

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Enhancing Cognitive Assessment through Multimodal Sensing:A Case Study Using the Block Design Test

Permalink

<https://escholarship.org/uc/item/8bd948mg>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors

Cha, Seunghwan

Ainooson, James

Chong, Eunji

et al.

Publication Date

2020

Peer reviewed

Enhancing Cognitive Assessment through Multimodal Sensing: A Case Study Using the Block Design Test

Seunghwan Cha^{*1}, James Ainooson^{*1}, Eunji Chong², Isabelle Soulières³, James M. Rehg²,
and Maithilee Kunda¹ (mkunda@vanderbilt.edu). ^{*}These authors contributed equally to this work.

¹Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA.

²School of Interactive Computing, Georgia Tech, Atlanta, GA, USA.

³Department of Psychology, University of Quebec at Montreal, Quebec, Canada.

Abstract

Many cognitive assessments are limited by their reliance on relatively sparse measures of performance, like per-item accuracy or reaction time. Capturing more detailed behavioral measurements from cognitive assessments will enhance their utility in many settings, from individual clinical evaluations to large-scale research studies. We demonstrate the feasibility of combining scene and gaze cameras with supervised learning algorithms to automatically measure key behaviors on the block design test, a widely used test of visuospatial cognitive ability. We also discuss how this block-design measurement system could enhance the assessment of many critical cognitive and meta-cognitive functions such as attention, planning, progress monitoring, and strategy selection.

Introduction

The meat of the matter is often how a patient solves a problem or approaches a task rather than what the score is.

(Lezak et al., 2012, *Neuropsychological Assessment*, p. 160)

Consider a cognitive assessment like the block design test (BDT), as shown in Figure 1 (top), in which a person has to reconstruct a given visual design using red and white blocks. Suppose you administer the BDT to two participants, and they both get the same score in terms of accuracy (items correctly built) and reaction time. Do these scores imply that your two participants have similar visuospatial cognitive ability?

Now suppose you watch each participant as they perform the test. The first participant methodically places each block exactly once, making no errors as they complete each item. The second participant (who is working at a much more frenetic pace—same per-item reaction time, remember) places each block many times, continually checking and changing and re-checking each placement, but finally obtaining with the correct answer in the end. Now would you say that your two participants have similar visuospatial cognitive ability?

We might quibble about the meaning of the word “ability” here, but certainly there are significant cognitive differences between the two participants, however we might label them.

Invariably, when a cognitive assessment boils a person’s behavior down to one or two scores, there is information being lost about *how* that person performed the assessment (Milberg et al., 2009; Poreh, 2012; Kunda, 2019). Of course, a person’s cognitive processes are not directly observable. However, we can get clues about these processes by obtaining fine-grained observations about a person’s externally observable behaviors while they are taking the assessment. Examples of such behaviors include: patterns of eye gaze; reaction

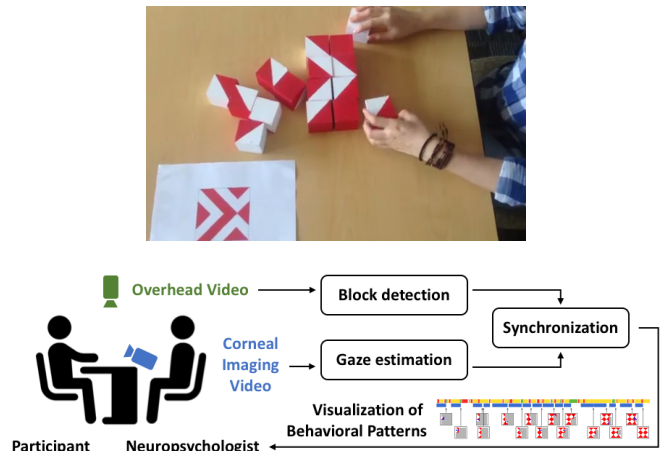


Figure 1: Top: Solving a BDT-like item. Bottom: Multimodal sensing to measure significant behaviors on the BDT.

times (i.e. per-item or even within-item); types of errors; language (e.g. talking to oneself while solving an item); and even affective or physiological characteristics like facial expression, heart rate, and skin conductance.

Many research efforts aim to measure fine-grained behaviors on cognitive assessments using computerized setups that can easily record reaction times for every key press, mouse movement logs, and eye gaze data. However, many important cognitive tests are not amenable to being computerized, including (but not limited to) tests in the domain of visuospatial reasoning that draw heavily upon a person’s motor functions and/or abilities to reason about physical objects.

Here, we examine the feasibility of using multimodal sensing to obtain detailed behavioral measurements from a non-computerizable cognitive assessment, namely the block design test (BDT). Our contributions are: (1) We show how data from an overhead camera plus a corneal-image-based gaze measurement system, with manual annotations, suffice for capturing a comprehensive record of a person’s BDT performance. (2) We demonstrate that standard supervised learning algorithms can be used to classify a person’s block placements (>95% accuracy) and gaze targets (~70% accuracy). (3) We provide examples from our participant study of the kinds of process-level observations enabled by our multimodal measurement approach.

Why Block Design Is a Good Test Case

The block design test (BDT) is a widely-used assessment of visuospatial cognitive ability. The BDT is found on many standardized IQ tests and has been used to study learning disabilities like dyslexia, neurodevelopmental conditions like autism, general child development, cognitive decline during aging or after stroke, cultural differences in cognition, relationships between spatial ability and STEM learning, etc.

In its standard format, the BDT is scored in terms of accuracy and reaction time for each completed target design. However, many research studies have observed that subtle patterns of behavior on the BDT provide intriguing clues about a person's cognitive processes.

For example, patterns of gaze between the target design and the block construction area can indicate individual differences in strategy (Hoffman et al., 2003; Rozenchwajg & Corroyer, 2002; Rozenchwajg et al., 2005; Rozenchwajg & Fenouillet, 2012). The errors a person makes while solving BDT items have been studied in terms of particular sequences of block placements (Joy et al., 2001; Rozenchwajg & Corroyer, 2002; Toraldo & Shallice, 2004), incorrect placements of blocks (Ben-Yishay et al., 1971; Hoffman et al., 2003; Jones & Torgesen, 1981; Joy et al., 2001; Schatz et al., 2000; Troyer et al., 1994), and qualitative types of errors (Akshoomoff et al., 1989; Akshoomoff & Stiles, 1996; Joy et al., 2001; Kramer et al., 1991, 1999; Schatz et al., 2000; Troyer et al., 1994; Zipf-Williams et al., 2000).

Despite the known value of measuring a person's patterns of gaze, block placements, errors, etc., on the BDT, such information is rarely collected in practice, mainly due to the difficulty of recording such data accurately and in real time while also administering the test (Milberg et al., 2009). In fact, the original BDT scoring system published in 1920 included tallying the number of block placements made, (Kohs, 1920), but by 1932, this method was deemed too cumbersome for practice and was dropped (Hutt, 1932).

For all of these reasons, the BDT is an ideal test case on which to evaluate our multimodal sensing approach.

System Design and Task Setup

We designed our system to use two sensors. First, an overhead camera recorded a top-down view of the table used for test administration; these images were used for detection of block placements. Second, a corneal imaging system recorded images of the world reflected in the cornea of the participant; these images were used for estimation of gaze targets. Due to the limited field of view of our corneal imaging system, participants were asked to use a chin rest to eliminate large head movements. The system and task setup is illustrated in Figure 2.

We collected data from undergraduate computer science students ($n = 7$) who were not members of our research team. All necessary IRB approvals were obtained for this study. Each participant in our study was asked to complete 17 different block design items. Of these items, 13 were taken from

the standard Wechsler block design test and the remaining 4 were slightly more complex designs created for our study. All analyses reported in this paper use only the final 6 out of 17 items, as these were the most difficult and showed the greatest variability in participant performance.

Block Detection from Overhead Camera

The overhead camera was a Kinect RGB-D camera positioned near ceiling height. We did not use depth information for the analyses presented in this paper, though the inclusion of depth information could help with automated detection of hands and blocks in future work.

Overhead videos were manually annotated using the ELAN software tool. Annotations included individual block placement locations and block faces (e.g., empty, white, red, etc.). Annotations also included transition periods during which any block was in motion. For our pilot study, annotations were completed by a single annotator, as there were no significant issues with ambiguity in labeling.

Automated Block Detection: Methods

As shown in Figure 3, our system for automated block detection took overhead videos as input and produced frame-level block placement labels as output.

Hand detection and smoothing. Based on initial experiments, we determined that a key source of noise for the problem of frame-level block placement detection was participants' hands and/or the blocks they were carrying occluding portions of the construction area. To address this issue, we experimented with methods for automated detection of hands so that we could filter out these frames. Among the many available options, we chose to use the Single Shot Multibox Detector (SSD) MobileNet network available from the Tensorflow

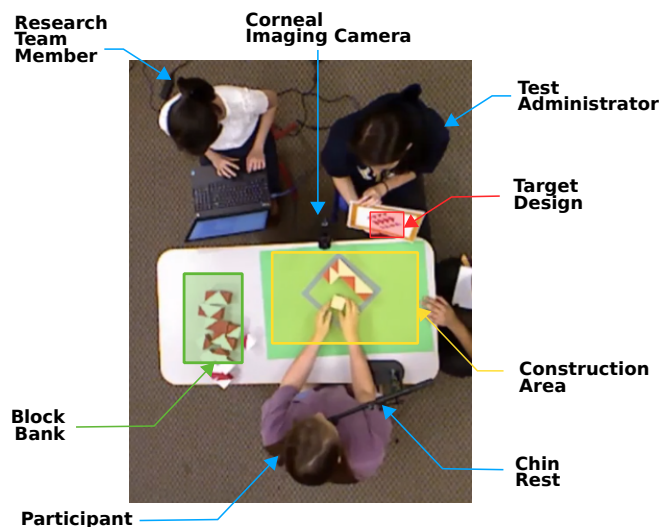


Figure 2: System setup. View shown is from overhead camera, which records information for measuring block placements. Corneal imaging camera, located across table from participant, records information for measuring gaze.

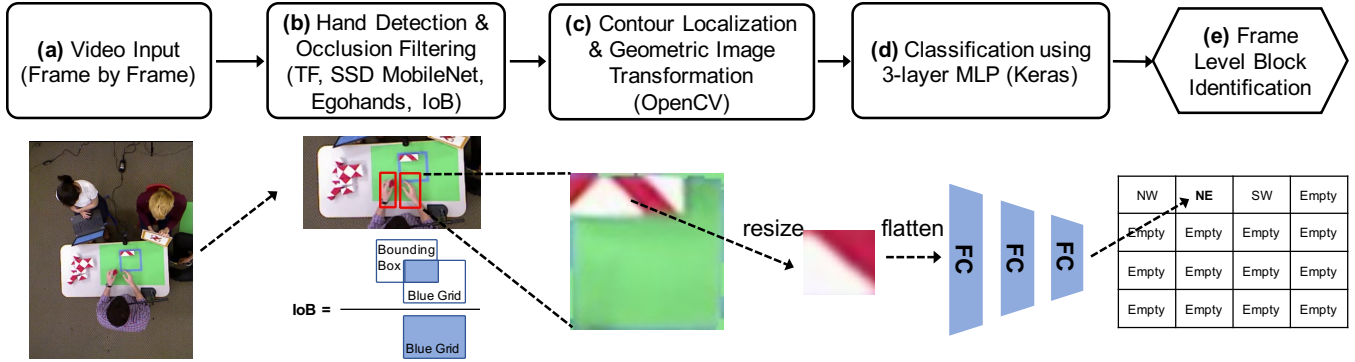


Figure 3: Automated detection of block placements using overhead video. **(a)** Inputs are individual frames. **(b)** Regions occluded by participants’ hands are filtered out through hand detection using the Tensorflow Object Detection API (Huang et al., 2017). The SSD Mobilenet network (Liu et al., 2016) was pre-trained on COCO and re-trained for hand detection on the Egohands Dataset (Bambach et al., 2015). Frames having Intersection over Blue (IoB) greater than 0.3 were filtered out. **(c)** Blue contour was localized and rectified through OpenCV geometric image transformation functions (Bradski, 2000). **(d)** Rectified image of blue contour area is divided into $n \times n$ block-sized sub-images, which are fed through neural network classifier to obtain the final block label, as shown in **(e)**.

Object Detection API (Huang et al., 2017), pre-trained on the COCO dataset (Liu et al., 2016) and then re-trained for hand detection using the EgoHands dataset (Bambach et al., 2015).

Then, for frames with a large overlap of hands over the construction area, we set the current frame’s block labels to be equal to those of the previous frame, i.e., a smoothing operation to interpolate over frames with occlusion from hands.

Locating the construction area. Next, we identified the construction area in each frame. This area was outlined with colored tape on the tabletop to simplify the vision processing in this initial pilot study. In future studies, we will remove this simplification, as it is not a standard procedure for the BDT. Though the overhead camera was fixed in our study, movements of the physical table and of the green tabletop sheets meant that the position of the construction area could (and often did) change from item to item.

The blue contour was located using blue HSV thresholds and processed to correct for rotations using standard transformations from the OpenCV library (Bradski, 2000). Finally, the cropped and rotated image of the construction area was divided into $n \times n$ sub-images, depending on the size of the given block design test item (either 3×3 or 4×4). Each sub-image was then fed into a classifier, as described below.

Block classification. We explored several techniques to classify each block sub-image into categories identifying the top-most block face and its orientation.

RGB Averaging. Each block-sized sub-image was first divided into four diagonal quadrants. Then, each quadrant’s pixel values were averaged in each respective RGB channel and compared with a threshold value of 140 to determine the quadrant color. The threshold was set empirically based on initial experiments. Identifying the color within a segment of each quadrant uniquely determines the overall block label.

K-Means Clustering. We clustered pixel values within each color channel to acquire the dominant RGB value in each

quadrant (Kanungo et al., 2000). We tested this approach with $K = 1$ and $K = 4$, and kept the threshold of 140.

K-Nearest Neighbors (KNN). We used synthetic red, white, and green color images (10 each) as training data for this approach. Then, color histograms of each color channel and the corresponding color were recorded to train a KNN classifier (Cover & Hart, 1967).

Multi-Layer Perceptron. We also trained a Multi-Layer Perceptron (MLP) for block label classification. Data instances were flattened to a 1D tensor and was fed into a three-layer fully connected MLP. Dropout was inserted in between every layer to reduce overfitting. ReLU was used as the activation function for the first two layers, and a softmax function was used for the last fully-connected layer. The network was optimized using Stochastic Gradient Descent with an initial learning rate of 0.01. Training proceeded for 10 epochs.

Results. Table 1 shows accuracy across all of our methods. For all methods except MLP, results are reported as accuracy over all frames from the dataset. For MLP, accuracy is obtained by 7-fold cross validation in which all of the data from a single participant was held out for evaluation during each fold. Clearly, the data-driven MLP approach performs best.

Table 1: Accuracy results for different approaches.

Category	Method	Accuracy
Color Determination	RGB Averaging	0.68
	K-Means Clustering (k=1)	0.68
	K-Means Clustering (k=4)	0.64
	K-Nearest Neighbors	0.67
Hand Filtering	Block in Motion	0.73
	Hand Bounding Box	0.69
Post Processing	Smoothing with RGB Averaging	0.81
ML-based	MLP (+ hand bounding box)	0.96



Figure 4: Sample corneal images from our study. Dark eyes (left) produce reflections across the entire iris, while lighter eyes (right) produce reflections just over the pupil.

Gaze Estimation Using Corneal Imaging

Eye gaze is another behavioral measure that would be useful to obtain on the block design test (BDT). Because the BDT involves manual interactions with physical objects, we cannot use standard monitor-based eye trackers to measure gaze. For such tasks, head-mounted gaze trackers are currently the primary, commercially available option. While head-mounted gaze trackers are valuable in many settings, special usability concerns can arise in applications involving children or people with sensory sensitivities (Sasson & Elison, 2012).

Thus, we investigated gaze tracking using *corneal imaging*, which captures images of the world reflected in a person’s cornea (Chong et al., 2017; Nakazawa & Nitschke, 2012; Nishino & Nayar, 2006). Corneal imaging requires relatively little calibration, and the image provides a wide field of view that moves along with the participant’s head. Most importantly, participants do not have to wear any equipment.

For our study, we used a single high resolution camera equipped with a lens that had a narrow field of view and a shallow depth of field. A second wide angle depth camera was used to estimate the distance between the eye and the corneal imaging camera, to enable auto-focusing. To reduce large head movements, we asked participants to use a chin rest. In future studies, a bank of corneal image cameras could instead be used to capture a wider field of view.

Videos of the corneal image recordings were annotated with the ELAN Annotation Tool by members of our research team. The annotation involved marking intervals of frames with the *gaze target* (block bank, construction area, target design, or “other”), *blink state* of the eye (blink vs. non-blink), and *gaze event* (saccade or smooth pursuit).

Unlike annotations of the overhead video to identify block placements, the corneal imaging annotations were both time consuming and subject to ambiguity. Among four raters, reliability estimates for various BDT items, calculated using Cohen’s kappa, ranged from 0.66 to 0.80.

We expect that reliability could be improved in several ways. First, while annotations were defined as time intervals over video, reliability was computed frame by frame, and so a more appropriate reliability metric should also use intervals. Second, our annotation process unfolded in mostly a feed-forward fashion; we did not yet have annotators meet to discuss and resolve discrepancies. With these modifications, it is likely that videos from a corneal-imaging-based system could be annotated to provide more reliable gaze estimates.

Table 2: Results for automated gaze classification. Corneal images were classified as blink vs. non-blink. Non-blink images were further classified according to gaze target, i.e., block bank, construction area, target design, or “other.” Results are reported as proportion of frames correctly classified.

Image-based blink classification	Accuracy	Precision
Within-participant, across BDT item	0.81	0.59
Across participants	0.65	0.41
Image-based gaze classification		
Within-participant, across BDT item	0.69	0.55
Across participants	0.51	0.45
Geometric gaze classification with KNN		
Within-participant, across BDT item	0.77	0.66
Across participants	0.61	0.54
Geometric gaze classification with MLP		
Within-participant, across BDT item	0.71	0.41
Across participants	0.64	0.37

Automated Gaze Estimation: Methods and Results

We explored two different methods for automatically estimating gaze targets from our corneal imaging videos.

Method 1: Using image information with a CNN. We used a pre-trained convolutional neural network (CNN) on the images obtained directly from the corneal imaging camera. The images fed to the classifier were not unwrapped and the other visible parts of the eye in the frame remained intact.

Two CNNs were trained: one for predicting gaze target locations and one for detecting blinks. All classifiers took in the raw image pixels as captured from the corneal imaging camera without any modifications. A pre-trained copy of the Inception v3 (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016) convolutional neural network running with Tensorflow v1.9 (Abadi et al., 2016) was used as the classifier.

Method 2: Using geometric information. Next, we directly considered the geometric information extracted when the ellipse model was fit onto the frame, using an existing ellipse fitting approach (Chong et al., 2017). The ellipse parameters consists of five separate values which represent the geometry of the ellipse fit onto the limbus in the eye frame. For this approach, we used corneal image data from just two participants; future work will incorporate additional data.

We fed ellipse parameters directly into the classifiers as feature vectors. We evaluated two classifiers: k-nearest neighbours (KNN), with $k = 5$, and multi layer perceptrons (MLP), with three hidden layers having 50, 100 and 50 nodes respectively. The KNN implementation was obtained from the scikit-learn (Pedregosa et al., 2011) toolkit and the MLP implementation was from Tensorflow (Abadi et al., 2016).

Results. We evaluated all classifiers with two different training and testing approaches: 1) a within-participant, across-BDT-item leave-one-out approach, and 2) an across-participant leave-one-out approach. Results are summarized

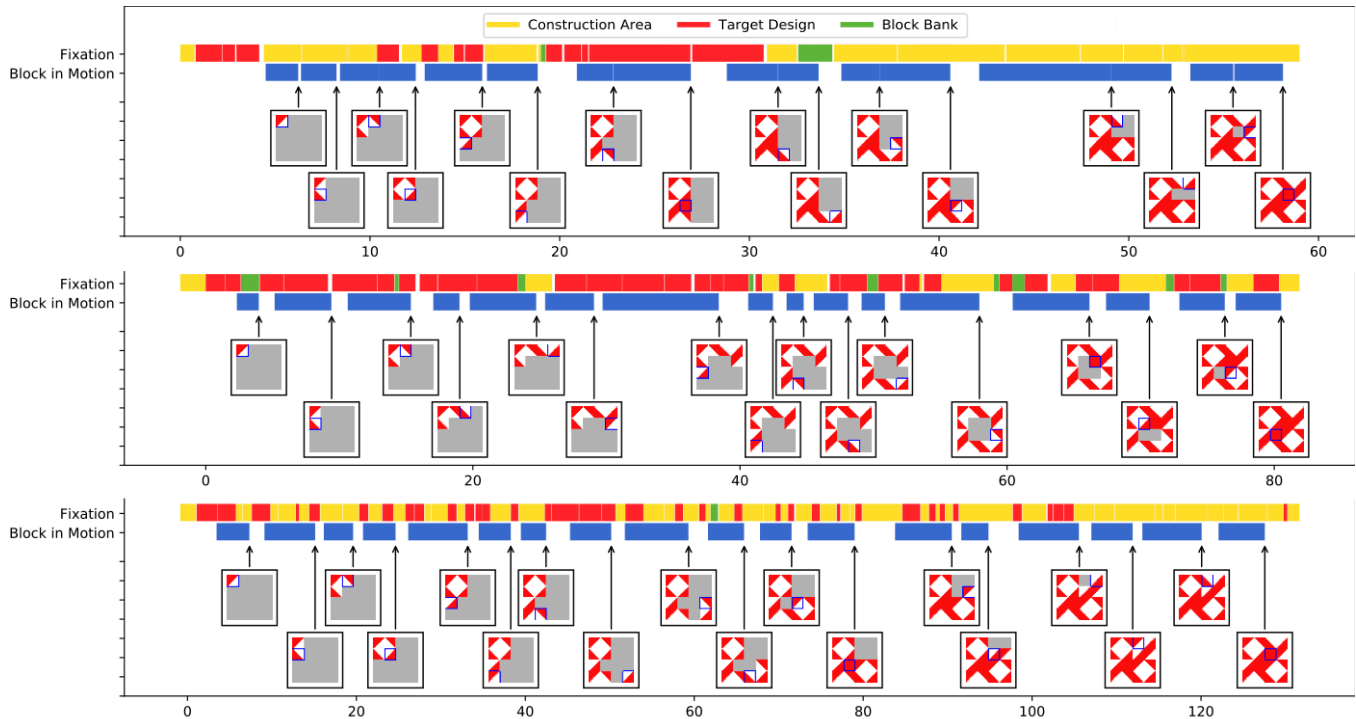


Figure 5: Sample visualizations from the same BDT item across three participants in our study.

in Table 2. While evaluating within-participant yielded better results, testing across participants more closely matches likely real-world applications in terms of generalizability.

Synchronization

In order to combine block placement data and gaze data, the video feeds from both cameras were first synchronized. The corneal imaging camera ran at about 15 frames per second, and the overhead camera ran at 30 frames per second.

Synchronization was enabled by using clapper boards during the study. Before each BDT item was attempted, a clapper board was clapped two or three times over the construction area while the participant looked on. (This was done for each individual BDT item because the corneal imaging camera was set up to make separate recordings for each.) Synchronization was done by matching the frames where the clapper hits in both video streams and dropping excess frames from the faster overhead video feed. Figure 5 shows manual annotations from the two sensor streams plotted together.

Discussion: Examples of Observations

There are many interesting observations that can be made using the outputs from our system, e.g., from the types of visualizations shown in Figure 5. (Note that because our study recruited a fairly homogeneous sample of undergraduate CS students, the variations we see here likely underestimate the magnitudes and types of individual differences we would see in a more general population.)

For example, the first and third participants complete the

design using a four-quadrant spatial pattern, while the middle participant uses an outside-to-inside spatial pattern. About halfway through, the top participant stops looking at the target design altogether, which suggests effective recruitment of working memory to remember the design for the rest of the trial; this participant is also the fastest to complete the design.

Additional questions that a researcher or clinician might query from this kind of BDT administration include:

- Does a participant show consistency in strategy across items, or do they switch, especially if they are having a hard time completing items with their initial strategy?
- If a participant makes an error, how soon do they detect it?
- Does a participant look at the target design to verify and get feedback on what they have just done? Could progress monitoring be taught as a new strategy?
- How long does a participant spend planning before making their first block move?

Furthermore, in addition to human-generated observations, this approach opens the doors for using machine learning and data mining techniques to discover new behavioral patterns that might be significant for many types of research, including on learning or neurodevelopmental conditions, education, foundational psychology and cognitive science, and more.

In conclusion, we have shown feasibility of a multimodal sensing system for measuring detailed behaviors on the block design test (BDT), including hardware design and setup, initial results from automated measurement algorithms, and examples of “cognitively significant” observations. Future work

will include removing simplifying assumptions from our task setup, working to increase the accuracy of our automated measurement algorithms, and investigating the use of data mining to discover interesting behavioral patterns in data collected from various populations.

Acknowledgments

We thank Mohamed El-Banani, Yongkoo Kang, Mika Munch, Emeke Nkadi, and Richard Stauffer for contributions to this project in designing test stimuli and helping to carry out the participant study. This research was supported by NSF awards #1029679 and #1730044.

References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., . . . others (2016). Tensorflow: a system for large-scale machine learning. In *OsdI* (Vol. 16, pp. 265–283).
- Akshoomoff, N. A., Delis, D. C., & Kiefner, M. G. (1989). Block constructions of chronic alcoholic and unilateral brain-damaged patients: A test of the right hemisphere vulnerability hypothesis of alcoholism. *Archives of Clinical Neuropsychology*, *4*(3), 275–281.
- Akshoomoff, N. A., & Stiles, J. (1996). The influence of pattern type on children’s block design performance. *J. International Neuropsychological Society*, *2*(5), 392–402.
- Bambach, S., Lee, S., Crandall, D. J., & Yu, C. (2015, December). Lending a hand: Detecting hands and recognizing activities in complex egocentric interactions. In *The IEEE international conference on computer vision (iccv)*.
- Ben-Yishay, Y., Diller, L., Mandelberg, I., Gordon, W., & Gerstman, L. (1971). Similarities and differences in block design performance between older normal and brain-injured persons: A task analysis. *J. Abnormal Psychology*, *78*, 17–25.
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb’s Journal of Software Tools*.
- Chong, E., Nitschke, C., Nakazawa, A., Rozga, A., & Rehg, J. M. (2017). Noninvasive corneal image-based gaze measurement system. *arXiv preprint arXiv:1708.00908*.
- Cover, T., & Hart, P. (1967, January). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27. doi: 10.1109/TIT.1967.1053964
- Hoffman, J. E., Landau, B., & Pagani, B. (2003). Spatial breakdown in spatial construction: Evidence from eye fixations in children with Williams syndrome. *Cognitive Psychology*, *46*(3), 260–301.
- Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., . . . Murphy, K. (2017, July). Speed/accuracy trade-offs for modern convolutional object detectors. In *CVPR* (p. 3296–3297).
- Hutt, M. (1932). The Kohs block-design tests: A revision for clinical practice. *J. Applied Psychology*, *16*(3), 298–307.
- Jones, R. S., & Torgesen, J. K. (1981). Analysis of behaviors involved in performance of the block design subtest of the WISC-R. *Intelligence*, *5*(4), 321–328.
- Joy, S., Fein, D., Kaplan, E., & Freedman, M. (2001). Quantifying qualitative features of block design performance among healthy older adults. *Archives of Clinical Neuropsychology*, *16*(2), 157–170.
- Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C., Silverman, R., & Wu, A. Y. (2000). *An efficient k-means clustering algorithm: Analysis and implementation*.
- Kohs, S. (1920). The block-design tests. *J. Experimental Psychology*, *3*(5), 357–376.
- Kramer, J., Blusewicz, M., Kaplan, E., & Preston, K. (1991). Visual hierarchical analysis of block design configural errors. *J. Clin. and Exp. Neuropsychology*, *13*(4), 455–465.
- Kramer, J., Kaplan, E., Share, L., & Huckeba, W. (1999). Configural errors on WISC-III block design. *J. International Neuropsychological Society*, *5*(6), 518–524.
- Kunda, M. (2019). AI and cognitive testing: A new conceptual framework and roadmap. In *Proc. 41st annual meeting of the cognitive science society* (p. 2065–2070).
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment, fifth edition*. Oxford University Press, USA.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In *ECCV* (pp. 21–37).
- Milberg, W. P., Hebben, N., Kaplan, E., Grant, I., & Adams, K. (2009). The boston process approach to neuropsychological assessment. *Neuropsychological assessment of neuropsychiatric and neuromedical disorders*, 42–65.
- Nakazawa, A., & Nitschke, C. (2012). Point of gaze estimation through corneal surface reflection in an active illumination environment. In *ECCV* (pp. 159–172).
- Nishino, K., & Nayar, S. K. (2006). Corneal imaging system: Environment from eyes. *International Journal of Computer Vision*, *70*(1), 23–40.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
- Poreh, A. M. (2012). *The quantified process approach to neuropsychological assessment*. Psychology Press.
- Rozenwajg, P., Cherfi, M., Ferrandez, A., Lautrey, J., Lemoine, C., & Loarer, E. (2005). Age related differences in the strategies used by middle aged adults to solve a block design task. *Int. J. Aging & Human Dev.*, *60*(2), 159–182.
- Rozenwajg, P., & Corroyer, D. (2002). Strategy development in a block design task. *Intelligence*, *30*(1), 1–25.
- Rozenwajg, P., & Fenouillet, F. (2012). Effect of goal setting on the strategies used to solve a block design task. *Learning and Individual Differences*, *22*(4), 530–536.
- Sasson, N. J., & Elison, J. T. (2012). Eye Tracking Young Children with Autism. *Journal of Visualized Experiments*.
- Schatz, A., Ballantyne, A., & Trauner, D. (2000). A hierarchical analysis of block design errors in children with early focal brain damage. *Dev. Neuropsychology*, *17*(1), 75–83.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z.

- (2016). Rethinking the inception architecture for computer vision. In *CVPR* (pp. 2818–2826).
- Torraldo, A., & Shallice, T. (2004). Error analysis at the level of single moves in block design. *Cognitive Neuropsychology*, *21*(6), 645–659.
- Troyer, A., Cullum, C., Smernoff, E., & Kozora, E. (1994). Age effects on block design: Qualitative performance features and extended-time effects. *Neuropsychol.*, *8*, 95-99.
- Zipf-Williams, E., Shear, P., Strongin, D., Winegarden, B., & Morrell, M. (2000). Qualitative block design performance in epilepsy patients. *Archives of Clinical Neuropsychology*, *15*, 149–157.