

Are We Speaking the Same Language?
When is a Mild Impairment a Mild Impairment?

Linda S. DeYoung
Newport Language, Speech & Audiology Center, Inc., Laguna Hills, California

Although test batteries for aphasia have been published that assign severity levels in relation to performance on each subtest, such as the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass and Kaplan, 1983) and The Porch Index of Communicative Ability (PICA) (Porch, 1967), it is felt that these systems have not been adapted at large to assigning severity levels beyond their respective test batteries. Since we frequently describe patients according to degrees of severity (mild, moderate, and severe) and describe their progress with regard to changes in these severity levels, clinical discussions among our colleagues are greatly hampered if our definitions of severity levels are dissimilar. The purpose of this study is to determine if speech pathologists are applying similar guidelines in judging severity levels.

METHOD

Twenty-seven speech pathologists with diverse graduate backgrounds (75% were PICA trained), representing all regions of the United States were asked to rate severity levels of impairment reflecting patient performance. Number of correct and incorrect responses to common aphasia testing tasks were provided under three different conditions. Under each condition, the speech pathologists were given a copy of the survey rating sheet (Table 1) to fill in their judgments of severity. Under the first condition, severity judgments were made based on the information given on this survey sheet only. For the second and third conditions, a videotape of the clinician and patient completing these tasks listed in the survey sheet was presented. The clinicians were asked to rejudge half of the tasks (Numbers 1, 3, 5, 8, and 9). The number of correct and incorrect responses remained the same. In the second condition, all responses were either complete, accurate, and prompt (a "15" response, according to the PICA scoring system) or inaccurate (a "6" on the PICA). In the third condition, correct responses also included the addition of distortions delays, incompletions, repeats, self-corrections, and cues ("15-8" "8" responses on the PICA) while perseverations, rejections, no responses, related, minimal, and unintelligible responses ("7" and below on the PICA) comprised error responses. This format provided a comparison of degree of agreement in determining severity levels when given testing data only vs. observing the patient's behavior in addition to receiving the test data, contrasting simple and complex response situations.

RESULTS

Tables 2, 3, and 4 display the data under each condition. The results indicate that whether given test data only or the addition of the patient to observe during testing, there is a remarkable lack of agreement among speech pathologists in rating a patient's level of impairment under these conditions. On almost every task the speech pathologists' judgments ranged across 3 or 4 severity categories. For instance, when a patient missed half the questions (3/6) testing paragraph comprehension, interpretations varied from stating that this constituted mild involvement all the way to severe deficit. However,

Table 1. Survey sheet provided to each speech pathologist to record their severity judgements.

VERY SHORT SURVEY TO IMPROVE OUR COMMUNICATIONS

Is what you consider a "mild" impairment the same thing another speech pathologist would call a mild impairment?

Please rate the level of severity on the following questions to find out.

Please put a checkmark by your choice of severity:

| TASK | (No. Correct) Score | LEVEL OF IMPAIRMENT | | | | |
|--|------------------------|---------------------|------|----------|------------|--------|
| | | Minimal | Mild | Moderate | Mod-severe | Severe |
| 1. Identifies object/picture when named | 11/12 | | | | | |
| 2. Comprehends simple yes/no questions | 10/20 | | | | | |
| 3. Comprehends simple yes/no questions | 17/20 | | | | | |
| 4. Comprehends simple yes/no questions | 15/20 | | | | | |
| 5. Comprehends paragraph length material, tested through questions about paragraph | 3/6 | | | | | |
| 6. Comprehends paragraph length material | 5/6 | | | | | |
| 7. Comprehends paragraph length material | 4/6 | | | | | |
| 8. Comprehends written sentences | 7/10 | | | | | |
| 9. Reads & comprehends simple paragraphs | 3/4 | | | | | |
| 10. Reads & comprehends simple paragraphs | 1/4 | | | | | |
| 11. The following writing sample was generated for the kite flying picture: Man by the chimney with the dock pond. Kite in the tree by the boy. | | | | | | |

Table 2. Speech pathologists' ratings of severity given the following scores and tasks only (N = 27).

| TASK | SCORE | JUDGED LEVEL OF IMPAIRMENT | | | | | | | |
|--|-------|----------------------------|---------|-----|---------|-----|---------|-----|---|
| | | Min | Min-Mld | Mld | Mld-Mod | Mod | Mod-Sev | Sev | |
| 1. Identifies object/picture when named | 11/12 | 7 | 2 | 16 | | 2 | | | |
| 2. Comprehends simple yes/no questions | 10/20 | | | | | 1 | 11 | 15 | |
| 3. Comprehends simple yes/no questions | 17/20 | | | 8 | | 17 | 2 | | |
| 4. Comprehends simple yes/no questions | 15/20 | | | 1 | | 9 | 17 | | |
| 5. Comprehends paragraph length material, tested through questions about paragraph | 3/6 | | | 1 | | 5 | 16 | 5 | |
| 6. Comprehends paragraph length material | 5/6 | 5 | | 14 | 2 | 5 | 1 | | |
| 7. Comprehends paragraph length material | 4/6 | | | 7 | 1 | 15 | 4 | | |
| 8. Comprehends written sentences | 7/10 | | | 2 | 1 | 19 | 5 | | |
| 9. Reads and comprehends simple paragraphs | 3/4 | 1 | | 6 | 2 | 17 | 1 | | |
| 10. Reads and comprehends simple paragraphs | 1/4 | | | | | | 8 | 19 | |
| 11. The following writing sample was generated for the kite flying picture: Man by the chimney with the dock pond. Kite in the tree by the boy. | | | | | | 1 | 17 | 8 | 1 |

Table 3. Second condition: Speech pathologists' ratings of severity with addition of a videotape presentation.

| TASK | SCORE | JUDGED LEVEL OF IMPAIRMENT | | | | | | |
|--|-------|----------------------------|---------|-----|---------|-----|---------|-----|
| | | Min | Min-Mld | Mld | Mld-Mod | Mod | Mod-Sev | Sev |
| 1. Identifies object/picture when named | 11/12 | 12 | 2 | 13 | | | | |
| 3. Comprehends simple yes/no questions | 17/20 | | 1 | 10 | 6 | 10 | | |
| 5. Comprehends paragraph length material, tested through questions about paragraph | 3/6 | | | | 2 | 7 | 15 | 3 |
| 8. Comprehends written sentences | 7/10 | | | 1 | 4 | 19 | 3 | |
| 9. Reads and comprehends simple paragraphs | 3/4 | | 1 | 7 | 7 | 12 | | |

Table 4. Third condition: Speech pathologists' ratings of severity with addition of a videotape presentation displaying more varied responses, such as perseverations, cues, delays, and self corrections.

| TASK | SCORE | JUDGED LEVEL OF IMPAIRMENT | | | | | | |
|--|-------|----------------------------|---------|-----|---------|-----|---------|-----|
| | | Min | Min-Mld | Mld | Mld-Mod | Mod | Mod-Sev | Sev |
| 1. Identified object/picture when named | 11/12 | | | 2 | 11 | 11 | 3 | |
| 3. Comprehends simple yes/no questions | 17/20 | | | | 1 | 23 | 2 | 1 |
| 5. Comprehends paragraph length material, tested through questions about paragraph | 3/6 | | | | | 6 | 16 | 5 |
| 8. Comprehends written sentences | 7/10 | | | | | 17 | 10 | |
| 9. Reads and comprehends simple paragraphs | 3/4 | | | 4 | 6 | 15 | 2 | |

the majority of responses usually fell into two adjacent severity levels. In the third condition (Table 4), there was an increase in agreement, suggesting a higher consensus as a patient's severity increased.

DISCUSSION

These findings indicate that we need to be cautious in describing our patients according to levels of impairment since misinterpretations may occur. This could lead to inaccurate suggestions for treatment. Furthermore, a patient's progress could be misinterpreted with a change in clinicians. Sometimes diagnostic and progress reports are written in terms of levels of impairment (e.g., checklists of statement such as "improved from a moderate to a mild deficit for understanding simple sentences"). Consequently if a patient comes up for review shortly after a change in clinicians, the new clinician may note no gains if her or his criteria for defining severity levels are more stringent. A patient making noteworthy gains could be terminated from treatment prematurely, or the reverse could happen -- a patient not making gains could be maintained unnecessarily in treatment. Conclusions can be drawn needlessly and repeatedly just because we are operating with a subjective set of terminology. Since our field is continually striving for objective and scientific advancements, this study's findings should be addressed.

The subjective use of severity terminology can be easily rectified in several ways. One solution is that test or treatment scores or descriptions could be provided in conjunction with the severity levels. However, this may hamper efficiency. Secondly, a key defining severity levels could be included on report and treatment forms. Another solution would be to develop one guideline for rating criteria accepted and advocated by ASHA for universal usage. Rating criteria based on a percentile scale could be easily adapted to both testing and treatment data to increase objectivity in both realms. Refer to Table 5 for an example. Aphasic-type responses may include delays, cues, self-corrections, etc. If the presence of these behaviors affects communication efficiency and is not addressed in the patient's score, the level of impairment could be lowered an additional severity level. For example, if a patient scores 90%, reflecting a mild disorder, but delays were exhibited on one-third of the responses, this would reflect a mild-to-moderate disorder, according to Table 5.

Table 5. Example of changing scores to percentages to equate to a severity rating scale for increased objectivity.

| Score in % | Level of Impairment |
|------------|---------------------|
| 96 - 100 | Minimal |
| 90 - 95 | Mild |
| 80 - 89 | Mild-Moderate |
| 70 - 79 | Moderate |
| 60 - 69 | Moderate-Severe |
| Below 60 | Severe |

Decrease severity by another level for unaccounted reductions in communication efficiency that affect one-third or more of the answers.

In summary, speech pathologists may be using differing criteria when judging levels of severity, often without specification. Perhaps some individuals are comparing aphasic people to aphasic people while others are comparing aphasic people to normal people. These differences may be affecting our objectivity and communication. Objectivity would be enhanced if we established standards for rating levels of severity.

REFERENCES

- Goodglass, H. and Kaplan, E. The Assessment of Aphasia and Related Disorders, Second Edition. Philadelphia, PA: Lea and Febiger, 1983.
- Porch, B.E. Porch Index of Communicative Ability. Palo Alto: Consulting Psychologists Press, 1967.

DISCUSSION

- C: In your method, I think that the statement that you had on your survey sheet about mild impairments probably reduced your variability considerably. It would be interesting if you took out that statement about what you're after and just let them go ahead and make their own ratings. I think that the variability would be even greater.
- A: Thank you.
- Q: Did you give them any other instructions or just have them read the survey sheet and then go ahead and rate it?
- A: That was it. I wanted to try to imitate the real clinical situation as much as possible, because so often when we take on a patient from somebody else, all we're given is a report that states that he's mildly involved, moderately involved, etc., and we operate off that kind of information, or when I go in for quality control purposes and in doing peer reviews, I look at a clinician's chart and I say "I don't consider that a moderate deficit. That person missed over two-thirds of the responses." Yet I couldn't go back to that clinician and say "that's wrong" because in my mind that was wrong, but where are the standards stating what is wrong? For certain tests we have criteria like that, but for treatment and for other situations, clinicians have not taken these criteria beyond their respective test batteries.
- Q: Your instructions were fairly meager. Do you think it's possible that people were providing ratings for different reasons? That is, do you think it's possible that some people were thinking, "well if someone has 3 of 4 errors on auditory comprehension of paragraphs, then his overall severity level is -- whatever"? Then someone else might have been interpreting the rating to be his severity level for that task.
- A: I did explain to them, take it only for each task. Because they did ask me questions when I gave out the survey sheets, and I did say I'm giving you this because I think there is a lot of variability amongst us and I did have people say "well it's hard to do, out of context, for just that task." I got that kind of feedback from people, and I said, "well - just pretend this is the only information you have on this person." And for instance with a score of 3 of 4 correct on this task, etc., and rate it on that.

- Q: So you did instruct them that this task is specific?
- A: Yes.
- Q: You're in position to look at a lot of speech-language pathologist's approaches to describing severity, given the size of your department. I was wondering from a standpoint of utilization review, and from reimbursement review, do you have a sense about whether this kind of inconsistency between clinicians also exists within a single clinician in terms of describing the improvement or change for a patient from admission to your program to discharge from your program? And if it has had any impact on PRO or reimbursement review in terms of patients attaining goals and there being an appropriate description of changing the severity level of the patient?
- A: I see that occur in both ways, because I often ask after looking at different clinician's evaluations and treatment reports. I have seen some that have a consistent guideline in their mind that they always use; and others I see real variability and then I ask them about it; and they may not really know where they are coming from when they say "it's just a kind of gut feeling" for them, where they haven't made some quantifiable standard. There is some variability in some people and in others not. I have found it a very individualistic kind of thing, maybe dependent upon people's previous supervision, their own experiences, and their training.
- C: While the focus of your paper is very important when considering variability between or lack of agreement between clinicians I think there is also an issue about lack of consistency within some clinicians. That presents a real problem in terms of the evaluation by third party of documentation and reimbursement for a single case, and that I hope we can also address.
- A: That's a good point, because there isn't always conformity among clinicians. Some of them have variability.
- Q: I'm concerned about the validity of your procedures, and I think your data support my concern. It seems to me that what you're asking the people to do is not what people traditionally do in a clinical environment; and your data suggest that as you move closer to what people actually do in a clinical environment, their agreement does get better -- as you move from the questionnaire to the videotape to the videotape with more complete responses.
- A: I agree with you after I did it the first time with just the sheet only, I felt that that wasn't replicating the real clinical situation enough; and that's why I thought I should do a videotape too. Do you have a suggestion on how to make it more valid? Do you think maybe having the patient right there and the clinician right there and we all go through the whole activity that that would be better than showing the videotape of it?
- C: I'm not prepared at this point to design a procedure, but it seems to me that the procedure should require that the judge be able to see very precisely the materials to which the patient is responding and should have access to all the information to which a clinician typically has access when they make those judgments in an actual clinical situation. I don't know what would happen. It may be that judgments would still be very unreliable, but I think it's premature to extend your findings to actual clinical situations, because of that dissonance between what you did and what reality is in the clinic.

- A: On the second and third condition, they did see the materials. The only thing they aren't provided with, of course, is the case history, which includes the medical history. And if you think that may perhaps influence our judgments too, then that's something really to consider.
- C: I suspect there's still some distance between even that condition and the conditions under which such judgments are ordinarily made, and I'd like to see that distance decrease, if possible.
- A: Good point.