Evaluation of an Ecological Interface Design–Driven Augmentative and Alternative Communication Interface

Kaela Shea, Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada, Olivier St-Cyr, Faculty of Information, University of Toronto, Toronto, Ontario, Canada, and Tom Chau, Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

This study evaluated the change in usability, mental workload and information transfer rate associated with an augmentative and alternative communication (AAC) interface designed through ecological interface design (EID). The design and development process is detailed in Shea et al. (2021). Digital AAC interfaces are considered high-tech interventions for individuals who experience complex communication needs (e.g., from etiologies such as cerebral palsy) and enable users to select language options from a visual display. Interface usability, mental workload and information transfer rate collectively influence users’ communication. Ten AAC-naïve participants engaged in three semi–scripted conversations (verbal, AAC-mediated commercial interface, and EID interfaces) with an actor. Augmentative and alternative communication interfaces were accessed through a single switch pathway. Information transfer rate, error rate, heart rate variability and subjective workload performance measures were recorded for every trial. During AAC-mediated trials, interface interactions were also documented. The EID AAC interface presented improved communication in 5 out of 7 performance measures (p < .05). The EID AAC interface was associated with a significantly higher information transfer rate, lower error rate, less time elapsed between switch activations, less switch activation per word communicated and lower subjective workload. The application of EID to an AAC interface can lead to a significantly improved communication experience.

Keywords
interface evaluation, assistive technologies, ecological interface design, children, speech production

Introduction

Augmentative and Alternative Communication

Augmentative and alternative communication (AAC) is a common intervention for individuals experiencing complex communication needs that may arise from autism spectrum disorder (ASD), cerebral palsy, brain–stem stroke, or acquired central nervous system injury, among other conditions (Light & Mcnaughton, 2012). Electronic speech generation is a form of high-tech AAC technology that enunciates machine-generated words (Light, McNaughton, et al., 2019). Typical electronic speech generation devices, an example from AssistiveWare B.V. (2019) is shown in Figure 1, comprise an interface display of possible word selections among which the user navigates and selects through an access pathway. An AAC interface typically arranges possible word selections into specific taxonomies. Vocabulary requirements depend greatly on environmental factors such as the physical context, the partners in the conversation (Baxter et al., 2012; Fern et al., 2005), as well as personal characteristics such as literacy level, developmental stage (Beukelman et al., 1991) and personal preferences (Thistle & Wilkinson, 2015). To communicate a word during a conversation, an AAC user navigates the interface and makes selections through an access pathway. Configurations of access pathways depend on the voluntary control commanded by users. For example, single switch...
access provides interface control through a single binary input; sensors and transducers allow for varied pathways that leverage a user’s abilities to generate the binary input (i.e., humming, blinking, pressing a button or volitional cortical potentials). To support control through binary input, the display sequentially highlights areas of the screen in which items can be activated. This software functionality is available within accessibility settings of many electronic devices (i.e., Switch Control in Accessibility within iOS Settings or Switch Access in Services in Accessibility within Android Settings). With two binary inputs, the movement and activation of the highlighted area are both controlled independently. However, with a single switch input, the speed and pattern of the movement are predetermined through software settings.

Despite its many potential benefits, AAC can constrain a user’s expressive communication as systems often only afford messages with incomplete context and inadequate detail (Hoag et al., 2008). Slow communication is a frustrating challenge for AAC users that creates negative interactions and perceptions (Baxter et al., 2012). Users of strictly text–based devices in particular experience more restrictive information transfer rates compared to image-based or combined image and text-based systems (Baxter et al., 2012). Despite this, Hoag et al. (2008) shows that the relevance of a newly formed message is preferred by communication partners over reduced elapse time of a pre-recorded message. For these reasons, usability of AAC systems impacts an individual’s capacity to participate in day-to-day activities and may be enhanced through empirically driven design (Light, Wilkinson, et al., 2019).

**Food Ordering Use Case: Example Commercial AAC Interface**

Figure 1 depicts an example home screen of the Proloquo2Go AAC interface. The following
example depicts how a user of the commercial AAC device might order a sandwich. In the AAC system’s idle state, the user’s interface shows a home screen of a grid of colourful word and category cells with a blank message bar above.

As ‘buy’ is not a regularly used word that appears on the home screen, the user searches for and selects the ‘Actions’ category. Words and categories within the ‘Actions’ category replace some of the words on the home screen. The user searches for and selects the word cell corresponding to the verb ‘buy’. The system enunciates the word ‘buy’, which now also appears in the message bar. The interface then automatically reverts to the home screen. The user now searches for the ‘Food’ category, and once selected, the display updates with words associated with food and selects the noun ‘sandwich’, which is enunciated by the system. At the completion of the phrase the customer presses on the word bar. The system articulates the selected words again, that is, ‘buy sandwich’. The server asks for clarification in the form of a yes/no question. The AAC user searches and selects the Small Words Category. Words within the ‘Little Words category replace some of the words on the home screen. The customer searches and selects either ‘yes’ or ‘no’ and selects the word cell. The user can identify specific items for their sandwich within the category ‘Food’ and its subcategories (not shown in Figure 1).

**Ecological Interface Design**

Ecological Interface Design (EID) is a framework derived from cognitive work analysis (CWA; Vicente, 1999) that relays work domain constraints through the interface to allow the user to focus cognitive resources on such tasks as decision making and problem solving (Vicente & Rasmussen, 1992). Ecological Interface Design applies two phases of CWA, namely, work domain analysis (WDA) and worker competency analysis (WCA), to inform user interface design (Burns & Hajdukiewicz, 2004). The framework identifies and maps work domain constraints, as well as constraints for user perception, to interface display elements (Holt et al., 2015). Work domain analysis maps the goals of the AAC system to the functionality of the system components (i.e., the unit forms of English) to identify constraints on the system and the information required to understand system control. Worker competency analysis identifies the information necessary to support the user behaviours required to fulfil system goals (Vicente & Rasmussen, 1992).

The EID framework considers the bounds of system actions imposed by the constraints of the work domain and user capacity to inform user interface design. As a result, EID yields interface designs that optimize human cognition and support users’ response to novel, unplanned, events (McIlroy & Stanton, 2015; Naikar, 2013). Ecological Interface Design is increasingly used to model complex sociotechnical systems (St-Cyr et al., 2013). Ecological Interface Design implementation in intent-driven systems, such as AAC, can yield superior user performance compared to alternatively designed interfaces (Bennett & Flach., 2019; Euerby & Burns, 2014).

**Ecological Interface Design Application to an AAC Interface**

Recently, Shea et al. (2021) applied EID to develop a prototype AAC interface (refer to Figure 2) to support the communication goals of an AAC user. Through WDA the information required to provide sufficient understanding of the AAC system, and the function needed to support user control were identified, while the behaviours necessary to convey a communication act were delineated through WDA.

The analysis revealed that: a communication act using an AAC interface typically requires the user to indicate their intention to communicate, plan language, search the interface through navigation acts and visual search, and select their desired words and letters; the communication partner must observe the communication intent, recognize the context of communication, and understand the meaning of the enunciated words; the commercial AAC interface design (e.g., Figure 2) fails to support behaviours required to perform the necessary actions of communication. Certain skill-, rule- and knowledge-based behaviours were flagged as being unsupported (Table 1). In the application of EID, information
requirements identified through WDA were used to inform the development of graphic forms and the process view of the EID AAC interface. Graphical forms of the EID user-interface, such as the emotion grid, were designed to map information to the user with little need for processing (see Figure 2). Also, colours and graphic forms were carefully chosen to minimize attention capture. The hierarchical relationships within the information requirements guided information presentation of the layout of EID AAC viewports and structure of the system’s navigation. The EID user interface deployed word-type category buttons to present abstract information as an aggregate of low-level units of English language. Further, the EID user-interface considered the grammatical structure of English, that is, words were organized syntactically instead of semantically to increase the information communicated to the user through the interface. The organization of the unit forms of language was relatively flat and monopolized the interface to avoid deeply structured menus. Additionally, category menus (i.e., word type category buttons, see Figure 2) were always present to maintain visual momentum in support of user navigation. Uniquely, an interface for the communication partner was also introduced to facilitate interaction with a non-speaking AAC user.

**Food Ordering Use Case: EID AAC Interface**

Figure 2 shows an annotated home screen of the EID AAC interface. The following example depicts how a user of the commercial AAC device might order a sandwich. In the system home screen, the user’s interface is mostly a blank grey screen with the purple and blue bars containing the Emotion Grid, Initiate Speech Button, Question-Statement Toggle, Yes and No Word Cells, and Word Type Category Buttons.

The customer selects the Word Type Category containing verbs. The customer’s interface is populated with verbs clustered by themes. The customer searches and locates the verb ‘buy’ and presses the word cell. The system enunciates the word ‘buy’ and the system returns to the home.

*Figure 2. Annotated image of the EID user interface designed by Shea et al. (2021).*
These actions are repeated with the noun ‘sandwich’. If the server has clarifying questions the customer would press the Yes or No Word Cells accessible from the system home screen. The user can identify specific items for their sandwich within the Word Type Category containing nouns.

**Interface Evaluation**

The EID interface was contrasted with a commercial interface design (Shea et al., 2021). As a framework applied within EID, the skill-, rule- and knowledge (SRK) taxonomy breaks down system information processing and resulting knowledge states to analyse the behaviour supported by the interface (Kilgore & St-Cyr, 2006). Through SRK taxonomy, Table 1 and Table 2 highlight discrepancies between EID and commercial interface support of necessary AAC user behaviours. The EID application to a linguistic domain of an AAC interface identified unrecognized barriers to communication that were subsequently addressed through EID interface development.

**Current Study**

This article discusses the EID interface performance through the lens of the SRK taxonomy. To empirically evaluate performance of the EID display, an experimental paradigm was
developed to simulate a real-world conversation between two people. Augmentative and alternative communication interface-related performance measures were information transfer per conversation, error rate, usability and mental workload. During AAC-mediated trials, participant behaviour measures of time elapsed between switch activations and switch activations per word were also recorded.

Mental workload is modulated by the degree to which a necessary behaviour is supported by the interface and is strongly associated with performance ability (Marinescu et al., 2017). Poorly supported navigation, for example, exacts high working memory demands on the user (Light & Mcnaughton, 2012). Several physiological measures provide objective information on mental workload and are sensitive to instantaneous changes (Charles & Nixon, 2019). Heart-rate variability is often measured simply through a blood-volume pulse sensor on a digit (e.g., Charles & Nixon, 2019). Other common physiological measures of mental workload, such as eye-blinks and pupil size were not considered due to the visual nature of study tasks. Additionally, an accepted standard for evaluating subjective mental workload in literature is the NASA Task Load Index (NASA-TLX) (Hart & Staveland, 1988) while perceived usability of interfaces is commonly assessed via the system usability score (SUS) (Tullis & Albert, 2013). We anticipated that, compared to a commercial AAC alternative, the prototype EID interface would generally improve system usability and reduce mental workload.

Methods

This research compiled with the American Psychological Association Code of Ethics and was approved by the Research Ethics Boards of the University of Toronto and Holland Bloorview Kids Rehabilitation Hospital. The usability, mental workload, and information transfer rate of the EID interface were evaluated using three conversational trials: a speech trial to provide baseline physiological measurements; a control trial involving a commercial AAC interface; and an experimental trial using the EID interface.

Instrumentation and Experimental Set-up

The iOS application, Proloquo2Go (Figure 1), was used for the control trial and an iOS implementation of the EID interface (Figure 2) was used for the experimental trial. The communication partner interface was not considered in this study such that the communication partner and participants remained oblivious to the type of trial. The AAC software was hosted on an Apple iPad 2 iOS 9. Scanning style, scanning speed, dwell time, number of scanning loops before system time-out, and tap actions were set up equivalently for single switch access in the accessibility settings of each device.

Single switch access was chosen as the access pathway for the AAC interfaces. Although some users can access AAC interfaces directly (e.g., touching the screen), the vast majority of alternative access pathways designed to accommodate users without this level of motor control are based on single switch technology (Tai et al., 2008). Participants controlled both AAC interfaces with button presses using the Orby button-style switch (Origin Instruments, n.d.) (Figure 3). Switch activations (i.e., button presses) were either navigation events or word selections, and were recorded through Thought Technology ProComp Infiniti Encoder at a sampling rate of 2048 Hz (Thought Technology Ltd. 2018).

Heart rate of the participant was recorded using a Thought Technology blood volume pulse sensor (BVP) (Thought Technology Ltd. 2018). The BVP sensor was fastened to the index figure of the participant’s left hand with a Velcro strap (refer to Figure 3). Heart rate was sampled at 2048 Hz.

Participants

Participants consisted of 10 typically developed adult volunteers (23 ± 4°years; eight females) who were able to converse fluently in English but with no prior experience with AAC interfaces. Augmentative and alternative communication-naïve participants were used to remove the potential of a learning bias associated with using a familiar interface. Participants provided informed written consent. The study protocol was approved by the local research ethics board.
| Information processing step | Resultant knowledge State | Skill-based behaviour | Rule-based behaviour | Knowledge-based behaviour |
|-----------------------------|---------------------------|----------------------|---------------------|--------------------------|
| Indicate intent to communicate | Whether the communication partner is attending | Directly perceive the ‘Initiate Speech Button’ | Directly perceive indication function through iconic mapping | Reason when to activate the button to gain communication partner attention |
| Plan communication | Linguistic and non-linguistic elements to realize the communicative goal | Directly perceive linguistic and non-linguistic options | Directly perceive syntactic patterns associated with linguistic communication | Reason the elements of the communicative goal that can be communicated linguistically or non-linguistically |
| Search of AAC interface | Location of buttons and toggles | Directly perceive ‘Emotion Grid’, ‘Question-Statement Toggle’, ‘Yes and No word cells’ and ‘Word Type Category Buttons’ | Directly perceive syntactic relationships through propositional mapping of “Word Type Category Buttons”. Directly perceive communication act and emotion representative of communication goal through iconic mapping. | Reason the grammar category comprising a word, the emotion quadrant, or communication act that progresses the communicative goal |
| Select category | Words available in the selected category | Directly perceive desired category | Directly perceive that cell selection will expand ‘Word Cell Button’ options in viewport | Update understanding of category |
| Select communicative unit | Word that is uttered | Directly perceive desired word | Indirectly perceive word selection will convey semantic meaning to communication partner | Reason comprehension of the communication partner and update the language plan, if necessary, to achieve communication goal |
Study Design

In this within-subject design, participants completed three trials; each participant underwent one speech-based trial and two AAC-mediated trials. The order of AAC interface presentation was counterbalanced through assigning the participants to two groups. Group 1 completed the control (conversing with the commercial AAC) prior to the experimental (conversing the EID AAC interface) condition, while Group 2 did the reverse. Participants within groups were matched on educational level, age, and sex. Immediately prior to an AAC-mediated trial, participants received a brief description of the interface (Appendix B) and underwent a 2-minutes self-guided orientation.

In all three trials, participants held a semi-scripted conversation with a contracted actor who had no prior acquaintance with the participants. The actor had specific training to normalize facial expressions and conversational flow between participants. The actor was not aware of which device was being used in a trial. The actor’s script was designed by the actor and a research team-member to balance the number of switch activations required to access anticipated words between the two interfaces and the two alternative script pathways (Appendix A). The semi-scripted scenario was also designed to mirror a scenario to which all participants would have been exposed in their lived experiences (Appendix A). The actor’s conversational turns followed the script while participants were independent in their conversational choices. Participants were given agency to simulate the pressure of language planning and decision making encountered in an uncontrolled setting. All trials were stopped after 10 minutes, even if the conversational script remained incomplete, to standardize trial length. No trials completed the script in less than 10 minutes.

Measures

Several performance measures were computed: short-term heart rate variability; the time elapsed between switch activations; the number of switch activations per word; the number of words communicated per trial; and, the number of errors (incorrect communicated words as determined by the conversational partner) per trial. Short-term heart rate variability served as a physiological indicator of mental workload. Specifically, the root-mean sum-of-squared differences (RMSSD) between adjacent inter-beat intervals, $x_i$ and $x_{i+1}$, $i = 1, \ldots, n - 1$ over one-minute windows were determined from the recorded BVP (equation (1)) (Charles & Nixon, 2019).

$$RMSSD = \sqrt{\frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{n}}$$  

where $n$ is the number of interbeat intervals within a one-minute window.

The NASA TLX was completed by participants immediately post-trials. Participants rated six dimensions of perceived mental workload (mental demand, physical demand, temporal demand, performance, effort, and frustration), each on a 21-gradation scale between low and high. Participants

Figure 3. Orientation of participant, actor and instrumentation during data collection tasks
also indicated the relative importance of each factor for the task they performed through a series of pairwise factor comparisons; the weighted sum of the scores for each of the six factors yielded the overall rating of subjective mental workload (Hart & Staveland, 1988). Following AAC-mediated trials, the SUS was also administered to capture each participant’s perspective of system usability (Brooke, 1996; Tullis & Albert, 2013).

Data Analysis

Performance measures were compared independently between the two AAC interfaces. Interparticipant analyses were conducted with a nonparametric paired Wilcoxon test (Python 3.8) for each measure separately, as assumptions of normality were not upheld. Heart rate variability data were normalized relative to each participants’ speech-based baseline trial by subtracting RMSSD values.

Comparisons of ordinal/ratio variables (behavioural, physiological and psychophysical measures) were conducted through linear separability modelling. In addition to performance measures listed in Table 3, presentation order of the interfaces was included as a feature. Linear machine learning architecture was applied to model the structure associated with the two classes (use of a commercial AAC interface or use of the EID AAC interface) from which observations were measured and considered features of each class. The logistic regression algorithm was chosen due to its popularity in modelling features associated with a discrete dependent variable (Hastie et al., 2017; Pedregosa et al., 2011). Performance of the logistic regression model is contingent on the distinctiveness the two interface use profiles.

The logistic regression model was run in a Shuffle-Split cross-validation framework (a.k.a. Monte Carlo cross validation) to optimize data availability. The area under the receiver operating characteristics curve (AUC) was derived from tenfold cross-validation as a measure of model performance. The ROC curve depicts the trade-off between sensitivity (true positive rate) and specificity (true negative rate) of model predictions. AUC of 0.8–0.9 is considered good, whereas 0.9 to 1 is considered excellent (Bradley, 1997). Models were further evaluated by estimating their accuracy through permutation tests. The permutation test for classifier performance generates random permutations of the class label vector while the feature matrix retains its original order such that existing dependency between features and labels is removed. The test generates n permutations and trains the machine learning model with the original features and each permuted label vector. The test output describes the fraction of permutations for which model performance improves compared to that trained with the original feature matrix and class-label vector. Permutation tests mitigate the risk of false discoveries of significant effects. This is a particular source of

Table 3: Performance metrics in AAC-mediated trials. Numerical entries are means and standard deviations across participants. Significance between measures for the commercial and EID interfaces were determined independently for each metric through a nonparametric paired Wilcoxon paradigm *(p < .05), **(p < .01), and ***(p < .005)
concern with a low quantity of sampled data. Permutation test scores provide information on the dependency between features and classes captured by a classifier (Ojala & Garriga, 2010).

Backward recursive elimination (BRE) and forward stepwise regression (FSR) were applied to consider the influence of features on logistic regression model performance. Following the elimination or addition of a feature, a permutation test of 1000 iterations and AUC were used to determine the significance of each feature to the interface use profile. To visualize the structures of the optimized, select-feature model compared with the model trained on all features, linear discriminant analysis (LDA) was employed. Linear discriminant analysis applies supervised probabilistic modelling to identify attributes that account for the greatest variance between known binary response variables. In this model, the binary response variables are the two classes: use of a commercial AAC interface or use of the EID AAC interface. We implemented eigenvalue decomposition to identify linear separability between covariances calculated for each response variable. The performance of the LDA classifier reflects the linear separability of the data being modelled, that is, the ability to correctly assign membership of the data points using a linear decision boundary (Pedregosa et al., 2011).

Results

Significant differences were found across participants, between the commercial and EID AAC interface use performance in 5 of 7 measures (see Table 3). Figure 4 shows time elapsed between switch activations during the conversational trials of the commercial or EDI AAC interfaces for one participant. For this participant, the EID interface seemed to have required only episodic selections with supra-5s elapse times.

The EID interface required significantly fewer switch activations to communicate each word. It should be noted that the variance in the average number of activations per word was much higher for the commercial prototype than for the EID interface. The consistency in the number of selections for each word during EID interface use is evident in Figure 5. The more correct words communicated using the EID interface relative to the commercial interface is mostly due to participants progressing further in the semi-scripted conversation during EID interface use. Errors were uncommon throughout both types of AAC trials and were committed deliberately in the face of navigational uncertainty, to backtrack in their interface navigation. In both interfaces, the selection of a word resulted in the immediate return to the home screen. For the commercial interface, this action removed the need to navigate back through multiple menu levels using the ‘back’ button or to locate the ‘home’ button (see Figure 1). For the EID interface, the action eliminated the need to wait for a timeout that returned the system to the home screen after inactivity (Shea et al., 2021).

The subjective workload and system usability measures reflected participant preference for the EID interface over the commercial alternative. Although the difference between the mean SUS scores measured after the commercial and EID interface uses were not significant. Ecological interface design interface use resulted in significantly lower subjective mental workload measurements ($p < .025$) from the NASA TLX. No influence of interface use was found through HRV measures across participants (Table 3).

The logistic regression classifier was able to accurately identify the interface associated with participant use profiles at a rate of 81.6% and AUC value of 0.87. The probability score obtained through the permutation test ($p < .01$) suggested that the logistic regression classifier captured dependency between features and classes to maximize classification accuracy. The most salient features for accurate model identification were determined to be: (1) the number of selections per word communicated; (2) the total number of words communicated; and (3) the number of erroneous words communicated. Figure 6 shows the impact on classification accuracy when each feature was independently omitted. Figure 7 shows the linear separability of the classes through optimal feature projection of an LDA model trained on all the features and an LDA model trained on the three most salient features as identified through BRE and FSR. As shown by the lack of overlapping features in Figure 7, linear separability, and thus accurate class identification through linear machine learning architecture, was
achieved by both models. The number of selections per word, number of words communicated, and number or erroneous words communicated were the variables which distinctly identified the interface in use.

**Discussion**

We compared conversational trials of a commercial AAC interface and a recently proposed EDI-designed alternative. To effectively realize a communication goal, AAC users must
initiate communication, plan language, search the interface, navigate the interface, and perform a selection (Akcakaya et al., 2014). Augmentative and alternative communication interfaces must support SRK-based behaviours for each stage of communication. Table 1 and Table 2 highlight the discrepancies in behavioural support between the commercial and EID interfaces. Using the EID-designed interface, participants experienced lower subjective mental workload ($p < .05$), reduced elapse time between navigation actions ($p < .05$), reduced number of switch activations required per word communicated ($p < .005$), increased information transfer rate ($p < .01$), and a decrease in erroneous words communicated ($p < .01$). Observations made throughout the trials provided insight into the different behaviours supported by the EID and commercial interfaces. Collectively, behavioural measures (e.g., quantity and time elapsed between switch activations), subjective mental workload, and system usability results suggest a better user experience with the EID interface for communication in a simulated conversation scenario. Descriptive statistics applied independently to analyse each performance measure are supported by the performance of a binary classification machine learning model that considered the overall performance profile of interface use. Due to experimental restrictions, word type category buttons and viewport organisation were the main EID design elements implemented.

### Elapsed Time

The time elapsed between switch activations encompasses the time for language planning, interface information processing, and cursor scanning time. As scanning speed and dwell time were equivalent between the devices, the significantly reduced elapsed time observed in the EID interface can be attributed to interface design. The greater variance in elapsed times with the EID interface can be explained by considering the time series for time elapsed between switch activations. Considering Figure 4, we observe that the interface navigation behaviour was different between interfaces. While there exists consistently moderate to long elapsed times throughout commercial interface navigation, participants using the EID interface required long elapsed times at the initiation of a response selection, but these were followed by rapid subsequent selections.

Noted in communication actions of ‘Plan Language’ and ‘Search AAC Interface’ in Table 2, the EID interface uniquely included design elements that directly supported SRK-behaviours for language planning and navigation. The word type category buttons (annotated in Figure 2) of the EID interface afforded a grammatical overview of language units located in different viewports which served as a representation of the syntactic limits for phrase formation. To effectively navigate an interface, it is important for users to maintain a mental model and understand the hierarchical structure of the system through support of knowledge-based behaviour (García et al., 2014; Rasmussen, 1983). The word type category buttons integrated multiple levels of system information, as well as provided visual momentum to support skill-based behaviour for navigation events. For example, when a participant navigated to a word category, the word type category buttons remained visible with a visual indication of the selected category. A lack of rule-based behavioural support for category navigation results in procedural traps (Vicente & Rasmussen, 1992). The commercial interface did not include emergent features reflective of the hierarchical nature of language and participants were subjected to procedural traps of navigating through several categories prior to finding the viewport of their desired word. The impact of this design difference was shown in the participants’ performance. The consistent, 5°s–15°s time elapsed between switch activations during commercial interface navigation, shown in Figure 4, suggested a static level of familiarity with categories, navigation options and word orientation in viewports. Participants’ familiarity with the interface did not increase during the trial; they resorted to exploratory navigation throughout commercial interface use. In contrast, with the EID interface (see Figure 4), participants planned their navigation acts. The support for interface search and navigation behaviours can be attributed to the word type category buttons that served as emergent features and a transparent taxonomy (Holt et al., 2015).
Control through low bandwidth access pathways such as brain-computer interfaces can be very challenging for the user (Silva et al., 2008). Thus, the number of switch activations needed to communicate a word is an important metric for characterizing ease of access. The number of switch activations required to locate and select the desired word was significantly higher with the commercial interface (\( p < 0.005 \)). The highest number that a participant required was 35 switch activations to communicate a single word (see Figure 5). Due to the multilevel menu structure of the commercial interface, each navigational goal required multiple switch activations to reach a viewport displaying the desired word option. Other taxonomies that similarly distill large datasets into deep hierarchical structures have been found to be time-consuming to search, require several decisions about thematic relationships and entail complicated navigational paths (García et al., 2014).

The greater consistency in the number of switch activations per word during EID interface use across participants (Table 3) suggests that similar strategies for EID interface navigation were adopted across participants due to a functional-level control of the AAC interface (Hajdukiewicz & Vicente, 2002). This finding
highlights the knowledge-based behavioural support for language planning through hierarchical organisation of linguistic elements of the EID interface. The demonstrated knowledge-, rule- and skill-based behavioural support from word type category buttons allowed participants to plan language and circumvent procedural traps to achieve significantly higher communication rates (number of correct words per trial as shown in Table 3).

**Mental Workload**

No significant differences were found in measures of physiological workload. This finding contrasts with changes observed in participants’ subjective mental workload. A possible explanation for the lack of measured change in physiological mental workload is that, as participants were AAC-naïve, the activity of using an AAC interface to mediate a conversation for the first time may have induced a high mental workload. Due to steps taken to counterbalance AAC interface presentation, the resultant physiological workload measure would yield no significant difference between the interfaces. Physiological signals are generally not sensitive to small changes in mental workload (Charles & Nixon, 2019). Conversely, the NASA TLX survey prompts participants to reflect on their experiences through controlled questions (Hart & Staveland, 1988). Thus, the NASA TLX survey was more suited to capture subtle mental workload changes. Overall, changes to mental workload observable through the NASA TLX survey (refer to Table 3) were not accompanied by significant physiological responses when the interfaces were navigated by typically developed adults through single switch access. In single switch users with severe disabilities, the physical and mental effort (e.g., attention, visual scanning, and timed action) associated with each switch activation would be significantly higher. Consequently, the mental workload and usability experiences of these users may be more responsive to changes in interface design (Light, McNaughton, et al., 2019).

**Limitations and Future Work**

To fully understand the efficiencies provided by the EID application in AAC interface design, we must examine the deployment of the combined user and communication partner displays (Shea et al., 2021) and assess the support provided for language planning and initiating non-situational topics. Currently, paediatric users of AAC interfaces minimally initiate communication spontaneously (Andzik et al., 2016) and generally cannot choose a topic of communication beyond that of the immediate activity or prompt (Ferm et al., 2005). Consequently, these children experience limitations in participation and inclusion in education and social interactions (Andzik et al., 2016; Ferm et al., 2005).

The communication partner interface presents the intention to initiate speech, the goal of the communication act, the effect of the communicator and completeness of phrase formation. In displaying these components of communication, the additional interface supports the communication partner in understanding the wishes of the AAC users and the context of communication. The value of such a communication-partner interface will be evaluated in future research.

Limitations exist in the experimental paradigm that impacts the statistical power of the data analysis, and generalizability of findings. The high variance in subjective measures (i.e., SUS, and NASA TLX) suggests a larger sample size may have yielded more conclusive findings.

This study presents an initial evaluation comparing the EID AAC interface to a single representation of current commercial AAC interfaces. It is important for future research to investigate the impact of the EID application to AAC interfaces for individuals who use AAC technology to supplement or replace their primary communication pathway. Although we discuss the expected implications of design decisions and empirical findings for AAC users and various access pathways, future research must consider populations from the target demographics for a comprehensive conclusion on the potential of EID application to AAC systems.

**Conclusion**

Barriers in AAC interfaces limit participation and inclusion for people who experience complex...
communication needs. This study empirically evaluated an AAC interface previously developed through the systematic and scientific design process of EID. Our findings suggest that the EID interface has potential to improve information transfer rate ($p < .01$), and mental workload ($p < .025$) compared to a commercially available AAC interface.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Brain Canada and the Kids Brain Health Network Training Awards in Developmental Neuroscience Research Competition.

ORCID iD

Kaela Shea https://orcid.org/0000-0002-9629-3121

References

Akacayaka, M., Peters, B., Moghadamfalahi, M., Mooney, A., Orhan, U., Barry, O., Erdogmus, D., & Fried-Oken, M. (2014). Non-invasive brain-computer interfaces for augmentative and alternative communication. IEEE Reviews in Biomedical Engineering, 7, 31–49. https://doi.org/10.1109/RBME.2013.2295097.

Andzik, N. R., Chung, Y.-C., & Kranak, M. P. (2016). Communication opportunities for elementary school students who use augmentative and alternative communication. Augmentative and Alternative Communication, 32(4), 272–281. https://doi.org/10.1080/07434618.2016.1241299.

AssistiveWare, B.V. (2019). Proloquo2Go. Retrieved March 7, 2019 from https://www.assistiveware.com/products/proloquo2go.

Baxter, S., Enderby, P., Evans, P., & Judge, S. (2012). Barriers and facilitators to the use of high-technology augmentative and alternative communication devices: A systematic review and qualitative synthesis. International Journal of Language & Communication Disorders, 47(2), 115–129. https://doi.org/10.1111/j.1460-6984.2011.00090.x.

Bennett, K. B., & Flach, J. (2019). Ecological interface design: Thirty years of refinement, progress, and potential. Human Factors, 61(4), 513–525. https://doi.org/10.1177/0018720819835990.

Beukelman, D., McGinnis, J., & Morrow, D. (1991). Vocabulary selection in augmentative and alternative communication. Augmentative and Alternative Communication, 7(3), 171–185. https://doi.org/10.1080/07434619111231275883.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. Pattern Recognition, 30(7), 1145–1159. https://doi.org/10.1016/s0031-3203(96)00142-2.

Brooke, J. (1996). SUS: A quick and dirty usability scale. Burns, C. M., & Hajdukiewicz, J. R. (2004). Ecological interface design. CRC Press.

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. Applied Ergonomics, 74, 221–232. https://doi.org/10.1016/j.apergo.2018.08.028.

Euerby, A., & Burns, C. M. (2014). Improving social connection through a communities-of-practice-inspired cognitive work analysis approach. Human Factors, 56(2), 361–383. https://doi.org/10.1177/0018720813494410.

Fern, U., Ahl´ensen, E., & Bj¨orck–kessons, E. (2005). Conversational topics between a child with complex communication needs and her caregiver at mealtime. AAC: Augmentative and Alternative Communication, 22(1), 19–41. https://doi.org/10.1080/074346104142331270507.

Garcia, P. A. G., Martin–Moncunill, D., Sánchez–Alonso, S., & Fermosho Garcia, A. 2014. A usability study of taxonomy visualisation user interfaces in digital repositories. Online Information Review 38(2):284–304.

Hajdukiewicz, J. R., & Vicente, K. J. (2002). Designing for adaptation to novelty and change: Functional information, emergent feature graphics, and higher-level control. Human Factors, 44(4), 592–610. https://doi.org/10.1518/00187202024969800.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA–TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology, 52(C), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hastie, T., Tibshirani, R., & Friedman, J. (2017). The elements of statistical learning data mining, inference, and prediction (2nd ed.). Springer.

Hoag, L. A., Bedrosian, J. L., McCoy, K. F., & Johnstown, D. E. (2008). Hierarchy of conversational rule violations involving utterance-based augmentative and alternative communication systems. Augmentative and Alternative Communication, 24(2), 149–161. https://doi.org/10.1076/aac.0840238288.

Holt, J., Bennett, K. B., & Flach, J. M. (2015). Emergent features and perceptual objects: Re-examining fundamental principles in analogue display design. Ergonomics, 58(12), 1960–1973. https://doi.org/10.1080/00140139.2015.1049217.

Kilgore, R., & St-Cyr, O. (2006). The SRK inventory: A tool for structuring and capturing a worker competencies analysis. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(3), 506–509. https://doi.org/10.1177/154193120605000362.

Light, J., & McNaughton, D. (2012). Supporting the communication, language, and literacy development of children with complex communication needs: State of the science and future research priorities. Assistive Technology, 24(1), 34–44. https://doi.org/10.1080/10400435.2011.684717.

Light, J., McNaughton, D., Beukelman, D., Fager, S. K., Fried–Oken, M., Jakobs, T., & Jakobs, E. (2019). Challenges and opportunities in augmentative and alternative communication: Research and technology development to enhance communication and participation for individuals with complex communication needs. Augmentative and Alternative Communication, 35(1), 1–12. https://doi.org/10.1076/aac.2018.1556732.

Light, J., Wilkinson, K. M., Thiessen, A., Beukelman, D. R., & Fager, S. K. (2019). Designing effective AAC displays for individuals with developmental or acquired disabilities: State of the science and future research directions. Augmentative and Alternative Communication, 35(1), 42–55. https://doi.org/10.1076/aac.2018.1558283.

Marinescu, A., Sharples, S., Ritchie, A., L´opez, T., McDowell, M., & Morvan, H. (2017). Physiological parameter response to variation of mental workload. Human Factors, 60(1), 51. https://doi.org/10.1177/0018720817733101.

Melroy, R. C., & Stanton, N. A. (2015). Ecological interface design two decades on: Whatever happened to the SRK taxonomy? IEEE Transactions on Human–Machine Systems, 45(2), 145–163. https://doi.org/10.1109/thms.2014.2369372.

Naikar, N. (2013). Work domain analysis. Taylor & Francis Group.

Ojala, M., & Garriga, G. C. (2010). Permutation tests for studying classifier performance. Journal of Machine Learning Research, 11(62), 1833–1863.

Origin Instruments. (n.d.). Orby switch. Retrieved March 4, 2021 from https://www.orin.com/access/orby/.
Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, M., Weiss, R., Vanderplas, J., Cournapeau, D., Varoquaux, G., Gramfort, A., Thirion, B., Dubourg, V., Passos, A., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. Journal of Machine Learning Research, 12(85), 2825–2830.

Rasmussen, J. (1983). Skills, rules, and knowledge: signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, SMC–13(3), 257–266. https://doi.org/10.1109/tsmc.1983.6313160.

Shea, K., St-Cyr, O., & Chau, T. 2021. Ecological design of an augmentative and alternative communication device interface. Journal of Cognitive Engineering and Decision Making, 15(4): 175–197. https://doi.org/10.1177/15553434211029530

Silva, J., Torres-Solis, J., Chau, T., & Mihailidis, A. (2008). A novel asynchronous access method with binary interfaces. Journal of NeuroEngineering and Rehabilitation, 5(1), 1–19. https://doi.org/10.1186/1743-0003-5-24.

St-Cyr, O., Jamieson, G. A., & Vicente, K. J. (2013). Ecological interface design and sensor noise. International Journal of Human Computer Studies, 71(11), 1056–1068. https://doi.org/10.1016/j.ijhcs.2013.08.005.

Tai, K., Blain, S., & Chau, T. (2008). A review of emerging access methods for individuals with severe motor impairments. Assistive Technology, 20(4), 204–211. https://doi.org/10.1080/10400435.2008.10131947.

Thistle, J. J., & Wilkinson, K. M. (2015). Building evidence-based practice in AAC display design for young children: Current practices and future directions. AAC: Augmentative and Alternative Communication, 31(2), 124–136. https://doi.org/10.3109/07434618.2015.1035798.

Thought Technology Ltd. (2018). BioGraph Infini. Montreal.

Tullis, T., & Albert, B. (2013). Measuring the user experience (2nd Edition). Elsevier.

Vicente, K. J (1999). Cognitive work analysis: Toward safe, productive, and healthy computer-based work. Lawrence Erlbaum Associates, Inc.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. IEEE Transactions on Systems, Man and Cybernetics, 22(4), 589–606. https://doi.org/10.1109/21.156574.

Appendix A

Trial Script

Hello, how are you doing today?
Welcome to Bloorview Café!
Would you like to order food? (yes/no)
Would you like a salad or a sandwich?
[Picked Salad]
Excellent! Would you like lettuce or spinach?
You can pick 4 vegetables to go into your salad
Would you like Vinaigrette dressing? (yes/no)
Would you like Caesar dressing? (yes/no)
OR
[Picked Sandwich]
Excellent! Would you like meat on your sandwich? (yes/no)
Turkey, chicken or ham?
What 4 vegetables would you like on your sandwich?

Would you like mustard? (yes/no)
Would you like to drink? We have water, wine, beer, juice or milk.
What size would you like? We have small, medium or large?
For dessert, we have cake, pie or ice cream – which would you like?
Great! That will be $8.50, would you like to pay by cash, debit or credit?

Appendix B

Script for Interface Introductions

Both interfaces will be operated using what we call single-switch access. You will indicate a selection with a press of the button when the desired section is highlighted. If scanning times-out, press the button to resume scanning.

1.1.1 Proloquo2Go

For this trial, please answer all questions using the interface.

In this interface you see boxes with words and pictures representing different words, or different categories of words. If the top right corner of the box is filled in, there is grammar support that will open upon box selection. Grammar support will provide different variation of the displayed words. Categories are represented with tab at the top left corner of the box. Selecting a category will show a screen with words of the category, as well as words that may be used alongside. Words that are similar have the same colour box outline.

Navigation between screens is possible through selection and a back button.

1.1.2 EID Prototype Interface

For this trial, please answer all questions using the interface.

In this interface you see an overview of the word layout: a ‘yes’ and ‘no’ button, as well 5 word categories. Examples of the words that are within each category are shown in each box. Selecting a category will populate the lower half of the screen with the words of each category.

Navigation between screens is possible through selection, and allowing a screen to times-out. When the screen times out it will revert to the word layout overview.
Kaela Shea: Kaela Shea is a PhD Candidate in Biomedical Engineering at the University of Toronto researching at Holland Bloorview Kids Rehabilitation Hospital in Dr. Tom Chau’s PRISM Lab. Her current research focuses on improving access to assistive communication. Kaela has a B.Eng. in Biomedical Engineering, MASc in Mechatronics Engineering, and a Collaborative Specialization in Public Health Policy.

Olivier St–Cyr: Olivier St–Cyr is an Assistant Professor, Teaching Stream in the Faculty of Information at the University of Toronto, in Toronto, Canada. He is the liaison for the User Experience Design concentration. His research interests lie in the areas of Human–Computer Interaction (HCI), Human Factors Engineering (HFE), and EID. Prior to joining the University of Toronto, he spent 8 years working in industry on HCI related projects. Olivier holds an Honours BA in Computer Science and Psychology, a MASc in Systems Design Engineering, and a PhD in Industrial Engineering.

Tom Chau: Tom Chau is Vice President of Research, Holland Bloorview Kids Rehabilitation Hospital, the Raymond Chang Foundation Chair in Access Innovations, and Professor in the Institute of Biomaterials and Biomedical Engineering at the University of Toronto. He holds a doctorate in pattern analysis and machine intelligence. His lab focuses on decoding the physiological manifestations of communicative intent among children and youth who are non-speaking.