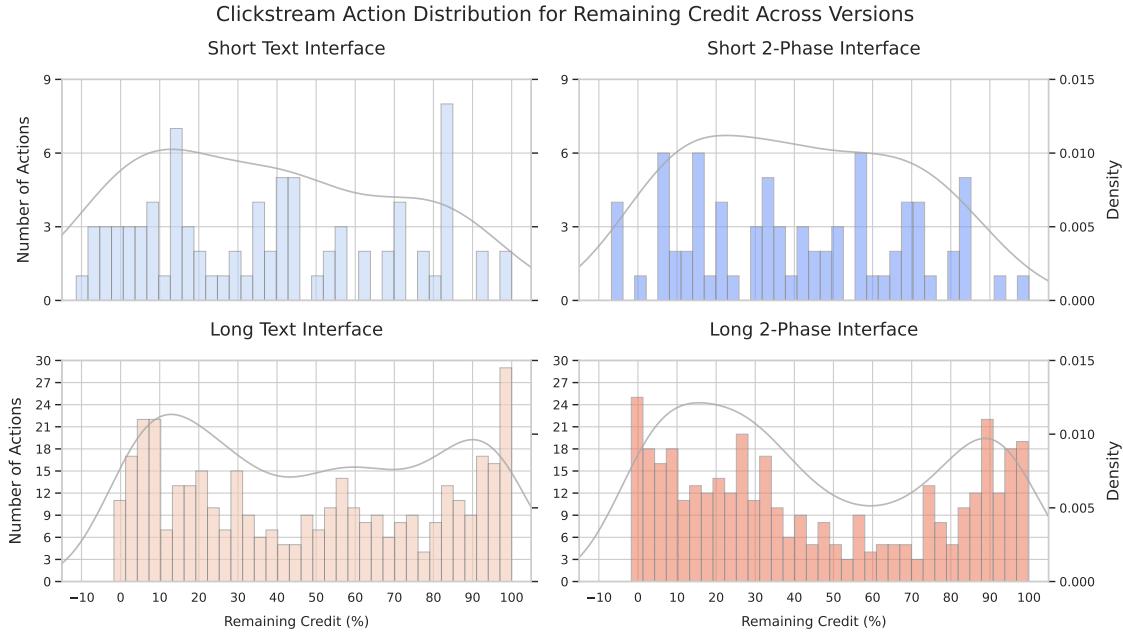


Fig. 15. This plot counts the number of voting actions when there are x percentages of credits remaining. A KDE plot is provided to help better understand the action distribution.

We plotted all actions against the time to complete them. Revisiting the KDE curve in the second row in Figure 15 and the curve of the second row in Figure 16 show a stronger bimodal distribution for small vote adjustments across interfaces. In fact, the bimodal distribution is more pronounced in the two-phase interface. This suggests that participants make small adjustments both at the beginning and toward the end of the QS. However, the two-phase interface shows more frequent and faster edits towards the end. Visually, dots are more clustered in the long two-phase interface for small vote adjustments compared to the long text interface. The Mann-Whitney U Test on the time spent on small vote adjustments showed significant differences ($U = 13037, p < 0.001$), with a small effect size (Rank-biserial: 0.227, Cohen's d : 0.195) and a power of 0.381. Based on the KDE plots in the first row of Figure 16, participants also made more large vote adjustments early on that spread more equally compared to the text interface. This indicates that participants had a clearer idea of how to distribute their credits across the options.

In interviews, five participants highlighted the importance of the interface's flexibility and their use of an incremental, iterative approach. All these participants used the two-phase interface. While this doesn't mean participants using the text interface didn't take an iterative approach, it highlights that the two-phase interface encouraged iterative and incremental updates. As one participant pointed out:

I like the fact that it remembers everything that you know [...] that's very important is that it's an iterative process. ↗ S019 (LI)

To summarize, participants spent more time on the two-phase interface compared to the text interface in both short and long surveys. Across the two-phase interfaces, organization time remained consistent. While voting time did not differ between interfaces for the short survey, participants voted more quickly on the two-phase interface in the long survey, confirming the hypothesis that the two-step design enhances decision-making efficiency. Voting

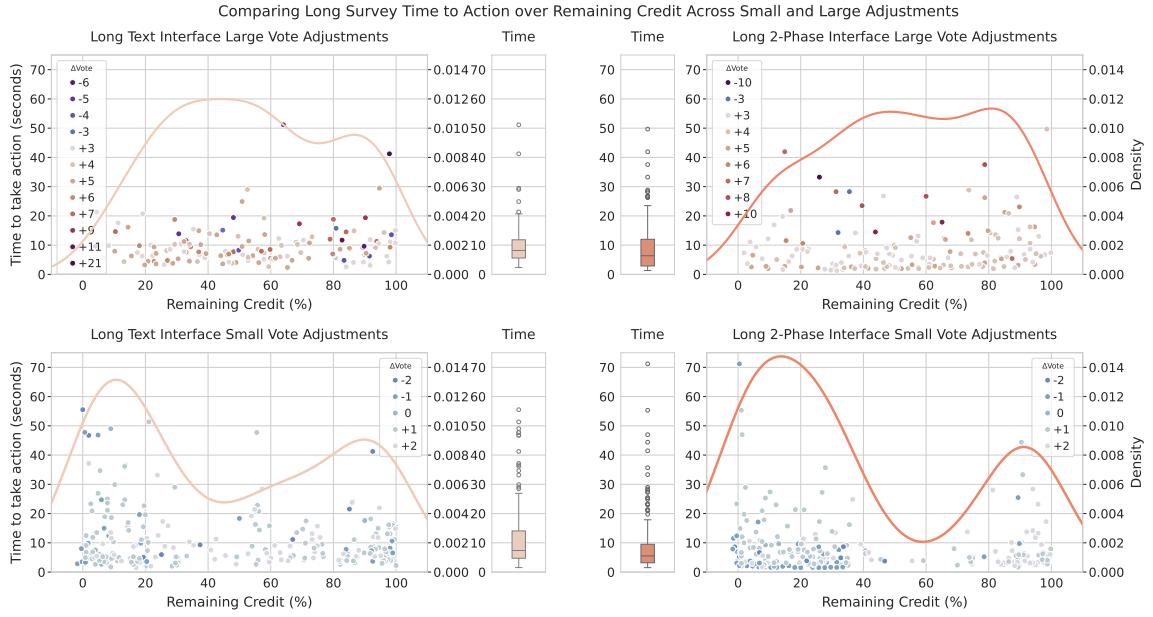


Fig. 16. This plot further separates participants' interaction behavior based on the number of votes participants adjusted. We observed a bimodal interaction pattern across long QS when small vote adjustments are made.

behaviors indicated more frequent actions when the budget was abundant and nearly exhausted, particularly in the long two-phase interface. Additionally, the analysis revealed more frequent and faster small vote adjustments towards the end of the QS in the two-phase interface, demonstrating an iterative and incremental approach.

7 Discussion and Future Works

This study proposed a two-phase interface for QS to study how the number of survey options and interfaces influence cognitive load and behavior in social resource allotment context. Section 5 and 6 revealed that longer surveys did not increase cognitive load, contrary to expectations.

In the discussion section, we interpret results related to cognitive load and survey respondent behaviors. We focus on three key topics: what two-phase interface elements influenced behavior, how these elements supported preference construction, and remaining design challenges. Additionally, we provide recommendations for using QS and suggest design improvements.

We concluded that the two-phase interface prevented satisficing behaviors and promoted more strategic, holistic thinking, unlike the text-based interface, which leans more toward operational tasks. Behavioral analysis showed that long two-phase interface participants made frequent, small updates, shifting their cognitive load focus.

7.1 Interpretation of results: Two-phase interface limits satisficing during cognitive overload

Participants using the long two-phase interface reported slightly higher cognitive load compared to those using the long text interface. In comparison, the short two-phase interface resulted in a lower cognitive load. This suggests that the two-phase interface alone does not inherently increase cognitive load.

1249 One explanation is that the two-phase interface reduces cognitive load in the short QS but not in the longer one. In
1250 longer surveys, the interactive nature of the two-phase interface may require participants to perform more operations
1251 without significantly altering their decision-making processes. However, our findings suggest that participants using
1252 the two-phase interface in longer surveys responded more efficiently. Across both short and long surveys, participants
1253 using the two-phase interface demonstrated iterative and efficient fine-tuning preferences, indicating deeper survey
1254 engagement. Therefore, we reject the claim that the two-phase interface increases cognitive load in the long survey.
1255

1256 An alternative explanation is that while the two-phase interface limits satisficing during cognitive overload, the
1257 complexity of the long QS may have counteracted this effect, preventing a full reduction in cognitive load. This suggests
1258 that long QS participants resorted to satisficing behaviors to manage their cognitive overload. Qualitative results support
1259 this explanation: participants using the two-phase interface engaged in broader, more strategic considerations, while
1260 those using the text interface focused on operational tasks. For instance, participants using the long two-phase interface
1261 experienced less cognitive burden from precise voting and managed their time better. Overall, the two-phase interface
1262 participants demonstrated deeper and more critical thinking, as they engaged in more strategic decision-making.
1263

1264 In contrast, participants using the long text interface satisficed due to cognitive overload from having to decide
1265 and allocate more credits for more options. However, fewer participants in this group reported high cognitive load
1266 compared to those in the short-text interface. This group also experienced the least temporal demand and frustration
1267 with operational tasks, despite spending a similar amount of time per option as participants using the text interface.
1268 These counterintuitive findings pointed to the remaining plausible explanation that long-text interface participants
1269 used satisficing to manage the overload.
1270

1271 **In summary**, the interactive components of the two-phase interface likely prevented mental shortcuts [53, 9, 10,
1272 11], resulting in a cognitive *shift* towards deeper reflection and decision-making. Thus, while QS with many options
1273 might lead to satisficing, the two-phase interface redirects participants' attention, facilitating the decision-making
1274 process that promotes comprehensive preference construction. In the following section, we examine these specific
1275 elements that guided participants to achieve this.
1276

1277 7.2 Bounded rationality and interface design

1278 Bounded rationality emerged as a core theme in participants' responses, highlighting how cognitive limitations lead to
1279 sub-optimal decision-making due to their inability to process all available information [9]. This often led to *satisficing*
1280 behaviors, where participants settle for *good enough* but not *optimal* decisions [91].
1281

1282 For instance, S036 described making meaningful choices while acknowledging the limits of their effort:
1283

1284 [...] you thought of enough things, you know, and so it wasn't the most effort I could put in because again, that would have been
1285 diminishing returns. I tried to think of enough things [...] and then move on. [...] I felt like that (the response) was satisfied, but not
1286 perfect. Cause perfect is not a reality. S036 (ST)
1287

1288 This illustrates typical satisficing decision-making, where participants settle for suboptimal choices. In long QS, the
1289 significant increase in decision points—due to the numerous vote and credit options—was so overwhelming to some
1290 that even features like pre-calculated vote-credit values in drop-down menus provided relief, as they helped avoid the
1291 cognitively demanding task of searching for bounds.
1292

1293 Participants often relied on *heuristics* [92] and *defaults* [64], behaviors less common among two-phase interface
1294 participants. Presenting one option at a time reduced reliance on defaults and encouraged deeper reflection, as illustrated
1295 by S013, who highlighted how the organization phase supported their option preferences construction:
1296

1301 [...] it (organization phase) gives you time to just focus on that single thing and rank it based on how you feel at that moment.

1302 ↗ S013 (SI)

1303

1304 Conversely, text interface participants, like S003, described how default placements influenced their decisions:

1305 Honestly, if medical research [...] was the first one I saw, I think it would automatically give it a lot more.

↗ S003 (ST)

1306

1307 The three key elements of the organization phase—presenting options one at a time, grouping them into categories,
1308 and enabling drag-and-drop—worked together to structure participant preferences. These elements aligned with
1309 cognitive strategies like *problem decomposition* [93] and *dimension reduction* which reduces cognitive overload. Bounded
1310 rationality illuminates the importance of decision-making support interfaces rather than being a critique of human
1311 behaviors. One participant explained how the organization phase broke down complex decisions into manageable steps:
1312

1313 [...] being able to have a preliminary categorization of all the topics. First, it introduced me to all the topics, [...] to think about
1314 and process [...] being able to digest all the information prior to actually allocating the budget or completing the quadratic survey.

1315 ↗ S009 (LI)

1316

1317 Participants using the two-phase interface, especially in the long version, organized options along dimensions such
1318 as topics (e.g., health vs. humanitarian) and preferences (positive vs. negative) before voting. Others expressed that
1319 the upfront introduction of all options and the ability to rank and group them helped manage their cognitive load
1320 effectively. In contrast, almost half of the participants using the long text interface, like S028, expressed a desire for
1321 features that can help reduce the decision space when responding to the QS, further supporting the importance of these
1322 organizational design elements:

1323 Because with this many (options), especially when I'm thinking ... Ok, where was (the option)... Where was (the option) you know?

1324 Oh, that's right. Maybe I could give another up another upvote to the, you know whatever [...]

↗ S028 (LT)

1325

1326 This quote reflected participants' need to manually track and revisit options, which occupies cognitive load, without
1327 a more structured interface.

1328

1329 In summary, individual's bounded rationality encouraged participants to exhibit *satisficing* behaviors, *heuristics*,
1330 and *defaults* when responding to QS. Showing all options upfront, one at a time, and repositioning options based on partic-
1331 ipants' rough preferences prevented participants from using defaults and heuristics. The two-phase organization actively
1332 scaffolds participants' decision-making process, supporting efficiency *problem decomposition* and *dimension reduction*.
1333 Together, these elements in the two-phase interface design prevented *satisficing* behaviors and supported participants
1334 in making more informed decisions through a more strategic planning and holistic thinking process.
1335

1336

1337 7.3 Construction of Preference on QS

1338 Completing QS is a series of difficult decision tasks Lichtenstein and Slovic [6]. Svenson [60]'s differentiation and
1339 consolidation theory helps explain how participants process these decisions. The decision process begins with dif-
1340 ferentiation, where participants identify differences and eliminate less favorable options, followed by consolidation,
1341 which strengthens their commitment to selected choices. This theory aligns with how the two-phase interface helps
1342 participants decompose options into categories, effectively reducing decision complexity.

1343 Participants started by constructing preferences in situ, especially regarding options they hadn't previously consid-
1344 ered:

1345

1346 [...] 'Oh, there are other aspects that I never care about.' And actually ... some people care <an option>. Sure. Why? Why (should) I
1347 spend money on that?

1348 ↗ S037 (LI)

1349 Manuscript submitted to ACM

1353 Those using the text interface, lacking the interactive tools, found it challenging to facilitate differentiation as S025
 1354 noted:

1355 I would like to be able to like, click and drag the categories themselves so I could maybe reorder them to like my priorities. [...] make
 1356 myself categories and subcategories out of this list ... If I could organize it.

Q S025 (LT)

1358 In contrast, the two-phase interface allowed participants to express at least one dimension of differentiation more
 1359 easily. The drag-and-drop feature helped blend this differentiation into the consolidation phase. Not only does partic-
 1360 ipants drag-and-drop options post voting to reflect and assure a correct vote allocation, it also enables participants,
 1361 like S039, to make fine-grain comparisons between options:

1363 I think the system was actually really helpful because I could just drag them. [...] I can really compare them, I can drag this one up
 1364 here, and then compare it to the top one [...]

Q S039 (SI)

1366 The bi-modal behavior observed in the long interactive interface participants provided in the results aligns with the
 1367 differentiation and consolidation framework. Participants in the two-phase interface began differentiating options earlier
 1368 allowing them to later adjust fine-grain votes. The faster and smaller vote updates indicated participants consolidating.
 1369 The less prominent bi-modal behavior from the long text interface participants implied that the interface guided this
 1370 decision framework as participant 037 explained:

1373 I only start from the positive one [...] I finish everything ... and then I move to the second part (the neutral box). [...] I want to focus
 1374 on these and make sure that resources are at least they get the attention they want. And if there's surplus and they can move to the
 1375 second part.

Q S037 (LI)

1376 These evidences explain how the organization phase and the drag-and-drop features supported differentiation and
 1377 consolidation, scaffolding a decision-making framework enabling deeper engagement.

1379 **In summary**, participants construct their preferences as they complete QS. We observed behaviors and qualitative
 1380 insights that align with the differentiation and consolidation theory in decision-making. Our interface scaffolded many
 1381 of the differentiation stages through pre-voting organization and some consolidation phases through drag-and-drop,
 1382 explaining how the two-phase approach supports preference construction to yield effective QS responses.

1385 7.4 Opportunities for better budget management

1386 Budget management is a recurring theme in participant interviews. While they appreciated the automatic calculation
 1387 feature in modern QV interfaces, we identified three challenges for future QS interfaces: *cognitive load*, *the cold-start*
 1388 *problem*, and *navigating between budget, votes, and outcome*.

1390 **7.4.1 Automatic calculation is critical.** Over one-third of participants ($N = 14$) from all four experiment conditions
 1391 emphasized the importance of automated calculation for deriving costs and tracking expenditures. For example:

1394 I thought I have [...] (to) do all the numbers or calculations myself as a part of checking my ability of doing mathematics. [...] I said
 1395 that credit summary to be very specific. The credit summary section was really wonderful in doing all the calculations on that end.

Q S005 (LT, keeping track of spent)

1397 The quotes marked the importance that QS must be facilitated by computer-supported interfaces.

1399 **7.4.2 The coldstart problem.** We notice from the study that one of the biggest challenges for participants is deciding
 1400 'how many votes' to start with. This challenge pertains to the initial vote, not the relative vote. Some participants began
 1401 by equally distributing their credits to all options and then made adjustments. Others established 1, 2, and 3 votes as
 1402 starting points. A small handful surprisingly used the tutorial's example of 4 upvotes as their anchor.

1405 This arbitrary voting decision echoes discussions in prior literature about the existence of an absolute value for
 1406 individuals. Coherent arbitrariness [94], similar to the anchoring effect in marketing, refers to participants' willingness
 1407 to allocate votes, which can be influenced by an arbitrary value. However, the ordinal utility remains intact among the
 1408 set of preferences.
 1409

1410 7.4.3 *Navigating Between Budget, Votes, and Actual Impact.* The third challenge is participants' confusion between
 1411 budget, votes, and outcomes, despite understanding their definitions. One participant stated:
 1412

1413 [...] get rid of the upvote column or just get rid of the word upvote and just really focus on the money column. Listen. You're an
 1414 organization or your participant. You have X amount of dollars you need to. You can only distribute X amount of dollars to these
 1415 causes. So you have to figure out which ones get the most, which ones don't get as much. [...]
 1416

1417 Interviewer: [...] Do you feel that the more votes you're giving to a cause you're actually spending more on it?
 1418 Yeah.

1419 Q S003 (ST)

1420 Participants like S003 bypassed the quadratic formulation, directly translating votes to resource allocation. While this
 1421 does not invalidate the power of the quadratic mechanism, it causes frustration and friction for participants to construct
 1422 a clear picture of how to make voting decisions. Future interfaces should better communicate these relationships to
 1423 facilitate respondents' trade-offs.
 1424

1425 In summary, while the interface supports budget management through automated cost calculation, participants still
 1426 face cognitive load from managing the budget. The cold-start problem and the confusion between budget, votes, and
 1427 actual impact are open questions for future research. These challenges highlight the need for better budget management
 1428 support to complete the QS interface.
 1429

1430 7.5 Quadratic Survey Usage, Design Recommendations and Future Work

1431 With a deeper understanding of how survey respondents interact with QS and the sources of cognitive load, we
 1432 recognize that while this current interface may not significantly reduce cognitive load, it represents a crucial step
 1433 toward constructing better interfaces to support individuals responding to QS. In this subsection, we outline usage
 1434 and design recommendations applicable to all applications using the quadratic mechanism and highlight directions for
 1435 future work.
 1436

1437 7.5.1 *Usage Recommendation: QS for Critical Evaluations.* Our study highlighted the complex cognitive challenges
 1438 and in-depth consideration required when ranking and rating options using QS, even in a short survey. Similar to
 1439 survey respondents needing to make trade-offs across options, researchers and agencies seeking additional insights and
 1440 alignment with respondent preferences must ensure that survey respondents have the cognitive capacity to complete
 1441 such surveys rigorously. QS should be designed for critical evaluations, such as investment decisions, or situations
 1442 where participants have ample time to think and process the survey. For instance, revealing the options ahead of time
 1443 can aid in preference construction.
 1444

1445 7.5.2 *Design Recommendations.*

1446 Use Organization Phases for Quadratic Mechanism Applications. Our study demonstrated that preference construc-
 1447 tion can shift from operational to strategic and higher-level causes. An additional organizational phase with direct
 1448 manipulation capability allows survey respondents to engage in higher-level critical thinking. We believe this approach
 1449 should extend beyond QS to other ranking-based surveying tools, such as rank-choice voting and constant sum surveys.
 1450 Further research should examine how implementing such functionality alters survey respondents' mental models.
 1451

1457 *Facilitate Differentiation through Categorization, Not Ranking.* Participants in our study were less inclined to provide
1458 a full rank unless necessary. The final 'rank' of option preferences often emerged as a byproduct of their vote allocation,
1459 constructed in situ. Therefore, for survey designs to be effective in constructing preferences, it is more important to
1460 facilitate differentiation than to focus on direct manipulation solely for fine-tuning. Emphasizing categorization can
1461 better support participants in articulating their preferences.
1462

1463
1464 *7.5.3 Future Work: Support for Absolute Credit Decision.* Deciding the absolute amount of credits in QS is highly
1465 demanding. Designing interfaces and interactions that address the cold start challenge and help participants decide the
1466 absolute vote value while considering ways to limit direct influences remains an open question. Future research should
1467 explore innovative solutions to support participants in making these complex decisions effectively.
1468

1469 By implementing these recommendations and pursuing future research directions, we can improve the usability and
1470 effectiveness of QS and other quadratic mechanism-powered applications, ultimately aiding respondents in making
1471 more informed and accurate decisions.
1472

1473 1474 8 Limitations

1475 Evaluating the QS interface is challenging due to its novelty. During the study, we identified several limitations that
1476 require further research.
1477

1478 *Understanding results influence on decision-makers.* Further research is required to understand how the QS interface
1479 impacts decision-makers and broader societal resource distributions. Since QS is still in its early stages, we prioritize its
1480 widespread adoption and usage before attempting a comprehensive assessment of its influence on decision-making.
1481 Future studies will examine how decision-makers interpret and use QS data, as well as its broader implications for
1482 societal decisions.
1483

1484 *Individual differences in cognitive capacity.* Variations in individual cognitive capacity influenced participants' cog-
1485 nitive scores. For example, participants with more experience in decision-making might be able to manage multiple
1486 options more effectively. A within-subject study could clarify cognitive load shifts, but deconstructing established
1487 preferences and altering options further complicates this. Thus, we opted for this in-depth, between-subject study,
1488 although the small sample size may introduce noise that distorts the actual cognitive load. Future research should
1489 quantify the impact of different QS interfaces.
1490

1491 *Limited experience with QS.* Participants had no prior experience with the QS interface. Following a tutorial and
1492 quiz, participants proceeded to complete tasks using the QS interface. While participants understood the QS mechanics,
1493 familiarity with the interface still influences strategies and cognitive load. As quadratic mechanisms become more
1494 prevalent, future research can compare novices and experts.
1495

1496 *Duration between clicks to represent decision-making.* Click duration may include time spent considering other options,
1497 so it must be treated as an approximate measure of decision-making time. For instance, deciding between two options
1498 may take longer for the first option and less time for the second. Despite its limitations, this approach provides valuable
1499 insights into decision-making within our experimental constraints.
1500

1509 9 Conclusion

1510 Surveys enable decision-makers to aggregate crowd opinions. In this study, we use QS to elicit individual responses in
 1511 the context of social resource allotment. After multiple design iterations, we propose an two-phase interface for QS. We
 1512 then examined its influence on individuals' cognitive load and behaviors when faced with societal issues of varying
 1513 lengths. In a 2x2 between-subject study, we had participants experience either a long or short QS using a text-based or
 1514 two-phase interface. NASA-TLX questionnaires and interviews revealed that participants using the two-phase interface
 1515 for a long QS demonstrated a more comprehensive and critical evaluation of societal issues, despite not experiencing a
 1516 lower cognitive load. Participants using the long text interface experienced cognitive overload, which led to satisficing
 1517 behaviors or mental shortcuts. Analyzing click-stream data, we identified that participants made fine-grain iterations
 1518 using the long two-phase interface when credits were low. We demonstrate that a two-phase, organize-then-vote
 1519 interface can scaffold the complex decision-making process, helping individuals express their opinions for collective
 1520 societal decisions. Through the iterative design process and detailed interviews, we identified future directions and
 1521 design recommendations for collective decision-making applications using the quadratic mechanism.
 1522

1523 References

- [1] Ziang Xiao, Sarah Mennicken, Bernd Huber, Adam Shonkoff, and Jennifer Thom. 2021. Let Me Ask You This: How Can a Voice Assistant Elicit Explicit User Feedback? *Proceedings of the ACM on Human-Computer Interaction*, 5, CSCW2, (Oct. 2021), 1–24. doi: [10.1145/3479532](https://doi.org/10.1145/3479532).
- [2] Martin Pielot and Mario Callegaro. 2024. Did You Misclick? Reversing 5-Point Satisfaction Scales Causes Unintended Responses. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, (May 2024), 1–7. doi: [10.1145/3613904.3642397](https://doi.org/10.1145/3613904.3642397).
- [3] Ti-Chung Cheng, Tiffany Li, Yi-Hung Chou, Karrie Karahalios, and Hari Sundaram. 2021. "I can show what I really like": Eliciting Preferences via Quadratic Voting. *Proceedings of the ACM on Human-Computer Interaction*, 5, (Apr. 2021), 1–43. doi: [10.1145/3449281](https://doi.org/10.1145/3449281).
- [4] Theodore Groves and John Ledyard. 1977. Optimal Allocation of Public Goods: A Solution to the "Free Rider" Problem. *Econometrica*, 45, 4, 783–809. JSTOR: [10.2307/1912672](https://doi.org/10.2307/1912672). doi: [10.2307/1912672](https://doi.org/10.2307/1912672).
- [5] David Quarfoot, Douglas von Kohorn, Kevin Slavin, Rory Sutherland, David Goldstein, and Ellen Konar. 2017. Quadratic voting in the wild: real people, real votes. *Public Choice*, 172, 1–2, 283–303.
- [6] Sarah Lichtenstein and Paul Slovic, eds. 2006. *The Construction of Preference*. (1. publ ed.). Cambridge University Press, Cambridge.
- [7] Adam Rogers. 2019. Colorado Tried a New Way to Vote: Make People Pay—Quadratically | WIRED. *Wired*, (Apr. 2019). Retrieved June 22, 2024 from.
- [8] Internet Team. [n. d.] Taiwan Digital Minister highlights country's use of technology to bolster democracy in FT interview. https://www.roctaiwan.org/uk_en/post/6295.html (). Retrieved June 13, 2024 from.
- [9] Herbert A. Simon. 1955. A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69, 1, 99–118. JSTOR: [10.2307/1884852](https://doi.org/10.2307/1884852). doi: [10.2307/1884852](https://doi.org/10.2307/1884852).
- [10] John W. Payne, James R. Bettman, and Eric J. Johnson. 1988. Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 3, (July 1988), 534–552. doi: [10.1037/0278-7393.14.3.534](https://doi.org/10.1037/0278-7393.14.3.534).
- [11] Amos Tversky and Daniel Kahneman. [n. d.] Judgments of and by Representativeness.
- [12] Erik J Engstrom and Jason M Roberts. 2020. *The Politics of Ballot Design: How States Shape American Democracy*. Cambridge University Press.
- [13] Bert Weijters, Elke Cabooter, and Niels Schillewaert. 2010. The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27, 3, (Sept. 2010), 236–247. doi: [10.1016/j.ijresmar.2010.02.004](https://doi.org/10.1016/j.ijresmar.2010.02.004).
- [14] N. D. Kieruj and G. Moors. 2010. Variations in Response Style Behavior by Response Scale Format in Attitude Research. *International Journal of Public Opinion Research*, 22, 3, (Sept. 2010), 320–342. doi: [10.1093/ijpor/edq001](https://doi.org/10.1093/ijpor/edq001).
- [15] Vera Toe poel, Brenda Vermeeren, and Baran Metin. 2019. Smiley, Stars, Hearts, Buttons, Tiles or Grids: Influence of Response Format on Substantive Response, Questionnaire Experience and Response Time. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 142, 1, (Apr. 2019), 57–74. doi: [10.1177/0759106319834665](https://doi.org/10.1177/0759106319834665).
- [16] Habiba Farzand, David Al Baiaty Suarez, Thomas Goodge, Shaun Alexander Macdonald, Karola Marky, Mohamed Khamis, and Paul Cairns. 2024. Beyond Aesthetics: Evaluating Response Widgets for Reliability & Construct Validity of Scale Questionnaires. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, (May 2024), 1–7. doi: [10.1145/3613905.3650751](https://doi.org/10.1145/3613905.3650751).
- [17] Christian Jilek Paula Gauselmann Yannick Runge and Tobias Tempel. 2023. A relief from mental overload in a digitalized world: How context-sensitive user interfaces can enhance cognitive performance. *International Journal of Human-Computer Interaction*, 39, 1, 140–150. eprint: <https://doi.org/10.1080/10447318.2022.2041882>. doi: [10.1080/10447318.2022.2041882](https://doi.org/10.1080/10447318.2022.2041882).

- 1561 [18] A Norman Donald. 2013. *The Design of Everyday Things*. MIT Press.
- 1562 [19] Michael Xieyang Liu, Aniket Kittur, and Brad A. Myers. 2021. To reuse or not to reuse? A framework and system for evaluating summarized knowledge. *Proc. ACM Hum.-Comput. Interact.*, 5, CSCW1, (Apr. 2021). doi: [10.1145/3449240](https://doi.org/10.1145/3449240).
- 1563 [20] Benjamin Lafreniere, Andrea Bunt, and Michael Terry. 2014. Task-centric interfaces for feature-rich software. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design* (OzCHI '14). Association for Computing Machinery, New York, NY, USA, 49–58. doi: [10.1145/2686612.2686620](https://doi.org/10.1145/2686612.2686620).
- 1564 [21] Soomin Kim, Jinsu Eun, Joseph Seering, and Joonhwan Lee. 2021. Moderator chatbot for deliberative discussion: Effects of discussion structure and discussant facilitation. *Proc. ACM Hum.-Comput. Interact.*, 5, CSCW1, (Apr. 2021). doi: [10.1145/3449161](https://doi.org/10.1145/3449161).
- 1565 [22] Amy X. Zhang and Justin Cranshaw. 2018. Making sense of group chat through collaborative tagging and summarization. *Proc. ACM Hum.-Comput. Interact.*, 2, CSCW, (Nov. 2018). doi: [10.1145/3274465](https://doi.org/10.1145/3274465).
- 1566 [23] Steven P Lalley, E Glen Weyl, et al. 2016. Quadratic voting. Available at SSRN.
- 1567 [24] Eric A Posner and E Glen Weyl. 2018. *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press.
- 1568 [25] Ryan Naylor et al. 2017. First year student conceptions of success: What really matters? *Student Success*, 8, 2, 9–19.
- 1569 [26] Charlotte Cavaille and Daniel L Chen. [n. d.] Who Cares? Measuring Preference Intensity in a Polarized Environment.
- 1570 [27] Vitalik Buterin, Zoë Hitzig, and E. Glen Weyl. 2019. A Flexible Design for Funding Public Goods. *Management Science*, 65, 11, (Nov. 2019), 5171–5187. doi: [10.1287/mnsc.2019.3337](https://doi.org/10.1287/mnsc.2019.3337).
- 1571 [28] Luis Mota Freitas and Wilfredo L. Maldonado. 2024. Quadratic funding with incomplete information. *Social Choice and Welfare*, (Feb. 2024). doi: [10.1007/s00355-024-01512-7](https://doi.org/10.1007/s00355-024-01512-7).
- 1572 [29] Tobin South, Leon Erichsen, Shrey Jain, Petar Maymounkov, Scott Moore, and E. Glen Weyl. 2024. Plural Management. SSRN Scholarly Paper. Rochester, NY, (Jan. 2024). doi: [10.2139/ssrn.4688040](https://doi.org/10.2139/ssrn.4688040).
- 1573 [30] 2023. Gov4git: A Decentralized Platform for Community Governance. (Mar. 2023). Retrieved June 13, 2024 from.
- 1574 [31] 2024. RadicalxChange/quadratic-voting. RadicalxChange. (May 2024). Retrieved June 17, 2024 from.
- 1575 [32] [n. d.] Read the Whitepaper | Gitcoin. <https://www.gitcoin.co/whitepaper/read/>. (). Retrieved June 17, 2024 from.
- 1576 [33] [n. d.] About RxC. <https://www.radicalxchange.org/wiki/about/>. (). Retrieved June 17, 2024 from.
- 1577 [34] yehjxraymond. 2024. Yehjxraymond/qv-app. (Mar. 2024). Retrieved June 17, 2024 from.
- 1578 [35] Syavash Nobarany, Louise Oram, Vasanth Kumar Rajendran, Chi-Hsiang Chen, Joanna McGrenere, and Tamara Munzner. 2012. The design space of opinion measurement interfaces: exploring recall support for rating and ranking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2035–2044.
- 1579 [36] Paul Van Schaik and Jonathan Ling. 2007. Design parameters of rating scales for web sites. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 14, 1, 4–es.
- 1580 [37] Jing Wei, Weiwei Jiang, Chaofan Wang, Difeng Yu, Jorge Goncalves, Tilman Dingler, and Vassilis Kostakos. 2022. Understanding how to administer voice surveys through smart speakers. *Proc. ACM Hum.-Comput. Interact.*, 6, CSCW2, (Nov. 2022). doi: [10.1145/3555606](https://doi.org/10.1145/3555606).
- 1581 [38] Aman Khullar et al. 2021. Costs and benefits of conducting voice-based surveys versus keypress-based surveys on interactive voice response systems. In *Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies (Compass '21)*. Association for Computing Machinery, New York, NY, USA, 288–298. doi: [10.1145/3460112.3471963](https://doi.org/10.1145/3460112.3471963).
- 1582 [39] Soomin Kim, Joonhwan Lee, and Gahgene Gweon. 2019. Comparing Data from Chatbot and Web Surveys: Effects of Platform and Conversational Style on Survey Response Quality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland UK, (May 2019), 1–12. doi: [10.1145/3290605.3300316](https://doi.org/10.1145/3290605.3300316).
- 1583 [40] Martin Feick, Niko Kleer, Anthony Tang, and Antonio Krüger. 2020. The virtual reality questionnaire toolkit. In *Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 68–69.
- 1584 [41] Bert Weijters, Kobe Millet, and Elke Cabooter. 2021. Extremity in horizontal and vertical Likert scale format responses. Some evidence on how visual distance between response categories influences extreme responding. *International Journal of Research in Marketing*, 38, 1, (Mar. 2021), 85–103. doi: [10.1016/j.ijresmar.2020.04.002](https://doi.org/10.1016/j.ijresmar.2020.04.002).
- 1585 [42] Vera Toe poel and Frederik Funke. 2018. Sliders, visual analogue scales, or buttons: Influence of formats and scales in mobile and desktop surveys. *Mathematical Population Studies*, 25, 2, (Apr. 2018), 112–122. doi: [10.1080/08898480.2018.1439245](https://doi.org/10.1080/08898480.2018.1439245).
- 1586 [43] Jonathan N. Wand, Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Michael C. Herron, and Henry E. Brady. 2001. The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. *The American Political Science Review*, 95, 4, 793–810. Retrieved Dec. 16, 2023 from JSTOR: [3117714](https://doi.org/10.1215/0003055X-95-4-793).
- 1587 [44] Dana Chisnell. 2016. Democracy Is a Design Problem. 11, 4.
- 1588 [45] 2015. Designing usable ballots | Center for civic design. <https://civicdesign.org/fieldguides/designing-usable-ballots/>. (June 2015). Retrieved June 17, 2024 from.
- 1589 [46] Whitney Quesenberry. 2020. Opinion | Good Design Is the Secret to Better Democracy. *The New York Times*, (Oct. 2020). Retrieved June 17, 2024 from.
- 1590 [47] Sarah P. Everett, Kristen K. Greene, Michael D. Byrne, Dan S. Wallach, Kyle Derr, Daniel Sandler, and Ted Torous. 2008. Electronic voting machines versus traditional methods: improved preference, similar performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. Association for Computing Machinery, New York, NY, USA, (Apr. 2008), 883–892. doi: [10.1145/1357054.1357195](https://doi.org/10.1145/1357054.1357195).

- 1613 [48] Seunghyun "Tina" Lee, Yilin Elaine Liu, Ljilja Ruzic, and Jon Sanford. 2016. Universal Design Ballot Interfaces on Voting Performance and
 1614 Satisfaction of Voters with and without Vision Loss. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16).
 1615 Association for Computing Machinery, New York, NY, USA, (May 2016), 4861–4871. doi: [10.1145/2858036.2858567](https://doi.org/10.1145/2858036.2858567).
- 1616 [49] Kathryn Summers, Dana Chisnell, Drew Davies, Noel Alton, and Megan McKeever. 2014. Making voting accessible: designing digital ballot
 1617 marking for people with low literacy and mild cognitive disabilities. In *2014 Electronic Voting Technology Workshop/Workshop on Trustworthy
 1618 Elections (EVT/WOTE 14)*.
- 1619 [50] Shaneé Dawkins, Tony Sullivan, Greg Rogers, E. Vincent Cross, Lauren Hamilton, and Juan E. Gilbert. 2009. Prime III: an innovative electronic
 1620 voting interface. In *Proceedings of the 14th International Conference on Intelligent User Interfaces* (IUI '09). Association for Computing Machinery,
 1621 New York, NY, USA, (Feb. 2009), 485–486. doi: [10.1145/1502650.1502727](https://doi.org/10.1145/1502650.1502727).
- 1622 [51] Juan E. Gilbert, Jerome Dunbar, Alvitta Ottley, and John Mark Smotherman. 2013. Anomaly detection in electronic voting systems. *Information
 1623 Design Journal (IDJ)*, 20, 3, (Sept. 2013), 194–206. doi: [10.1075/ijd.20.3.01gil](https://doi.org/10.1075/ijd.20.3.01gil).
- 1624 [52] Frederick G. Conrad, Benjamin B. Bederson, Brian Lewis, Emilia Peytcheva, Michael W. Traugott, Michael J. Hanmer, Paul S. Herrnson, and Richard
 1625 G. Niemi. 2009. Electronic voting eliminates hanging chads but introduces new usability challenges. *International Journal of Human-Computer
 1626 Studies*, 67, 1, (Jan. 2009), 111–124. doi: [10.1016/j.ijhcs.2008.09.010](https://doi.org/10.1016/j.ijhcs.2008.09.010).
- 1627 [53] Kahneman Daniel. 2017. *Thinking, Fast and Slow*.
- 1628 [54] Sheena S Iyengar and Mark R Lepper. 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and
 1629 social psychology*, 79, 6, 995.
- 1630 [55] Duane F Alwin and Jon A Krosnick. 1985. The measurement of values in surveys: A comparison of ratings and rankings. *Public Opinion Quarterly*,
 1631 49, 4, 535–552.
- 1632 [56] N. T. Feather. 1973. The measurement of values: Effects of different assessment procedures. *Australian Journal of Psychology*, 25, 3, (Dec. 1973),
 1633 221–231. doi: [10.1080/00049537308255849](https://doi.org/10.1080/00049537308255849).
- 1634 [57] Peter Coy. 2019. A New Way of Voting That Makes Zealotry Expensive - Bloomberg. *Bloomberg*, (May 2019). Retrieved Dec. 16, 2023 from.
 1635 2022. Quadratic Voting Frontend. Public Digital Innovation Space. (Jan. 2022). Retrieved Dec. 16, 2023 from.
- 1636 [58] Henry Montgomery. 1983. Decision Rules and the Search for a Dominance Structure: Towards a Process Model of Decision Making. In *Advances
 1637 in Psychology*. Vol. 14. Elsevier, 343–369. doi: [10.1016/S0166-4115\(08\)62243-8](https://doi.org/10.1016/S0166-4115(08)62243-8).
- 1638 [60] Ola Svenson. 1992. Differentiation and consolidation theory of human decision making: A frame of reference for the study of pre- and post-decision
 1639 processes. *Acta Psychologica*, 80, 1-3, (Aug. 1992), 143–168. doi: [10.1016/0001-6918\(92\)90044-E](https://doi.org/10.1016/0001-6918(92)90044-E).
- 1640 [61] Fritz Strack and Leonard L. Martin. 1987. Thinking, Judging, and Communicating: A Process Account of Context Effects in Attitude Surveys. In
 1641 *Social Information Processing and Survey Methodology*. Recent Research in Psychology. Hans-J. Hippler, Norbert Schwarz, and Seymour Sudman,
 1642 editors. Springer, New York, NY, 123–148. doi: [10.1007/978-1-4612-4798-2_7](https://doi.org/10.1007/978-1-4612-4798-2_7).
- 1643 [62] John Sweller. 2011. Cognitive Load Theory. In *Psychology of Learning and Motivation*. Vol. 55. Elsevier, 37–76. doi: [10.1016/B978-0-12-387691-1.0002-8](https://doi.org/10.1016/B978-0-12-387691-1.0002-8).
- 1644 [63] Robert Münscher, Max Vetter, and Thomas Scheuerle. 2016. A Review and Taxonomy of Choice Architecture Techniques. *Journal of Behavioral
 1645 Decision Making*, 29, 5, 511–524. doi: [10.1002/bdm.1897](https://doi.org/10.1002/bdm.1897).
- 1646 [64] Richard H. Thaler and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. *Nudge: Improving Decisions about
 1647 Health, Wealth, and Happiness*. Yale University Press, New Haven, CT, US, x, 293.
- 1648 [65] Christopher D Wickens and Anthony D Andre. 1990. Proximity compatibility and information display: Effects of color, space, and objectness on
 1649 information integration. *Human factors*, 32, 1, 61–77.
- 1650 [66] Jon A Krosnick, Charles M Judd, and Bernd Wittenbrink. 2018. The measurement of attitudes. In *The Handbook of Attitudes*. Routledge, 45–105.
- 1651 [67] Jerry P Timbrook. 2013. *A Comparison of a Traditional Ranking Format to a Drag-and-Drop Format with Stacking*. PhD thesis. University of Dayton.
- 1652 [68] Duncan Rintoul. [n. d.] Visual and animated response formats in web surveys: Do they produce better data, or is it all just fun and games?, 126.
- 1653 [69] Susan C. Herring and Ashley R. Dainas. 2020. Gender and Age Influences on Interpretation of Emoji Functions. *ACM Transactions on Social
 1654 Computing*, 3, 2, (June 2020), 1–26. doi: [10.1145/3375629](https://doi.org/10.1145/3375629).
- 1655 [70] Weidong Huang, Peter Eades, and Seok-Hee Hong. 2009. Measuring Effectiveness of Graph Visualizations: A Cognitive Load Perspective.
 1656 *Information Visualization*, 8, 3, (Sept. 2009), 139–152. doi: [10.1057/ivs.2009.10](https://doi.org/10.1057/ivs.2009.10).
- 1657 [71] Robert Ferber. 1952. Order Bias in a Mail Survey. *Journal of Marketing*, 17, 2, 171–178. JSTOR: [1248043](https://doi.org/10.2307/1248043). doi: [10.2307/1248043](https://doi.org/10.2307/1248043).
- 1658 [72] M. P. Couper. 2001. Web survey design and administration. *Public Opinion Quarterly*, 65, 2, 230–253. doi: [10.1086/322199](https://doi.org/10.1086/322199).
- 1659 [73] 2023. Charity Navigator. <https://www.charitynavigator.org/index.cfm?bay=search.categories>. (May 2023). Retrieved Dec. 16, 2023 from.
- 1660 [74] William F. Moroney and Joyce A. Cameron. 2019. *Questionnaire Design: How to Ask the Right Questions of the Right People at the Right Time to Get
 1661 the Information You Need*. Human Factors and Ergonomics Society, (Feb. 2019).
- 1662 [75] Thomas L. Saaty and Kirti Peniwati. 2013. *Group Decision Making: Drawing Out and Reconciling Differences*. RWS Publications, (Nov. 2013).
- 1663 [76] Thomas L. Saaty. 1987. Principles of the Analytic Hierarchy Process. In *Expert Judgment and Expert Systems*. Jery L. Mumford, Ortwin Renn,
 1664 Lawrence D. Phillips, and V. R. R. Uppuluri, editors. Springer, Berlin, Heidelberg, 27–73. doi: [10.1007/978-3-642-86679-1_3](https://doi.org/10.1007/978-3-642-86679-1_3).
- 1665 [77] George A. Miller. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological
 1666 Review*, 63, 2, 81–97. doi: [10.1037/h0043158](https://doi.org/10.1037/h0043158).
- 1667 Manuscript submitted to ACM

- [78] Thomas L Saaty and Mujgan S Ozdemir. 2003. Why the magic number seven plus or minus two. *Mathematical and computer modelling*, 38, 3-4, 233–244.
- [79] Alexander Chernev, Ulf Böckenholt, and Joseph Goodman. 2015. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25, 2, (Apr. 2015), 333–358. doi: [10.1016/j.jcps.2014.08.002](https://doi.org/10.1016/j.jcps.2014.08.002).
- [80] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183.
- [81] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50, 9, (Oct. 2006), 904–908. doi: [10.1177/154193120605000909](https://doi.org/10.1177/154193120605000909).
- [82] Brad Cain. 2007. A review of the mental workload literature. *DTIC Document*.
- [83] Qin Gao, Yang Wang, Fei Song, Zhizhong Li, and Xiaolu Dong. 2013. Mental workload measurement for emergency operating procedures in digital nuclear power plants. *Ergonomics*, 56, 7, (July 2013), 1070–1085. doi: [10.1080/00140139.2013.790483](https://doi.org/10.1080/00140139.2013.790483).
- [84] Susana Rubio, Eva Díaz, Jesús Martín, and José M. Puenté. 2004. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, 53, 1, 61–86. doi: [10.1111/j.1464-0597.2004.00161.x](https://doi.org/10.1111/j.1464-0597.2004.00161.x).
- [85] Judith S. Olson and Wendy A. Kellogg, eds. 2014. *Ways of Knowing in HCI*. Springer, New York, NY. doi: [10.1007/978-1-4614-0378-8](https://doi.org/10.1007/978-1-4614-0378-8).
- [86] Evan Polman. 2010. Why are maximizers less happy than satisficers? Because they maximize positive and negative outcomes. *Journal of Behavioral Decision Making*, 23, 2, 179–190. doi: [10.1002/bdm.647](https://doi.org/10.1002/bdm.647).
- [87] Barry Schwartz, Andrew Ward, John Monterosso, Sonja Lyubomirsky, Katherine White, and Darrin R. Lehman. 2002. Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology*, 83, 5, 1178–1197. doi: [10.1037/0022-3514.83.5.1178](https://doi.org/10.1037/0022-3514.83.5.1178).
- [88] John W. Payne, James R. Bettman, and Eric J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press, Cambridge. doi: [10.1017/CBO9781139173933](https://doi.org/10.1017/CBO9781139173933).
- [89] Anuj K. Shah, Eldar Shafir, and Sendhil Mullainathan. 2015. Scarcity frames value. *Psychological Science*, 26, 4, 402–412.
- [90] Ernst-Jan de Bruijn and Gerrit Antonides. 2022. Poverty and economic decision making: a review of scarcity theory. *Theory and Decision*, 92, 1, (Feb. 2022), 5–37. doi: [10.1007/s11238-021-09802-7](https://doi.org/10.1007/s11238-021-09802-7).
- [91] Gerd Gigerenzer and Daniel G. Goldstein. 1996. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 4, 650–669. doi: [10.1037/0033-295X.103.4.650](https://doi.org/10.1037/0033-295X.103.4.650).
- [92] Amos Tversky and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science*, 185, 4157, 1124–1131. Retrieved June 21, 2024 from JSTOR: [1738360](https://doi.org/10.1126/science.1738360).
- [93] Herbert A. Simon. 1996. *The Sciences of the Artificial*. (3rd ed ed.). MIT Press, Cambridge, Mass.
- [94] Dan Ariely, George Loewenstein, and Drazen Prelec. 2003. “Coherent Arbitrariness”: Stable Demand Curves Without Stable Preferences*. *The Quarterly Journal of Economics*, 118, 1, (Feb. 2003), 73–106. doi: [10.1162/00335530360535153](https://doi.org/10.1162/00335530360535153).

A Voting Interface Breakdown

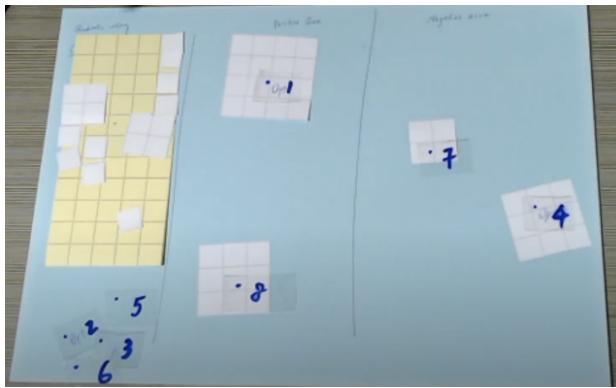
Designs that shifted voter decisions: For example, states without straight-party ticket voting (where voters can select all candidates from one party through a single choice) exhibited higher rates of split-ticket voting [12]. Another example from the Australian ballot showing incumbency advantages is where candidates are listed by the office they are running for, with no party labels or boxes.

Designs that influenced errors: Butterfly ballots increased voter errors because voters could not correctly identify the punch hole on the ballot. Splitting contestants across columns increases the chance for voters to overvote [46]. On the other hand, Everett et al. [47] showed the use of incorporating physical voting behaviors, like lever voting, into graphical user interfaces.

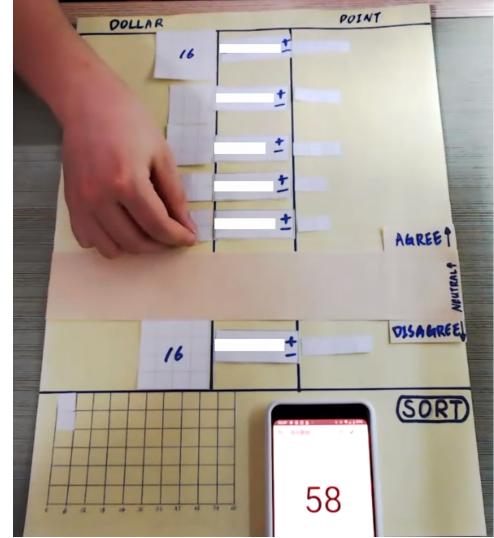
Designs that incorporated technologies: Other projects like the Caltech-MIT Voting Technology Project have sparked research to address accessibility challenges, resulting in innovations like EZ Ballot [48], Anywhere Ballot [49], and Prime III [50]. In addition, Gilbert et al. [51] investigated optimal touchpoints on voting interfaces, and Conrad et al. [52] examined zoomable voting interfaces.

1717 B Interface design process

1718 In this section, we outline the design process leading to our final interface. As mentioned in the paper, our design
 1719 iteration is based on existing QV applications in the wild.
 1720



1721
 1722
 1723
 1724
 1725
 1726
 1727
 1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741 (a) In this paper prototype, issues are denoted by different numbers that
 1742 appear on mouseover. Pretest respondents can move options anywhere in
 1743 the two sections of the interface, one denoting positive and one negative.
 1744 The blocks represent the cost for each option, with no indication of the
 1745 number of current votes. The credits are shown in the yellow box on the
 1746 left.



1747
 1748 (b) This paper prototype separates the positive and negative areas with a 'band' at the center. Undecided options
 1749 are placed inside this band. The cost and the votes on
 1750 both sides of the interface are denoted by small blocks.
 1751 The budget is shown in the yellow box below the inter-
 1752 face with a numerical counter.

Fig. 17. Initial paper prototypes designed for QS interface

1753 **B.0.1 Prototype 1: Ranking-Vote.** Considering that relative preference is often through ranking items, we tested whether
 1754 ranking options before voting would help establish an individual's relative preference in our prototype 1. This prototype
 1755 allowed respondents to reposition options before voting. Pretests revealed that respondents rarely moved the options
 1756 and questioned the necessity of full ranking, as it did not influence their QS submission. Additionally, many were
 1757 unaware that options were draggable until shown. This insight indicates that full ranking is unnecessary for establishing
 1758 relative preferences. Therefore, we decided to ask respondents to select a subset of options instead of requiring a full
 1759 rank among all options.

1760 **B.0.2 Prototype 2: Select-then-Vote.** Based on feedback from Prototype 1, instead of *allowing* individuals to rank options,
 1761 Prototype 2 implemented a two-phase process that *intentionally* asks respondents to select options to express opinions
 1762 before voting. As shown in Figure 19, survey respondents selected their preferred options (Figure 19a), and the interface
 1763 positioned these options at the top of the list for voting (Figure 19b). We identified several issues during the prototype 2
 1764 pretest: many respondents marked most options as 'options they care about,' which undermined the design's purpose.
 1765 Additionally, the lack of clear distinction between selected and unselected options confused respondents about the
 1766 Manuscript submitted to ACM

1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786

What societal issues need more support?

Please express your opinion using this survey mechanism as described above. You have a total of \$324 for the following 9 issues. You do not need to use up all your budget, but you cannot exceed \$324.

If you think that an issue needs more support, you can rate the issue higher. Vice versa, you can rate the issue lower if you think it requires less support.

<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Pets and Animals (Animal Rights, Welfare, and Services; Wildlife Conservation; Zoos and Aquariums)	Your ratings cost \$9 You rated this option +3
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Human Services (Children and Family Services; Youth Development; Shelter, and Crisis Services; Food Banks, Food Pantries, and Food Distribution; Multipurpose Human Service Organizations; Homeless Services; Social Services)	Your ratings cost \$16 You rated this option +4
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Arts, Culture, Humanities (Libraries, Historical Societies and Landmark Preservation; Museums; Performing Arts; Publishing and Media)	Your ratings cost \$4 You rated this option -2
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Education (Early Childhood Programs and Services; Youth Education Programs and Services; Adult Education Programs and Services; Special Education; Education Policy and Reform; Scholarship and Financial Support)	Your ratings cost \$36 You rated this option +6
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Environment (Environmental Protection and Conservation; Botanical Gardens, Parks and Nature Centers)	Your ratings cost \$4 You rated this option -2
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Health (Glosses, Disorders, and Disciplines; Patient and Family Support; Treatment and Prevention)	Your ratings cost \$4 You rated this option -2

Summary

You have spent \$73 and you have \$251 remaining

Fig. 18. A Ranking-Vote Prototype: The goal of this prototype is to test whether ranking options prior to voting help establish an individual's relative preferences and reduce effort when voting. Each option is draggable to position in a specific location amongst the full list of options. Votes can be updated using the buttons to the right of the interface with vote count and costs to the right of the interface. A summary box is placed sticky to the bottom of the screen.

1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820

This is a playground designed to help you understand how to use Quadratic Survey.

There is a limited budget to purchase the food for dinner party tonight. Your friend is asking for your preference of the type of food to get for the dinner party tonight. Please complete the survey below.

Step 1: What is important to you?

In this step, please elect the options that you cared about to the left of the column.

All Options	Options You Care About
American	Roman
Japanese	Chinese
Mexican	

Step 2: Quadratic Voting

BACK TO STEP 1

You will be given a budget of a certain amount of dollars. You will use the available money to rate the options. If you think more resources should be allocated to a certain issue, you can rate the option higher using the +1. If you believe that less resources should be allocated on a certain issue, you can rate the option lower using the -1. If you are neutral on an issue, you can choose to not rate the option. You are allowed to rate the options positively or negatively.

Based on the intensity of your opinions, you can rate each issue positively and negatively. The stronger your opinion is, the higher the rating you can put on one option. Note that the cost of the ratings would increase quadratically. In other words, rating of X will cost X^2 (square of X) dollars. The table shows the cost for ratings of 1 to 10 as an example. You can rate higher than 10 or lower than -10 if the budget allows you to do so.

Rating	1	2	3	4	5	6	7	8	9	10
Cost in dollars against budget	1	4	9	16	25	36	49	64	81	100

You cannot exceed the budget, but you do not have to use up all the budget either. You can see your total budget you have and the amount of dollars you have spent already in the "Summary" section below. The interface will provide real-time calculation of the remaining budget you have, the accumulated ratings the current options have received and the dollar spent for each option. The interface also provides a drag and drop feature to help you complete the survey.

<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Chinese Orange chicken and rice	Your ratings cost \$4 You rated this option +2
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Italian Pasta and bread	Your ratings cost \$9 You rated this option -3
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	American Burgers, fries and ribs	Your ratings cost \$0 You rated this option 0
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Japanese Sushi and udon	Your ratings cost \$0 You rated this option 0
<input type="button" value="+1 rating"/>	<input type="button" value="-1 rating"/>	Mexican Tacos and burrito	Your ratings cost \$0 You rated this option 0

Summary

You have spent \$13 and you have \$37 remaining

(a) Options are dragged and dropped to the 'Option You Care About' box.

(b) The previous step collapses showing all voting options.

Fig. 19. A Select-then-Vote Prototype: The goal of this prototype is to nudge participants to focus on a subset of options to vote, rather than ranking all of them. This prototype introduces a two-step voting process. As shown in Fig. 19a, the first step involves selecting options for further consideration. Important options are placed at the top of the list for voting shown in Fig. 19b, but options can be placed anywhere on the list if desired. The rest of the controls remain the same as the previous prototype.

necessity of Step 1. Thus, we need a clearer distinction and connection between the two phases to effectively construct relative preferences.

1821

1822 **What societal issues need more support?**

1823

1824 Please express your opinion using this survey mechanism as described above. You have a total of \$324 for the following 9 issues. You do not need to use up all your budget, but you cannot exceed \$324.

1825 If you think that an issue needs more support, you can rate the issue higher. Vice versa, you can rate the issue lower if you think it requires less support.

1826 **I don't know**

- 1 Pets and Animals (Animal Rights, Welfare, and Services; Wildlife Conservation; Zoos and Aquariums)
- 1 Arts, Culture, Humanities (Libraries, Historical Societies and Landmark Preservation; Museums; Performing Arts; Public Broadcasting and Media)
- 1 Health (Diseases, Disorders, and Disciplines; Patient and Family Support; Treatment and Prevention Services; Medical Research)
- 1 Religious Activities (Religious Activities; Religious Media and Broadcasting)
- 1 Welfare (Wounded Troops Services; Military Social Services; Military Family Support)

1827 **Positive**

- 1 Education (Early Childhood Programs and Services; Youth Education Programs and Services; Special Education; Education Policy and Reform; Scholarship and Financial Support)

1828 **Negative**

- 1 Environment (Environmental Protection and Conservation; Botanical Gardens, Parks and Nature Centers)

1829 **Neutral**

- 1 International (Development and Relief Services; International Peace, Security, and Affairs; Humanitarian Relief Supplies)
- 1 Homeless Services (Children's and Family Services; Youth Development, Shelter, and Crisis Services; Food Banks; Food Pantries, and Food Distribution; Multipurpose Human Service Organizations; Homeless Services; Social Services)

1830 **Next**

1831

1832

1833

1834

1835

1836

1837

1838

1839

1840

1841

1842 (a) The Organization Interface: Options are shown initially in the first bin labeled as 'I don't know.' Survey respondents can then drag and drop these options into the latter bins: Positive, Neutral, or Negative. Only the details of each option are shown on this interface.

1843

1844

1845

1846

1847

1848 Fig. 20. Organize-then-Vote Prototype: The goal of this prototype is to encourage participants at deriving finer grain categories among options before voting. Survey respondents first organize their thoughts into categories, then vote on the options.

1849

1850

1851

1852 *B.0.3 Prototype 3: Organize-then-Vote.* Figure 20 shows the last prototype where we built on the previous takeaway by providing finer-grain groupings and creating a clear connection between option organization and voting position. Specifically, we provided three categories: Positive, Negative, and Neutral. Initially, respondents see all options under the section labeled 'I don't know,' which includes only the option descriptions. We ask respondents to move these options into the categories below. Voting controls and information appear on each option once respondents move to the subsequent page, forming a clear connection between option groups, positions, and voting controls.

1853 Feedback indicated that survey respondents are comfortable with the two-phase organize-then-vote design, demonstrating it as a central strategy for our interface development. However, several areas for enhancement were identified:

1854 First, the dragging and dropping mechanism in the organization phase is cumbersome and may inadvertently suggest a ranking process, contrary to our intentions. Second, placing unorganized options at the top of the voting list is

1855 counterintuitive. Third, the voting controls are disconnected from the option summaries, dividing attention between

1856 the left and right sides of the screen. These insights guided refinements in the final two-phase interface, adhering to the

1857 two-phase organize-then-vote design framework.

1858

1859

1860

1861

1862

1863

1864

1865

1866

1867

1868

1869 **C List of Options**

1870

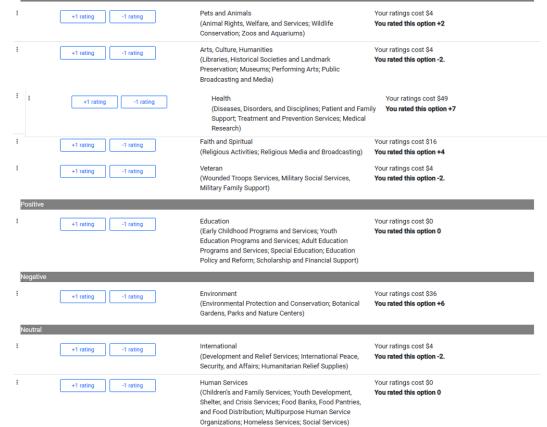
1871 We provide the full list of options presented on the survey.

1872 Manuscript submitted to ACM

What societal issues need more support?

Please express your opinion using this survey mechanism as described above. You have a total of \$324 for the following 9 issues. You do not need to use up all your budget, but you cannot exceed \$324.

If you think that an issue needs more support, you can rate the issue higher. Vice versa, you can rate the issue lower if you think it requires less support.



Summary
You have spent \$117 and you have \$207 remaining

(b) The Voting Interface: Voting controls appear on the left side of each option, showing the current votes and associated costs on the right. A budget summary is stuck at the bottom of the screen.

- 1873 • **Animal Rights, Welfare, and Services:** Protect animals from cruelty, exploitation and other abuses, provide
1874 veterinary services and train guide dogs.
- 1875 • **Wildlife Conservation:** Protect wildlife habitats, including fish, wildlife, and bird refuges and sanctuaries.
- 1876 • **Zoos and Aquariums:** Support and invest in zoos, aquariums and zoological societies in communities through-
1877 out the country.
- 1878 • **Libraries, Historical Societies and Landmark Preservation:** Support and invest public and specialized
1879 libraries, historical societies, historical preservation programs, and historical estates.
- 1880 • **Museums:** Support and invest in maintaining collections and provide training to practitioners in traditional
1881 arts, science, technology, and natural history.
- 1882 • **Performing Arts:** Support symphonies, orchestras, and other musical groups; ballets and operas; theater
1883 groups; arts festivals; and performance halls and cultural centers.
- 1884 • **Public Broadcasting and Media:** Support public television and radio stations and networks, as well as
1885 providing other independent media and communications services to the public.
- 1886 • **Community Foundations:** Promote giving by managing long-term donor-advised charitable funds for indi-
1887 vidual givers and distributing those funds to community-based charities over time.
- 1888 • **Housing and Neighborhood Development:** Lead and finance development projects that invest in and
1889 improve communities by providing utility assistance, small business support programs, and other revitalization
1890 projects.
- 1891 • **Jewish Federations:** Focus on a specific geographic region and primarily support Jewish-oriented programs,
1892 organizations and activities through grantmaking efforts
- 1893 • **United Ways:** Identify and resolve community issues through partnerships with schools, government agencies,
1894 businesses, and others, with a focus on education, income and health.
- 1895 • **Adult Education Programs and Services:** Provide opportunities for adults to expand their knowledge in a
1896 particular field or discipline, learn English as a second language, or complete their high school education.
- 1897 • **Early Childhood Programs and Services:** Provide foundation-level learning and literacy for children prior
1898 to entering the formal school setting.
- 1899 • **Education Policy and Reform:** Promote and provide research, policy, and reform of the management of
1900 educational institutions, educational systems, and education policy.
- 1901 • **Scholarship and Financial Support:** Support and enable students to obtain the financial assistance they
1902 require to meet their educational and living expenses while in school.
- 1903 • **Special Education:** Provide services, including placement, programming, instruction, and support for gifted
1904 children and youth or those with disabilities requiring modified curricula, teaching methods, or materials.
- 1905 • **Youth Education Programs and Services:** Provide programming, classroom instruction, and support for
1906 school-aged students in various disciplines such as art education, STEM, outward bound learning experiences,
1907 and other programs that enhance formal education.
- 1908 • **Botanical Gardens, Parks, and Nature Centers:** Promote preservation and appreciation of the environment,
1909 as well as leading anti-litter, tree planting and other environmental beautification campaigns.
- 1910 • **Environmental Protection and Conservation:** Develop strategies to combat pollution, promote conservation
1911 and sustainable management of land, water, and energy resources, protect land, and improve the efficiency of
1912 energy and waste material usage.

- **Diseases, Disorders, and Disciplines:** Seek cures for diseases and disorders or promote specific medical disciplines by providing direct services, advocating for public support and understanding, and supporting targeted medical research.
- **Medical Research:** Devote and invest in efforts on researching causes and cures of disease and developing new treatments.
- **Patient and Family Support:** Support programs and services for family members and patients that are diagnosed with a serious illness, including wish granting programs, camping programs, housing or travel assistance.
- **Treatment and Prevention Services:** Provide direct medical services and educate the public on ways to prevent diseases and reduce health risks.
- **Advocacy and Education:** Support social justice through legal advocacy, social action, and supporting laws and measures that promote reform and protect civil rights, including election reform and tolerance among diverse groups.
- **Development and Relief Services:** Provide medical care and other human services as well as economic, educational, and agricultural development services to people around the world.
- **Humanitarian Relief Supplies:** Specialize in collecting donated medical, food, agriculture, and other supplies and distributing them overseas to those in need.
- **International Peace, Security, and Affairs:** Promote peace and security, cultural and student exchange programs, improve relations between particular countries, provide foreign policy research and advocacy, and United Nations-related organizations.
- **Religious Activities:** Support and promote various faiths.
- **Religious Media and Broadcasting:** Support organizations of all faiths that produce and distribute religious programming, literature, and other communications.
- **Non-Medical Science & Technology Research:** Support research and services in a variety of scientific disciplines, advancing knowledge and understanding of areas such as energy efficiency, environmental and trade policies, and agricultural sustainability.
- **Social and Public Policy Research:** Support economic and social issues impacting our country today, educate the public, and influence policy regarding healthcare, employment rights, taxation, and other civic ventures.

D Cognitive Demand Tables

Here we provide the full qualitative analysis table.

Table 2. This table lists all the causes participants mentioned as contributing to their Mental Demand. The shaded cells represent the percentage of participants citing each source of mental demand, allowing for comparison within columns. The abbreviations are: ST (Short Text Interface), SI (Short Two-phase Interface), LT (Long Text Interface), and LI (Long Two-phase Interface). Short and Long refer to the sum across both interfaces; Text and Inter refer to the sum across both survey lengths. We include Sparklines for comparisons across these experiment groups.

[Mental Demand]	Total	Version				Experiment Conditions			
		ST	SI	LT	LI	Short	Long	Text	Inter
Budget Management	14	3	3	5	3	6	8	8	6
Budget within limited credit	5	2	2	1	0	4	1	3	2
Track remaining credits	10	2	2	3	3	4	6	5	5
Maximize credit usage	8	2	3	2	1	5	3	4	4
Operational	12	3	2	4	3	5	7	7	5
Strategic	7	2	4	1	0	6	1	3	4
Preference Construction	39	10	9	10	10	19	20	20	19
Determining relative preference	16	4	4	5	3	8	8	9	7
Option prioritization	17	6	4	3	4	10	7	9	8
Precise resource allocation	30	9	6	9	6	15	15	18	12
Narrow - Consider a few options/personal causes	23	6	6	8	3	12	11	14	9
Broad - Considering all options or higher order values	23	5	5	4	9	10	13	9	14
Demand from Experiment Setup	24	6	6	6	6	12	12	12	12
Many options on the survey	6	0	0	3	3	0	6	3	3
QS Mechanism	4	2	0	2	0	2	2	4	0
Recalling experience or understanding options	20	5	6	4	5	11	9	9	11
Justification or Reflection on response	8	2	2	1	3	4	4	3	5
External Factors	12	3	1	4	4	4	8	7	5
Demand due to Interface	8	2	2	0	4	4	4	2	6
Increase	4	1	1	0	2	2	2	1	3
Decrease	4	1	1	0	2	2	2	1	3

D.1 Mental Demand Table

D.2 Physical Demand Table

Table 3. Physical Demand Causes: Most participants expressed little or no physical demand. Results reflected that participants in the long two-phase interface required more actions, hence the higher mention of mouse usage as a source.

[Physical]	Total	Version				Experiment Conditions			
		ST	SI	LT	LI	Short	Long	Text	Inter
Reading	4	0	2	1	1	2	2	1	3
Mouse	16	3	5	2	6	8	8	5	11
Vertical Screen	4	1	0	1	2	1	3	2	2
None/Little	32	8	9	8	7	17	15	16	16

D.3 Performance Table

Table 4. Performance Causes: Most causes are shared across experiment conditions. We provided qualitative interpretations of their own performance assessments.

[Performance]	Total	Version				Experiment Conditions			
		ST	SI	LT	LI	Short	Long	Text	Inter
Operational Action	13	2	3	3	5	5	8	5	8
Budget Control	6	1	1	2	2	2	4	3	3
Preference Reflection	6	1	1	2	2	2	4	3	3
Limited Resources	5	1	2	1	1	3	2	2	3
Social Responsibility	8	2	2	2	2	4	4	4	4
Decision maker	7	1	2	2	2	3	4	3	4
Outcome Uncertainty	7	1	2	2	2	3	4	3	4
Performance Assessment									
Did their best	8	2	1	3	2	3	5	5	3
Feel Good	17	3	5	3	6	8	9	6	11
Good Enough	10	2	2	3	3	4	6	5	5

2081 **D.4 Temporal Demand Table**

2082

2083 Table 5. Temporal Demand Sources: Decision-making and Operational Tasks are the main causes. Participants framed their decision-
2084 making sources differently.

2085 [Temporal]	Total	Version			Experiment Conditions				
		ST	SI	LT	LI	Short	Long	Text	Inter
Budget Management	4	0	1	1	2	1	3	1	3
Decision Making	15	5	2	3	5	7	8	8	7
Affirmative	9	0	2	2	5	2	7	2	7
Negative	8	5	1	2	0	6	2	7	1
Operational	16	5	6	3	2	11	5	8	8
Task completion	8	2	2	3	1	4	4	5	3
Being efficient	8	3	4	0	1	7	1	3	5

2100 **D.5 Frustration Table**

2103 Table 6. Frustration Sources: needs to be updated with some new terms definitions for some of the columns.

2104 [Frustration]	Total	Version			Experiment Conditions				
		ST	SI	LT	LI	Short	Long	Text	Inter
Strategic	17	4	4	5	4	8	9	9	8
Higher-level	11	3	2	3	3	5	6	6	5
x Conflict between personal preference and broader society and common values	6	1	1	2	2	2	4	3	3
x Trade-offs among all options	8	3	1	2	2	4	4	5	3
Lower-Level	10	3	3	2	2	6	4	5	5
x Conflict between personal preference and	4	1	2	0	1	3	1	1	3
x Trade-offs among a few options	8	2	2	2	2	4	4	4	4
Operational	15	4	5	2	4	9	6	6	9
Credit management	6	2	3	1	0	5	1	3	3
Adhering to the Quadratic Mechanism	5	2	1	1	1	3	2	3	2
Deciding number of votes for an option	4	2	0	0	2	2	2	2	2
Making multiple decisions	3	2	0	0	1	2	1	2	1
Understanding Option	4	0	3	0	1	3	1	0	4
None/Little	16	4	5	5	2	9	7	9	7