AI supporting AAC Pictographic Symbol Adaptations

E.A. DRAFFAN[a,1], Mike WALD[a], Chaohai DING[a] and Yuanyuan Yin[a]
[a] *University of Southampton, UK*
ORCiD ID: E.A. Draffan https://orcid.org/0000-0003-1590-7556

Abstract. The phenomenal increase in technological capabilities that allow the design and training of systems to cope with the complexities of natural language and visual representation in order to develop other formats is remarkable. It has made it possible to make use of image to image and text to image technologies to support those with disabilities in ways not previously explored. It has opened the world of adaptations from one picture to another in a design style of a user's choosing. Automated text simplification alongside graphical symbol representations to enhance understanding of complex content is already being used to support those with cognitive impairments and learning difficulties. Symbol sets have become embedded within applications as dictionaries and look up systems, but the need for flexibility and personalization remains a challenge. Most pictographic symbols are created over time within the bounds of a certain style and schema for particular groups such as those who use augmentative and alternative forms of communication (AAC). By using generative artificial intelligence, it is proposed that symbols could be produced based on the style of those already used by an individual or adapted to suit different requirements within local contexts, cultures and communities. This paper explores these ideas at the start of a small six-month pilot study to adapt a number of open licensed symbols based on the symbol set's original style. Once a collection has been automatically developed from image to image and text descriptions, potential stakeholders will evaluate the outcomes using an online voting system. Successful symbols will be made available and could potentially be added to the original symbol set offering a flexible personalized approach to AAC symbol generation hitherto not experienced by users.

Keywords. artificial intelligence, pictographic symbols, cognitive impairment, augmentative and alternative communication, symbol adaptations

1. Introduction

Although it is thought that globally 0.05-1.55% of the population have a form of intellectual disability [1], the data around the various difficulties related to cognitive impairment has proved hard to capture worldwide. Many of those with these types of disabilities may also have written and spoken receptive and expressive language impairments requiring additional support to aid understanding, speech and literacy skills. The types of support may include assistive technologies such as simple message systems combined with some speech, gestures or signing, others will use text to speech for reading digital content aloud. Where there are more complex communication needs (CCN) the use of speech generating devices with pictographic symbols may be introduced to encourage communication, language and literacy [2]. Pictographic symbols can also be used alongside simplified text to help explain complex printed information or embedded within digital systems to enhance the understanding of online content such as web pages.

[1] Corresponding Author: E.A. Draffan, ead@ecs.soton.ac.uk

Augmentative and alternative communication (AAC) and literacy skill strategies tend to be personalized and customized to suit user needs, but this takes time and is not always easy to achieve if suitable pictographic symbols or images are not available and in particular if they do not suit local contexts, cultures and communities. There is also the issue of graphic design skills to create the symbols which cannot happen in a 'just in time' fashion without considerable cost and effort. AAC users tend to use one symbol set in early childhood and this is then developed to align with their needs over time. However, adults who have a sudden traumatic brain injury affecting cognition and/or loss of communication skills such as a stroke with aphasia, often respond better to photographs and symbols to suit their age and milieu, that are adapted to work with visual scene displays [3]. These usually represent daily life activities and emotions. This is where it is particularly hard to find appropriate freely available multicultural and multilingual representations.

This paper is the result of the beginnings of a short six-month project that still has two months to run. The aim has been to explore ideas around the use of generative artificial intelligence (AI) models for open license symbol-to-symbol style transformation based on a repository of freely available AAC pictographs. The importance of developing images in a style that matches already developed symbols allows users to recognize them more easily as part of their personalized symbol set. It has also been found that those who have cognitive disabilities may have difficulty understanding concepts that involve visual representations and remembering their meaning. However, when people are completing an action in an image they tend to find these images easier to understand. For example, a picture for 'baking' with a cake beside an oven works better when a person is putting the cake into the oven [4]. For that reason, as part of the following methodology, a symbol set that has representations of actions with people has been chosen for the adaptation of symbols.

2. Methodology

Initially an English core vocabulary of 100 commonly used words was chosen and matched to the visual image of a pictograph from the Mulberry symbol set. A visual description of each symbol was provided to test the type of prompt that might be needed to successfully create training data for symbol generation in the style of the symbol set. These words were selected as being examples of frequently used pronouns, verbs, adverbs, adjectives and prepositions plus greetings rather than concrete nouns that tend to be considered as fringe vocabulary that may be easier to portray such as a dog, man etc. During the pilot phase DALL·E 2[2] was used and each time a prompt was written the word 'symbol' was also added to provide an indication of the style required. Prompts were adapted as the images appeared until a reasonable representation of a potential symbol was created. The prompt for Figure 1 was a five-pointed pink star symbol.

Figure 1. illustrates how simple text prompts for nouns can produce relatively acceptable symbols e.g. star compared to a symbol in the Mulberry set (last star).

[2] https://openai.com/product/dall-e-2

This was not in the perfect style of Mulberry when one considers the width of the black outline and specific pink color based in the original symbol set schema. At this stage no image to image styling had been provided as training data and DALL·E 2 is not an open system for creating specifically personalized data training resources. Table 1 contains samples of symbols chosen from the Mulberry set with their initial visual description to start the process of training and to learn more about what makes a good prompt when the subject matter can be rather abstract. Making improvements to the prompts is ongoing and includes the need for additional adjectives, specific colors plus an indication of positioning and an awareness that symbols tend to have no backgrounds and need to have defined outlines.

Table 1. Example image descriptions to aid recognition acting as potential prompts.

Symbol Label	Image visual description	Part of Speech	Definition
go	front view of a man wearing grey trousers, black shoes and a pale blue top standing with an arrow going away from his body from left to right	verb	change location; move, proceed
good	left hand thumbs up	adjective	having desirable or positive qualities especially those suitable for a thing specified
good-bye / goodbye	back view of a man wearing grey trousers, black shoes and a pale blue top standing with an outstretched right arm held up with a hand waving	noun	a farewell remark
happy	front view of the face of a man with brown hair smiling	adjective	enjoying or showing or marked by joy or pleasure
have	palm of an upturned right hand with a small red ball in the middle	verb	have or possess, either in a concrete or an abstract sense
he	front view of two male faces and one man standing to the side wearing grey trousers, black shoes and a pale blue top with an arrow pointing down at his head	pronoun	the male person or animal being discussed
help	the back of two hands coming together but not touching	verb	assist, aid, give help or assistance; be of service
here	pale blue dotted outlined square with a black thin lined arrow pointing down to the top and a black outlined square in the background.	adverb	in or at this place; where the speaker or writer is
hi or hello	front view of the torso of man with brown hair wearing a red top with his right arm raised palm forward	noun	an expression of greeting
new	pale pink rectangle with thin black lines radiating out from the sides	adjective	not of long duration; having just (or relatively recently) come into being

As the co-design aspects of the project evolved, the participation of image design experts and AAC professionals supporting those with cognitive impairments and CCN, provided input. In particular they helped with the criteria for the evaluation of sample symbols. These included visual aspects such as: Does the symbol reflect the concept required and the label or gloss suggested? Are design features (for instance color, contrast, shape and outline) consistent with the original symbol set? Where there are combined symbols (such as an action with an arrow to denote direction, as is seen with

the word ‘go’ in Table 1) are the elements logical and based on the schema and style of the original symbol set? The process involved comparisons being made to ensure consistency as well as a good match (Figure 2).

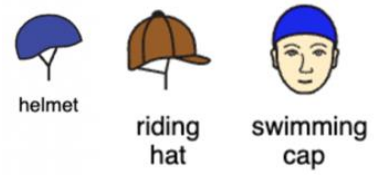


Figure 2. Comparison of Mulberry Symbols for head gear with one having a face and the others just showing the hat which may affect the training data

Those symbols that were a good fit with similar patterns and features as part of the symbol set schema were used (with their labels) as training data for the image recognition algorithms. Open AI models, such as Generative Adversarial Networks [5] and Stable Diffusion Models as [6] were used to investigate how style transformation and generation of AAC symbols can be supported by deep learning and unsupervised artificial neural networks. Uploading the pictographic symbols to the system to work alongside image prompts or visual descriptions will be part of the next stage in generating adaptations and testing the process.

Finally, as a quick and simple evaluation process due to time constraints a link to an online voting system, that had been developed for previous AAC symbol design projects [7] will be sent to five experts known to work in the field. The aim will be to evaluate the acceptance of the pictographic symbols that result from the use of the AI technologies. Five experts have been asked to look at each symbol and select a checkbox based on a 1-5 scale Likert scale, where 1 is 'completely unacceptable' and 5 is 'completely acceptable'. The criteria require a selection based on:

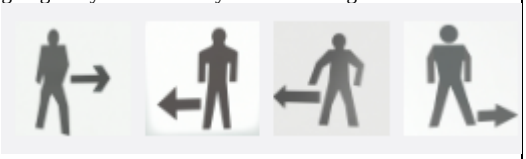
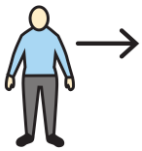
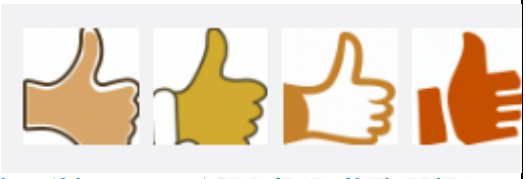
- feelings about the symbol as a whole (immediate reaction, iconicity, transparency, ease of recognition)
- represents the word or phrase (referent/concept)
- color contrast (outline and colors used for clarity and visually impaired)
- cultural sensitivity (not liable to offend, appropriate and relevant)

Participants will also be asked to add a general comment on ideas not mentioned in the list above, such as size or background etc.

3. Findings and Discussion

Initially attempts to make adaptations to symbols were based on text to symbol systems such as DALL·E 2 with short phrases or sentences used as prompts. It was found that even when the subject had adjectives and prepositions with the symbol guide most outcomes were unacceptable unless the symbol was a very simple object. So, where it was possible to have simple accurate phrases successes occurred, but not in the style of the Mulberry Symbols. There was also the complication that unless the system captured the fact that a symbol or outline was required, the image tended to appear in one color or came with a background (Table 2).

Table 2. Samples of text to symbol without the image to image support so lack the style input

go	front view of a man wearing grey trousers, black shoes and a pale blue top standing with an arrow going away from his body from left to right Sentence changed to A pictograph symbol of a man standing wearing grey trousers, black shoes and a pale blue top with an arrow going away from his body from left to right  https://labs.openai.com/e/eZLvmTpyxrt3IS9inkZerjkK	verb	change location; move, proceed Mulberry symbol 
good	left hand thumbs up Sentence changed to A pictograph symbol of a cream colored left hand thumbs up  https://labs.openai.com/e/KW1ifFa27cofdjJ5b2RVlTn2	adjective	having desirable or positive qualities especially those suitable for a thing specified Mulberry symbol 
good-bye / goodbye	back view of a man wearing grey trousers, black shoes and a pale blue top standing with an outstretched right arm held up with a hand waving Sentence changed to A pictograph symbol of a back view of a man wearing grey trousers, black shoes and a pale blue top standing with an outstretched right arm held up with a hand waving  https://labs.openai.com/e/tkIQuED0r0QP9datqtUlvvQQ	noun	a farewell remark Nothing in Mulberry so maybe reverse hello 

The Mulberry Symbol set has over 3,500 images, but within each category, such as parts of speech or topic categories such as professions, animals, buildings etc. the style, color and composition vary as mentioned in the methodology. Additional symbol sets could have been incorporated within the training data, but once again the variations in schemas or rules for design differ to such a degree that they could further skew results. Further investigation is needed to see how these effects can be ameliorated in the coming months and the results will be forthcoming at the conference.

Some applications using Stable Diffusion provide an interface that makes it possible to adapt images on an individual basis by using various parameters, examples

include img2img[3] and pix2pix[4]. This is achieved by controlling the amount of 'denoising strength' required when working out how much change is needed in order to make adaptations to an original symbol so that it can represent a different meaning or look. There is also the 'Classifier Free Guidance (CFG) scale' which is an indicator of how much the model should take account of the text prompt and finally it is possible to set the number of steps needed to achieve an acceptable image, but once again results are variable. It is possible to make several iterations but they do not necessarily achieve better results.

Figure 3. Mulberry symbol of a hand that needed to be adapted to a 'hand with a flat palm holding a small red ball' for the verb 'to have' which is not present in the symbol set.

There is the chance that by using concepts such as those described by Ruiz et al [8] it would be possible to solve some of the issues experienced in the early trials. The authors' Dreambooth application appears to not only use personal images, but also set them into different positions with objects, this could help with combination symbols such as a man mowing a lawn for 'to mow'. The examples from Dreambooth tend to use photographs and have backgrounds, but the latter can be removed and if the style if the Mulberry symbol can be picked up in an accurate way there could be a breakthrough for personalizing some symbols in an easier way.

4. Conclusion and Future Recommendations

At present the results indicate that the generation of open licensed symbols using image to image with supporting text prompts using AI models requires further work. It is possible to achieve basic concrete symbols from the stable diffusion models, but to develop acceptable results in the style of a particular symbol set is much harder. However, with time it is felt that improvements can be made to the proposed models so that AI can support AAC pictographic symbol adaptations. There may need to still be some human intervention to tweak the final versions for publication, but perhaps more simply than in the past. By the end of this project the intention is that this research will provide more informed future recommendations for the support of AI automated AAC symbol adaptations in a target style with the support of prompts.

Acknowledgements

The authors of this paper would like to thank the Web Science Institute Stimulus Fund at the University of Southampton for their support and all those collaborating on this innovative project especially Global Symbol colleague and the expert design and AAC specialists for their knowledge and evaluations.

[3] https://huggingface.co/spaces/fffiloni/stable-diffusion-img2img
[4] https://huggingface.co/spaces/timbrooks/instruct-pix2pix

References

[1] McKenzie, K., Milton, M., Smith, G. and Ouellette-Kuntz, H., Systematic review of the prevalence and incidence of intellectual disabilities: current trends and issues. Current Developmental Disorders Reports, 3, pp.104-115. (2016)

[2] Light, J., McNaughton, D. and Caron, J., New and emerging AAC technology supports for children with complex communication needs and their communication partners: State of the science and future research directions. Augmentative and Alternative Communication, 35(1), pp.26-41. (2019)

[3] Light, J., Wilkinson, K.M., Thiessen, A., Beukelman, D.R. and Fager, S.K., 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), pp.42-55. (2019)

[4] Thiessen, A., Brown, J., Beukelman, D., Hux, K. and Myers, A., Effect of message type on the visual attention of adults with traumatic brain injury. American journal of speech-language pathology, 26(2), pp.428-442. (2017)

[5] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B. and Bharath, A.A., Generative adversarial networks: An overview. IEEE signal processing magazine, 35(1), pp.53-65. (2018)

[6] Nichol, A.Q. and Dhariwal, P., Improved denoising diffusion probabilistic models. In International Conference on Machine Learning (pp. 8162-8171). PMLR. (2021)

[7] Draffan, E.A., Wald, M., Zeinoun, N. and Banes, D., Democratisation of AAC Symbol Choices Using Technology. In AAATE Conf. (pp. 355-362). (2017)

[8] Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M. and Aberman, K., 2022. Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. arXiv preprint arXiv:2208.12242. (2023) https://arxiv.org/pdf/2208.12242.pdf (Accessed 20 May 2023)