The Trust in Voting Systems (TVS) Measure

Claudia Ziegler Acemyan, Rice University, USA*
Philip Kortum, Rice University, USA
Frederick L. Oswald, Rice University, USA

ABSTRACT

It is essential to democracy that voters trust voting systems enough to participate in elections and use these systems. Unfortunately, voter trust has been found to be low in many situations, which could detrimentally impact human-computer interactions in voting. Therefore, it is important to understand the degree to which voters trust any specific voting method. Voting researchers have developed and used measures of overall trust in technology; yet researchers have long argued that trust in systems is domain-specific, implying that system-specific measures should be used instead. To address this latter point, this paper describes the development of a psychometrically reliable and validated instrument called the trust in voting systems (TVS) measure. The TVS not only allows researchers to understand group mean differences in trust across voting systems; it also allows researchers to understand individual differences in trust within systems—all of which collectively serves to inform and improve voting systems.

KEYWORDS

HCI Design Impacting Trust, Measurement in Human-Computer Interaction, Measuring Trust, Trust in Technology, Users’ Perceptions of Trust, Voting Systems

INTRODUCTION

Many people do not trust technology or the people who develop and maintain technological systems. Invasiveness is one reason: Some email users view Gmail skepticaly, for instance, because software algorithms scan the text of users’ emails to display relevant advertisements. Likewise, many consumers are concerned about Apple’s iPhone because of its FaceID and embedded fingerprint scanner that create the potential for users’ biometric information to be handed over to the NSA (Brundage, 2013). In the voting context, voters and election officials may worry about whether voters perceive invasiveness and, conversely, whether they find the voting process to be transparent. On this latter point, black-box voting systems may record and transmit voter selections without any verification, likely leading voters to wonder whether voting systems are secure, reliable, and accurate (Robertson, 2005).

Such concerns were amplified around the 2016 U.S. Presidential election, when candidates and voting security experts expressed concern about the integrity of elections and voting systems (Cylance Team, 2016; Halderman, 2016; Montanardo & Naylor, 2017; Wallach, 2016, 2017). Companies and product developers making and selling voting systems often reassure buyers and users about the
security and specific operations of their systems, yet often little is done to verify these claims with data and physical evidence (Robertson, 2005; Ulanoff, 2016). Now in 2020, as the world combats the COVID-19 pandemic, the United States is grappling with how to vote in its upcoming elections, which includes the November 2020 presidential election. One option that has been proposed is to replace in-person voting with voting-by-mail. However, voter fraud and corruption of mail-in voting methods is of high concern, should they replace access to the polling stations (Korecki & Cadelago, 2020). This summarizes today’s state of affairs in which there is a great range of heterogeneity in people (voters, election stakeholders) and systems (electronic, traditional, mail) interacting to influence voter trust.

Generally, people’s trust in systems is found to influence system use, acceptance, and reliance (Acemyan & Kortum, 2012). For instance, consumers who do not trust their email applications or cell phones might choose to limit the information given and received from them, change how they otherwise interact with them, or avoid using them altogether. Unlike consumer products, however, U.S. voters have no choice on election day but to use—or not use—the system provided at their assigned polling station (Acemyan & Kortum, 2015, 2016). If voters do not find voting systems to be trustworthy, they may decide to abstain from voting, or fail to express their beliefs freely through their ballot selection(s) (Ansper, Heiberg, Lipmaa, Overland, Van Laenen, 2009; Oostveen & van den Besselaar, 2004). At a broader level, they may distrust the overall election outcome(s) or decide to avoid voting in the future (Acemyan & Kortum, 2012).

Furthermore, once voter trust in a voting system is lost, it is difficult to regain (Ansper et al. 2009; Loeber 2008). Even if an improved voting system is established in response to voters’ mistrust, their previous negative system impressions may prevent them from using it (Ansper et al.). Thus, it is essential to democratic elections that voters always have and maintain a fundamental level of trust in voting systems, enough so that they will participate, uninhibitedly express their beliefs, and have confidence in the final tallies. This involves simultaneously preventing adverse human-computer interactions and user perceptions, while promoting positive ones.

To address issues of security and trust in elections, voting security experts have proposed secure, trustworthy, end-to-end (e2e) methods to make certain voters can cast votes as intended, and votes are counted as cast (e.g., Adida, de Marneffe, Pereira, & Quisquater, 2009; Bell et al., 2013; Benaloh et al., 2012; Popovienic & Carback, 2010; Ryan et al., 2009). As one example, the developer of the e2e system, Helios, reasonably assumed voters would be more likely to trust a reliable, accurate, cryptographically secure system than to trust potentially corrupt election officials (Adida, 2008). Their cryptographic auditing technique allows any interested party to audit the voting system and verify that it is operating as designed, and that election tallies cannot be modified or faked without detection. But in addition to voting technologies that intend to create trustworthy systems, voters themselves must also actually trust the system. Researchers and practitioners can measure trust in a voting system immediately after interacting with it.

Having said this, psychometrically developed and validated instrument to measure voters’ trust in voting systems currently do not exist. To date, only a few studies have explored voters trust in voting systems. In a conceptual paper, Hoffman, Lawson-Jenkins, and Blum (2004) proposed a general trust model for distributed computer systems. Factors that they anticipated would impact trust in an e-voting system included security (e.g., control of the cast votes database), usability (e.g., voters being able to use the system), privacy (e.g., keeping vote selections secure while voting), and user expectations (e.g., users’ anticipation of what the system should do and how they will interact with it).

Related to his model, an empirical paper by Oostveen and van den Besselaar (2005) asserted that most e-voting system requirements are largely trust related. They argued the systems should be secure, verifiable, and guarantee voter privacy. In their small-scale field experiment, 179 self-selected participants (15-72 years old) used one of two voting systems in a mock election; they then responded to items measuring trust in privacy of the system and trust in the security of the system. Psychometric analysis and factor analysis supported high reliability of these two scales as well as an empirical distinction between them. These authors also analyzed the scale scores to understand
whether voters’ self-reported trust in a system will predict how they vote. The researchers found a small effect, such that the more-trusted system had greater minority opinions expressed on ballots, highlighting that trust in the privacy of a voting system can have an important impact on voting and election outcomes.

Acemyan and Kortum studied the relationship between voter trust and subjective usability in voting systems (2012), first asking a sample of 30 eligible voters (i.e., U.S. citizens who are 18 years or older) to vote with three different ballots in a mock election, where each ballot reflected a different level of perceived usability. After voting with each ballot, participants responded to four trust items: (1) Overall, I trust the voting system; (2) I am confident in the reliability of the voting system; (3) I am confident in the security of the voting system; and (4) I am confident in the accuracy of the voting system. Although the sample and the number of items are both relatively small (like Oostveen’s), and although the within-subjects design of the study might create contrast effects as a demand characteristic that magnifies the real effects of trust, these results provided tentative evidence that trust in voting systems monotonically increases with increases in perceived system usability. The item content of this measure at least covered multiple constructs of trust (reliability, accuracy, and security), in addition to overall trust, across three voting methods. However, multiple items for each construct would be required to investigate multiple, reliable dimensions of trust.

Reflecting another approach to measuring trust, van den Besselaar, Oostveen, De Cindio, and Ferrazzi (2003) administered to 44 e-voters a post-voting trust measure comprising 60 unpublished items. Factor analysis identified three clusters of items: trust in safety, secrecy (privacy), and accountability. Although the study was based on a small sample and only disclosed a subset of the items, this important work is of value because it was among the first data-driven studies on trust in voting systems. As in previous studies, their study also supports the need to explore other constructs of trust in addition to security, autonomy, and transparency—a need that we pursue in the current work.

As a final example of trust measurement, Chiang (2009), used structural equation modeling (SEM) to identify how the two factors of security and attitudes towards e-voting directly impacted 281 students’ trust in a new Taiwanese e-voting system. Security of the e-voting system was measured through three items reflecting multiple constructs found in other studies; attitudes toward using the e-voting system were measured by two similar items; and trust in the system was measured through three items. This large-sample SEM analysis of Chiang’s research model revealed relationships between security and trust, perhaps because of content relatedness (not causal relationships). The SEM also showed little evidence of the attitude construct being independent of the security constructs, perhaps because the two items both reflected the intent to vote. This work seems more useful for its general approach than for its specifics. In particular, it further supports how system security perceptions can impact users’ trust in a system. The work also suggests that additional trust-related constructs, additional items and measure development, and additional data related to multiple voting systems are needed to better understand these key relationships.

It could be argued that researchers seeking to measure trust in voting systems could use a general trust-in-technology measurement, albeit to a limited extent. However, our review of the trust literature reveals that the research community does not even have a widely accepted and psychometrically validated measure of general trust in systems—voting or otherwise. Moreover, it has been cogently argued that trust in technology is domain specific (Bhattacherjee, 2002; Montague, Kleiner, & Winchester, 2009), meaning that a domain-specific measure of trust would be required for future research to test this argument empirically. If it is true for voting, it means that general trust measures have limited value in conceptually covering and/or empirically predicting trust in specific voting technologies—such as being able to cast a ballot as intended or ensuring a ballot is counted as cast. Thus, a specific measure of trust in voting systems is potentially of great benefit to both research and practice.

We have developed a Trust in Voting Systems (TVS) measure using psychometric methods; the TVS includes a range of relevant and important voting-related constructs. In the spirit of open
science and practice, the TVS measure and its items are free to use (and further refine if needed). Researchers and practitioners can use resulting TVS scores to understand how trust differs across voting systems, across voting populations, and within voters over time. Furthermore, TVS scores can be usefully correlated with external variables, such as voter individual differences (e.g., voting experience, political affiliation, perceived civic duty, voter personality) and general demographic features (e.g., average income or demographics by county) to illuminate other factors that might impact trust in voting systems. Overall, the TVS will extend previous research by providing a psychometrically reliable instrument that can be used to improve researchers’ understanding of voter trust in a particular voting system.

DEVELOPMENT OF THE TRUST IN VOTING SYSTEMS (TVS) MEASURE

The Trust in Voting Systems measure was developed in five phases. The first phase entailed defining the multiple constructs of trust in a voting system. Second, item content was generated to represent each of these constructs, resulting in subscales of the overall trust measure. Third, we administered this large set of items to a reasonably large sample of undergraduates who are eligible U.S. voters. Fourth, the TVS was psychometrically analyzed and refined, such that only the most statistically reliable items were retained. The fifth and final stage was a validation study, where scores on the refined TVS measure were analyzed to understand how the level of trust changes across voting systems used by study participants. The subsequent sections of this paper reflect this sequence of the TVS measure-development process.

Defining Trust

Measuring trust in a voting system first requires an operational definition of the overarching construct. An iterative approach was used to develop the definition of trust in voting systems, which aimed to encapsulate the numerous characteristics of the voting system and the voter that, individually and in combination, might impact trust in the voting method. To ensure our definition of trust in a voting system was credible, it was shown to experts in the areas of voting and psychological measurement. These experts offered feedback, and then the definition was further revised and re-submitted to them for review. This process continued until all reviewers were satisfied with the final definition:

*Trust in voting systems refers to a voter’s beliefs that a voting method has the system attributes to enable a voter to anonymously cast a ballot with intended selections and to have these cast ballots counted accurately—ultimately resulting in fair election outcomes.*

Every clause in the above sentence is essential for explaining trust in voting systems, because the construct of trust in voting systems is multidimensional: “A voter’s beliefs” indicates that trust is a combination of facts, feelings, and opinions held by an individual about a voting system. “System attributes” refers to the features of a system that allow a voter to be instrumental in reaching his/her voting goals. The system must also have attributes that make it usable so a voter can easily complete his/her ballot and cast it. It is also important for a voting system to be usable, so voters can vote for the person he or she wants (a voter should feel reassured that this is the case). “Anonymously cast a ballot” meets the goal of preserving the United States’ democratic process, because votes traced back to an individual can lead to voter coercion or vote-buying. Hence, a voting system that can be trusted must protect a voter’s identity.

Regarding “intended selections,” the voting method cannot contain some type of mechanism that either prevents or switches intended selections for particular candidates and propositions. For this reason, election officials, voters, and/or any interested third party need to be able to have “ballots counted accurately” and determine that malicious acts are not taking place. And finally, “fair election
outcomes” refers to the voting system used during an election being the sole determinant of the outcome of a race; the system must be perceived by voters as such.

Naturally, trust in a voting system does not guarantee that the system is functioning as designed, but the operating assumption here is that an accurate and reliable voting system will likely engender voter trust. Certainly, it is possible for a voting system to appear as though it is working properly, yet the votes cast on it would not change the outcome of a predetermined election. One example is the 2013 presidential election in Azerbaijan where the Election Commission accidentally released the vote tally before the election had started (Fisher, 2013). In this case, the votes cast did not determine the election outcome. Hence, trust in the election was tainted—as was trust in any subsequent voting system and process.

**Development of Subscales and Items**

After the definition for trust in voting systems was finalized, the Trust in Voting Systems (TVS) subscales and associated items were proposed for each construct. Similar to the iterative development method used to establish the definition of trust in voting systems, the same experts reviewed and commented on the subscales and their items until they reached consensus. This process served to refine and improve the measure’s initial content validity, thereby providing a high-quality pool of items that makes subsequent psychometric refinement more meaningful.

The TVS subscales that we developed reflect all constructs of trust in voting methods, as found in our definition above: i.e., the reflection of intended selections on cast ballots, votes counted accurately, votes being anonymous, and system security, usability, reliability, accuracy, and transparency. Two subscales were added to the measure to capture stable individual differences in trusting others (Goldberg et al., 2006) and an individual’s propensity to trust technology (Oostveen & van den Besselaar, 2004), as these general constructs are also likely to impact a voter’s specific trust level in a voting system. An additional scale was included in the instrument to measure overall trust in the voting system. Correlations between specific subscales and this overall scale can indicate the relative importance and influence of the specific constructs in relation to general trust.

Each subscale contains enough items to capture the full breadth of its respective construct. To develop each subscale, we first adapted items from previous trust measures to fit the voting context; then we wrote a set of new items to complete the full item pool (see Table 1 for each item, scale, and subscale). In the initial TVS development phases, some subscales comprised more items than others; but through the process of balancing construct representation and psychometric refinement—along with the practical need for a relatively short TVS measure—the finalized subscales contained the same number of items. Other considerations, in addition to content validity, were also taken into account when developing the items. For instance, all items were positively worded, because negatively worded items can add to cognitive fatigue (or confusion) of the test taker and lead to inaccurate responses (and we felt this potential drawback outweighed any gains from alternating the direction of the wording).

To use the TVS and obtain a TVS trust score for any voting system, users first need to vote with a voting method (in a mock election or actual election). Second, after voting, users need to complete the TVS by responding to each of the items on a 5-point Likert scale. Third, scale scores are calculated by system evaluators/researchers for each TVS subscale and overall trust. The composite TVS trust score is a unit-weighted composite across the constituent TVS scales (see Appendix A), with scores ranging from 0 to 100. This is an intuitive range of scores, similar to a school grade, or a percentage; it also mirrors the 0-100 range of scores used in other widely accepted human factors and human-computer interaction measures, such as the System Usability Scale (Brooke, 1996) and the NASA-TLX (Hart & Staveland, 1988).

Below we describe the further refinement of the TVS measure. Specifically, the TVS was used in a refinement and validation study that manipulated accuracy of vote selections shown on a final review screen.
Participants
In exchange for optional, partial class credit, this study recruited 183 undergraduate students who were eligible to vote in U.S. elections, i.e., they were 18 years old or older and U.S. citizens. Of these, 67% were female, and the mean age was 19.2 years (SD = 1.1, median = 19). Although the sample was young and thus not fully representative of the U.S. voting population, it was comprised of real voters representing an important segment of the voting population—meaning that if they found a system to be untrustworthy, other voters may also. Related to this age demographic, 62% percent indicated that they voted in a national election; 50% had voted in one election; and 12% had voted in two or more elections. The participant data contributed both to the preliminary TVS item pool that was then psychometrically refined, and to the use of scores from the refined TVS in the voting system experiment.

Design
The voting experiment provided the data on all preliminary items of the TVS scale that would then be psychometrically refined. The voting experiment was a between-subjects design, where the independent variable was the accuracy of the system, where there were 0, 1, 2, 3, or 4 errors displayed on the review screen (i.e., the voter’s actual selection was changed to a different one). The primary dependent variables were the composite TVS trust score, the overall trust in voting systems score (OTVS), and the 10 subscale scores that contributed to the TVS trust score by capturing the various constructs of trust.

Materials
Voting System
The electronic voting system presented one voting race or proposition at a time in an electronic bubble ballot format. The screen format was typical: The top of the screen presented the title of a race or a proposition, and the selections were placed immediately below it. To vote in the race, voters clicked on the bubble to the left of the intended selection. After completing their ballots, voters were shown a screen in which they could review their choices before casting a vote. Although some electronic voting systems provide the option to go back, this option was not provided, meaning that voters would have to cast their ballot “as is,” instead of being able to correct a vote. This study’s electronic voting system was technically supported by an online platform; however, it was not apparent to users that this was the case. Instead, the voting system appeared to be an electronic voting system presented on a computer. This voting format and presentation method is the same as that used by other vote-changing studies (see Acemyan, Kortum, & Payne, 2013; Campbell & Byrne, 2009; and Everett, 2007); importantly, it does not affect the primary dependent variable of this study, which is whether a noticed change to one or more of the candidate selections impacts trust.

Ballot
There were four races and propositions in total: President and Vice President; Speaker of the House; whether political contributions should be regulated; and whether at least 25% of electricity should be generated from a renewable source. Because the study was conducted via computer and was simulating voting as closely as possible, it was essential that voters not require additional information—such as a voter’s guide—to make a candidate selection decision. To increase the realism of the mock election, the races included on the ballot were designed so that any voter should have some prior knowledge about them (e.g., the study was run during the fall of a presidential election year, and we used the actual presidential candidates). Similarly, propositions were written to be realistic, so voters did not require specific prior knowledge to decide.
Survey

The online survey, which was immediately administered following the mock election, presented the TVS measure (93 items) to assess multiple constructs of trust (see Table 1). Subjects were asked to respond to each of these on a 5-point Likert scale, from 1 = strongly disagree to 5 = strongly agree. Items were presented in a cyclical order, meaning that one item from each subscale was presented before presenting another item from the same subscale. Each screen presented a subset of items where once completed, another subset of items was presented, and participants could not go back and change their previous item responses. The survey concluded by asking basic questions about system performance (i.e., did they notice system inaccuracies and/or changed votes), demographics, and prior voting experience.

Procedure

The study began with the IRB-approved informed consent process; then, participants voted with the electronic voting system in a mock election. Next, participants completed an online survey that addressed various features and perceptions of the voting method that they just used to vote. Last, they were debriefed, thanked for their participation, and assigned partial class credit.

Psychometric Refinement and Validation of the TVS

The online survey required participants to respond to every item, and although participants could leave the study at any time, all participants ultimately completed the survey. Thus, no data were missing.

Before analyzing the data, we inspected the distributions of item scores across all conditions, which were wide-ranging and roughly normally distributed, with no apparent ceiling or floor effects. Additionally, all correlations between items were positive, and given that all items were scored in the same direction, this is to be expected and, at least in this initial pass, indicative of reliable variance in the items and scales.

Next, we arrived at the final version of the TVS subscales by deleting a subset of the preliminary item pool through psychometric refinement. More specifically, items retained in each subscale were the ones that contributed the most to internal-consistency reliability (i.e., would produce largest drop in Cronbach’s alpha if dropped). In the end, each TVS subscale had four items and high levels of reliability (above .80). The goal of establishing four items per subscale was motivated by balancing multiple aims (a) assuring reasonable content coverage of the relevant trust construct, (b) realizing a relatively short measure, and (c) allowing for appropriate confirmatory factor analyses (CFAs) in future work (e.g., having just three items per factor only allows one to fit a one-factor model, but with four items, you also can test the one-factor model, with 2 degrees of freedom). Similar balancing can be found in the development of short multidimensional measures in the domains of personality (see Donnellan, Oswald, Baird, & Lucas, 2006) and working memory (see Oswald, McAbee, Redick, & Hambrick, 2015).

See Table 1 for the list of the retained items making up each subscale; see Table 3 for the descriptive statistics (mean and standard deviation) for each refined subscale; and see Table 4 for correlations between all refined TVS scale scores, along with the Cronbach’s alphas (internal-consistency reliability estimates) for each scale.

In addition to refining the internal structure of each TVS subscale, we also used factor analysis to summarize the relationships between the refined subscales. Specifically, the scree plot of the unrotated factors from an orthogonal principal axis exploratory factor analysis (EFA) suggested that two factors would effectively summarize the 10 subscales (excluding the overall trust subscale), where the first two unrotated eigenvalues respectively accounted for 71.6% and 10.7% of the reliable variance. Extracting two factors, then varimax-rotating the pattern matrix of factor loadings, suggested that one factor represented the two TVS scales that measure *subjective*, personal perceptions of trust that are more independent from the actual system used (i.e., trust in others, propensity to trust in technology, and vote is anonymous), whereas the second factor represented the TVS scales that measure more
Table 1. TVS Subscales and Items (“S” = subscale associated with subjective trust, and “O” = subscale associated with objective trust)

| Overall Trust in the Voting System (OTVS) | System Reliability (O) |
|-----------------------------------------|------------------------|
| I trust the voting method               | I am confident that the voting system is dependable |
| The voting system is trustworthy        | The voting system never fails me |
| I have no reason to distrust the voting method that I just used to vote | I am confident that the voting system always works |
| Overall, my trust in the voting system is high | The voting system is reliable |

| Trust in Others (S) | System Security (O) |
|---------------------|---------------------|
| I believe that others have good intentions | I am confident that the voting method is not exposed to harm |
| I believe that people are basically moral | The voting system can't be tampered with |
| I believe in human goodness | The voting system is secure |
| I trust people | Voting with the method is riskless |

| Propensity to Trust in Technology (S) | System Usability (O) |
|-------------------------------------|---------------------|
| In general, I trust technology      | Using the voting method, I was able to cast a vote that I was confident in |
| I am confident in technology        | The voting system is great |
| I am totally comfortable working with technology | I would rely on this voting method in the future |
| Things will be fine when I use technological systems | The voting system works the way I want it to work |

| Cast Ballot Reflects Intended Selections (O) | Votes Are Accurately Counted (O) |
|-----------------------------------------------|---------------------------------|
| I believe my vote was cast as I intended      | I think the vote I cast is accurately counted |
| My final ballot reflected the people and propositions that I wanted to vote for | My cast ballot is included in the total number of votes |
| Every choice I made while using the voting method is represented as a vote on my cast ballot | I believe that every one of my ballot selections is counted correctly |
| My ballot reflects my intentions and beliefs | I don't believe that someone would intentionally leave out my vote |

| Vote is Anonymous (S) | System Accuracy (O) |
|-----------------------|---------------------|
| My vote is anonymous  | The voting method is accurate |
| My ballot cannot be traced back to me | The voting method is precise |
| Unless I tell someone, there is no way for anyone to ever know who I voted for | I am confident that the voting system is always correct |
| Only I know who I voted for | The voting system does what it is supposed to do |

| System Transparency (O) |
|-------------------------|
| I am confident that the voting method is transparent |
| I am confident that a voter can verify the voting system is working as expected in a fair election |
| I am confident that a voter can verify their vote |
| The voting system does not hide anything from me |
objective, system-driven facets of trust (i.e., the remaining TVS subscales, such as system accuracy, reliability, and usability). See Table 2 for the rotated factor loadings associated with each scale.

We then created subjective and objective composite scores based on subscales that had the highest factor loading on their respective factors (see Tables 1 and 2). Note that to create each composite, subscale scores were simply added up (i.e., unit-weighted; using factor loadings as weights is inappropriate, because those weights apply to the latent factor scores, not to the observed scale scores). The correlation between these two scale composites was \( r = .63 \) – a relatively high value, yet lower than the constituent composite reliabilities, and therefore low enough to support both conceptual distinctiveness, as well as empirical distinctiveness in correlating with a variety of external variables (e.g., voter personality and voting outcomes for a given voting system).

In summary, the two factors identified in this research paper are both robust and representative: Robust because the factor analysis produces subjective and objective components that are both reliable yet distinct from one another; representative because these factors summarize a wider array of constructs relevant to trust in voting that have not been examined in the literature to-date.

Next, the subjective and objective composite subscales were correlated with the scale reflecting overall trust in the voting method (OTVS). These results revealed that the objective composite correlated very highly, \( r = .93 \), with overall trust—much higher than the comparable correlation between the subjective trust composite and overall trust, where \( r = .64 \). These two correlations were statistically different from one another (Fisher’s \( z = -10.79, p < .001 \)). This pattern of findings suggests that when participants were providing overall ratings for their trust in the voting system, they were influenced by objective aspects of trust, such as voting system usability and accuracy, more than personal, subjective facets, such as their trust in others, or their propensity to trust in technology. This makes sense as a function of our experiment, which manipulated the technical aspects/performance of the system, not the personal aspects.

After psychometric refinement to create the finalized TVS measure, we then conducted a follow-up validation study using the measure. In particular, we explored whether measured trust in voting changes across levels of accuracy in a voting system, we analyzed the data of participants who either noticed the errors on their review screen or who were in conditions in which there were no errors displayed on the review screen. Those 31% of participants (\( n = 183 \)) who had errors on their review screen but did not notice them were excluded from this part of the analysis, because their perceptions of trust in the voting system was an “ignorance is bliss” effect, where trust did not change across

| Refined Scale                        | Factor 1: Subjective, Personal Perceptions | Factor 2: Objective, System-Driven Facets |
|--------------------------------------|--------------------------------------------|------------------------------------------|
| Trust in Others                      | .68                                        | .13                                      |
| Propensity to Trust in Technology    | .64                                        | .44                                      |
| My Vote is Anonymous                 | .61                                        | .37                                      |
| Cast Ballots Reflect Intended Selections | .28                                   | .89                                      |
| Votes are Accurately Counted         | .51                                        | .77                                      |
| System Security                      | .44                                        | .74                                      |
| System Usability                     | .31                                        | .90                                      |
| System Reliability                   | .32                                        | .92                                      |
| System Accuracy                      | .35                                        | .91                                      |
| System Transparency                  | .50                                        | .73                                      |

Note: Highest loading for each scale is in boldface.
conditions, $F(3, 53) = 0.15, \text{MSE} = 165.69, p = .93$. The effect of many (e.g., 50% to 80%) subjects failing to identify inaccuracies on the voting review screen is not unusual, as it has consistently been found in previous research (Acemyan et al., 2013; Campbell & Byrne, 2009; Everett, 2007).

Figure 1 shows that average TVS scores decreased until two inaccuracies are introduced into the system, and then the average score levels off. This pattern of means is both statistically and practically significant ($F(4, 121) = 29.26, \text{MSE} = 174.28, p < .001, \eta^2_p = .49$), providing initial criterion-related validity information in support of the TVS measure. A follow-up comparison of means using Tukey-Kramer HSD reveals that trust in a perfectly accurate system is higher ($M = 65.8$) than when there are inaccuracies in the system. As for the inaccurate voting systems, trust tends to be higher when there is one incorrectly displayed selection ($M = 45$) than when three or four selections are switched (the means are 29.1 and 31.3, respectively, and Cohen’s $d$s against the no-inaccuracy condition are generally considered large, $d = .87$ and $d = .74$, respectively). Despite trust means appearing to be

| Table 3. Descriptive Statistics for Each Final TVS Subscale |
|----------------------------------------------------------|
| **Refined Scale**                                      | **Mean (SD)**  |
| Overall Trust in the Voting System (OTVS)             | 3.1 (1.15)    |
| **Subjective Trust**                                   |               |
| Trust in Others                                       | 3.46 (0.78)   |
| Propensity to Trust in Technology                      | 3.65 (0.85)   |
| My Vote is Anonymous                                  | 3.11 (0.93)   |
| **Objective Trust**                                    |               |
| Cast Ballots Reflect Intended Selections              | 3.26 (1.37)   |
| Votes are Accurately Counted                          | 3.38 (0.99)   |
| System Security                                        | 2.64 (0.91)   |
| System Usability                                       | 3.17 (1.25)   |
| System Reliability                                     | 2.88 (1.10)   |
| System Accuracy                                        | 3.06 (1.14)   |
| System Transparency                                    | 2.96 (0.94)   |

| Table 4. Correlations between Refined TVS Scales |
|--------------------------------------------------|

Note. $N = 183$. Internal consistency reliability coefficients (Cronbach’s $\alpha$s) are italicized on the main diagonal.
higher with three incorrect selections versus its neighbors (two or four), this is not a statistically or practically significant difference.

In summary, in this instrument validation study, trust in voting systems was highest when the review screen was 100% accurate in its display of voter selections. When the system is characterized by lower accuracy rates, the level of trust tended to decrease accordingly, and the trust scores hit a floor effect when there were two or more inaccuracies in the system.

Figures 2 and 3 examine objective and subjective trust scores that were derived from the factor analysis of the 10 TVS subscales. Figure 2 shows that the pattern for objective trust scores is similar to the pattern for TVS scores, and the pattern is also statistically significant ($F(4, 121) = 34.02$, $MSE = 128.93$, $p < .001$, $\eta_p^2 = .53$). From Figure 3, it appears that subjective trust tends to be highest when the system is completely accurate and is generally lower in all conditions with one or more incorrect selections ($F(4, 121) = 4.26$, $MSE = 7.72$, $p = .003$, $\eta_p^2 = .12$). However, a comparison of means using Tukey-Kramer HSD reveals that trust in a perfectly accurate system is higher ($M = 14.94$) than when there are three inaccuracies in the system ($M = 11.97$); there was no evidence to support other statistically significant differences. These findings are not unexpected, because earlier in this paper, we showed how the data support objective trust (or voter trust in the system’s properties and performance) impacting overall trust in the voting system more than does subjective trust (or an individual’s overall inclination to trust technology and people).

**DISCUSSION: TVS DEVELOPMENT, REFINEMENT, AND VALIDATION**

This paper describes how we developed an instrument to assess voters’ trust in any voting system with which voters have immediate, direct experience. This instrument, called the Trust in Voting Systems (TVS) measurement, it is believed to be one of the first multi-dimensional, reliable, and valid trust-in-voting-systems measurements developed using psychometric methods. This work was made
possible through the collaboration between human factors/human-computer interaction researchers and psychometricians. Due to this unique combination of expertise, the TVS is now ready to be further refined, if needed, and used to assess voter trust in specific voting methods with which voters have direct experience. The measurement development methods presented in this paper can also be used to develop other human-computer interaction and human factors instruments to assess subjective, self-reported aspects of human-computer/system interactions.

Throughout the psychometric refinement process for the Trust in Voting Systems measure, analyses revealed a useful and novel finding: that voting system attributes—such as usability, security, and reliability—were more closely related to voters’ overall trust ratings than to personal, subjective factors such as trust in others or propensity to trust in technology. The current project then applied the refined TVS in an experiment measuring trust in one specific e-voting system with several manipulated levels of accuracy. The intent of this study was to begin validating the instrument through actual use. These TVS data indicated an initial decrease in average trust, leveling off once two or more inaccuracies are introduced into the system. This is an important finding, because it examines whether voter trust is gradual or an all-or-nothing concept with respect to a voting system. Similarly, just as we found reliable differences in trust across levels of system inaccuracy, we found reliable differences in levels of voter trust, both within and across the experimental conditions of the e-system: The TVS composite trust scores were wide-ranging, from 5.6 to 93.1 in the current sample.

Figure 2. Box plot for level of objective trust in the voting system as a function of system accuracy
(M = 49.5, SD = 22.8), with a high proportion of this variance being reliable, given the psychometric information provided here. Future studies involving the TVS will further ensure that the measure can reliably capture the full range of possible scores when measuring both voting-system differences and individual differences.

Many future research directions are possible. For instance, the level of TVS scores to be considered high, medium, and low is largely unknown in any absolute sense, and relative to other voting samples and systems. Obtaining norms or benchmarks for practical significance would help voting researchers, practitioners, and policymakers understand and use TVS scores. This would be especially helpful in studies of a single voting system, which could be compared to relevant norms already established in other voting systems. Similar norms have been established for adjective rating scales associated with the System Usability Scale in the HF/HCI setting (Bangor, Kortum, & Miller, 2009) and by SAT and ACT scores familiar to the college admissions setting.

Besides examining system differences, future research could also explore individual differences further. For instance, one could correlate objective and subjective dimensions measured with a single system with other variables, such as voting perceptions and voting outcomes within a system. This type of research could lead to a more integrated model of voting that analyzes and predicts outcomes,
with scale main effects, system main effects, scale-by-system interactions (and reporting relevant $R^2$ values).

Note that future research might also recruit large and representative samples of the U.S. voting population. The current project sampled undergraduate students who were eligible U.S. voters. Obviously, these participants tend to be younger, more highly educated, and more technologically savvy than U.S. voters from the general population. Our findings remain useful to this demographic, and our pattern of findings may generalize even further (even if specific estimates may differ in other types of voter samples). Examining different samples across time, age, demographics, educational background, and countries may inform future refinements of the TVS. Additional studies across large and diverse samples are always helpful in supporting and understanding the reliability, validity, and fairness of any specific measure.

In offering a final topic for future research: the multiple constructs and the two factors found in the TVS are applicable to any voting system, including paper ballots, lever machines, and the cryptographically secure, voter-verifiable systems like Helios (Adida, 2008) and Prêt à Voter (PaV; Lundin & Ryan, 2008). We focused on one particular electronic voting system, yet different types of voting systems likely correspond to different types of errors and thus different types of trust.

CONCLUSION

As described, the development of the Trust in Voting Systems (TVS) measure is conceptually multidimensional, psychometrically refined, and empirically validated; the measure addresses a critical need in the voting literature and in practice when assessing voting systems. The TVS builds upon the previous trust research by incorporating additional substantive constructs that researchers found to impact trust, such as perceptions of security and anonymity (e.g., Besselaar, Oostveen, De Cindio, & Ferrazzi, 2003), but it also expands upon this body of work by including additional constructs, such as system usability and trust in others. The TVS measure opens up a multitude of other avenues for impactful research regarding voting systems and trust. For example, it could be used to identify systematically how trust is impacted by specific system features, such as reliability, aesthetics, security logos, and voting system instructions. The more systems and features examined—both within and across research studies, and both within and across voters—the better one can understand and appreciate the full spectrum of voter trust.

Data and analyses from the TVS can be deployed to design, implement, and evaluate trustworthy voting methods, which includes user interfaces; shape voting policy; and increase voters’ confidence that their vote will impact election outcomes. We earnestly hope that future researchers will make use of the TVS in future research and to measure trust in any voting system. If researchers all use the TVS, then their data can be more easily compared and meta-analyzed across systems. An experiment-ready version of the TVS, and its scoring rubric, is included in the Appendices.

ACKNOWLEDGMENT

This research was supported in part by NSF grant [withheld for review]. We would like to thank Peter Ryan and his colleagues for contributing their time and insights in answering our questions about Prêt à Voter.
REFERENCES

Acemyan, C. Z., & Kortum, P. (2012). The relationship between trust and usability in systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 1842–1846. doi:10.1177/1071181312561371

Acemyan, C. Z., & Kortum, P. (2015). Does the polling station environment matter? The relation between voting machine layouts within polling stations and anticipated system usability. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59*(1), 1066–1070. doi:10.1177/1541931215591299

Acemyan, C. Z., Kortum, P., Byrne, M. D., & Wallach, D. (2014). Usability of voter verifiable end-to-end voting systems: Baseline data for Helios, Prêt à Voter, and Scantegrity II. *Journal of Election Technology and Systems, 2*, 26–56.

Acemyan, C. Z., Kortum, P., Byrne, M. D., & Wallach, D. S. (2015a). Users’ mental models for three end-to-end voting systems: Helios, Prêt à Voter, and Scantegrity II. *International Conference on Human Aspects of Information Security, Privacy, and Trust, 463-474. doi:10.1007/978-3-319-20376-8_41*

Acemyan, C. Z., Kortum, P., Byrne, M. D., & Wallach, D. S. (2015b). From error to error: Why voters could not cast a ballot and verify their vote with Helios, Prêt à Voter, and Scantegrity II. *USENIX Journal of Election Technology and Systems, 3*(2).

Acemyan, C. Z., Kortum, P., & Payne, D. (2013). Do voters really fail to detect changes to their ballots? An investigation of voter error detection. *Proceedings of the Human Factors and Ergonomics Society, 57*(1), 1405-1409. doi:10.1177/1541931213571314

Adida, B. (2008). Helios: Web-based open-audit voting. *Proceedings of the 17th USENIX Security Symposium, 17*, 335-348. Retrieved from https://www.usenix.org/legacy/event/sec08/tech/full_papers/adida/adida.pdf

Adida, B., de Marneffe, O., Pereira, O., & Quisquater, J.J. (2009). Electing a university president using open-audit voting: Analysis of real-world use of Helios. *EVT/WOTE, 9*.

Ansper, A., Heiberg, S., Lipmaa, H., Overland, T. A., & Van Laenen, F. (2009). *Security and trust for the Norwegian e-voting pilot project E-valg 2011. In Identity and Privacy in the Internet Age.* Springer.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjectival rating scale. *Journal of Usability Studies, 4*(3), 114–123.

Bell, S., Benaloh, J., Byrne, M. D., DeBeauvoir, D., Eakin, B., Fisher, G., . . . Pereira, O. (2013). STAR-Vote: A secure, transparent, auditable, and reliable voting system. *The USENIX Journal of Election Technology Systems, 1*(1), 18-37. Retrieved from https://arxiv.org/pdf/1211.1904.pdf

Benaloh, J., Byrne, M., Kortum, P., McBurnett, N., Pereira, O., Stark, P. B., & Wallach, D. S. (2012). STAR-Vote: A secure, transparent, auditable, and reliable voting system. Retrieved from https://arxiv.org/pdf/1211.1904.pdf

Bhattacherjee, A. (2002). Individual trust in online firms: Scale development and initial test. *Journal of Management Information Systems, 19*(1), 211–241. doi:10.1080/07421222.2002.11045715

Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan & B. A. Thomas (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Brundage, S. (2013, December 17). Why we don’t trust technology companies. *Scientific American, 310*(1). http://www.scientificamerican.com/article/why-we-dont-trust-technology-companies/

Burton, C., Culnane, C., Heather, J., Peacock, T., Ryan, P. Y., Schneider, S., . . . Xia, Z. (2012). Using Pret a Voter in Victorian state elections. *Electronic Voting Technology Workshop/Workshop on Trustworthy Elections*. Retrieved from https://www.usenix.org/system/files/conference/evtwote12/evtwote12-final9_0.pdf

Byrne, M. D., Greene, K. G., & Everett, S. P. (2007). Usability of voting systems: Baseline data for paper, punchcards, and lever machines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 171-180). ACM. doi:10.1145/1240624.1240653

Campbell, B. A., & Byrne, M. D. (2009). Now do voters notice review screen anomalies? A look at voting system usability. *Proceedings of the 2009 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE ‘09), 1-14*. Retrieved from https://www.usenix.org/legacy/events/evtwote09/tech/full_papers/campbell.pdf
Carback, R., Chaum, D., Clark, J., Conway, J., Essex, A., Herrnson, P. S., . . . Vora, P. L. (2010). Scantegrity II Municipal Election at Takoma Park: The first e2e binding governmental election with ballot privacy. *Proceedings of the 19th USENIX Security Symposium*, 19.

Chiang, L. (2009). Trust and security in the e-voting system. *Electronic Government, an International Journal*, 6(4), 343-360.

Cylance Team. (2016). *Cylance Discloses Voting Machine Vulnerability*. https://blog.cylance.com/cylance-discloses-voting-machine-vulnerability

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, 18(2), 192–203. doi:10.1037/1040-3590.18.2.192 PMID:16768595

Everett, S. P. (2007). *The Usability of Electronic Voting Machines and How Votes Can Be Changed Without Detection* (Doctoral dissertation). Rice University. Retrieved from http://chil.rice.edu/alumni/petersos/EverettDissertation.pdf

Fisher, M. (2013, October 9). Oops: Azerbaijan released election results before voting had even started. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/worldviews/wp/2013/10/09/oops-azerbaijan-released-election-results-before-voting-had-even-started/

Fox, J. E., & Boehm-Davis, D. A. (1998). Effects of age and congestion information accuracy of advanced traveler information systems on user trust and compliance. *Transportation Research Record: Journal of the Transportation Research Board*, 1621(1), 43–49. doi:10.3141/1621-06

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96. doi:10.1016/j.jrp.2005.08.007

Halderman, A. J. (2016). *Want to Know if the Election was Hacked? Look at the Ballots*. A Medium Corporation. https://medium.com/@jhalderm/want-to-know-if-the-election-was-hacked-look-at-the-ballots-c61a6113b0ba#o0worzxvd

Hart, S. G., & Staveland, L. E. (1988). *Development of NASA=TLX (Task Load Index): Results of empirical and theoretical research*. Retrieved from https://humanfactors.arc.nasa.gov/groups/TLX/downloads/NASA-TLXChapter.pdf

Hoffman, L. J., Lawson-Jenkins, K., & Blum, J. (2004). Trust beyond security: An expanded trust model. *Communications of the ACM*, 49(7), 95–101. doi:10.1145/1139922.1139924

Korecki, N., & Cadelago, C. (2020). Democrats fear for November after Wisconsin voting spectacle. *Politico*. Retrieved from https://www.politico.com/news/2020/04/13/democrats-fear-november-wisconsin-voting-spectacle-179585

Liptak, A. (2017). Virginia is replacing some of its electronic voting machines over security concerns. A demonstration at Defcon helped prompt the decision. *The Verge*. Retrieved from https://www.theverge.com/2017/9/10/16284108/virginia-replacing-direct-recording-electronic-voting-machine-cybersecurity

Loeber, L., & Dutch Electoral Council. (2008). E-voting in the Netherlands: From general acceptance to general doubt in two years. *Electronic Voting*, 13(1), 21–30.

Lundin, D., & Ryan, P. Y. (2008). Human readable paper verification of Prêt à Voter. In S. Jajodia & J. Lopez (Eds.), *Computer Security – ESORICS 2008: Proceedings of the 13th European Symposium on Research in Computer Security, Malaga, Spain, October 6-8, 2008* (pp. 379-395). Berlin, Germany: Springer Berlin Heidelberg. doi:10.1007/978-3-540-88313-5_25

Montague, E. N., Kleiner, B. M., & Winchester, W. W. III. (2009). Empirically understanding trust in medical technology. *International Journal of Industrial Ergonomics*, 39(4), 628–634. doi:10.1016/j.ergon.2009.01.004

Montanardo, D., & Naylor, B. (2017). *On Intelligence and Russian Hacking, Are Trump and His Team Missing the Point?* NPR. https://www.npr.org/2017/01/06/508520414/on-intelligence-and-election-hacking-trump-and-his-team-continue-to-miss-the-poi
Oostveen, A., & van den Besselaar, P. (2005). Trust, identity, and the effects of voting technologies on voting behavior. *Social Science Computer Review, 23*(3), 304–311. doi:10.1177/0894439305275852

Oostveen, A. M., & Besselaar, P. V. (2009). Users’ experiences with e-voting: A comparative case study. *International Journal of Electronic Governance, 2*(4), 357–377. doi:10.1504/IJEG.2009.030527

Oostveen, A., & van den Besselaar. (2004). Security as belief: User’s perceptions on the security of electronic voting systems. In *Electronic voting in Europe-technology, law, politics and society* (pp. 73-82). Retrieved from http://twoday.net/static/evoting/files/E-Voting-in-Europe-Proceedings.pdf#page=72

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods, 47*(4), 1343–1355. doi:10.3758/s13428-014-0543-2 PMID:25479734

Popoveniuc, S., & Carback, R. (2010). Clearvote: An end-to-end voting system that distributes privacy between printers. *Proceedings of the 9th Annual ACM Workshop on Privacy in the Electronic Society*, 119-122. doi:10.1145/1866919.1866937

Prêt à Voter. (n.d.). Retrieved from http://www.pretavoter.com

Randell, B., & Ryan, P. Y. (2006). Voting technologies and trust. *IEEE Security and Privacy, 4*(5), 50–56. doi:10.1109/MSP.2006.140

Robertson, S. (2005). Voter-centered design: Toward a voter decision support system. *ACM Transactions on Computer-Human Interaction, 12*(2), 263–292. doi:10.1145/1067860.1067866

Ryan, P. Y., Bismark, D., Heather, J., Schneider, S., & Xia, Z. (2009). Prêt à Voter: A voter-verifiable voting system. *IEEE Transactions on Information Forensics and Security, 4*(4), 662–673. doi:10.1109/TIFS.2009.203323

Ulanoff, L. (2016). *In ballots we trust: E-voting, hacking and the 2016 election*. Mashable. http://mashable.com/2016/09/26/election-voting-system-analysis/#76DPUSF8KSql

Van den Besselaar, P., Oostveen, A. M., De Cindio, F., & Ferrazzi, D. (2003). *Experiments with e-voting technology: Experiences and lessons*. Retrieved from http://www.social-informatics.net/Bologna2003.pdf

Wallach, D. (2016). *Committee on Science, Space, & Technology*. Full Committee Hearing—Protecting the 2016 Elections from Cyber and Voting Machine Attacks. Congress.

Wallach, D. (2017). Up for debate: Is Harris County’s vote safe from the Russians? A computer-security expert says no. The county clerk says yes. *Houston Chronicle*. https://www.houstonchronicle.com/local/gray-matters/article/Is-Harris-County-s-voting-system-ready-for-the-11297533.php
APPENDIX A

Using the TVS

Beginning on the next page is the Trust in Voting Systems (TVS) measure. The TVS should be administered immediately after a participant has used the voting system being evaluated.

Scoring the TVS

The TVS yields two scores of primary concern. The first is the TVS score. TVS scores are cumulative and range from 0 to 100. To calculate the TVS, a scale score is calculated for each scale, and then 1 is subtracted from each scale score; then 10 scale scores are added together; and finally, this sum is multiplied by 2.5. Stated mathematically:

\[
TVS = \{ \left( \frac{I_{t2} + I_{t13} + I_{t24} + I_{t35}}{4} \right) - 1 \} + \left( \frac{I_{t3} + I_{t14} + I_{t25} + I_{t36}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t4} + I_{t15} + I_{t26} + I_{t37}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t5} + I_{t16} + I_{t27} + I_{t38}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t6} + I_{t17} + I_{t28} + I_{t39}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t7} + I_{t18} + I_{t29} + I_{t40}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t8} + I_{t19} + I_{t30} + I_{t41}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t9} + I_{t20} + I_{t31} + I_{t42}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t10} + I_{t21} + I_{t32} + I_{t43}}{4} \right) - 1) \]
\[
+ \left( \frac{I_{t11} + I_{t22} + I_{t33} + I_{t44}}{4} \right) - 1) \} \times 2.5
\]

For reference, the following is a list of items associated with each scale.

**Trust in Others:** It2, It13, It24, It35  
**Propensity to Trust in Technology:** It3, It14, It25, It36  
**Cast Ballot Reflects Intended Selections:** It4, It15, It26, It37  
**Votes Are Accurately Counted:** It5, It16, It27, It38  
**Vote is Anonymous:** It6, It17, It28, It39  
**System Security:** It7, It18, It29, It40  
**System Usability:** It8, It19, It30, It41  
**System Reliability:** It9, It20, It31, It42  
**System Accuracy:** It10, It21, It32, It43  
**System Transparency:** It11, It22, It33, It44

The second is the Overall Trust in Voting Systems score (OTVS). This overall trust score has a range of 1 to 5. It is calculated by averaging items 1, 12, 23, and 34:

\[
\text{Overall Trust in the Voting System (OTVS) score} = \frac{I_{t1} + I_{t12} + I_{t23} + I_{t34}}{4}
\]

The OTVS scale score can be used to assess users’ overall trust in the voting system quickly, and how different subscales of the TVS correlate with OTVS can provide information about the nature of overall trust.
APPENDIX B

Trust in Voting Systems (TVS) Instrument

Check the box that reflects your immediate response to each statement. Don’t think too long about each. Make sure you respond to every statement. If you don’t know how to respond, check box “3.”

1. I trust the voting method

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

2. I believe that others have good intentions

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

3. In general, I trust technology

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

4. I believe my vote was cast as intended

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

5. I think the vote I cast is accurately counted

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

6. My vote is anonymous

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

7. I am confident that the voting method is not exposed to harm

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

8. Using the voting method, I was able to cast a vote that I was confident in

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

9. I am confident that the voting system is dependable

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

10. The voting method is accurate

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |
11. I am confident that the voting method is transparent

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

12. The voting system is trustworthy

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

13. I believe that people are basically moral

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

14. I am confident in technology

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

15. My final ballot reflected the people and propositions that I wanted to vote for

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

16. My cast ballot is included in the total number of votes

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

17. My ballot cannot be traced back to me

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

18. The voting system can’t be tampered with

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

19. The voting system is great

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

20. The voting system never fails me

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |

21. The voting method is precise

| Strongly disagree | Strongly agree |
|-------------------|---------------|
| 1                 | 2             |
| 3                 | 4             |
| 5                 |               |
| Question                                                                 | Strongly disagree | Strongly agree |
|-------------------------------------------------------------------------|-------------------|---------------|
| I am confident that the vote was counted correctly                       | 1 2 3 4 5         |
| I have no reason to distrust the voting method                           | 1 2 3 4 5         |
| I believe in human goodness                                              | 1 2 3 4 5         |
| I am totally comfortable working with technology                        | 1 2 3 4 5         |
| Every choice I made while using the voting method was represented as a vote on my cast ballot | 1 2 3 4 5         |
| I believe that every one of my ballot selections is counted correctly    | 1 2 3 4 5         |
| Unless I tell someone, there is no way for anyone to ever know who I voted for | 1 2 3 4 5         |
| The voting system is secure                                              | 1 2 3 4 5         |
| I would rely on this voting method in the future                         | 1 2 3 4 5         |
| I am confident that the voting system always works                       | 1 2 3 4 5         |
| I am confident that the voting system is always correct                  | 1 2 3 4 5         |
| Question                                                                 | Strongly disagree | Strongly agree |
|-------------------------------------------------------------------------|-------------------|---------------|
| 33. I am confident that a voter can verify their vote                  | 1 2 3 4 5         |               |
| 34. Overall, my trust in the voting system is high                      | 1 2 3 4 5         |               |
| 35. I trust people                                                      | 1 2 3 4 5         |               |
| 36. Things will be fine when I use technological systems               | 1 2 3 4 5         |               |
| 37. My ballot reflects my intentions and beliefs                        | 1 2 3 4 5         |               |
| 38. I don’t believe that someone would intentionally leave out my vote | 1 2 3 4 5         |               |
| 39. Only I know who I voted for                                          | 1 2 3 4 5         |               |
| 40. Voting with the method is riskless                                  | 1 2 3 4 5         |               |
| 41. The voting system works the way I want it to work                   | 1 2 3 4 5         |               |
| 42. The voting system is reliable                                       | 1 2 3 4 5         |               |
| 43. The voting system does what it is supposed to do                    | 1 2 3 4 5         |               |
| 44. The voting system does not hide anything from me                    | 1 2 3 4 5         |               |
Claudia Ziegler Acemyan is an Adjunct Assistant Professor in the Department of Psychological Sciences at Rice University. She is also a human factors and human-computer interaction consultant at her company, Post Hoc, LLC. Her research has focused on topics like usability of voting systems, user-centered design, errors, incident evaluations, and measurement development.

Philip Kortum is an Associate Professor in the Department of Psychological Sciences at Rice University. His research focuses on human factors.

Frederick L. Oswald is an industrial-organizational psychologist and Professor in the Department of Psychological Sciences at Rice University. Learn more about his research at https://workforce.rice.edu.