Response Variability in a Low-Functioning Individual with Autism: Practical and Theoretical Implications

Erin J. Pickett[1], Nia M. Billings[1], Laura V. Van Droof[1], and Barry Gordon[1,2]

[1]Cognitive Neurology/Neuropsychology, Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD; [2]Department of Cognitive Science, The Johns Hopkins University, Baltimore, MD

JOHNS HOPKINS MEDICINE
SCHOOL OF MEDICINE

IMFAR 2010, Philadelphia, PA

Background

In addition to difficulty with initial learning, it is also widely recognized that low-functioning individuals with autism exhibit considerable variability in demonstrating what they presumably have learned (e.g., Koegel & Koegel, 1987). Such variability has many practical consequences, such as difficulty in obtaining valid assessments. It also may have a more profound theoretical significance – inconsistent responding may arise, in whole or in part, from an intrinsic weakness and/or variability in the underlying neural representations of knowledge in such individuals, as has been suggested as an explanation for the variability in naming performance across sessions that has been found in subjects with aphasia (Small et al.,1995).

Objective

These experiments were initiated to try to identify the origin of response variability found in earlier sets of experiments, which were conducted to assess our subject's knowledge of visual and auditory stimuli. Our aim was to establish the magnitude of response variability for well-known stimuli, in a simple response paradigm, in familiar and unchanging circumstances, in one extensively studied low-functioning individual with autism. We reasoned that inconsistent responding found under these circumstances would have a greater chance of being due to intrinsic variability, rather than to extraneous, adventitious, factors.

Methods

Participant

DL (not his real initials), an 18-year-old, nonverbal, low-functioning young man with autism

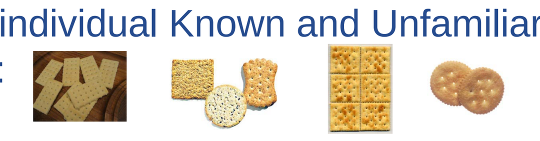
- met criteria for diagnosis of autism on the ADI-R
- consistently passed items through 26 months of age on the General Cognitive Index of the Bayley
- developmental ages of 2.5 and 3.5 on the Block Design and Object Assembly subtests of the WPPSI-III
- age equivalent score of 1 year 1 month on ROWPVT

Setting

All sessions took place in the dining room of DL's home, typically after he arrived home from school. Sessions generally occurred four days per week and lasted 45-60 minutes. Each session was led by one of three experienced, Master's level, therapists (one speech-language pathologist and two behavioral clinicians) on a rotating schedule.

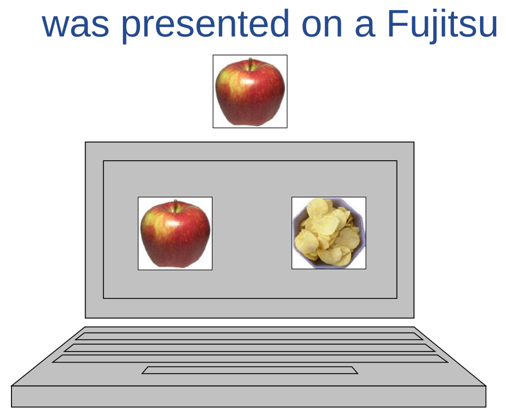
Stimuli

Known Items: One set of 24 object types (e.g., bed, car, pretzel). These were categorized as being well-known to DL on the basis of parental report, and because he had achieved an accuracy of at least 85% with each type on previous tests. DL had been extensively tested on these items over a period of six months prior to the study reported here.
Unfamiliar Items: Three sets of 24 objects that DL had been exposed to only rarely, if at all (e.g., cacomistle, daikon, scabbard).
Four different color photographs of each individual Known and Unfamiliar item were used to prevent boredom, e.g.:



Task

A two-alternative forced-choice (2AFC) picture-to-picture matching task was presented on a Fujitsu LifeBook T1010 Tablet PS using E-Prime. In each trial, two color photographs appeared on the laptop screen, one on the left side and one on the right. The experimenter held a hard-copy stimulus photo above the screen. Using the mouse, DL's task was to choose the photo on the screen that matched the stimulus photo. All sessions were videotaped with two cameras.



Conditions

Condition	Targets	Distracters
1	Known items	Known items
2	Unfamiliar items	Unfamiliar items
3	Unfamiliar items	Known items
4	Known items	Unfamiliar items

Procedure

Experiment 1: Each of the four conditions was presented in succession (e.g., Condition 2 was run after Condition 1 was completed). Each of the 24 target items was presented twice per block (two different photos), once on each side of the screen, totaling 48 trials per block. Two blocks were run per session for five sessions, totaling 10 blocks per condition. Within one session, the two blocks were the same, i.e. the same photos of the same targets in the same position in the same order. However, the distracters paired with each target were different, because they were randomly selected from a list of candidate items. Between sessions, the order of the targets and the two photos used for each target were quasi-randomized, with a limit of no more than three targets consecutively presented on the same side. After the completion of each block, DL was allowed to watch a short video clip.

Experiment 2: Experiment 2 was conducted after all four conditions of Experiment 1 were completed. The procedure was the same as Experiment 1 but with feedback given to DL based on his performance on each individual trial.

Correct: 

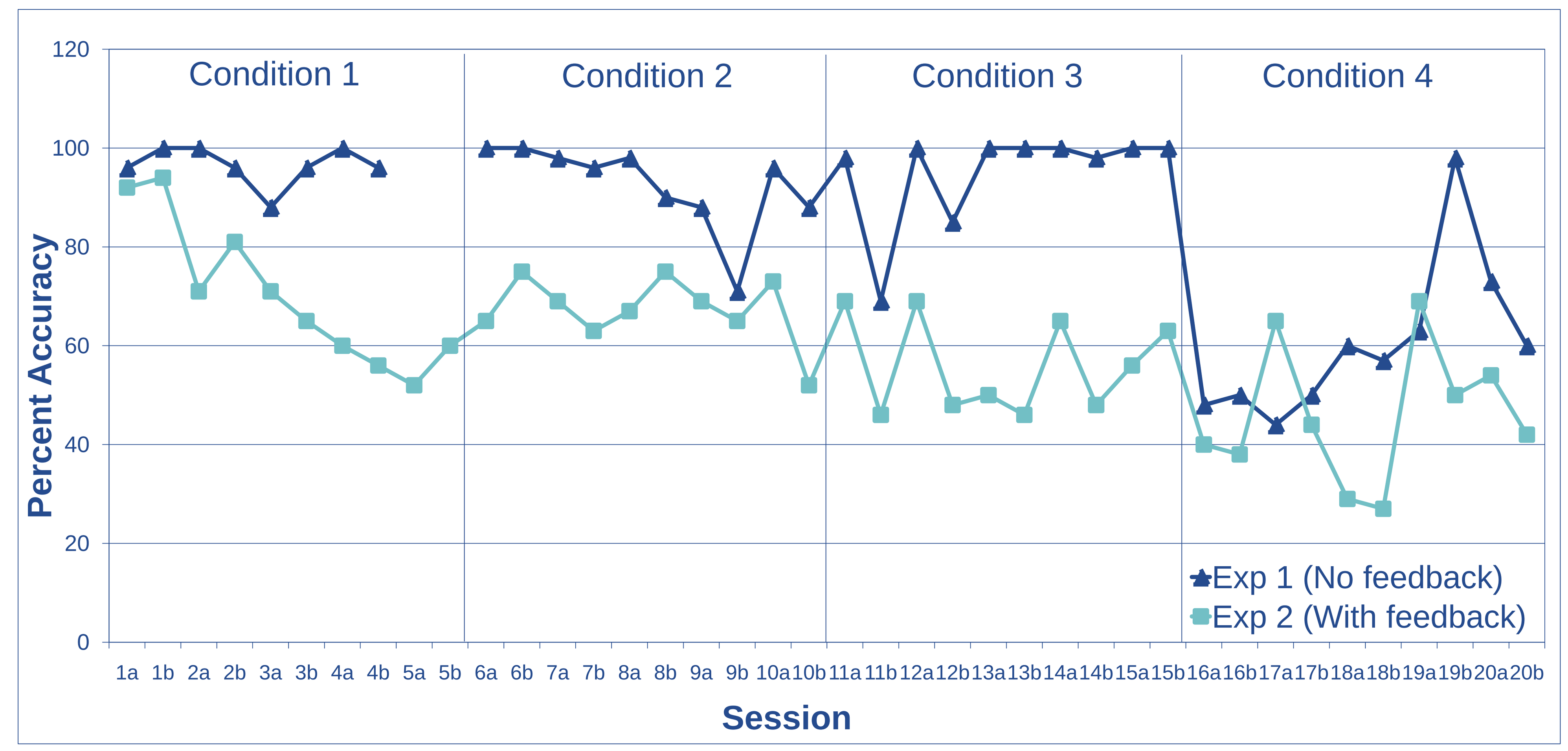
Incorrect: 

Results

Condition:	1	2	3	4
Average Percent Correct				
Exp 1	96 (370/384)	92 (443/480)	95 (456/480)	60* (289/479)
Exp 2	70 (337/480)	67 (322/480)	56** (268/480)	46*** (219/480)
Largest Percent Accuracy Change in Same Session				
Exp 1	8	17	29	35
Exp 2	10	21	23	21
Percent Accuracy on Left vs. Right Side				
Exp 1	95 vs. 98	98 vs. 87	98 vs. 92	76 vs. 44
Exp 2	80 vs. 60	84 vs. 50	81 vs. 30	57 vs. 34
Percent Left Side Chosen				
Exp 1	48	55	53	66
Exp 2	60	67	75	68

*Significantly different from conditions 1, 2, and 3, $p < .0001$; **Significantly different from condition 1, $p < .05$; ***Significantly different from conditions 1 and 2, $p \leq .001$ (One-way ANOVA, Tukey's post hoc test)

DL demonstrated the ability to perform competently on the 2AFC matching task, achieving 100% accuracy on 30% of the blocks in Experiment 1. However, his performance varied considerably and tended to worsen with each successive condition in both Experiments 1 and 2. DL's accuracy was the lowest in condition 4 in both experiments, when targets were known and distracters were unfamiliar. Overall, DL had higher accuracy rates without feedback (Exp 1 mean=85%) than with feedback (Exp 2 mean=60%), $p < .0001$ (Independent Samples T-Test, equal variances assumed).



Discussion and Conclusions

The Two-Alternative Forced Choice (2AFC) task format was used as a measure of sensitivity independent of response bias (Green & Swets, 1966), but this inference is only strictly true if the subject does not have a bias toward any of the alternatives prior to the task (Green & Swets, 1966). Spatial and item biases are well-recognized in individuals such as DL, and were therefore specifically examined. DL did show spatial bias. By condition 4 of Experiment 1, he had a clear bias to choosing the left side of the screen, and his preference for the left side continued throughout Experiment 2. In general, DL was more accurate on targets appearing on the left side. However, DL did not demonstrate any reliable item biases on these experiments. Therefore, we are limited in how rigorously we can use DL's results with the 2AFC tasks an indices of his true "strength" of representation of known and unknown items.

What we can be certain about is that his performance was quite variable, even with highly familiar stimuli on a simple, familiar task, in familiar circumstances. Our data do not yet let us fully identify the cause of this pronounced variability. His motivation and attention appeared to be good, based not only upon the observations of the skilled examiners present at the time, but even on later review of the videotaped sessions. The feedback we provided was not effective in guiding DL to maintain high rates of performance. It has been appreciated that error correction procedures may have differing effects on skill acquisition across different individuals (Smith et al., 2006). Feedback proved to be unexpectedly problematic in DL's case. It is possible that the variable performance found here represented, in part, boredom with the stimuli during testing. The experiments reported here were the last in a long set that had used the same stimuli. There is indeed some evidence for boredom, in that DL's responses were frequently poorer on the second session of the day then the first.

Whatever the basis or bases for DL's variable performance, its existence should perhaps be a cautionary message for educational and research efforts with these subjects. Had we not surveyed such familiar materials, had we not used such a relatively simple and relatively bias free response methodology, and if we had not run so many trials, we might have been misled into taking any single sample of DL's performance as representative of his true capabilities. But examining the broader picture shows that this would have been problematic.

More fundamentally, it is possible that at least part of DL's performance variability is arising from a basic weakness and/or intrinsic noise in his neural representations of even familiar items. There have been suggestions that the neural processing modules in individuals with autism may be less robust than those in normally developing individuals (e.g., Casanova, 2006; Cohen, 1994). Our data cannot resolve this question, but do provide some evidence for this possibility.

References

Casanova, M. et al. (2006). *Clinical Neuroscience Research*, *6*, 127-133.
Cohen, I.L. (1994). *Biological Psychiatry*, *36*, 5-20.
Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
Koegel, R.L. & Koegel, L.K. (1987). *Seminars in Speech and Language*, *8*, 241-256.
Small, S.L. et al. (1995). *Clinical Aphasiology*, *23*, 25-38.
Smith, T. et al. (2006). *Behavioral Interventions*, *21*, 245-263.

Acknowledgments

We express our gratitude to DL for his hard work and good humor as well as to his parents for their patience and hospitality. This work is supported by the Therapeutic Cognitive Neuroscience Fund and The Benjamin & Adith Miller Family Endowment on Aging, Alzheimer's Disease, and Autism Research.

Conflict of Interest: None