Behavior Research Methods
https://doi.org/10.3758/s13428-021-01583-9

Introducing BITTSy: The Behavioral Infant & Toddler Testing System

R. S. Newman[1,2] · E. A. Shroads[1] · E. K. Johnson[3] · J. Kamdar[1] · G. Morini[4] · K. H. Onishi[5] · E. Smith[1] · R. Tincoff[6]

Accepted: 18 March 2021

Abstract

This manuscript introduces BITTSy, the **B**ehavioral **I**nfant & **T**oddler **T**esting **Sy**stem. This software system is capable of running the headturn preference procedure, preferential looking, conditioned headturn, and visual fixation/habituation procedures. It uses only commercial-off-theshelf (COTS) hardware to implement the procedures in an affordable and space-efficient setup. The software package, example protocols and data sets, and manual are freely available and downloadable from go.umd.edu/BITTSy, making this entire set of procedures available to resource-limited labs. Researchers can easily use BITTSy at multiple sites in a uniform manner, resulting in a standardized, powerful research tool that can enhance cross-site research collaborations.

Keywords Infant · Looking time · Headturn preference procedure · Preferential looking · Conditioned headturn · Habituation · Software · Looking-while-listening · Visual fixation

Infancy is a crucial time for cognitive and linguistic development; even before infants and toddlers can say their first words, they are forming the skills that will serve as a foundation for their subsequent development (Panneton & Newman, 2012). Recent years have seen a growing appreciation of the importance of studying infants' early cognitive and linguistic abilities, particularly because infants' early skills have been found to be predictive of later developmental outcomes (e.g., Junge et al., 2012; Kooijman et al., 2013; Kuhl et al., 2005; Newman et al., 2016; Salley et al., 2013; Singh et al., 2012). Within this time frame, infant researchers have also identified a variety of technological barriers associated with the development and implementation of infant testing paradigms. We provide the only software system designed to run all of the major infant paradigms used for testing language, allowing for further groundbreaking discoveries about infants' developing linguistic and cognitive systems.

Much of the research examining infants' abilities relies on experimental paradigms that track infants' attention or visual gaze. These simple behavioral responses can reveal what infants know or have learned about the world around them. For example, the headturn preference procedure (HPP) (Colombo & Bundy, 1981; Fernald, 1985) and the visual fixation/habituation procedure (Leslie, 1984) use infants' natural tendency to maintain visual orientation towards an attended sound source to assess what infants have already learned about their language(s) or what they are able to learn in the laboratory. The conditioned headturn (also referred to as visually reinforced infant speech discrimination; Eilers et al., 1977, or as operant headturn; Kuhl, 1985) trains infants to look at a reinforcer when they hear a particular type of sound. Additionally, the preferential looking (Golinkoff et al., 1987; Golinkoff et al., 2013) and looking-while-listening (Fernald et al., 2008) paradigms measure

All authors after the first two listed in alphabetical order

✉ R. S. Newman
rnewman1@umd.edu

1 Department of Hearing and Speech Sciences, University of Maryland, College Park MD 20742 USA

2 Language Science Center, University of Maryland, College Park MD USA

3 Department of Psychology, University of Toronto, Mississauga Canada

4 Department of Communication Sciences & Disorders, University of Delaware, Newark DE USA

5 Department of Psychology, McGill University, Montreal QC, Canada

6 Psychology Department, The College of Idaho, Caldwell ID USA

Published online: 19 May 2021

infants' and toddlers' understanding of spoken language by comparing the amount of time the infant spends looking to a target versus a competing object (Spelke, 1979) (e.g., when they hear "Do you see the kitty?", whether they look longer at a cat than a dog). All of these approaches fundamentally rely on similar behaviors from an infant or toddler: the child's tendency to look towards, or attend to, particular items in the world. (For a comparison of such paradigms, see Johnson & Zamuner, 2010; Junge et al., 2020).

Studies using these paradigms have taught us much about what infants and toddlers recognize and understand about the speech input they experience. For example, such studies have shown that infants generally listen longer to items that are familiar in their auditory environment, including their own name (Mandel et al., 1995), their native language (Nazzi et al., 2000), and speech in an infant-directed speaking style (Cooper et al., 1997; Fernald & Kuhl, 1987; Werker & McLeod, 1989), implying that infants are regularly attending to and picking up on patterns in what they hear. These paradigms have been used to demonstrate both knowledge acquired in the home (as in early word recognition; Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999) and the ability to learn new information within the lab setting (Saffran et al., 1996; Werker et al., 1998). In short, these paradigms enable researchers to uncover essential knowledge for our understanding of infant and toddler language development.

The discoveries and advances from these behavioral/experimental paradigms are especially notable given the difficulties in using them. One major limitation for researchers is that these paradigms are supported by separate tools—each with its own independent update timelines, relying on different platforms/operating systems and, in some cases, using obsolete technology (i.e., software or equipment that is not compatible with current operating systems). This major limitation can pose problems for researchers looking to make use of these paradigms in new ways, or using multiple paradigms in combination with one another, as well as for new labs just starting up. It also limits cross-site collaborations and replications, as further described below. To address these problems, the current paper describes a new software program that can run all of these major infant testing paradigms using off-the-shelf hardware that can be easily set up by new and established labs around the world.

Limitations with current systems

Specialized hardware and systems with limited scope create significant constraints for the conditioned headturn procedure and the HPP, among others. As an example of a hardware constraint, HPP relies on flashing lights as attention-getters, as well as real-time input from a researcher—this generally necessitates having a physical connection between the computer running the test sessions and the test booth or lights, as well as needing an input panel or button box which researchers use to indicate the direction in which infants orient. Conditioned headturn has similar hardware limitations, as it relies on both a physical toy as a reinforcer and real-time input. There has been no readily available, off-the-shelf testing system for either of these paradigms, and so most laboratories using HPP or conditioned headturn have employed their own programmer and electrical engineering consultants to build a custom-designed system, often with little documentation. This has restricted the use of these approaches to those individuals with substantial funding and technical resources, and has reduced the likelihood of new researchers implementing the methodology, or of the methodology being used in lower-resourced institutions. It has also led to a lack of standardization, increasing the difficulty of cross-site collaborative research. This in turn has limited comparisons across populations located in distinct geographical areas (e.g., work comparing infants with different types of community language exposure), and may be a contributing factor to current concerns about replicability in the field (see, for example, The ManyBabies Consortium, 2020).

In addition, the specialized nature of software programs and hardware designed to implement different infant paradigms has led many labs to rely on a single, specific methodological approach, and thus to a fragmentation of methodological approaches across the field. The fact that most systems used in laboratories are able to run only one of these methods has hampered attempts to use convergent methods to address the same question. Laboratories that want to use a range of methodologies often need to set up separate testing systems for use with different methods. As an example, one of our labs previously had one room set up for HPP studies, and a second room for preferential looking studies, with different software and hardware in each. This has constrained labs with less physical space to specialize in a particular technique, and has limited the design of methodologies that combine multiple approaches (see also Hollich, 2006).

These specialized systems have also led to particular difficulties for early-career researchers, many of whom were trained in laboratories running paradigms using older hardware or software that is no longer easily available. These researchers have found themselves needing to shift away from well-established, traditional methods, such as HPP, because they are unable to set up such a system when they start their own laboratory. These early-career researchers frequently face delays setting up an independent research program, as they have to simultaneously learn a new method or shift to a different age of participants. Similar issues are faced by smaller laboratories with limited financial support, including laboratories in undergraduate-only institutions and in countries with less funding for behavioral research.

The solution: BITTSy

This is the first publication describing BITTSy (**B**ehavioral **I**nfant & **T**oddler **T**esting **Sy**stem), a newly developed software system designed to run multiple infant paradigms. BITTSy is free and downloadable, and is capable of running HPP, preferential looking, conditioned headturn, and visual fixation/habituation procedures, all with the same setup (see Figs. 1 and 2). BITTSy uses only commercial-off-the-shelf (COTS) hardware. Moreover, it comes with example protocols and data sets for each of these infant paradigms, making it easy to start testing infants in a new lab.

We designed BITTSy to be a complete, multi-paradigm infant experimental testing platform that can be easily set up with minimal technological experience; it requires no engineering design or development time. BITTSy is highly cost-effective and interfaces with standard, easily accessible off-the-shelf hardware. Experiments in BITTSy are created through highly customizable protocol files, allowing it to run a variety of infant testing protocols using the same standard design tools, while also being flexible enough to allow for the development of new testing approaches. BITTSy's easy setup facilitates consistency at multiple sites in a uniform manner, establishing a standardized research tool that will enhance the feasibility of cross-site research collaborations. The multi-paradigm design makes BITTSy an ideal program for researchers building a new lab, since the same hardware and software can run multiple protocols in a single space. BITTSy is thus an immediately available and incredibly powerful system for behavioral testing with infants and toddlers. Our hope is that this new system will (1) open the use of behavioral infant testing paradigms to early career researchers and smaller laboratories with limited financial support, including laboratories in undergraduate-only institutions and in countries with less financial support for behavioral research; (2) promote multi-researcher and multi-site collaborations, allowing for collaborative testing of low-incidence populations or populations located in different geographical regions; (3) facilitate growth in our knowledge of early language development across a broader range of populations with different socioeconomic, cultural, and linguistic characteristics; and (4) encourage creative combinations of existing paradigms to address new questions.

BITTSy has a great deal of flexibility in terms of how researchers can structure an experiment. It allows researchers to use flashing lights, digital images, or videos as attention-getters, to present stimuli on any given trial that are either audio-only, visual-only, or audio-visual, and to use video images or plug-in standard devices (such as disco lights) as reinforcers/rewards (e.g., for conditioned headturn). Researchers can fully determine which stimuli to present on which trials, whether trials are repeated, and whether stimuli are presented in a fixed order, randomized within blocks, or fully randomized. They can set the relative timing of all

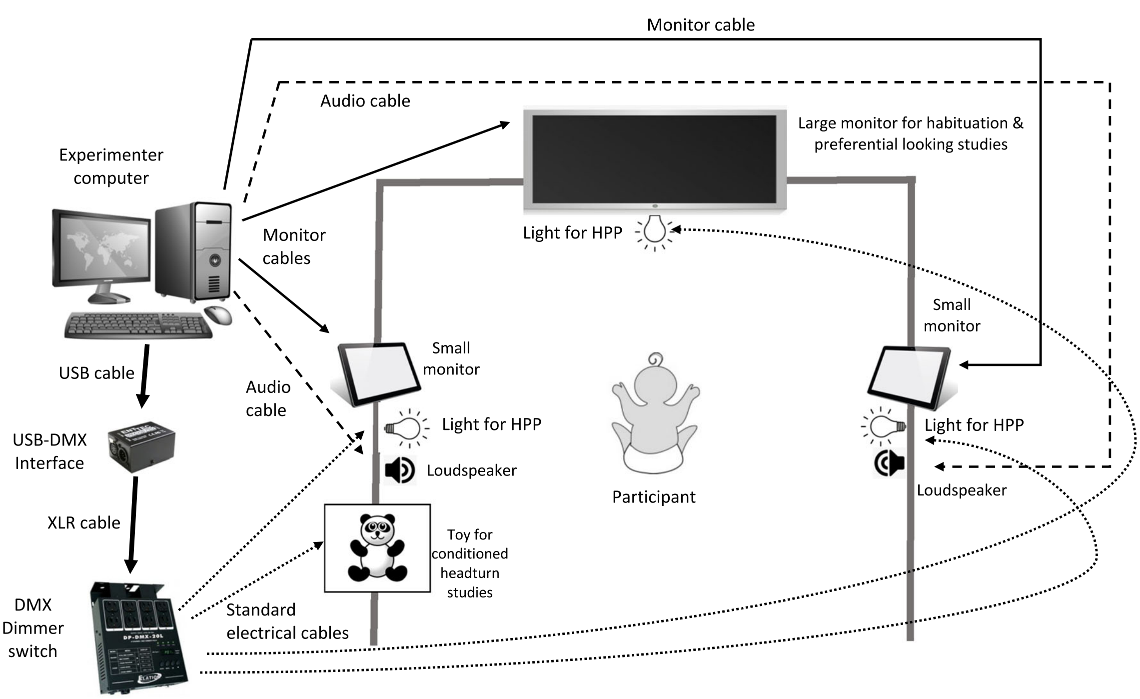

Fig. 1 A graphic display of a single BITTSy setup that would allow for implementing all of the main infant testing paradigms. Not shown is a video camera for recording test sessions

Fig. 2 A typical BITTSy setup. (The wide-angle camera lens makes the booth look curved, but it is actually a square with three walls and an open back.) The front panel has a large screen that can be used for central fixation or preferential looking studies; there are doors to the side that can close and cover the central monitor when not in use. Just under the monitor is a camera, and below that a central light, used for HPP studies. The two side panels (slightly distorted because of the use of a fisheye camera) each have a small monitor (for conducting HPP studies with video images) and a light (for standard HPP) mounted at the height of an infant's head while seated on his/her parent's lap. The setup also requires a monitor for the researcher to view the video feed from the camcorder recording the infant. The researcher's monitor can be located behind the booth or in an adjacent room. This particular setup could also run conditioned headturn studies with a video reinforcer, but is not set up for versions with a light-up toy reinforcer; the only addition required would be the reinforcer itself and the materials to incorporate it into the booth.

presentation events, and set up multiple phases, or sections, within an experiment. Researchers can specify whether individual stimuli, trials, and experimental phases continue for fixed amounts of time or continue until the child reaches a particular looking/listening criterion (such as in habituation). Researchers can code infant looking during testing, or offline from recordings; and if coding is done during testing, the results of that coding can influence, in real time, the next steps of the procedure. The full details of all of the options and numerous example BITTSy protocol script files are provided in an online user manual available at go.umd.edu/BITTSy.

The only required component for BITTSy is a PC running Windows 10. All other components (monitors, lights, speakers) may or may not be needed, depending on the type of study being run. The sections below discuss the primary functions within BITTSy, and how the software solves the paradigm limitations described above.

Presenting visual stimuli

All infant testing procedures have some means of attracting the infant's attention, usually via a visual and/or auditory attention-getter. For example, HPP uses infants' natural tendency to maintain visual orientation towards an attended sound source to test their ability to recognize familiar stimuli or aspects of their native language. One of the primary ways that the classic HPP differs from other procedures is in its use of flashing lights to direct infants' attention, rather than images on video monitors. Most researchers who use this method have testing rooms with lights that are hardwired to the computer via some form of timing board; this requirement for specialized wiring has made the paradigm difficult to set up for labs without engineering and technical support staff. Indeed, HPP has historically required individual labs to employ their own programmer and electrical engineering consultants to build the testing system. Conditioned headturn, with its use of a toy as a reinforcer, poses similar problems. This level of necessary support staff has limited the paradigm to only those labs with substantial funding and technical resources, and reduces the likelihood of new researchers employing this methodology. Moreover, even those labs with the resources to set up HPP have often been forced to continue to rely on outdated hardware, or have been unable to innovate or vary the structure of their experiments created by their single-purpose custom software. Differences in the capabilities and specifications of custom systems across labs have led to a lack of standardization and make cross-site collaboration very difficult. BITTSy avoids this problem by relying on a lighting system called DMX; DMX is an industry standard for digital communication networks that are used to control stage lighting, electronic billboards, Christmas light displays, etc. Because it is an industry standard, there are many different examples available on the market, making this approach both inexpensive and likely to remain current and available. The DMX control system replaces the specialized wiring, electrical engineering consultants, and custom software that required substantial funding and technical resources and limited researchers' access to these essential infant testing procedures.

A DMX system typically involves two small pieces of equipment: a USB interface box that connects (via USB) to the computer, and a dimmer pack that looks somewhat like a power strip. The two are connected via an XLR cable. The DMX system accepts commands from the interface box and uses them to turn on and off standard power outlets (see Fig. 3). There are a variety of interface boxes and dimmer packs on the market that can be

Fig. 3 An example DMX box (left) and dimmer switch (right)

used with BITTSy. With this combination of off-the-shelf hardware devices, BITTSy can send signals to any kind of device that can be plugged into a standard outlet. While we use BITTSy to make nightlight bulbs flash on and off, BITTSy could theoretically be used with any other powered device. For example, it can control small toys such as those used in conditioned headturn experiments.[1] Because most DMX boxes have two plugs per channel, two devices could be turned on with a single command (e.g., a plug-in toy plus a light that illuminates it). The two devices plus lights and cables that we use cost approximately US$200.

In addition to controlling lights, BITTSy can control multiple computer monitors—as many monitors as the computer running the system (and its video card) can support. While some infant testing methods rely on a single central monitor in front of the child (e.g., visual fixation, preferential looking), BITTSy's ability to control four monitors (the monitor from which the researcher controls the experiment and three others) also allows researchers to run an HPP using videos or images (presented on monitors to the left, center, and right sides of the room) rather than lights as the attention-getters. BITTSy leverages the Microsoft .NET Framework for image presentation, so it can present any kind of media file for which a Windows codec exists; this includes both still images (.jpg, .png, and .gif) and videos (.wmv and .mp4). By varying the complexity level of the images or movies, the paradigm can be adapted to participants across a range of ages (children older than a year are unlikely to be captivated by flashing lights in the same way as young infants). This approach can be used to present images as a reward for behavior (as in conditioned headturn studies and conditioned play audiometry), or to present images or videos as a part of the primary stimuli (as in preferential looking and visual fixation paradigms).

Presenting auditory stimuli

BITTSy can present sounds to as many speakers and audio channels as are present on the computer's sound card—including both stereo and surround sound. It can thus be used to present audio from multiple locations around a room. As with video, BITTSy can play multiple file types, currently supporting .wav and .mp3. For studies where complete audio-visual synchrony is required, we recommend integrating the audio as part of a movie file rather than as a separate sound file.

General study design

Stimuli in BITTSy can be presented in a set order, or can be selected randomly, either with or without replacement. Stimuli can also be selected based on hierarchical sets, with randomization occurring within sets. As an example, a study could include pictures of VEHICLES (including cars and planes) as well as ANIMALS (including cats and dogs), with multiple exemplars from each subcategory; BITTSy could be instructed to select one of the two superordinate categories randomly (e.g., ANIMALS), and then to randomly select a subcategory within that (e.g., dogs), and then to order the presentations of each exemplar before moving on to the other subcategory (cat) or the other superordinate category. There is theoretically no limit to the number of nesting levels you can use, although even our most complex designs have not needed more than four.

Experiments run using BITTSy can be hierarchically structured. An experiment can be subdivided into phases, or sections of a study, each of which can be made up of multiple blocks, and each block made up of multiple trials (and each trial can consist of multiple steps). For example, many studies with infants and toddlers involve a familiarization or training phase followed by a test phase; in BITTSy, these different phases can be entirely different in the types of stimuli, number of trials, methods, and timing aspects they employ. Indeed, because BITTSy can run multiple paradigms, it is possible to have the different phases each be typical of a different type of classic behavioral study (e.g., familiarizing infants via a visual habituation paradigm, but then testing using the HPP; see Hollich, 2006). This allows for a wide range of hybrid procedures, enabling researchers to ask questions that could not previously be addressed.

[1] The original conditioned headturn procedure used an animatronic toy monkey with cymbals that was lit up when it turned on. Since most modern toys are battery-operated, rather than plug-in, we have not found a comparable toy to exactly replicate the original studies. For this reason, we recommend using a different reinforcer than in the original studies: either short video clips presented over video monitors or a fun plug-in device such as a colorful disco light or a fan with streamers attached. That said, if you can find, or make, a toy monkey with cymbals that plugs into an outlet, BITTSy could control it appropriately.

Timing

BITTSy has a great deal of flexibility in terms of timing of events and how items are presented. Stimuli can be played for a set amount of time, until a particular event happens, such as the child looking away or a stimulus ending, or until a particular criterion is reached, such as the child reaching a preset criterion for looking time or for habituation. These criteria can also be combined. For example, a researcher can prepare a protocol file to play a stimulus until either the child looks away for 2 seconds or 30 seconds has elapsed, whichever comes first.

Protocol files

BITTSy works on the basis of "protocol files," plain text files that serve as a series of commands and specifications for the makeup of a study; these are essentially instruction sets for how to run a particular experiment or experimental session. The protocol file instructs the program on what stimuli to play, when to play them, how many trials should occur, how they should be randomized, etc. The order of commands has a certain expected structure, but they can be selected, combined, and ordered to create variations on HPP, preferential looking, visual fixation studies, and more. No previous expertise in programming languages is required to manipulate and create these protocol files; the files use a very simple and intuitive syntax that can be easily modified with very little experience, and all syntax terms are defined in the manual. Since these protocol files are simply text files, they can be easily shared across laboratories, allowing the same program to be run across sites. This allows for easy collaboration across testing sites; the only changes a new site would need to make is to correct the file paths (e.g., for the stimulus locations or output files on their computer). Moreover, once a protocol file for a study is created, it can be used as a template to create different variations of the study.

When running BITTSy, a researcher first loads a protocol file (see Fig. 4); BITTSy then checks the protocol file to ensure that the sound and video files required by the protocol are available and in their expected locations, that there are no typographical errors, and that the hardware is on and ready for use. The researcher then enters additional information about the specific session: the participant identifier and date of birth, the person running the session, and any additional notes that might be desired. BITTSy stores this information as part of the permanent data record, along with the date and time, the version of the program that was used, and the name and path of the protocol file.

BITTSy comes with sample protocol files for a number of standard study procedures, most modeled after published studies, which can be used as the basis for other experiments. Once a template has been created for an experiment, it is very easy to modify by exchanging the file names (e.g., for stimulus files), reordering stimuli in groups, changing the restrictions on how items are chosen, etc. Moreover, the protocol files have been designed to be highly intuitive—even a novice student researcher should be able to read a protocol file and understand what is happening at each point in the experiment. As one example, Fig. 5 shows the text of a protocol for an experiment which plays one randomly selected video at a time from a set of 20 videos. Each trial begins with an attention-getter video, which plays until the researcher presses the "C" key (to indicate the child is looking at the center monitor). Then that video is turned off, and the actual test trial starts; a new video is selected at random from a group of videos (referred to as "trial_videos," which were set up earlier in the protocol), and this video plays on the center display one time, until it ends. This entire process repeats for a total of 20 test trials (by looping back from step 7 to step 2 an additional 19 times). This could be a word recognition experiment if each video contained an image of an object paired with an auditory label. The intuitiveness of the sample protocols makes BITTSy a good choice for researchers who want a ready-to-go system and do not want to have to learn a new coding language to create experiments. But for those who want absolute control, BITTSy's detailed protocol files allow for complete customization and more detailed data output than any other infant system available.

Coding

Experiments can be run with fixed trial lengths, in which case the researcher does not need to code behavior in real time. However, studies that involve habituation, or studies where trial length depends on infant behavior, require live coding of the infants' behavior. BITTSy uses a standard keyboard for input; the researcher presses one key for looking towards a stimulus location (center, left, right), and another key to indicate that the child looked away. Which keys are used can be set by the researcher in advance; we recommend that the same keys be used consistently throughout the lab, but this is entirely at the user's discretion. BITTSy also allows the researcher to set criteria for the "minimum" duration for a look, as well as a minimum look away duration; these two parameters essentially set what the program considers as a single "look."

Timing reliability

In order to assess the timing of BITTSy, we conducted a series of tests, measuring the time between a researcher's key press on a keyboard indicating that a light should turn on and the actual turning on of that light, as measured using an optical sensor. The exact timing of the key press was measured via a

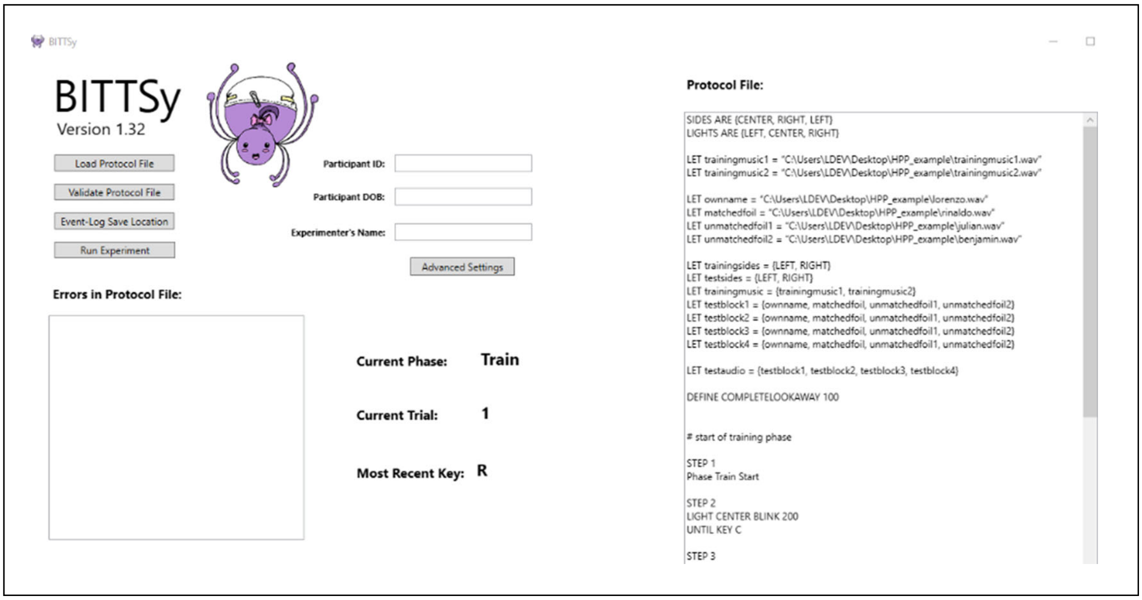

Fig. 4 BITTSy interface

microphone that picked up the click on the keyboard; both this and the output from the optical sensor were displayed on an oscilloscope to determine the time between them. Note that this test includes (1) the time required for the computer to register that a key press has taken place; (2) the time required for BITTSy to identify the event and initiate the subsequent response; and (3) the time for the light itself (here, an incandescent nightlight) to turn on sufficiently for the optical sensor to register it. This test was conducted over a series of 50 attempts; the time lag ranged from 36 to 156 ms, with an average of 61.5 ms (standard deviation 20.6 ms). All but five of the tests came in between 40 and 80 ms, suggesting that the general range of timing is likely to be within that range. Note, however, that this may be dependent on the particular keyboard selected, as some may have greater sensitivity than others. It will also depend on the experimenter, and their motor movements in pressing the key. Additionally, the range of timing could vary slightly from one computer to another, depending on things like RAM, CPU, and/or background processes running on the machine.

In general, the range of timing required for BITTSy processes is quite small relative to the types of events being measured (e.g., the time required for an infant to turn his/her head). However, it can vary slightly across different stimulus types; in most cases this is not perceptible, but for testing where audiovisual synchrony is critical, we recommend incorporating audio files into a movie file, rather than using separate audio and video files.

```
STEP 1
Phase Test Start
STEP 2
VIDEO CENTER attentiongetter LOOP
UNTIL KEY C
STEP 3
VIDEO CENTER OFF
STEP 4
Trial Start
STEP 5
LET vid = (TAKE trial_videos RANDOM)
VIDEO CENTER vid ONCE
UNTIL FINISHED
STEP 6
Trial End
VIDEO CENTER OFF
STEP 7
LOOP STEP 2
UNTIL 19 TIMES
STEP 8
Phase End
```

Fig. 5 Text from a protocol for a word recognition preferential looking experiment in BITTSy

Data and session logs

BITTSy creates a detailed log file of all events that occur within a session. This means that all events are logged, allowing full access to data, and a high degree of flexibility for creating new reports or for designing new ways to analyze the data. A second program, the reporting module, then takes information from the log file and summarizes it in a more user-friendly manner, ready to be used in data analysis programs (see Fig. 6). These summaries can be done over individual files (to obtain information about a particular subject's behavior) or over sets of files (to obtain summary information about all of the participants in a study). We have tried to anticipate common summary information that might be desired, but additional summary analyses and custom reports can be created from the detailed log file at any time. For example, one of our studies plays a series of different videos in sequence on each

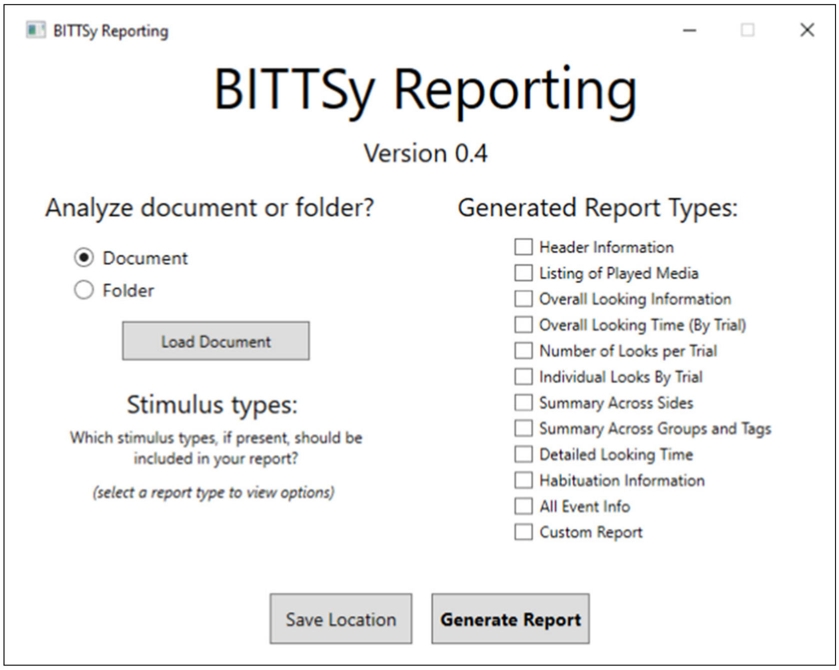

Fig. 6 BITTSy reporting module interface

trial; since the built-in reporting tracks looking to individual stimuli, we created a custom report that measures the looking across all of the stimuli in a single trial.

Technical support, maintenance, and enhancement

The BITTSy software is freely available for download. BITTSy comes with a complete manual, several test protocols to ensure the system is set up and running properly, and a variety of example study protocols and data sets (including ones for each of the primary infant testing paradigms), all available at go.umd.edu/BITTSy. There is also a series of troubleshooting FAQ pages, a Slack channel where users can communicate with one another (to ask questions, discuss best practices, etc.), and a webform for submitting support requests for functional extensions. The code is being maintained by the primary researchers in order to ensure consistency, and new versions will be shared with the community in the form of a version update. The code for the reporting module is also available online for individuals who wish to make modifications for additional reports; the website has a place where users can submit such new analysis codes for inclusion in a library in order to support data and procedure sharing.

Extended uses

The flexibility inherent in BITTSy also allows for the development of a wide array of extended uses. For example, while BITTSy was intended for testing infants and toddlers, we have also used it with both adults and dogs (see Fig. 7). The use of the DMX system for controlling lights means that it can run virtually any powered device; thus, it could be used in human olfaction research by plugging in scent diffusers or olfactometers. The variant of HPP using side monitors allows for testing of manual languages in approaches similar to aural ones (Kulsar et al., 2012), and the fact that the same software can run multiple paradigms allows for the use of creative combinations of tasks to examine interactions across skills (Hollich, 2006).

Comparisons across systems

In this section, we compare BITTSy with two other readily available infant testing solutions: HABIT2 and PyHab. Like BITTSy, HABIT2 (Oakes et al., 2019) is a freely available software package designed specifically for testing infants. Both HABIT2 and BITTSy can present video and audio stimuli over (at least) three monitors and record button presses from a human observer to indicate child looking behavior; and both allow for a wide variety of experimental choices regarding how stimuli are presented and randomized, timing and termination conditions, how the study is divided into phases, and how looks are defined. Both programs allow users to edit existing study designs, and come with a number of templates to allow a new user to begin testing quickly. HABIT2 relies on a press/release system for coding such that the experimenter uses preset keyboard keys to press and hold for looking and then release when the child looks away.

Fig. 7 BITTSy can be used with a variety of different participants; we have used it with human participants from 4 to 24 months of age, as well as with human adults and dogs

BITTSy relies on separate keys (of the researcher's choosing) for looking towards and away, but both use a standard keyboard for coding.

HABIT2 has several advantages. For example, (1) it can be run on both Mac and Windows systems, making it more flexible in terms of hardware use than BITTSy; (2) it uses a graphical user interface (GUI), potentially making it more user-friendly, particularly for novice users; and (3) it creates workspaces, which keep all of the settings, results, and log files from a particular set of studies together, and separate from those of other studies, potentially aiding in organization. HABIT2 can also be used without stimuli for reliability coding; BITTSy likewise has a sample protocol for this purpose, although it does require setting the trial duration (to match that of the original study) in advance.

However, unlike BITTSy, HABIT was designed primarily for habituation/visual fixation procedures. While it can also run preferential looking paradigms, it has no mechanism for running HPP or conditioned headturn. Thus BITTSy has an advantage in being able to run all of the paradigms that HABIT can run, plus conditioned headturn and HPP-based studies. Moreover, because HABIT has a set of fixed design choices (e.g., one monitor or three monitors), BITTSy allows for more flexibility in designing new types of experimental protocols. However, since HABIT2 is a rewrite of a powerful prior package (HABIT), it has the advantage of being widely known and used in the field, and thus researchers might consider it a better choice if they are focusing solely on habituation studies.

Table 1 Comparison of BITTSy, HABIT2, and PyHab

	BITTSy	HABIT2	PyHab
Can run habituation experiments	Yes	Yes	Yes
Can run preferential looking/looking-while-listening experiments	Yes	Yes	Yes
Can run HPP experiments	Yes	No	Limited to monitors
Can run conditioned headturn experiments	Yes	No	No
Number of monitors (not including experimenter's)	As many as the video card supports (but most current paradigms use 0, 1, or 3)	1 or 3	1 or 3
Includes prepackaged experiments	Yes	Yes	Yes
Has GUI interface for running experiments	Yes	Yes	Yes
Has GUI interface for building experiments	No, but comes with protocols that are easy to modify	Yes	Yes (Can be done via GUI or through programming)
Operating system	Windows 10	Windows 7/10 and Mac OS	Windows, macOS or Linux
Freely available	Yes	Yes	Yes
Open source	Only analysis program	No	Yes

PyHab (Kominsky, 2019) is an open source program designed for conducting looking-time-based studies with infants. It works with PsychoPy (Pierce, 2007), and can run studies using one (visual fixation), two (preferential looking), or three (HPP) monitors. Like HABIT, it uses a press/release system for coding, and has a GUI interface for basic experimental designs, although its real strength comes from the ability to modify the programs in Python. The GUI interface is quite user-friendly, with drag-and-drop experiment components.

Advantages of PyHab include the fact that it is multiplatform, running on Windows, Mac, and Linux systems, and is easily modifiable by anyone who already knows Python. Moreover, PyHab stores all files needed for an experiment (the protocol, data, stimuli, program launcher, etc.) in self-contained experiment folders, making it easy to share studies and contribute them to open-source repositories. As the author notes, "the capabilities of Habit and PyHab are very similar" (p.118), with a primary difference being the open-source nature of the latter. However, PyHab cannot control lights, and thus cannot be used for classic HPP studies, or for conditioned headturn studies. Moreover, while PyHab's interface is highly intuitive for basic behavioral/experimental paradigms, studies involving more complicated or innovative experimental designs require programming in Python, which might be a barrier for experimenters with little to no experience with Python programming.

Below we summarize this comparative information in Table 1.

Summary

In summary, we designed BITTSy to be an eminently flexible system, capable of replicating many current infant and child testing paradigms, while maintaining ease of use and setup. It uses only off-the-shelf hardware, and yet can be used to conduct a wide range of study designs. For stimulus presentation, it can turn lights and other electronic devices on and off, and can present any combination of static images, dynamic movies, and sound files, across multiple video screens and speakers. The software is free and downloadable at go.umd.edu/BITTSy, with a clear manual, setup guide, and support group. BITTSy is thus an incredibly powerful behavioral testing system ready for researchers to use as they seek further groundbreaking discoveries about infants' developing linguistic and cognitive systems.

Acknowledgments The software described herein is available for free download from go.umd.edu/BITTSy.

All authors participated in the creation of this manuscript.

Funding This work was funded by NSF award BCS1152109 to the University of Maryland.

Declarations

Conflicts of Interest None of the researchers have any conflicts of interest.

References

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253–3258.

Colombo, J. A., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development, 4*, 219–223.

Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development, 20*(4), 477–488.

Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech & Hearing Research, 20*, 766–780.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279–293.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Language acquisition and language disorders: Vol. 44. Developmental psycholinguistics: On-line methods in children's language processing* (Vol. 44, pp. 97–135). Amsterdam: John Benjamins Publishing Company.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23–45.

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science, 8*(3), 316–339.

Hollich, G. (2006). Combining techniques to reveal emergent effects in infants' segmentation, word learning, and grammar. *Language & Speech, 49*(1), 3–19.

Johnson, E. K., & Zamuner, T. S. (2010). Using infant and toddler testing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 73–94). John Benjamins Publishing Company.

Junge, C., Everaert, E., Porto, L., Fikkert, P., de Klerk, M., Keij, B., & Benders, T. (2020). Contrasting behavioral looking procedures: A case study on infant speech segmentation. *Infant Behavior & Development*. https://doi.org/10.1016/j.infbeh.2020.101448

Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental Science, 15*(4), 463–473.

Kominsky, J. F. (2019). PyHab: Open-source real time infant gaze coding and stimulus presentation software. *Infant Behavior & Development, 54*, 114–119.

Kooijman, V., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive brain signals of linguistic development. *Frontiers in Language Sciences, 4*, 1–13.

Kuhl, P. K. (1985). Methods in the study of infant speech perception. In G. Gottlieb & N. A. Krasnegor (Eds.), *Measurement of audition and* *vision in the first year of postnatal life: A methodological overview*. Ablex.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the 'Critical period'. *Language Learning & Development, 1*(3&4), 237–264.

Kulsar, S., Seal, B. C., Baird, C., & DePaolis, R. A. (2012). *Head-Turn Sign Preferences in Babies of Deaf and Babies of Hearing Parents.* Paper presented at the XVIII Biennial International Conference on Infant Studies, Minneapolis.

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception, 13*, 287–305.

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science, 6*(5), 314–317.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43*, 1–19.

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language, 43*(5), 1158–1173. https://doi.org/10.1017/S0305000915000446.

Oakes, L. M., Sperka, D., DeBolt, M. C., & Cantrell, L. M. (2019). Habit2: A stand-alone software solution for presenting stimuli and recording infant looking times in order to study infant development. *Behavior Research Methods, 51*, 1943–1952.

Panneton, R., & Newman, R. S. (2012). Development of speech perception. In L. A. Werner, R. R. Fay, & A. N. Popper (Eds.), *Springer handbook of auditory research: Human auditory development* (Vol. 42, pp. 197–222). Berlin: Springer.

Pierce, J. W. (2007). PsychoPy—Psychophysics software in python. *Journal of Neuroscience Methods, 162*(1–2), 8–13.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Salley, B., Panneton, R., & Colombo, J. (2013). Separable attentional predictors of language outcome. *Infancy, 18*(4), 462–489.

Singh, L., Reznick, J. S., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science, 15*(4), 482–495.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology, 15*, 626–636.

The ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Psychological Science, 3*(1).

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science, 10*(2), 172–175.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34*(6), 1289–1309.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*(2), 230–246.