



VOIS: A framework for recording Voice Over Internet Surveys

Teresa Ristow¹ · Ivan Hernandez¹

Accepted: 1 December 2022 / Published online: 25 January 2023
© The Psychonomic Society, Inc. 2023

Abstract

Verbal data provide researchers insight beyond that offered by text-based responses, including tone, reasoning elaboration, and experienced difficulty, among other processes. Additionally, it offers a less cognitively taxing way for participants to provide long responses. Verbal data collection methods are found in a variety of fields, mostly conducted in lab-based settings or requiring specialized hardware. Restricting verbal protocols to lab-based settings can have several drawbacks, including smaller sample sizes, biased populations, reduced adoption, and incompatibility with potential social distancing requirements. No method currently exists for researchers to collect verbal data within major online survey collection platforms. The current paper offers a user-friendly approach for collecting verbal data online, where a researcher can copy and paste JavaScript code into the desired survey platform. By providing a framework that does not require any advanced programming ability, researchers can collect verbal data in a scalable way using familiar modalities.

Keywords Experimental design · Verbal protocol · Voice data · MTurk · Survey platform

“Verbal protocol” describes a type of method for collecting and analyzing voice recordings from participants to gain insight beyond more traditional and observationally based techniques (Zainal Abidin et al., 2009). Verbal protocols provide oral records of participants’ thoughts during thinking-aloud processes while participants work on a specified task or after completion of a task (Kasper, 1998). This elaboration can provide insight into a variety of cognitive psychological tasks such as problem-solving, knowledge acquisition, memory tasks, learning processes, judgement, human-computer interaction, reading and writing, and decision-making (Biggs et al., 1993; Kraemer & Ummelen, 2004; Van Someren et al., 1994). Incorporating recording requires minimal changes from a survey design task because audio is recorded passively. The only change is that participants are instructed to narrate their thoughts, rather than leave them internal during task completion. Therefore, incorporating voice recording into study designs offers additional information into a variety of social and cognitive processes, in a way that generalizes easily across study designs. The limitation to using this methodology is the technical issue of

having a recorder present during the study, especially when the study is done remotely across hundreds of participants, as is common in survey data collection.

Usage of verbal protocols across behavioral fields

Think-aloud tasks

Verbal protocol methods are often used in *think-aloud* studies. These studies require participants to speak aloud any words that come to mind as they complete a task, while researchers record participants’ voices (Charters, 2003). Additionally, participants can record their thoughts retrospectively after the completion of a task. Concurrent and retrospective verbal protocols provide distinct, but equally useful methods of obtaining data on thought processes (Crutcher, 1994). Tasks relevant for think-aloud designs can include memorization tasks, problem-solving tasks, puzzles, and economic games, which all involve complex cognitive processes leading to a final decision or answer (Ericsson, 2006).

Voice data from verbal protocol methods offer important data to supplement observations and task results. Thinking is a sequence of thoughts used to get from one processing

✉ Teresa Ristow
tristow@vt.edu

¹ Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060, USA

activity to another often manifested in subvocalization (Ericsson & Simon, 1998). Therefore, recording thoughts provides insight into unobservable cognitive processes required to complete certain commonly researched tasks (Austin & Delaney, 1998). Although researchers could ask participants to write their thoughts, voice recording provides a less intrusive, timely, and more natural way for participants to express their thought process. Thus, the verbalization of thoughts is seen as one of the most direct methods to study deliberative cognitive processes for complex tasks (Kraemer & Umellen, 2004). This benefit stems from the immediacy between the task being done and the verbal report of the action. Participants typically have no time to exert the cognitive effort required to adjust or manipulate their verbal report of their actions during or after a task (Van Someren et al., 1994).

Methods to collect vocal data

In order to capture vocal data in experimental studies, researchers utilize a variety of techniques, mechanisms, and hardware. When the experimental setting is in a lab, typically the materials and recording equipment are provided to the participant by the experimenter. Commonly, the materials used for studies within a laboratory setting include an apparatus that consists of a microcomputer equipped with a megahertz processor, a sound card for storage of the recording files, and headphones or a headset with a microphone (Wirth et al., 2000). Laboratories that may not have access to computers or all the equipment suggested for verbal protocol data collection can opt for simply audio recording a participant as they complete a think-aloud protocol, which the researchers or assistants can later transcribe for analysis (Hoffman et al., 2009). As a means to avoid the tedious task of transcribing vocal recordings, researchers may instead use two-way radio and single earbud-microphone systems to enter verbal data as it is being spoken, which is traditionally called “bug-in-ear” (BIE) technology (Goodman et al., 2008). However, many of these methods and devices used in the lab limit the researcher from obtaining larger sets of data because of the configuration and demonstration of the technology required for each participant.

Additionally, researchers may utilize specifically developed devices for accessing and collecting vocal data outside of the laboratory setting. The Electronically Activated Recorder (EAR) is a device designed to be used exclusively for vocal recording out of the lab. It consists of a microcassette tape recorder and a tie-clip microphone, as an unobtrusive means for at-home vocal data collection (Mehl et al., 2001). Specifically, the EAR was created to be used with an experience sampling methodology (ecological momentary assessment, EMA), where participants fill out questionnaires several times a day in their own homes. The EAR can help researchers gain insight into the thought process of those

participants while they fill out surveys using a think-aloud protocol. This can provide vocal data in addition to data collected via EMA, which uses several types of tools (such as cell phones or fitness trackers) that collect psychological or physiological data moment-by-moment in real-world situations. While EAR technology avoids the pitfalls that come from using vocal protocol collection devices within the lab, such as increased researcher effort and scalability of data collected, there are still some major concerns with the EAR. The most prevalent issue with the EAR is that participants may feel hesitant to record vocal data outside of the lab due to privacy concerns. Additionally, participants are required to set up the equipment on their own at home. This independence can subsequently lead to a loss of data if any participants have technical difficulties, which is more common outside of the lab, even if they are taught how to use the device in the lab.

In addition to the EAR device, other technology is utilized to collect vocal data online and outside of the lab. While the EAR functions with the same recording mechanisms as devices used in a laboratory setting, there are methods that use online capabilities to record vocal data outside of the lab. A popular method of online voice protocol recording is Voice over Internet Protocol (VoIP) (Goode, 2002). VoIP is a stand-alone Internet application that records high-quality voice data over the Internet. However, if it is used with an online survey platform, it must be opened in separate windows which may be cumbersome and contain problems with timing for participant vocal recording. Additionally, VoIP may suffer from low voice quality depending on an individual participant’s Internet speed or router connection. Another option for online vocal protocol data collection uses a telephone to place a call and receive voice over the Internet. This method utilizes the VoIP technology and applies it to a telephone system (Jones et al., 2000). This allows participants who may not have access to or be familiar with online recording systems to simply call a number and have their voice recorded for the researchers to access as an electronic mail message. While this means of vocal data collection is easier for participants unfamiliar with computers, it is inefficient and requires considerable setup on the side of the researcher. A more dated option that exists for at-home voice data collection is using a two-way call during the voice protocol task with Network Voice Protocol (NVP-II). This technology avoids the Internet entirely and simply records vocal data via telephone (Cohen et al., 1981). All of the current methods lack a way to effectively download or access more than one participant’s recording at a time. This may lead to large amounts of time spent with data entry post vocal data collection. While the NVP-II technology is the oldest mentioned, the most recent device used to collect voice data is still only from

an article published in 2009. This shows that much of the published means of collecting voice data during verbal protocol tasks are fairly dated and have not been updated by academics for over 10 years. Despite the dated methods of collecting voice data, there are still a variety of constructs that can be measured by these means, however ineffective.

More recently, developments in technology have facilitated online and updated means of utilizing vocal protocols to access vocally recorded data. One example is when Cauldron Science introduced Gorilla in 2016 which provides researchers with a tool to develop and build their own online experiments (Cauldron Science, 2016). This platform provides more experiment-based research options than more traditional survey platforms. This includes options to collect participant audio data, which is stored in a compressed WAV file. However, this option is closed-source and may require additional payment. Gorilla also may be more complex than many researchers may need in terms of their provided options. This could require a time investment from the researcher to familiarize themselves with what may be an unfamiliar experiment platform.

Another recent technological advancement for collecting voice data is via PsychoJS. PsychoJS is an open-source way for researchers to present participants with stimuli and recorded verbal responses online. This technology also utilizes compressed WAV files to avoid issues with large file sizes. While PsychoJS is a great option for researchers looking to implement experimental designs using voice protocol in an inexpensive and efficient way, it does require coding knowledge (PsychoJS, 2021). This may be an issue to many researchers with behavioral science backgrounds, who do not typically get training or frequently use coding languages to collect data from experiments.

Phonic Inc. also provides researchers with a survey platform designed around voice recording and video (Phonic, 2019). Along with their available survey platform, Phonic Inc. makes available Phonic.ai. Phonic.ai is a more widget-based vocal recording option and allows for the collection of audio data from outside sources or platforms. The stand-alone survey platform option is the most user-friendly option from Phonic Inc., but does require that an entirely new survey be created to use vocal recording. Phonic.ai allows for integration with outside survey platform options but does require that researchers pay to utilize these features, and cost can be a large factor for those without funding options. While Phonic.ai is a viable option for many researchers for integrating vocal recording with existing surveys, there is limited information on how to integrate HTML code provided with those platforms, and there is a paywall to access the full functionality for Phonic.ai. Additionally, there is not much information available on how researchers handle and

access the recorded data, which can be confusing to those new to using vocal protocols.

One option for online vocal recording is through addpipe.com, offered as their product, Pipe Audio and Video Recorder (Pipe Services S.R.L., 2015). This vocal recording option is useful for those looking to integrate vocal recording across multiple devices and operating systems. The Pipe Audio and Video Recorder uses an HTML5 audio and video recorder that is built mainly for developers. While this could enable the product to be integrated with many existing surveys on different platforms, it requires extensive coding and web development knowledge. It would also require researchers to sign up and pay for a subscription to access a secure server in order to save the recorded audio data and ensure participant data is safe.

Finally, there are several options for more experimental-based online vocal protocols. Some of these include: PsychoPy's Pavlovia, PsyToolkit, and OpenSesame's OSWeb extension that can publish online experiments through JATOS (OSWeb, 2020; PsychoPy., 2018; PsyToolkit, 2022). These options are successfully and commonly used for experimental psychology applications. This does pose issues for researchers unfamiliar with using more experimental-based applications to record vocal data. Specifically, all the previously mentioned options are separate study platforms that would require researchers to build their study from scratch on the platform in order to record participants' voices. This would not enable the full integration of vocal recording into any existing survey platforms the researcher may be more comfortable using or already pay a subscription to. Overall, many of these more updated and online-based methods of voice data collection solve some of the issues with older methods, but they present their own limitations.

Constructs captured by voice data and their applications

Tasks using verbal protocol methods can measure many different cognitive constructs, including (1) effort during a task through pauses in verbal protocol or verbalized effort, (2) common errors or deficits in thinking while completing a task, and (3) improvements in thinking processes documented across tasks as a participant learns (Kasper, 1998; Zainal Abidin et al., 2009). The ability to measure a wide variety of constructs allows for verbal protocol methods to be effectively used across many different fields and in conjunction with many different tasks. Below we describe how different psychological and other fields examine these broad constructs using voice data.

Human factors studies examine how humans interact and use products or computers in their environment, in an effort to make that process as efficient and ergonomic as possible. Human factors testing uses psychological concepts and

integrates those with how people interact with their current environment. A verbal protocol facilitates human factors research by offering insight into a participant's immediate experience with a product's design as they interact with it. This insight allows human factors researchers to find flaws or means of improvement with a specific product of interest (Deffner et al., 1990). More recently, software engineering research incorporates verbal protocol methods to analyze the difference between expert and novice problem-solving on software comprehension and programming tasks (Hughes & Parkes, 2003). In a more niche application, environmental regulation efforts utilize verbal protocols to understand the value of conservation efforts at various costs in willingness-to-pay tasks. This allows environmental researchers to better implement effective interventions (Schkade & Payne, 1994).

Industrial/organizational psychologists explore managers' problem-solving processes during high-risk organizational decision-making, using a verbal protocol design during work sample tasks (Isenberg, 1986). A verbal protocol can also be used in industrial/organizational psychology to review structured and unstructured interviews as they are given to an applicant. Industrial/organizational psychologists implement a verbal protocol in conjunction with EMA methods in order to prompt individuals to elaborate on what they are experiencing at a specific moment during the workday. EMA and verbal protocols have also been used together across medical fields to document instances of addictive or harmful behaviors and corresponding thought processes as they happen (Litt et al., 1998). Industrial/Organizational psychology, in addition to various other fields, uses verbal protocols as a more participant-friendly way to collect diary data. This allows participants to speak out loud and record their diary entries rather than write, which may take more effort for a participant (Gouveia & Karapanos, 2013).

The diverse usage of verbal protocol methods allows for its further implementation in a variety of fields outside of human factors and Industrial/Organizational psychology. For example, a verbal protocol is used to study how architects collaborate and go about completing design projects, allowing researchers to gain insight into the creative process (Zainal Abidin et al., 2009). Researchers also implement a think-aloud protocol within different fields of engineering. Specifically, using student participants during their studying of course material and working on design projects. By having these students describe their thought processes, it is possible to take a more in-depth look at how products, systems, or interventions are designed and conceptualized by electrical, mechanical, biomedical, environmental, civil, and industrial engineers (Atman & Bursic, 1998). In more humanistic-based fields, nursing and special education teaching both utilize aspects of verbal protocol methods in order to gain insight into a variety of constructs. The nursing field uses a think-aloud process while participant nurses care

for patients, allowing researchers to differentiate between experts' and novice nurses' thought processes during patient care (Hoffman et al., 2009). These differences in thought processes enable researchers to then understand how expert vs novice nurses make decisions differently which can better assist in the training development of newly hired nurses. Researchers focused on special education teachers utilize voice protocol methods to provide immediate feedback to novice teachers after think-aloud tasks or during lessons with students (Goodman et al., 2008). These voice recordings of special education teachers can also be collected for analysis by researchers post think-aloud protocols.

The broad range of fields and uses of vocal protocol collection emphasizes the variety of constructs of interest in vocal data. Many of these constructs collected by more traditional methods may actually be best collected by these more dated methods and not more updated means of vocal protocol collection. However, traditional methods do still limit researchers in their accessibility of vocal data collection. That is, many of the constructs reviewed may be difficult for researchers to access collection means without specialized equipment, resources, or training.

Analysis of voice data

After collecting voice recordings, researchers have a variety of options to quantify the unstructured data into constructs of interest. Commonly, researchers examine the content of the words spoken. Examples of content analysis include counting specific words or phrases, or examining how closely a participant's narrative matches a particular "thought pattern" that the researcher defines ahead of time (Chi, 1997). In addition to coding the spoken words, researchers can analyze the waveform of the recording to extract psychological relevant variables. One method of extracting meaning from vocal data is examining specific vocal indicators. Vocal indicators can be anything from tone, prosody, pitch, disfluencies, or pauses (Sondhi et al., 2015). Researchers can also calculate aggregate/composite variables from these vocal indicators to represent higher-level features (e.g., formants F1 and F2; Sondhi et al., 2015). These analytic methods illustrate the diverse range of options available to quantify verbal data, highlighting its potential to provide distinctive insights from traditional scale-based approaches.

Limitations of voice data collection

While vocal protocol methods can provide a source of rich data not easily accessed by other methods, there can be several limitations in its practical use. One limitation is implementing the vocal protocol method. Typically, participants

have to physically enter a lab setting supervised by researchers (Krahmer & Ummelen, 2004). This increased effort in data collection can potentially discourage participants from committing to being physically present for a study and thus can generate smaller sample sizes. Therefore, requiring participants to travel to a laboratory for voice data collection minimizes the study's scalability, increasing the difficulty of obtaining large sample sizes via current verbal protocol methods.

Additionally, verbal protocol methods can also be problematic for researchers because each lab study would need trained research assistants who are able to supervise and direct participants to record their voice data (Crutcher, 1994). This would mean that more people would need to be hired or involved in the lab, which can bring about a whole additional effort to increase lab personnel. All lab-based studies have this resource limitation, and psychological researchers are increasingly moving towards scalable online data collection, besides for vocal data collection.

In order to gain increased diversity and size in research samples, studies today tend to use online surveys. This is typically done through survey collection platforms such as Amazon's MTurk and Qualtrics (Brandon et al., 2014). These platforms provide a ready-to-use outlet for researchers to collect data outside of a laboratory setting. While these collection platforms offer a useful and scalable alternative to in-lab studies, there are some notable drawbacks. These limitations then lead researchers to sometimes opt for in-lab studies despite the own unique limitations of those methods as well.

Currently available solutions for collecting voice protocols online

Not much leeway has been made in addressing limitations of prior verbal protocol methods. This can be attributed to the risk in losing some of the richness in data provided beyond other observational methods. However, some improvements have been made in terms of better facilitating the richness of voice data. As technology of audio devices prospers, researchers benefit from these higher quality recording devices when collecting verbal data (Trickett & Trafton, 2009). The extra cost in using state-of-the-art recording technology is argued to be counteracted by additional data potentially provided. Better microphones or quality recording devices may pick up on small utterances of words, whispered words during thought, or even sighs or other non-language-based vocal indicators (Trickett & Trafton, 2009).

Other developments in verbal protocol methods have been in the realm of data analysis and not the collection method. Typical coding of verbal data requires counts of specific words, phrases, or the order they come in. This

is expanded on in attempts to recreate a mental model of thought through the coding of verbal protocol data (Zainal Abidin et al., 2009). This change in coding scheme allows researchers to try to generate a more comprehensive picture of the thought process enacted by participants. This coding scheme, however, is less universal and is best used in case study situations with researchers more familiar and trained on the protocol, because of its labor-intensive interpretations. Additionally, the vocal data are usually coupled with other methodologies such as behavioral observations or physiological measures. This integration of different sources of data requires considerable researcher effort that may not be matched with an increase in data richness. Researchers are relatively unaware of the limitations for collecting non-traditional data in online settings. Some recent developments include collecting mouse tracking data, but there is not yet a formal way to collect audio data easily in the same developed manner (Mathur & Reichling, 2019).

Currently available options to collect vocal data in conjunction with online surveys include integrating vocal protocol recording with a recording tape and microphone headset in-lab with survey data collection on a computer in front of the participant. During this vocal data collection, the participant will take the survey using a think-aloud protocol and describe their thought process while completing the survey (Wirth et al., 2000). The EAR device can also be combined with online survey data collection on survey platforms which will have a similar implementation as other recording devices but can be used during at-home verbal data collection (Mehl et al., 2001). VoIP is also available for researchers to combine with survey platform data collection but because it is a stand-alone Internet application, it must be opened in a separate window from the survey (Goode, 2002). In Keromytis (2009), options and usage guidance for VoIP are provided in the form of a road map for researchers to fully understand the capabilities of the system and how to use it in conjunction with an online survey. However, there is currently no survey platform that provides any means of voice recording or collecting vocal data within the survey itself, or as the participant completes the survey online. This lack of integration limits researchers using verbal protocol data in their ability to collect larger and more diverse datasets as well as to utilize online survey platform capabilities.

Proposed method

In order to provide an option not previously offered in the means of collecting vocal data, we propose VOIS (Voice Over Internet Surveys), a user-friendly method of collecting audio data by offering code that is compatible with common survey collection systems, and a way to convert the recorded data into audio files. VOIS can

then enable researchers who utilize online survey platforms to supplement their data collection methodology with vocal recording. This can be in the form of adding vocal recording to an existing survey, implementing VOIS with other question types available in survey platforms, or even creating a survey entirely around the collection of vocal data. Some of the reviewed existing methods for collecting vocal data provide unique contributions, regardless of their discussed limitations, beyond what is offered by VOIS. However, VOIS is structured to provide researchers with an opportunity not yet available for combining online survey platforms with the ability to collect vocal data from participants and potentially even expand on the constructs currently collected by more traditional methods.

The intended application of the proposed method is two-fold: (1) VOIS functions to break down barriers preventing many researchers from conducting voice protocol studies online, and therefore VOIS would be applied to research that needs the voice protocol design. (2) VOIS also opens up new possibilities for conducting voice protocol. This would allow VOIS to be applied to studies that are not exclusively voice protocol studies, but that could be enhanced by implementing VOIS to address potential research issues. While researchers may find numerous ways to utilize VOIS in their studies, the proposed method is useful for researchers who are looking for a simple way to collect voice data.

The first intended application of VOIS for vocal data collection is to assist researchers who may face barriers to collecting vocal data in more traditional methods or in the currently available online methods. These barriers can include the inability to access a lab space or more complex and involved tools that require time to learn how to properly use for vocal data collection. Without already having access to tools or equipment, researchers can also encounter barriers in the form of high costs involved in acquiring those resources. Other barriers could be associated with the currently available online methods used in place of more traditional in-lab methods. These include researchers possibly not having the funding for closed-source implementations for voice data collection, as well as a lack of researcher programming knowledge or background for some of the open-source options available. Thus, our proposed method circumvents these barriers, making the collection of voice data more accessible to researchers and even creating a possibility for researchers with experience in collecting alternative sources of data to explore supplemental uses of vocal data in an approachable way. We anticipate that this accessibility can be helpful specifically to researchers who may already be familiar with existing survey software and are looking for an approachable means to include voice recording in their research. This inclusion could be in several forms, such as

researchers adding vocal recording components to existing surveys as supplemental data collection.

In addition to breaking down barriers for vocal data collection, we intend VOIS to be useful to researchers looking to trial vocal recording in order to design new studies around this form of data collection or in addition to their currently used forms of data. Additionally, VOIS can be implemented within an experimental design on any survey software that a researcher may be most comfortable using. The availability of an option for vocal data collection that integrates with currently available survey platforms, unlike existing methods, can assist researchers in designing a study around their specific needs and available tools. Therefore, VOIS can allow researchers to improve upon potential research issues including a new way to verify that participants understood the study and that the responses have not been fabricated by third parties. VOIS can also replace more traditional survey response methods such as text input to reduce participant survey time and fatigue. Additionally, VOIS can address issues in measuring constructs in potentially more meaningful ways by capturing details related to vocal recording as opposed to self-report.

While VOIS may not be feasible as a substitute for every construct collected or every study design used via more traditional or in-lab methods, it can help bridge a gap in the benefits offered by both in-lab and online methods currently available. Thus, VOIS offers a risk-free way for researchers to trial vocal recording in their current and future studies. In providing a more user-friendly and accessible option that is not currently compatible with online survey methodology, we hope researchers can find an easy way to collect constructs of interest in their research.

Specifically, VOIS implements a vocal data collection via a copy-and-paste script that is compatible with any survey system that allows HTML code. This includes Amazon's Mechanical Turk (MTurk), Qualtrics, and self-hosted websites, as well as many others. This copy-and-paste approach is beneficial by not being tied to a specific platform and requiring no previous coding knowledge when implementing the method on third-party websites. Additionally, with the prevalence of online studies use in published academic articles as shown in Table 1, being compatible with even just MTurk covers a lot of online data collection landscape.

The VOIS method allows the researcher to copy and paste code into their survey, distribute the survey, obtain the audio data recorded within the main survey, and then upload it to the researcher's private server. This method is also compatible with every major operating system (Windows, Mac, Linux, Android, iOS), and most major browsers (Chrome, Firefox, Safari, Edge).

Table 1 Prevalence of online methods in social and applied journals

Journal	Year	No. of online studies	No. of MTurk studies	Total no. of studies	% of online studies	% of MTurk studies
JPSP	2019–2020	363	216	533	0.68	0.41
	2014–2015	233	137	427	0.55	0.32
	2009–2010	123	6	481	0.26	0.01
JAP	2019–2020	73	28	122	0.60	0.23
	2014–2015	73	9	176	0.41	0.05
	2009–2010	40	0	124	0.32	0.00
JCP	2019–2020	70	58	131	0.53	0.44
	2014–2015	77	38	151	0.51	0.25
	2009–2010	13	0	97	0.13	0.00

Journals examined were the *Journal of Personality and Social Psychology* (JPSP; $N_{\text{Articles}2019-2020} = 128$, $N_{\text{Studies}2019-2020} = 533$, $N_{\text{Articles}2014-2015} = 117$, $N_{\text{Studies}2014-2015} = 427$, $N_{\text{Articles}2009-2010} = 144$, $N_{\text{Studies}2009-2010} = 481$), the *Journal of Applied Psychology* (JAP; $N_{\text{Articles}2019-2020} = 76$, $N_{\text{Studies}2019-2020} = 122$, $N_{\text{Articles}2014-2010} = 123$, $N_{\text{Studies}2014-2015} = 176$, $N_{\text{Articles}2009-2010} = 97$, $N_{\text{Studies}2009-2010} = 124$), and the *Journal of Consumer Psychology* (JCP; $N_{\text{Articles}2019-2020} = 36$, $N_{\text{Studies}2019-2020} = 131$, $N_{\text{Articles}2014-2015} = 48$, $N_{\text{Studies}2014-2015} = 151$, $N_{\text{Articles}2009-2010} = 50$, $N_{\text{Studies}2009-2010} = 97$)

Overview of the proposed method

The proposed method is collected in a JavaScript file that the researcher links to in their survey. HTML-based surveys allow linking outside JavaScript files using the <script> tag. Within the script tag is a direct link to the VOIS JavaScript file, which handles the processing of the audio and sending of the data to the researcher’s storage. The VOIS JavaScript file uses the WebKit AudioContext feature found in major browsers (e.g., Chrome, Firefox, Safari, and Edge, as well as Android and iOS browsers)¹. The AudioContext feature can receive data from the participant’s microphone. This data is processed by the script into a 44,100 Hz 24-bit “WAV” formatted file, by processing the streaming data into chunks and converting the input into bytes. The WAV audio processing code incorporates elements from the JavaScript WZRecorder library, which is based on the Recorderjs library (Diamond, 2016; Dugger, 2018). By having the linked script perform the major computations and processing, the researcher then simply needs to paste the visual elements in their survey.

WAV files are uncompressed and so their file size can be large and taxing on bandwidth requirements for participants (~ 5 MB per minute). To make the method more compatible for users with limited data, at the expend of audio fidelity, we also have a version of the VOIS script that saves the audio in 160 kbps MP3 files (~1 MB per minute). The MP3 processing code incorporates elements from the WebAudioRecorder-js (Miyane, 2016) library to process the encoding

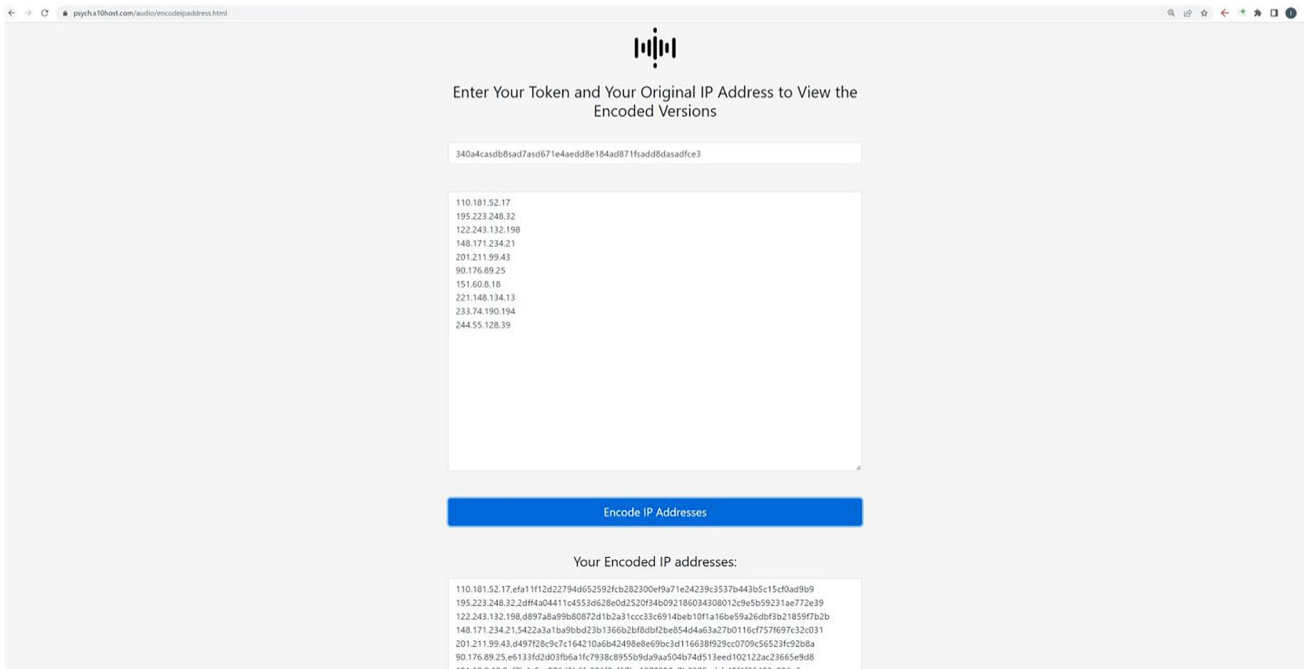
of captured audio into an MP3 file that is then uploaded to the researcher’s FTP server.

In addition to linking the VOIS JavaScript file, the researcher pastes seven lines of code (provided in the “Implementation Section”) that are responsible for the visual appearance of the voice protocol prompt. The code the researcher posts simply contains (1) the question text, (2) the recording button, (3) the time elapsed, (4) the audio player to hear the recorded clip, (5) the status of the transfer of the voice recording to the researcher’s server, (6) a data field to temporarily hold the recording data, and (7) a data field to hold the researcher’s access token. When pasted into the survey, the participant can see the instructions, click a record button, press the same button to stop the recording, and have the data sent to the researcher automatically. The data can be re-recorded by the participant and re-sent to the researcher if needed. The file sent to the researcher contains the participant’s IP address to be able to match the recordings with the rest of that participant’s responses. To protect participant privacy, the IP addresses are encoded using a SHA-256 hash of the participant’s IP address and the researcher’s access token. SHA-256 encryption is performed by the crypto-js library (Vosberg, 2021). A researcher can encode their collected IP addresses from the survey using the following site:

<https://psych.x10host.com/audio/encodeipaddress.html>

The screenshot below shows how users, after collecting survey data from a platform where the answers are identified by the original IP address, can enter (1) their VOIS access token, and the IP addresses from the survey platform. The site then shows how the IP addresses will be encoded, which can then be matched to the files saved in the researcher’s FTP server.

¹ <https://developer.mozilla.org/en-US/docs/Web/API/AudioContext>.



If a researcher does not wish to encode the IP address within the filenames that VOIS saves, they can simply deselect that option when initially generating their access token.

Benefits of using the proposed online voice protocol method versus lab-based options

Increased sample size

Our proposed alternative method of verbal protocol, as accessed via a common web-based survey platform, provides many benefits beyond how verbal protocols are currently used. One major benefit of VOIS above current verbal protocol methodology is the access to larger and more diverse participant samples than an in-lab setting, while still maintaining richness of data. The proposed method of online survey platform usage is especially relevant for researchers not affiliated with larger universities or institutions and lack access to the large populations those entities provide (Brandon et al., 2014). More independent researchers can then gain access to an otherwise unreachable population. By utilizing third-party survey methods, researchers can easily access data from a wider range of locations as well.

While the default may be for participants to complete online surveys on their laptops or desktops, many computers currently have integrated microphones that allow participants to easily take surveys with VOIS items included.

Also, in a now post-COVID-19 world, many computers are equipped with quality webcams and integrated microphones that participants are probably familiar with and comfortable using. The availability of this equipment indicates that VOIS can serve as a transition to new possibilities not as normalized or available prior to the pandemic. While participants can utilize computers to complete online surveys with VOIS, cell phone compatibility of most third-party surveys provides participants with a more efficient experience when recording vocal data. Participants who find it easier to complete the survey and provide data will be more inclined to do so, leading to more completed surveys and thus, more data.

Decreased researcher effort

Our method of collecting verbal data can also work with existing frameworks of online survey collection platforms to not only generate larger samples, but to also reduce researcher effort in doing so. The researcher would simply have to design the task or survey around the means for verbal data collection and implement it into their desired survey platform. Because VOIS is a copy-and-paste implementation, it is simple to add to an existing survey or task. Therefore, researcher effort can be reduced in terms of time spent in the lab and utilization of research assistants because of the online implementation of VOIS.

Increased generalizability of samples

Some other benefits of VOIS deal with the scalability of the data collection effort. Not only can larger samples be accessed with less research effort, VOIS also allows for remote collection. Because of the ability to access most web-based survey platforms on the phone or computer, every participant can record voice data from a more remote or comfortable locale. Because participants may be more comfortable recording vocal data in a private setting, this may potentially allow for richer content of the data so as to avoid lab effects.

Another benefit of being able to use verbal protocols in a web-based survey design is the ability to sample across large distances. Many survey platforms can be wide-reaching and access many different locations. This allows not only for a more diverse sample but also a larger sample that will, in turn, be more generalizable. While survey platforms may not provide an optimal solution for researchers who aim to recruit specific target populations, it is still possible to create advertisements, descriptions, or screening questions that ask for participation only from those meeting certain requirements or qualifications. Seeing that many surveys and studies posted on survey platforms such as MTurk can reach diverse and large populations that receive many views, they have the potential to include views from certain populations of interest.

Decreased participant effort

Another important benefit of web-based verbal protocol methodology via online survey platform integration is the decrease in participant effort. There may be some initial hesitation from participants to interface with VOIS as a response option when it comes to recording themselves due to the uncommon nature of the task and the perception of decreased anonymity. However, this limitation may be more of a concern with some of the more traditional methods of in-lab vocal data collection, where vocal recording is done as a stand-alone method. It is possible that over time, having participants record their voices as part of a survey response will become more natural and raise less concern related to participant reidentification or loss of anonymity. Additionally, with the ease of using a cell phone as a means for participants to complete surveys and thus record their voices using the VOIS integration, participants may actually find this easier than typing in a response. Because cell phone usage is part vocal, participants may find that less effort is required when completing a survey on their phone that uses VOIS, especially if they already have their phone with them and are used to talking into it.

The alternative to generalizing verbal data to a web-based design is having participants type, which may not

provide the same richness in data and may be more labor intensive, especially on a cell phone. Web-based surveys are easily accessible via cell phone or personal computer and thus, VOIS will be as well. This allows for integration with existing equipment that participants already have and doesn't require any specialized hardware such as badges or sensors, like some verbal protocol studies. To the same point as avoiding a lab-based study, not only will researchers not have to be present in the lab, but participants can provide vocal data from home using VOIS's implementation.

Simple verification of respondent participation

Voice data collection serves a simple, yet extremely effective countermeasure against data fabrication. Recording audio during an in person study conducted by a research assistant can provide many benefits (Gomila et al., 2017; Harrison & Krauss, 2002). The issue of data fabrication has become increasingly concerning, as research assistants or surveyors may submit data that was not in fact collected, but rather completely or partially created for efficiency or incentive reasons. Recording the participant and surveyor's voice can be helpful for verifying that both individuals are following the study's protocol. Having all respondents record their audio, even briefly, can also be used to promote data integrity, by making it more difficult for research assistants or surveyors to fabricate data across many respondents, a phenomenon known as "curbstoning" (for a review see Hernandez et al., 2022).

Prevalence of online psychology studies

In addition to the abovementioned benefits, the overwhelming move in psychological research, especially within universities, to utilizing survey platforms for data collections lays a strong foundation for the adoption of the proposed method. Even despite the current incompatibility with the collection of vocal data, psychological researchers frequently reap the benefits of online survey platforms for data collection in a variety of academic studies. Because of the ability to collect convenience samples, which are becoming increasingly relevant and necessary for generalization in social sciences, online survey platforms are now more mainstream than ever in psychological research (Boas et al., 2018). This prior know-how that many researchers have of generating and catering surveys to their research via survey platforms makes them especially user-friendly when implementing new protocol options.

A survey of 750 university human research ethics committees (HRECs) in the United States also revealed that Internet research in general, involving online or web surveys, is the type of methodology most often reviewed (94% of respondents) in submitted studies (Buchanan & Hvizdak, 2009). This indicates the growing prevalence of online

survey methods in academic research. Not only is this prevalence found with all academic research in universities, but it is specifically prevalent in social science fields and related research as a whole (Van Selm & Jankowski, 2006).

While social science, psychology, and applied researchers seem more likely to use online survey-based platforms via the Internet, there is an apparent increasing trend in studies adopting this methodology. Today, survey software packages and online surveys services make online survey research much easier and faster (Wright, 2005). With this ease of use, online survey companies are continuously updating and revising their services to offer more up-to-date and user-friendly platforms. This demonstrates the increased desirability and subsequent adoption of these services, especially in social, psychological, and applied research. In the same survey study by Buchanan and Hvizdak (2009), the overwhelming majority (94%) of respondents to the survey stated that online survey research was the main type of Internet research reviewed. Those respondents of 750 university HRECs also indicated that they were typically reviewing 0–5 Internet-related research protocols per month. Of the online surveys reviewed, nearly all fell into the exempt category of review, indicating that the nature of the data was not overly sensitive, nor were the data from vulnerable populations being surveyed.

However, many of these studies that analyze trends in online survey usage, especially within psychological fields, are relatively dated. The sheer number of the studies published in social, psychological, and applied journals each year nearly requires a yearly update to appropriately reflect the trends in online methodology. To address this issue and illustrate exactly how prevalent online study methods are in published research, we examined every article published in the previous year (September 2019–September 2020), 5 years ago (September 2014–September 2015), and 10 years ago (September 2009–September 2010) from three journals that study social, psychological, and applied fields: the *Journal of Personality and Social Psychology* (JPSP), the *Journal of Applied Psychology* (JAP), and the *Journal of Consumer Psychology* (JCP). In total we coded 389 JPSP articles, 296 JAP articles, and 134 JCP articles. We coded each study within an article for whether it utilized online study methodologies as stated in the methods section. This analysis coded a total number of 2242 studies (NJPS = 1441, NJAP = 422, NJCP = 379). We summarize the results of this review in Table 1. Based on these analyses, we find that online study methods are increasingly more common over time in popular social and applied psychological journals. Additionally, all studies analyzed in the last year reflect that at least 50% of the studies use online study methods.

While it is relevant to demonstrate the prevalence of online studies in major psychological journal articles, it is even more relevant to the proposed method to demonstrate

those studies that use online survey platforms. Specifically, it is common for researchers in psychological and related fields to access large and diverse populations on Amazon's MTurk online survey platform. With its low cost, high recruitment speed, pretesting availability, and both exploratory and current event research designs, MTurk is a popular choice in academic research for advanced online survey design platforms (Boas et al., 2018). In a literature search within 20 top Industrial/Organizational psychology journals using the search terms "Mechanical Turk" and "MTurk", Cheung et al. (2017) shows that 99 empirical papers out of all the papers within those journals from the year 2016 used at least one MTurk sample. Additionally, there is a steady increase in papers using MTurk samples since 2012, as shown by the same study. Another study looking at response rates for online survey platforms shows that MTurk collected a sample that consisted of 581 US respondents for a study that included a posting of a task examining the relationship between job fit and employee attitudes (Kraiger et al., 2019). Although the most recent source showing the frequency of MTurk use is from 2019, we updated the estimated prevalence of MTurk methods among popular social and applied psychological journals, similar to as discussed above. Again, we examined every article published in the previous year (September 2019–September 2020), 5 years ago (September 2014–September 2015), and 10 years ago (September 2009–September 2010) from three journals that study social, psychological, and applied fields: JPSP, JAP, and JCP. In total we coded 389 JPSP articles, 296 JAP articles, and 134 JCP articles. This time, we coded each study within an article for whether it utilized MTurk as a means for collecting study data, as stated in the methods section. This analysis coded a total number of 2242 studies (NJPS = 1441, NJAP = 422, NJCP = 379). We summarize the results of this review in Table 1. Based on these analyses, the prevalence of studies that use MTurk as a means for online data collection only increases over time in the popular social and applied psychological journals analyzed. These analyses also demonstrate that MTurk is used at least 23% of the time and up to 44% in the total number of most recent studies assessed. Due to the usage of online studies, specifically via MTurk services in published studies becoming increasingly more common, our implementation focuses mainly on MTurk and self-hosted websites.

Despite the growing prevalence of online studies, researchers still currently struggle to apply verbal protocol methodology to more modern options for generalizable data collection, such as survey collection platforms. These collection platforms are typically accessed via the Internet as open-source third-party survey implementations. Survey collection platforms allow researchers to input survey questions that participants view in a sequential fashion. Using these platforms, researchers can further control the survey's

design by altering the underlying HTML and JavaScript code, if they so desire. Third-party survey platforms also allow participants to access the survey via cell phone, as a more efficient and user-friendly means to survey completion. Many of these collection platforms allow for large sample sizes because of their accessibility and simplicity for both the researcher and participants (Boas et al., 2018).

While third-party collection platforms offer a scalable alternative to in-lab studies, there are some notable drawbacks. These platforms can only utilize written survey designs and provide no way to integrate vocal protocol collection methods or anything besides survey answers as data. This limits the types of studies in which larger and more diverse sample sizes can be readily accessed. The above limitation is a pervasive problem for researchers using vocal protocol collection methods. These researchers are then left out of the trend of increasing online survey platform adoption and remain left to collect vocal data in a laboratory setting. The proposed method aims to address the limitations of currently available in-lab and online methods of collecting voice data by providing an open-source option for researchers to include a means to collect voice data in a survey platform.

Implementing the proposed method

The proposed method can be implemented across a variety of survey platforms. However, those platforms must meet the requirement that they allow researchers to supply their own HTML code within the survey. Given the importance of customizing the appearance of one's survey and the fundamental nature of HTML to all webpages, we believe it is unlikely that this aspect will ever be removed from any major platform. By offering a method that works on major survey platforms such as MTurk and Qualtrics, as well as self-hosted websites, we hope to enable a variety of uses and applications for researchers to facilitate in vocal data collection online.

Obtain File Transfer Protocol credentials

Voice recording data requires more file storage than standard plain-text data. Therefore, researchers seeking to implement the voice protocol method remotely need a cloud-based storage option for saving their collected data. The most prevalent Internet standard for uploading data to servers remotely is the File Transfer Protocol (FTP). FTP is a standard communication protocol that anyone can use to transfer their computer files to a server. FTP users authenticate themselves with a username and password and offers secure transmission that protects the username and password, and encrypts the content. Unlike other cloud storage options like Google

Drive or Dropbox, it is not tied to a particular company, and is not subject to change by a single entity as it is a designated "Internet Standard" first published in 1971 as a Request for Comments, published by the Internet Engineering Task Force (IETF), which decides on global Internet protocols.

Researchers may already have access to FTP-accessible storage via their university. If researchers are not given storage by their university, FTP-compatible storage can be purchased for typically less than \$10 monthly from a web hosting service. A web hosting service allows individuals to make their website accessible via the World Wide Web. These companies provide space on a server to be leased by clients. Typically, web hosting services offer free FTP access to their clients because it provides a simple, universal way to transfer large files.

The proposed method requires the researcher to have the FTP credentials available to save files to their hosting service. The required FTP credentials are (1) the FTP domain, typically <ftp.domain.com>, (2) the username, and (3) the password. These are provided by the hosting service in their FTP instructions. When the researcher knows these credentials, they can create an access token, which will inform the VOIS software where to save the voice data without exposing the researcher's FTP information to the survey participants.

Creating an access token

Before implementing the VOIS protocol, researchers should create an access token, which will inform the system where to save the recorded voice files. The access token is linked to one's FTP account on a hosting domain. When a user provides the access token in the collection code, the system will use the associated credential to save the data. Therefore, the temporary voice file created by a user is sent directly to their hosting server, and no voice data is stored elsewhere.

To create an access token, a researcher simply goes to the token generation website in their web browser (<https://psych.x10host.com/audio/generatetoken.html>). They then enter their FTP server URL, their FTP username, and their FTP password (Fig. 1). Optionally, researchers can indicate a subdirectory if they are not able to save to the main folder or have another preferred saving destination. If a researcher desires to save in the main directory, then they can leave the subdirectory blank. The researcher also selects whether they want to encrypt the respondent's IP address so that the raw IP address is not saved in the filename. To match IP addresses from a survey to the filenames saved by VOIS, the researcher needs to encrypt the survey IP addresses using the following site:

<https://psych.x10host.com/audio/encodeipaddress.html>

After entering the information, the researcher clicks the "Generate Token" button, and the unique access token is created. The access token is a string of 30 randomly generated numbers and letters. With 26 possible letters and 10 possible

Fig. 1 Testing an access token associated with a researcher’s FTP credentials

digits, there are 36^{30} = forty-eight quattuordecillion eight hundred seventy-three tredecillion six hundred seventy-eight duodecillion possible combinations. Therefore, brute force guessing or accidentally using another person’s token is extremely unlikely. This token is pasted into the provided code in the following sections. If a researcher provides invalid FTP credentials, the website will inform them to correct the information and try again.

Verifying the access token

After receiving an access token, researchers may want to verify that they can save voice recordings to their FTP storage. We provide a website where a researcher can test VOIS, and see the resulting sound file saved in their storage (<https://psych.x10host.com/audio/testtoken.html>). Researchers visit the testing site after receiving their token. They then paste their access token in the input field. They can then record an audio clip using their computer (Fig. 2 left panel). After recording the audio clip, they click on the

“Test Your Access Token” button to send the voice recording to the FTP server associated with that token. Researchers are then instructed to check on their server to verify that the new file is there. The filename contains two parts: a string equivalent to the researcher’s IP address and “_test-question.wav”. After verifying that the file is saved in the FTP server, the researcher is ready to collect voice data online (Fig. 2 right panel).

Collecting voice data

Collecting a single recording To collect voice data, researchers simply need to paste the following code in their survey’s HTML editor. The researcher must replace the text that says, “ACCESS-TOKEN-GOES-HERE” with their own access token that they received from the first step. The researcher would also change “Voice Protocol Question 1 Text Goes Here” with their own prompt for the participants. In the following paragraph, each of the eight lines of code is described in further detail in terms of its purpose.

Fig. 2 The panel on the left shows the web page where researchers can paste their access token and record an audio clip to verify that voice recordings are correctly saved to their FTP storage. The panel

on the right shows the web page where researchers are then prompted to check their server for their recorded test audio file

```

<p>Voice Protocol Question 1 Text Goes Here</p>
<p><button type="button" id="record1">Start recording ◉</button></p>
<p>Duration: <span id="duration1">0ms</span></p>
<p><audio id="player1" controls></audio></p>
<p style="border-style:solid;text-align:center" id="uploadstatus1">No Audio Recorded</p>
<input type="hidden" id="voicedata1" name="voicedata1" value = "">
<input type="hidden" id="usertoken" value = "ACCESS-TOKEN-GOES-HERE">
<script src="https://psych.x10host.com/audio/vois.js"></script>

```

In the above code, the first line displays the question text or prompt to the participant. This prompt would be an instruction of what to record, such as the participants' thought process during a problem or a summary of one's experience. The second line displays a button that the participant clicks to begin the recording process. After clicking it, it changes into a "Stop recording" button, which the participant clicks to stop the program from recording any more audio. The third line displays how many milliseconds have elapsed since the beginning of the record. Researchers may desire to impose a suggested minimum or maximum length of response, and this duration feature can provide participants with the information needed to adhere to those guidelines. The fourth line displays an audio player so that the participant can listen to their recording after they press the "Stop Recording" button. The fifth line confirms to the participant whether they have recorded any audio. After pressing the "Stop recording" button, the status changes to "Uploading recording." After the recording finishes uploading to the researcher's FTP server, the status changes to "Audio Upload Complete." The sixth line is a hidden input that is not visible to the participant. It stores the audio data temporarily while it is being uploaded to the researcher's FTP server. The seventh line contains the researcher's access token information. This information is stored in a hidden field and is accessed by the VOIS software when uploading the audio data. The eighth line contains a link to the script that processes the recorded data and uploads it to the

appropriate server. The line should go at the very end of the survey. If there are multiple verbal prompts, then it should only be used once, after the final prompt.

Collecting multiple recordings If the researcher wants to collect multiple voice prompts, then the above code is simply copy–pasted multiple times in the survey, changing the number "1" to the number "2" in lines 2–6. That is, when the code indicates: id = "record1", the code should be changed to id = "record2" When the code says, id = "duration1", the code should be changed to id = "duration2". In total, six instances should be altered when creating a new question. Researchers can create up to 1000 questions using VOIS. As mentioned in the previous section, if using multiple questions, then line 8, containing the link to the vois.js script, should only appear once in the code, after the last question.

Example 1: Collecting voice data on MTurk

MTurk, the popular survey collection platform, creates surveys primarily via HTML elements. Therefore, researchers can paste the code provided above into its HTML editor, and simply need to change the access token to their own (Fig. 3). Researchers using MTurk should create a blank survey. Then in the "Design Layout" view, paste the code between the "crowd-form" tags. MTurk allows the researcher to preview the survey to verify the appearance (Fig. 4). When the code is pasted, the study will save the collected audio data to the researcher's FTP account with a filename of "[participant's IP address] question[question numbers].wav". Because MTurk automatically collects a participant's IP address, the answers to a survey can be linked to a participant's audio files.

Edit Project

Use the HTML editor below to design the layout of your task, which Workers will complete.

You can use any HTML, CSS, or JavaScript to customize your layout.

Or use [Crowd HTML Elements](#) to simply design your layout.

This layout is common for all of the tasks created with this project. You can define variables (eg $\$(variable_name)$) for data that will vary from task to task based on your CSV input file ([Learn more](#)).

```

1 <script src="https://assets.crowd.aws/crowd-html-elements.js"></script>
2 <crowd-form answer-format="flatten-objects">
3
4 <p>Question 1 Text Goes Here</p>
5
6 <p><button type="button" id="record1">Start recording ◉</button></p>
7 <p>Duration: <span id="duration1">0ms</span></p>
8 <p><audio id="player1" controls></audio></p>
9 <p style="width:200px;border-style:solid;text-align:center;" id="uploadstatus1">No Audio Recorded</p>
10 <input type="hidden" id="voicedata1" name="voicedata1" value = ""></input>
11 <input type="hidden" id="usertoken" value = "1524fc66ceae11267d8040c28ec84bbe123a1e3b4717498059ac8d1cf92">
12
13 <script src="https://psych.x10host.com/audio/vois.js"></script>
14
15 </crowd-form>
16

```

Fig. 3 Editing an MTurk survey to include the VOIS code within the design

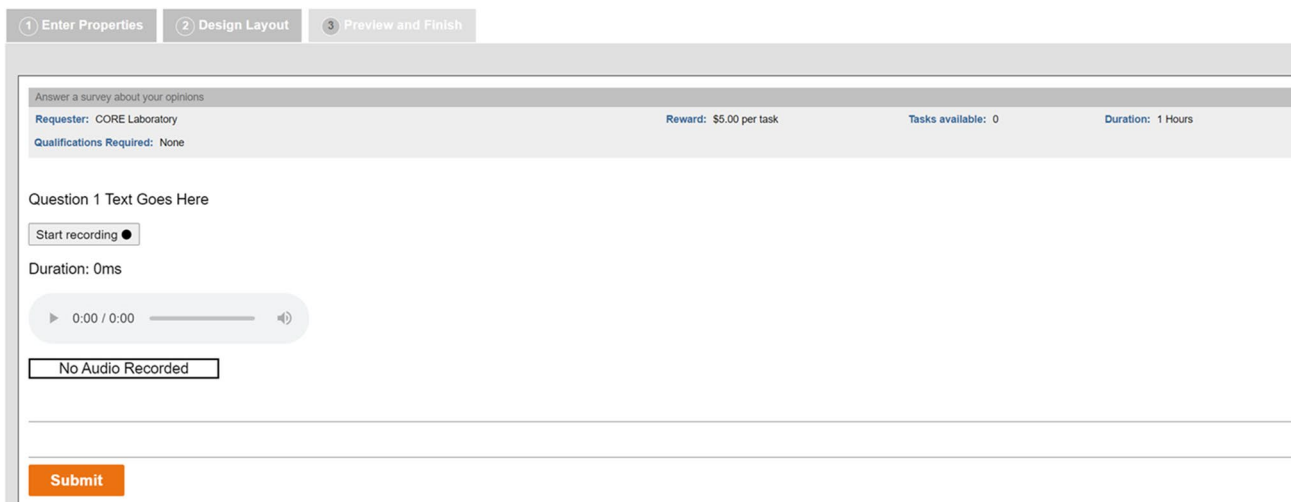


Fig. 4 Example of the final appearance of an MTurk survey that includes the VOIS code

Example 2: Collecting voice data on Qualtrics

Similar to MTurk, the popular survey collection platform, Qualtrics, allows researchers to edit the HTML of a given question and link JavaScript files. To implement the VOIS program, researchers should create a “text” question, without any response options. Then in the question, click on “HTML view” and paste the code for a single voice recording prompt (Fig. 5).

The researcher’s user token should be entered in the code. Qualtrics allows researchers to preview the survey, and the voice recording element should be usable within the preview view (Fig. 6).

For additional voice prompts, the researcher should create additional questions, and enter the code again, changing the 1’s to 2’s in the code for the second question, as described in the section on multiple voice recordings. For multiple recordings, it is important to only include the following part of the code only once, after the last question:

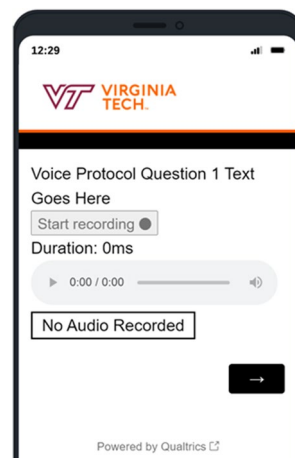


Fig. 6 Example of the final appearance of a Qualtrics survey that includes the VOIS code

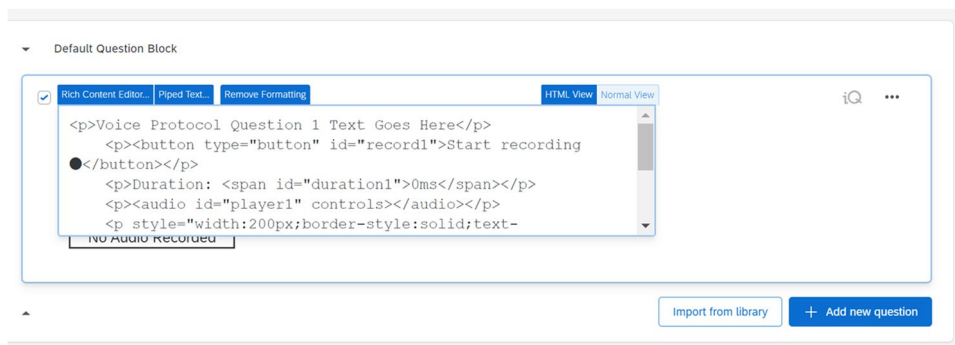


Fig. 5 Editing a Qualtrics survey to include the VOIS code within the design

```
<script src="https://psych.x10host.com/audio/vois.js"></script>
```

If the researcher desires to have VOIS save the audio recording as MP3s, they should use the following line instead:

```
<script src="https://psych.x10host.com/audio/voismp3.js"></script>
```

Example 3: Collecting voice data on a self-hosted website

Some researchers use self-hosted websites, which is a website where a researcher hosts and designs all of the questions using HTML and records the data using a server-side language (e.g., PHP, PERL, ASP). These websites are also compatible with the proposed method, and follow the same process. The researcher would create an HTML page, beginning with the `<html>` and `</html>` tags. Within the HTML tags the `<body>` `</body>` tags are inserted, which indicate the visible content of the page. Within the body tags would be the `<form>` and `</form>` tags that display the survey form elements. Within the form would go the survey question elements. That section is where the VOIS code would be copied and pasted. In the figure below, we show a survey with three text inputs and two voice recording prompts (Fig. 7). After pasting the code and creating the other survey elements, the HTML file can be hosted as usual with the

elements visible to the researcher (Fig. 8). The researcher would collect the participant's IP address using the script that processes the rest of the survey data collection, such as the PHP script.

Conducting a study with VOIS

To conduct an online study using VOIS, researchers would follow the instructions of their typical verbal protocol method. The only modification would be to direct participants themselves when to press the record button and when to end it. The code has a built-in audio player, which allows participants to confirm successful recording. Therefore, participants should be instructed to verify that the data recorded properly, using the playback option, before submitting or moving to the next page.

Researchers must collect the participant's IP address during the study, which is typically done by default on major survey platforms. Researchers are already advised to collect IP addresses, when possible, because of their ability to minimize repeat responses and to ensure the participant's location is within a given region. It is necessary to collect the IP address because the filenames of every recording begin with the IP address and contain the question number. This naming system allows researchers to match a participant's voice recording to their nonverbal responses on the survey. Therefore, it is essential that researchers verify that

```

1 <html>
2 <head>
3 <meta http-equiv="content-type" content="text/html; charset=utf-8">
4
5 </head>
6 <body>
7
8 <h1>Example Survey With Audio</h1>
9
10 <form action="savesurvey.php" method="post">
11
12 <p>Participant ID: <input id="participant_id" name="participant_id" value = ""></p>
13 <p>Participant Age: <input id="participant_age" name="participant_id" value = ""></p>
14 <p>Participant Gender: <input id="participant_gender" name="participant_id" value = ""></p>
15 <hr>
16
17 <p>Voice Protocol Question 1 Text Goes Here</p>
18 <p><button type="button" id="record1">Start recording ●</button></p>
19 <p>Duration: <span id="duration1">0ms</span></p>
20 <p><audio id="player1" controls></audio></p>
21 <p style="width:200px;border-style:solid;text-align:center;" id="uploadstatus1">No Audio Recorded</p>
22 <input type="hidden" id="voicedata1" name="voicedata1" value = "">
23 <input type="hidden" id="usertoken" value = "aaaaaaaaaaaaaaaaaaaa">
24
25 <hr>
26
27 <p>Voice Protocol Question 2 Text Goes Here</p>
28 <p><button type="button" id="record2">Start recording ●</button></p>
29 <p>Duration: <span id="duration2">0ms</span></p>
30 <p><audio id="player2" controls></audio></p>
31 <p style="width:200px;border-style:solid;text-align:center;" id="uploadstatus2">No Audio Recorded</p>
32 <input type="hidden" id="voicedata2" name="voicedata1" value = "">
33 <input type="hidden" id="usertoken" value = "aaaaaaaaaaaaaaaaaaaa">
34
35 <hr>
36
37 <p><input type="submit" ></p>
38 </form>
39 <script src="https://psych.x10host.com/audio/vois.js"></script>
40 </body>
41 </html>

```

Fig. 7 Editing a self-hosted HTML survey to include the VOIS code within the design

Example Survey With Audio

Participant ID:

Participant Age:

Participant Gender:

Voice Protocol Question 1 Text Goes Here

Start recording ●

Duration: 0ms

▶ 0:00 / 0:00 🔊

No Audio Recorded

Voice Protocol Question 2 Text Goes Here

Start recording ●

Duration: 0ms

▶ 0:00 / 0:00 🔊

No Audio Recorded

Submit

Fig. 8 Example of the final appearance of a self-hosted HTML survey that includes the VOIS code

the survey collecting the audio recording has IP address collection enabled.

Downloading audio data

After conducting the study, the researcher will have a collection of audio files saved on their server. The researcher can either leave the files on the server or download them all to their local computer. Each audio file begins with the participant's IP address. Therefore, researchers coding the content of the audio file can easily associate the file to the participant's other answers by sorting the survey file by IP address.

While there may be issues that could emerge with IP address stability in terms of identifying participants via a link with their IP address and voice data, this issue of stability is only of concern between sessions (Dennis et al., 2020). Stability of an IP address within a session is stable. Therefore, as long as researchers are trying to connect data within a session, there should be no issues in guaranteeing IP address stability. If a researcher provides a token for the participant to use in a single session, then their associated IP address could possibly be linked. However, to provide security for the respondent, VOIS utilizes a cryptographic hash of the participant's IP address, which is still unique for a single respondent, but reveals no information about their

specific IP address. The cryptographic hash of the IP address does so by providing a unique identifier without any overlap between the provided hash and password.

Discussion

In this article, we have presented VOIS, a collection of code written in HTML and JavaScript for collecting vocal data online. VOIS provides a tool for researchers that was previously unavailable and allows for vocal protocol methods to be implemented in third-party survey platforms. No current widely available methods for online vocal data collection via third-party survey platforms exists. Therefore, we provide the most user-friendly and accessible means to gather vocal data without sacrificing any data richness. VOIS offers a simple solution to allow researchers to collect vocal data outside of the lab, providing access to larger sample sizes. Therefore, this web-based verbal protocol implementation provides a more scalable method of collection and a less effortful participant and researcher experience. By introducing VOIS, we provide a previously unknown need for researchers interested in utilizing verbal protocol methods.

VOIS can be easily implemented into many third-party survey platforms such as MTurk and Qualtrics (or platforms like Prolific that link to outside surveys like Qualtrics), which allow for the addition of custom JavaScript and HTML code. VOIS uses a standard code across all platforms that the researcher can copy and paste into the desired survey platform; therefore, no prior programming experience is needed to benefit from this method.

Survey integration suggestions²

Outside of the suggested implementation of VOIS as an integration to existing surveys maintained on survey platforms or as a central component of a new study located on a survey platform, researchers may find additional and more creative uses for VOIS software. Due to the ease of use and accessibility of VOIS, researchers from many fields can now supplement their research with vocal data collection, especially using survey platforms they are already familiar with. VOIS allows researchers to integrate the collection of vocal data with other question types, which can create opportunities for unique usages of the proposed VOIS software. We provide several suggestions to researchers as to a few ways verbal data may be relevantly incorporated into their studies.

² We want to thank one of our manuscript reviewers, Dr. Aaron J. Moss, for the insightful suggestions of additional ways to incorporate our proposed method.

Implementing VOIS into a survey online could allow researchers to replace long written tasks with think-aloud tasks by asking participants to record their responses instead of typing them out. This could potentially increase the response rate for the question or manipulation due to decreased participant effort. Additionally, participants could complete the question more efficiently and the researcher could then include more survey items without too much additional participant effort or worrying about survey fatigue. Similarly, VOIS could be used in survey manipulations to gain insight into participant engagement and involvement. For example, researchers could ask participants to vocally respond to a survey manipulation and get an idea of engagement from the length and thoughtfulness of the participant's response.

Another way of using VOIS in existing survey platforms is as a comprehension check. This could be useful to researchers across a wide range of disciplines given common issues about data quality and response rates in online studies (Deutskens et al., 2004; Evans & Mathur, 2005). To do so, researchers could include an item that incorporates VOIS and asks participants to repeat back a set of survey instructions to verify understanding as a means of quality check. If the participant seems to vocalize instructions that seem different or seem that they misunderstood the methods, it could be a means to exclude their responses. Additionally, it may seem more natural for participants to verbally report back their understanding or interpretation of instructions as opposed to typing them out, where it may be tempting to just copy the provided instructions instead.

VOIS can also be useful to a wide variety of researchers as an additional think-aloud question at the end of a study to ask participants for feedback on the overall survey experience. Participants may not want to go through the effort or feel that it is not relevant to reach out to the researchers via email to express any issues with their survey experience. However, if this was included in the survey as an easy-to-respond-to vocal response question, participants might be more motivated to provide unique insight that could help researchers adjust their surveys for a better participant experience in the future. Using VOIS in this manner can help avoid researchers from continuing to include confusing tasks or having participants misinterpret aspects of the study by providing insight that the researcher may not otherwise have access to.

VOIS can also provide a previously unavailable way to measure creativity within a survey platform design or an already created survey. Traditional measures of creativity can be open-ended and thus hard to interpret or assess. Commonly used measures of creativity may also be dated and incompatible with implementation into an online survey design (Guilford, 1967). However, those creativity measures that are currently utilized in an online format can only be enhanced by VOIS. For example, creativity

speed tests (versus power tests) that measure how many items or responses a participant knows or comes up with in a certain amount of time can be more easily and quickly completed in a single vocal recording. This ensures that participants who are slower at typing are not at a disadvantage and thus provides the researchers with more variance in participant responses. In addition to creativity, vocal protocol via VOIS could be substituted for commonly typed measures like verbal reasoning. This substitution can assist researchers in a general methodological way by enabling them to cut participants off at a certain length of recording in order to keep all recorded data consistent and the same length, avoiding any unnecessary data preprocessing in the form of trimming audio. Researchers would also be able to control the exact amount of time participants spend on a task by only allowing participants a certain amount of recording time.

Another example of a construct that can be measured more effectively by utilizing VOIS is a participant's certainty. Certainty as a self-report measure can be easy to lie on or may not capture nuances of the true nature of a participant's certainty of a response. Specifically, VOIS allows for researchers to measure certainty through a participant's vocal behavior. These behaviors can capture specifics of a participant's response like their hesitation through "ums" or time spent gathering their thoughts, as well as their tonal porosity. Including these vocal behaviors in measuring certainty in an online study can potentially redefine how response certainty is measured.

Some additional replacements using VOIS to capture vocal recordings instead of typed or selected input include usage with situational judgment tests (SJTs). SJTs are used across a wide range of disciplines to measure various different constructs using participants' reported potential behavioral responses (Corstjens et al., 2017; Patterson et al., 2012). In responding to SJTs, participants are usually asked to evaluate several potential behavioral responses to a situation, all which vary in potential degree of plausibility or appropriateness in effectiveness (Webster et al., 2020). Having a participant respond vocally to SJTs can allow for more response variance as well as ease of responding and potentially more reflective survey responses of actual behavioral responses. In general, using VOIS as a means to capture vocal recoding in an online survey can allow for the measurement of constructs more effectively than previously available in online survey platforms. This can include the measurement of constructs such as microaggressions or covert discrimination in order to capture more authenticity in participants' responses. Constructs such as competency, teaching ability, warmth, or fluency can also be more effectively captured using VOIS implemented in an online survey. These types of constructs that are typically measured via self-report are better captured using vocal

recording due to the additional data of tone, vocal porosity, hesitation, or speed found in recorded speech, enabling a more subjective measurement of such constructs.

Finally, VOIS could be used as a means of data integrity validation for online surveys conducted using survey platforms—specifically, to deter instances of curbstoning, or survey data fabrication by third parties. Due to the difficulty of fabricating large amounts of participants' vocal recordings, potential fabricators would be required to actually put in more work to fabricate responses if VOIS was included as a response option. Additionally, alternative proposed methods using participants' responses to innocuous questions and matching response distributions to the expected known distribution of typical, non-fabricated responses can be facilitated using VOIS (Hernandez et al., 2021). The proposed method in Hernandez et al. (2021) suggests using a combination of common questions that participants can easily respond to in order to check for deviations from expected statistical distributions. These include asking participants for non-identifying, but easy to answer questions such as their address number, birthdate, or the last four digits of their phone number. If answering these questions is done via participant vocal response, this method of data integrity validation can be made more robust against instances of curbstoning through more accurate and efficient data authentication.

Limitations

Requiring an FTP hosting server

The current method records audio and stores the data as a WAV file on the researcher's FTP server. This method is intended to coincide with data retention plans that often require the data to be stored on password-protected systems.

Not all researchers may currently have an FTP hosting service. However, most researchers are able to obtain access to their own FTP hosting for free or at a minimal cost. Many sites exist that provide free FTP hosting. Often this hosting is limited in terms of total file storage size or upload size. Researchers who require more flexibility in file sizes can use paid hosting plans, which are typically offered for a monthly cost around US minimum wage. This expense is therefore roughly the cost of a single participant, and only needs to be used for the length of the data collection. Additionally, the most popular router brands (i.e., Netgear, Linksys, and Cisco) all provide the ability to turn the router into an FTP server. By turning this option on in the router settings page, researchers will receive the login URL needed to generate an access token (Fig. 9). This option, while free, may require some technical expertise and may not always apply, as it requires that the researcher has a static IP address (which most residential Internet accounts do not).

Routers that have ReadySHARE capabilities and are enabled as an FTP server store transferred files on a USB drive the researcher plugs into the router. Even without a router, software, such as FileZilla Server, IndiFTP, and Xlight are freely available, which turn any computer into an FTP server (though the computer should be one that remains on at all times and has a static IP address assigned to it). We recommend nontechnical researchers obtain a low-cost web host (e.g., X10premium.com) as it will be quicker and more reliable than a personally run solution.

Increased data file size

Audio data provides a richer view of a response compared to plain text, with the disadvantage of requiring more storage space to record. The proposed method saves the audio data as an uncompressed WAV format, offering maximum fidelity. As

The screenshot shows the Netgear WNDR3400v3 router's web interface. The 'BASIC' tab is selected, and the 'ReadySHARE' option is highlighted in the left sidebar. The main content area is titled 'USB Storage (Advanced Settings)'. It includes a table with columns for 'Enable', 'Access Method', 'Link', 'Port', and 'Admin Password Protection'. The 'FTP (via internet)' row is checked, with a port of 21 and 'Admin Password Protection' enabled. A warning message at the bottom states: 'Warning: We strongly recommend that you enable Admin Password Protection to secure your ReadySHARE data. If you did not change the default admin password, we also recommend that you change it to secure your network.'

Enable	Access Method	Link	Port	Admin Password Protection
<input type="checkbox"/>	Network Connection	\\readyshare	-	<input type="checkbox"/>
<input type="checkbox"/>	HTTP	http://readyshare.routerlogin.net/shares	80	<input type="checkbox"/>
<input type="checkbox"/>	HTTPS (via internet)	https://73.152.214.167/shares	443	<input checked="" type="checkbox"/>
<input type="checkbox"/>	FTP	ftp://readyshare.routerlogin.net/shares	21	<input type="checkbox"/>
<input checked="" type="checkbox"/>	FTP (via internet)	ftp://73.152.214.167/shares	21	<input checked="" type="checkbox"/>

Fig. 9 Enabling a household router with ReadySHARE to serve as an FTP server

a result, some data files will be noticeably larger. A minute of audio is approximately 5 megabytes. When collected on 200 participants (the sample size needed to have 80% power for detecting effect sizes of $r = .20$ at an alpha level of 5%), the survey size would be roughly 1 gigabyte. Therefore, researchers need to have at least 1 gigabyte of storage space on their hard drives per prompt, if a prompt takes 1 minute to record. To address this limitation, we also offer a version of VOIS that saves MP3 files instead of uncompressed WAV files. MP3 files use lossy compression to reduce file size, while retaining most of the psychoacoustic fidelity. When using a bitrate of 160 kilobytes per second, the audio is approximately 1 MB per minute. All that researchers need to do to save MP3s is to change the link from `vois.js` to `voismp3.js` in the `src` section of the script tag.

For example, instead of writing at the bottom of the page:

```
<script src="https://psych.x10host.com/audio/vois.js"></script>
```

The researcher would write:

```
<script src="https://psych.x10host.com/audio/voismp3.js"></script>
```

Additional effort for physiological data

Our implementation can only enable the additional collection of audio data. While the implementation can easily be supplemented by survey data in the same platform, several studies use other means of data collection in addition to verbal protocol methods. Therefore, researchers are limited to only collecting types of data supported by the survey platform's available question types and methodology; in addition to vocal data supported via VOIS implementation. If supplementary data is desired, researchers would have to couple it with additional means of collection. Supplementary data could include more physiological-based measures that may be of interest to researchers collecting vocal data. For example, researchers may want to supplement vocal data with the collection of eye movement, blood pressure, cortisol levels, heart rate, or other physiological measures that are not easily integrated into existing survey platforms.

These additional collection means that are not supported by VOIS may not be as easily applied to a survey platform and can be a potential drawback compared to more traditional verbal protocol methods.

Post-processing

Lastly, our implementation may provide difficulties to researchers in the usage of provided data. With the

implementation allowing for larger sample sizes and thus more vocal data, there may be issues in the analysis of a large amount of vocal data. Additionally, vocal data is coded in specific ways that can be labor- and time-intensive. Combined with the larger amount of vocal data, this could provide a potential drawback to researchers. However, there are many techniques of handling large amounts of data that exist. Deep neural networks offer speech-to-text analysis that rivals human transcription. These neural networks are provided as free and open-source software that allows researchers to provide the audio file and receive the transcribed text in seconds. Additionally, researchers can use this text-to-speech software without any programming ability.

Conclusion

To summarize, this article has presented VOIS, a method for collecting verbal protocol data via third-party survey platforms. The implementation of VOIS in the presented survey platforms allows researchers to collect vocal data in a more user-friendly and localized manner. Researchers can simply copy and paste the HTML code into the desired survey platform and follow the provided step-by-step instructions within the manuscript to adapt the protocol as needed. With the movement to more work being done remotely and the inability to currently adapt some in-lab study procedures, such as vocal recording protocols, to an at-home environment, our proposed method provides an advantage for researchers across fields.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13428-022-02045-6>.

Authors' contributions I.H. conceptualized the presented idea. T.R. developed the theoretical foundation and wrote most of the manuscript with support from I.H. I.H. developed the software code. T.R. created the tables and I.H. created the figures for visualization purposes. T.R. provided oversight and planning of the manuscript with assistance from I.H. T.R. and I.H. contributed to the final revisions and edits of the manuscript.

Data availability The materials generated and/or analyzed during the current study are available at the following Open Science Repository link: https://osf.io/msj7h/?view_only=52bd8769dcf54df7b2eaaa30b4927652

Code availability The code generated for this article is available at the following Open Science Repository link: https://osf.io/msj7h/?view_only=52bd8769dcf54df7b2eaaa30b4927652

Declarations

Conflicts of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval This is a code-based study where no human subjects were included. The Virginia Polytechnic Institute and State Univer-

sity Human Research Protection Program confirmed that no ethical approval is required.

Consent to participate This is a code-based study where no human subjects were included. Therefore, no informed consent to participate is necessary.

Consent for publication This is a code-based study where no human subjects were included. Therefore, no consent for publication is necessary.

References

- Atman, C. J., & Bursic, K. M. (1998). Verbal protocol analysis as a method to document engineering student design processes. *Journal of Engineering Education*, 87(2), 121–132. <https://doi.org/10.1002/j.2168-9830.1998.tb00332.x>
- Austin, J., & Delaney, P. F. (1998). Protocol analysis as a tool for behavior analysis. *The Analysis of Verbal Behavior*, 15(1), 41–56. <https://doi.org/10.1007/BF03392922>
- Biggs, S. F., Rosman, A. J., & Sergenian, G. K. (1993). Methodological issues in judgment and decision-making research: Concurrent verbal protocol validity and simultaneous traces of process. *Journal of Behavioral Decision Making*, 6(3), 187–206. <https://doi.org/10.1002/bdm.3960060303>
- Boas, T. C., Christenson, D. P., & Glick, D. M. (2018). Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics. *Political Science Research and Methods*, 1–19. <https://doi.org/10.1017/psrm.2018.28>
- Brandon, D. M., Long, J. H., Loraas, T. M., Mueller-Phillips, J., & Vansant, B. (2014). Online instrument delivery and participant recruitment services: Emerging opportunities for behavioral accounting research. *Behavioral Research in Accounting*, 26(1), 1–23. <https://doi.org/10.2308/bria-50651>
- Buchanan, E. A., & Hvizdak, E. E. (2009). Online survey tools: Ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics*, 4(2), 37–48. <https://doi.org/10.1525/jer.2009.4.2.37>
- Cauldron Science. (2016, October 1). *Gorilla: Web audio zone*. Retrieved December 2021, from <https://support.gorilla.sc/support/reference/task-builder-zones#webaudio>
- Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education: A Journal of Educational Research and Practice*, 12(2). <https://doi.org/10.26522/brocked.v12i2.38>
- Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. (2017). Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business and Psychology*, 32(4), 347–361. <https://doi.org/10.1007/s10869-016-9458-5>
- Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271–315. https://doi.org/10.1207/s15327809jls0603_1
- Cohen, D., Casner, S., & Forgie, J. W. (1981). A Network Voice Protocol (NVP-II). *University of Southern California/Information Sciences Institute*, 71. <https://www.rfc-editor.org/rfc/rfc741.html>
- Corstjens, J., Lievens, F., & Krumm, S. (2017). Situational judgement tests for selection. *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention*, 226–246. <https://doi.org/10.1002/9781118972472.ch11>
- Crutcher, R. J. (1994). Telling what we know: The use of verbal report methodologies in psychological research. *Psychological Science*, 5(5), 241–241. <https://doi.org/10.1111/j.1467-9280.1994.tb00619.x>
- Deffner, G., Snyder, H. L., Bittner Jr, A. C., Rhenius, D., & Sanderson, P. M. (1990, October). Verbal protocols as a research tool in human factors: Panel discussion. In: *Proceedings of the Human Factors Society Annual Meeting*, 34(16), 1145–1147. SAGE Publications. <https://doi.org/10.1177/154193129003401718>
- Dennis, S. A., Goodson, B. M., & Pearson, C. A. (2020). Online worker fraud and evolving threats to the integrity of MTurk data: A discussion of virtual private servers and the limitations of IP-based screening procedures. *Behavioral Research in Accounting*, 32(1), 119–134. <https://doi.org/10.2308/bria-18-044>
- Deutskens, E., De Ruyter, K., Wetzels, M., & Oosterveld, P. (2004). Response rate and response quality of internet-based surveys: An experimental study. *Marketing Letters*, 15(1), 21–36. <https://doi.org/10.1023/B:MARK.0000021968.86465.00>
- Diamond, M. (2016). Recorderjs, GitHub repository. Retrieved December 2021, from <https://github.com/mattdiamond/Recorderjs>
- Dugger, T. (2018). Kitura-Test, GitHub repository. Retrieved December 2021, from <https://github.com/troyibm/kitura-test>
- Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. *The Cambridge Handbook of Expertise and Expert Performance*, 223–241. <https://doi.org/10.1017/CBO9780511816796.013>
- Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, 5(3), 178–186. https://doi.org/10.1207/s15327884mca0503_3
- Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet Research*, 15(2), 195–219. <https://doi.org/10.1108/10662240510590360>
- Gomila, R., Littman, R., Blair, G., & Paluck, E. L. (2017). The audio check: A method for improving data quality and detecting data fabrication. *Social Psychological and Personality Science*, 8(4), 424–433. <https://doi.org/10.1177/1948550617691101>
- Goode, B. (2002, September). Voice over internet protocol (VoIP). In: *Proceedings of the IEEE*, 90(9), 1495–1517. <https://doi.org/10.1109/JPROC.2002.802005>
- Goodman, J. I., Brady, M. P., Duffy, M. L., Scott, J., & Pollard, N. E. (2008). The effects of “bug-in-ear” supervision on special education teachers' delivery of learn units. *Focus on autism and other developmental disabilities*, 23(4), 207–216. <https://doi.org/10.1177/1088357608324713>
- Gouveia, R., & Karapanos, E. (2013, April). Footprint tracker: supporting diary studies with lifelogging. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2921–2930. ACM. <https://doi.org/10.1145/2470654.2481405>
- Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, 1(1), 3–14. <https://doi.org/10.1002/j.2162-6057.1967.tb00002.x>
- Harrison, D. E., & Krauss, S. I. (2002). Interviewer cheating: Implications for research on entrepreneurship in Africa. *Journal of Developmental Entrepreneurship*, 7(3), 319.
- Hernandez, I., Ristow, T., Hauenstein, M. (2022). Curbing curbstoning: Distributional methods to detect data fabrication. *Psychological Methods*. <https://doi.org/10.1037/met0000403>
- Hoffman, K. A., Aitken, L. M., & Duffield, C. (2009). A comparison of novice and expert nurses' cue collection during clinical decision-making: Verbal protocol analysis. *International Journal of Nursing Studies*, 46(10), 1335–1344. <https://doi.org/10.1016/j.ijnurstu.2009.04.001>
- Hughes, J., & Parkes, S. (2003). Trends in the use of verbal protocol analysis in software engineering research. *Behaviour & Information Technology*, 22(2), 127–140. <https://doi.org/10.1080/0144929031000081341>

- Isenberg, D. J. (1986). Thinking and managing: A verbal protocol analysis of managerial problem solving. *Academy of Management Journal*, 29(4), 775–788. <https://doi.org/10.5465/255944>
- Jones, W. S., Cotton, T., & Holland, R. V. (2000). *U.S. Patent No. 6,141,341*. U.S. Patent and Trademark Office.
- Kasper, G. (1998). *Analysing verbal protocols*. *Tesol Quarterly*, 32(2), 358–362. <https://doi.org/10.2307/3587591>
- Keromytis, A. D. (2009, December). A survey of voice over IP security research. In: *Proceedings of the International Conference on Information Systems Security*, 1–17. Springer. https://doi.org/10.1007/978-3-642-10772-6_1
- Krahmer, E., & Ummelen, N. (2004). Thinking about thinking aloud: A comparison of two verbal protocols for usability testing. *IEEE Transactions on Professional Communication*, 47(2), 105–117. <https://doi.org/10.1109/TPC.2004.828205>
- Kraiger, K., Sanchez, D. R., & McGonagle, A. K. (2019). What's in a sample? Comparison of effect size replication and response quality across student, MTurk, and Qualtrics samples. https://www.sgu.ru/sites/default/files/samples_paper.pdf
- Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. *Health Psychology*, 17(1), 48. <https://doi.org/10.1037/0278-6133.17.1.48>
- Mathur, M. B., & Reichling, D. B. (2019). Open-source software for mouse-tracking in Qualtrics to measure category competition. *Behavior Research Methods*, 51(5), 1–11. <https://doi.org/10.3758/s13428-019-01258-6>
- Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, & Computers*, 33(4), 517–523. <https://doi.org/10.3758/BF03195410>
- Miyane, Y. (2016). *WebAudioRecorder.js*, GitHub repository. Retrieved December 2021, from <https://github.com/higuma/web-audio-recorder-js>
- OSWeb. (2020). *Running experiments online*. Retrieved December 2021, from <https://osdoc.cogsci.nl/3.2/manual/osweb/#jatos>
- Patterson, F., Ashworth, V., Zibarras, L., Coan, P., Kerrin, M., & O'Neill, P. (2012). Evaluations of situational judgement tests to assess non-academic attributes in selection. *Medical Education*, 46(9), 850–868. <https://doi.org/10.1111/j.1365-2923.2012.04336.x>
- Phonic. (2019). *Phonic documentation*. Retrieved December 2021, from <https://docs.phonic.ai/>
- Pipe Services S.R.L. (2015). *Video and audio recording clients and infrastructure*. Retrieved December 2021, from <https://addpipe.com/>
- PsychoJS. (2021). *GitHub repository*. Retrieved December 2021, from <https://github.com/psychoj/psychojs>
- PsychoPy. (2018). *PsychoPy: Now running studies online*. Retrieved December 2021, from <https://www.psychopy.org/#online>
- PsyToolkit. (2022). *About PsyToolkit*. Retrieved December 2021, from <https://www.psychotoolkit.org/>
- Schkade, D. A., & Payne, J. W. (1994). How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics and Management*, 26(1), 88–109. <https://doi.org/10.1006/jeem.1994.1006>
- Sondhi, S., Khan, M., Vijay, R., & Salhan, A. K. (2015). Vocal indicators of emotional stress. *International Journal of Computer Applications*, 122(15). <https://doi.org/10.5120/217805056>
- Trickett, S. B., & Trafton, J. G. (2009). A primer on verbal protocol analysis. *The PSI Handbook of Virtual Environments for Training and Education*, 332–346.
- Van Selm, M., & Jankowski, N. W. (2006). Conducting online surveys. *Quality and Quantity*, 40(3), 435–456. <https://doi.org/10.1007/s11135-005-8081-8>
- Van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical approach to modelling cognitive*. Academic Press.
- Vosberg, E. (2021). *crypto-js*, GitHub repository. Retrieved December 2021, from <https://github.com/brix/crypto-js>
- Webster, E. S., Paton, L. W., Crampton, P. E., & Tiffin, P. A. (2020). Situational judgement test validity for selection: A systematic review and meta-analysis. *Medical Education*, 54(10), 888–902. <https://doi.org/10.1111/medu.14201>
- Wirth, O., Chase, P. N., & Munson, K. J. (2000). Experimental analysis of human vocal behavior: Applications of speech-recognition technology. *Journal of the Experimental Analysis of Behavior*, 74(3), 363–375. <https://doi.org/10.1901/jeab.2000.74-363>
- Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication*, 10(3). <https://doi.org/10.1111/j.1083-6101.2005.tb00259.x>
- Zainal Abidin, S., Christoforidou, D., & Liem, A. (2009). Thinking and re-thinking verbal protocol analysis in design research. In: *Proceedings of the International Conference on Engineering Design, Vol. 2. Design Theory and Research Methodology*.

Open practices statement All code is available at the Open Science Repository, as described within the manuscript. The repository can also be located at: https://osf.io/msj7h/?view_only=52bd8769dcf54df7b2eaaa30b4927652

The raw bibliometric data on the frequency of online studies is available from the corresponding author upon reasonable request. No experiments were conducted or preregistered.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.