EMBEDDING EXPERT KNOWLEDGE: A CASE STUDY ON DEVELOPING AN ACCESSIBLE DIAGRAMMATIC INTERFACE

Stacy A. Doore
Colby College
sadoore@colby.edu

Justin K. Dimmel
University of Maine
justin.dimmel@maine.edu

Rose Xi
Bowdoin College
yxi@bowdoin.edu

Nicholas A. Giudice
University of Maine
nicholas.giudice@maine.edu

When students with blindness and visual impairment (BVI) are confronted with inaccessible visual graphics in the geometry classroom, additional instructional supports are often provided through verbal descriptions of images, tactile and haptic representations, and/or kinetic movement. This preliminary study examined the language used by instructional experts to describe geometry images to students with and without access to a visual instructional image. Specifically, we investigated expert descriptions of geometry diagrams for 1) spatial information, 2) instructional concept information, and 3) overall description structure (e.g., length, vocabulary, image part/whole order/relationships). We found that experts used nearly twice as many words to describe diagrams in the no visual access condition. We consider the double-edged nature of this result for supporting BVI learners in classrooms and chart possibilities for future research.

Keywords: Students with Disabilities, Geometry & Spatial Reasoning, Mathematical Representations

Introduction

There are approximately 12 million people with blindness or visual impairment (BVI) in the U.S., including over 600,000 school-aged children ages 5-18 (Erikson, Lee, von Schrader, 2021). Their success in school mathematics is hampered by inadequate accessible learning technologies and curricular resources, because representations such as graphs, tables, and diagrams are generally only available in visual formats. However, the information in such visual representations can be communicated through other perhaps equally valuable sensory channels (e.g., language, sound, haptics, tactile) (Giudice, 2018; Abrahamson, Flood, Miele, & Siu, 2019). The study reported here is one segment of a larger project whose purpose is to design an interface that will allow students with visual impairments to perceive geometry diagrams through a combination of audio and tactile sensory modalities. We ask: How can geometry diagrams be apprehended through non-visual sensory modalities? As a first step toward investigating this question, we report a case study of how expert users of geometry diagrams used *natural language (NL)* (Herzog & Wazinski, 1994) to describe a core set of 2D images.

Theoretical Framework

We conceptualized geometry diagrams from a systemic functional linguistics (SFL) perspective (O'Halloran, 2005), within which diagrams are mathematical texts that use spatial (e.g., orientation, size, position) and graphical (e.g., strokes, congruence markings, labels) resources systemically to communicate mathematical concepts, properties, and relationships (Dimmel & Herbst, 2015). One technique for representing diagrams with non-visual modalities is through natural language descriptions. Natural language is a term used in psychology, linguistics, and computer science for the communication and representation of any language that has evolved naturally in humans through use and diffusion (Lyons, 1981). NL descriptions of 2D diagrams are used in studies of spatial reasoning and patterning development (Clements, 2004).

Methods

Protocol design

Instructional experts were presented a protocol that consisted of five horizontally oriented digital pictures, each representing a different geometric diagram. All the images were chosen to provide typical examples of geometry diagrams. The geometry diagram prompts are provided in Figure 1.

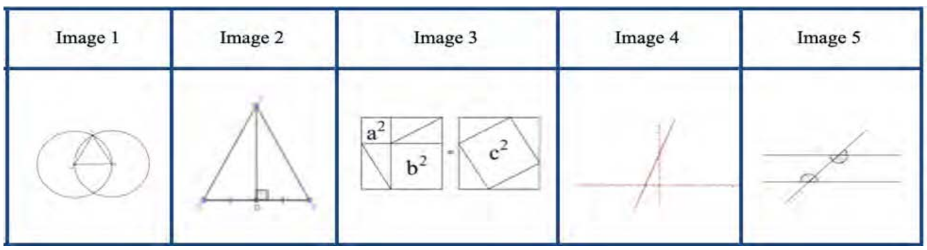


Figure 1: Diagrams used as prompts for expert NL descriptions

As a set, the diagrams were selected as realizations of a range of fundamental geometric relationships, including: *congruence* (images 1, 2, 3, and 5), *parallelism* (images 3 and 5), *perpendicularity* (images 2 and 3), *segment-angle relationships* (images 1and 2), and *incidence relationships* (each image). Images 1, 2, and 5 are examples of diagrams that might accompany typical proof problems (Herbst et al., 2009). Image 3 was selected as an example of a *proof without words,* and Image 4 was selected as a primitive example of a coordinate geometry diagram. The relationships described above are not exhaustive of the geometric concepts realized in the set of images, nor was the set of images intended to serve as a comprehensive set of primitive diagrams that would account for all possible diagrammatic variations. Rather, these images were selected because they entailed sufficient variation to provide a starting point for taking stock of how geometric concepts realized in diagrams could be described using natural language. Participants were given the following instructions and asked four questions for each of the diagrams used in the protocol:

Please briefly review the image and respond to the questions below for each image in the protocol. Record your responses to the questions for each image on a separate recording file.

- How would you describe this image to a student in your class?
- How would you describe the image to a student who you are talking to over a phone or who was listening to a podcast, who cannot see the image?
- What are the most significant mathematical concepts, relationships, or features that need to be described in this image?
- How would you convey the mathematical concepts, relationships, or features to a student who you are talking to over the phone who cannot see the image or to someone listening to a podcast?

Olanoff, D., Johnson, K., & Spitzer, S. (2021). *Proceedings of the forty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Philadelphia, PA.

The question prompts specifically did not mention that the student who could not see the image had a vision impairment so as not to bias the expert into focusing on the student's disability, as opposed to the instructional concepts that should be explained to any student who could not directly see the image that was the focus of instruction.

Participants, data, analysis

Four experienced mathematics teachers (university or secondary;1/F, 3/M; mean age: 45.8; mean years of experience: 12.3) completed the remotely administered study. Participants recorded their responses individually and at their convenience. The experts spent approximately 30 minutes to audio record their descriptions of 5 geometry diagrams. They were permitted to ask clarifying questions via email, if necessary, but none of the experts chose to stop the recording to ask follow-up questions.

The analysis of the description data used an NLP pipeline that included the processing of raw image descriptions and an annotation process based on techniques used in Cognitive Discourse Analysis (CODA) (Tenbrink, 2015). The annotation process coded specific elements of the image description such as the whole image description, the description of a specific part of the image, as well as identifying themes and relevant features in each *segment*, which was defined as "one coherent statement about a single item/space/topic." (Suwa & Tversky, 1997; Cialone, Tenbrink, & Spiers, 2017). An annotation review was then conducted segment by segment, counting the occurrences of each linguistic feature representing a specific annotation category. The relative frequencies were calculated for each image in relation to the overall number of words produced by each expert, the total number of words produced across experts, and the total identification of instructional or spatial concepts.

Results

80 raw image descriptions (4 experts x 5 diagrams x 4 description questions) were processed and analyzed for this case study. We present here the general language patterns found across expert descriptions based on several text-analysis metrics: Total number of words (n) and average number of words(M) across all expert descriptions for each image, variety of word types (unique) or range (R) of total words used by the experts, and average sentence length (M words per sentence) based on student visual access or non-visual access to each graphic. These aggregate descriptives are reported, by diagram (columns) and condition (rows), in Figure 2.

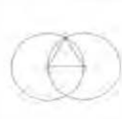
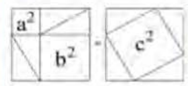
	Image 1	Image 2	Image 3	Image 4	Image 5
					
Visual Access	n= 293 words Unique =116 M= 73.25 R= 149 (18-167) M density = .73 M words per sentence = 21.4	n= 395 words Unique =138 M= 98.75 R= 101 (50-151) M density = .64 M words per sentence = 21.4	n= 1127 words Unique = 244 M= 108 R= 52 (77-129) M density = .39 M words per sentence = 23.8	n= 482 words Unique = 148 M= 120 R= 78 (87-165) M density = .53 M words per sentence = 14.27	n= 468 words Unique =155 M= 117 R= 103 (69-172) M density = .54 M words per sentence = 25.05
Non-Visual Access	n = 605 Unique = 189 M = 151.25 R= 188 (47-235) M density = .49 M words per sentence = 34.2	n = 560 Unique = 191 M = 186.66 R= 188 (63-328) M density = .55 M words per sentence = 25.8	n= 1880 words Unique = 342 M= 470 R= 511 (155-666) M density = .33 M words per sentence = 21.47	n= 412 words Unique = 170 M= 103 R= 23 (94-117) M density = 61.76 M words per sentence = 23.25	n= 1382 words Unique = 271 M= 345 R= 592(147-739) M density = 129.75 M words per sentence = 20.4

Figure 2: Summary statistics for expert description language use across the images.

Number of words in descriptions

The results reported in Figure 2 show that instructional experts used approximately two times the number of total words (n) in the non-visual access condition, including more unique words – i.e., words that were used only once in the description. This may suggest sighted instructional experts are using different description strategies (e.g., longer descriptions, a larger more diverse vocabulary set, etc.) for students without visual access to the image. This interpretation is consistent with the findings of a study using NL scene descriptions created by sighted participants for non-visual users (Doore, 2017).

Discussion

The primary findings from our analysis are that experts consistently used more words, sometimes as many as double, when describing diagrams for students without visual access to the diagram when compared to students with visual access to it. A critical question raised by these results is: *Do longer, more detailed natural language descriptions of diagrams result in more effective mental representations and conceptual comprehension for BVI students?* Future goals of this research include developing a controlled vocabulary to eventually generate automated descriptions for diagrams to be incorporated into a remote multimodal learning system that uses haptic, auditory, and NL supports for meeting the needs of BVI students in accessing diagrammatic information. The ultimate goal of the larger body of research is the development of a multimodal system that can act as an accessible personal learning environment (PLE) (Martindale & Dowdy, 2010) to support more BVI students to develop the academic skills and personal interests to enter the STEM professional pipeline as well as to increase their post-secondary attainment and employment success. This type of personal learning environment may someday provide a remotely- accessible, cloud-based system that will allow BVI students to direct their own learning.

Acknowledgments

The research reported here is supported by a *Cyberlearning for Work at the Human-Technology Frontier* grant (NSF solicitation 17-598) entitled "A Remote Multimodal Learning Environment to Increase Graphical Information Access for Blind and Visually Impaired Students (Award No: 1822800, PI: Giudice; CoPIs: Dimmel and Doore)". The views expressed in this report are the authors' and do not necessarily reflect the views of the National Science Foundation.

References

Abrahamson, D., Flood, V. J., Miele, J. A., & Siu, Y. T. (2019). Enactivism and ethnomethodological conversation analysis as tools for expanding Universal Design for Learning: The case of visually impaired mathematics students. *ZDM*, *51*(2), 291-303.

Bird, S., Klein, E., and Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media. ISBN 978-0-596-51649-9.

Cialone, C., Tenbrink, T., & Spiers, H. J. (2018). Sculptors, architects, and painters conceive of depicted spaces differently. *Cognitive science*, *42*(2), 524-553.

Clements, D. H. (2004). Geometric and spatial thinking in early childhood education. In D. H. Clements, J. Sarama, & A. M. Di Biase (Eds.), Engaging young children in mathematics: standards for early childhood mathematics education (pp. 267–298). Mahwah: Lawrence Erlbaum.

Dimmel, J. K., & Herbst, P. G. (2015). The semiotic structure of geometry diagrams: How textbook diagrams convey meaning. *Journal for Research in Mathematics Education*, *46*(2), 147–195. doi:10.5951/jresematheduc.46.2.0147

Doore, S. A. (2017). Spatial Relations and Natural-Language Semantics for Indoor Scenes.

Erickson, W., Lee, C., von Schrader, S. (2021). Disability Statistics from the 2018 American Community Survey (ACS). Ithaca, NY: Cornell University Yang-Tan Institute (YTI). Retrieved from Cornell University Disability Statistics website: www.disabilitystatistics.org

Giudice, N. A. (2018). Navigating without vision: principles of blind spatial cognition. In Montello, D. (ed.). *Handbook of behavioral and cognitive geography*. Edward Elgar Publishing.

Herzog, G., & Wazinski, P. (1994). VIsual TRAnslator: Linking perceptions and natural language descriptions. *Artificial Intelligence Review*, *8*(2-3), 175-187.

Lyons, J. (1981). *Language and linguistics*. Cambridge University Press.

Martindale, T., & Dowdy, M. (2010). Personal learning environments. Emerging technologies in distance education, 177-193.

O'Halloran, K. L. (2005). *Mathematical discourse: Language, symbolism, and visual images*. London, United Kingdom: Continuum.

Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. *Design studies*, *18*(4), 385-403.

Tenbrink, T. (2015). Cognitive discourse analysis: Accessing cognitive representations and processes through language data. *Language and Cognition*, *7*(1), 98-137.