Previous work on the assessment of cluttering severity has demonstrated a need for perceptual rating approaches, rather than objective and quantified behavioral measures, for effectively capturing the severity of cluttering. Among the issues reported were multidimensionality of this complex disorder, the role played by various qualitative aspects of cluttering (Myers & Bakker, 2011; Myers & Bakker, In Press), and lack of consistency regarding how specific objective measures such as speech rate or numbers of typical speech disfluencies impact perception of cluttering severity (Bakker, Myers, Raphael and St. Louis, 2011).

The Cluttering Severity Instrument (CSI) was developed to capture severity by means of a composite set of perceptual ratings for various speech and language dimensions associated with cluttering (Bakker and Myers, 2011). However, this process may not be expedient enough for routine use to capture moment-to-moment variations in cluttering severity during the course of therapy. The clinical value of an instrument such as the CSI, after all, is to provide clinicians with comprehensive estimates of cluttering severity at critical junctures of therapy: a) collecting a baseline; b) determining the need for treatment; c) evaluating the overall efficacy of current therapy approaches in order to make sound decisions about subsequent therapy strategies; and, finally, d) determining eligibility for dismissal. An additional value of the CSI, of course, is to allow clinicians to rate perceived severity of the various dimensions associated with cluttering.

The purpose of the Real-Time Continuous Perceptual Scoring tools was to develop and demonstrate expedient cluttering severity assessment options which track the perception of cluttering severity continuously throughout speech samples such as those typically recorded during therapy sessions. We anticipated that this methodology would provide a representation of the degree of impact a cluttering problem has on functional communication in naturalistic speaking situations during reading, monologue, or dialogue. These software programs were introduced and discussed for the first time at the Second World Conference on Cluttering, in Eindhoven, the Netherlands (Bakker and Myers, 2014).

Perceptual severity ratings of recorded samples are common in the assessment of communication disorders (e. g., the Boston Diagnostic Examination of Aphasia, Goodglass, Kaplan and Barresi, 2001). A range of options exist for post hoc perceptual ratings based on these recordings, such as Likert type rating scales, and more recently, visual analog scales. However, it could be argued that judgments regarding severity based on clinician's memory of the speech sample may lose sight of moment-to-
moment variations in severity that occur in ongoing running speech. Relying on post hoc general judgments, then, could make us prone to bias although this, to our knowledge, has not yet been determined through empirical research. For example, a few brief but very severe intervals with cluttering in the context of otherwise normally sounding speech could shift perception to a higher perceived severity level, especially if such intervals occur near the end of the sample.

There is another potential disconnect between assessments based on post hoc recollections of one's global perception, and assessments based on real-time moment-to-moment changes in speaking behaviors. Intervention usually addresses the client’s ability to make moment-to-moment changes in behaviors. The latter of course involves ongoing variations in perceived cluttering severity by the clinician, rather than depending on overall comprehensive estimates made at some moment following the sample.

We propose the development of a perceptual rating system based on real-time continuous perceptual tracking throughout a speech sample, to serve the needs of clinicians and researchers. A computer, or a smaller device such as a tablet or smart phone, when used for severity tracking, has the potential to retain all detailed changes in a perceptual rating trail and to derive information such as: (1) the mean overall severity based on the entire ‘history’ of perceived severity, (2) the variability (standard deviation) in perceived severity, and (3) the frequency distribution of severity levels (e.g. whether there were just a few very severe moments, or if the severity is rather comprised of many short but fairly mild intervals). It is hoped that this information helps reduce perceptual bias and, importantly, provides a fuller profile when collecting cluttering severity information about speech samples.

Three preliminary versions of real-time perceptual tracking have been completed, two of which were demonstrated at the 2nd World Conference on Cluttering (2014) and the downloads available on this website are the original versions that were introduced. A third version is nearly complete which, rather than a mouse, uses a joystick such as commonly used in computer games. Other input methodologies are still considered but have not yet been completed. The scoring tools may be used in the context of live speech, or prerecorded samples. These tools employ the use of a computer mouse (using either the “mouse hovering function” [see Figure 1], or a “mouse wheel” [see Figure 2]) for expressing perceived cluttering severity continuously as the sample proceeds.
Figure 1. Graphical representation of a scoring trail made visible following completion of a continuous perceptual rating session using the mouse hovering function. During perceptual rating, when moving the mouse cursor up and down, the graphical display area is masked by a green screen so users aren’t distracted and can focus on their observations. Means and standard deviations, based on the scoring trail data, are also visible in this screen as well as a frequency distribution of the rating levels.

Use of the version that employs the mouse hovering function can be compared to a continuously varying analog scale. Resolution of scoring results obtained with mouse hovering is defined in terms of pixels on the screen available for this process (full range, or 100%, consists of 200 pixels, which are the smallest graphical units on the computer screen). One pixel, then, approximates 100/200th, or .5% of change. Using the mouse hovering function, also, parallels how scoring would be accomplished using software designed for touch screen devices. The advantage is a high resolution in measured severity although the validity of this resolution would depend on what clinicians can reasonably and reliably accomplish using the system during actual cluttering severity ratings. Timing resolution can also be defined in pixels, with each pixel representing .1 second.
Figure 2. Graphical representation of a scoring trail made visible following completion of a continuous perceptual rating session using the mouse wheel function to select available Likert Type scores (compare to Likert Type rating scale used in many questionnaires). During perceptual rating the results are not visible. Quantified data collected from the scoring trail are however visible in this screen. The higher the profile, the more severe the cluttering perceived.

The mouse wheel function for scoring can be compared to a continuously varying Likert Type rating scale. The available rating range is now divided over 7 available steps in severity (ranging from 1 through 7).

Note that on most PCs the mouse wheel is set on moving three lines of text at a time with one click. In order for the mouse wheel application to move one number at a time it is necessary to change the mouse wheel setting in the Windows Control Panel from 3 to 1 line at a time. The image on the left shows the dialog box with this setting changed.

This integer discrete scaling option could have the advantage of being less prone to user errors in adjusting the ratings as well as improve reliability. If the mouse wheel is kept in the “clicking mode”,
Within one comfortable turn of the mouse there will be about 7 clicks, thus the selection of 7 possible steps in severity. The assumptions of improved resolution and reduced impact of operator-based adjustment errors of course are in need of empirical evaluation. Timing resolution of using the mouse wheel function is the same as when using the mouse hovering function. Please note that for mouse wheel scoring it doesn’t matter where the cursor is on the screen (this is the case only in Windows 7 and 8). This gives the user considerable freedom to pay attention to the client being evaluated without having to worry about the relationship between the mouse and changes on the computer screen. Unfortunately with Windows 10 the mouse behavior changed and we are presently looking for a solution to restore this function. In Windows 10, in other words, the Mouse Hover version is the only working option.

Figures 1 and 2 also demonstrate two versions of the quantified results (mean and standard deviation). One set of data applies 1 second buffering, or a “rolling average” of 1 s, to the rating trail thus averaging out random variability caused by user error in controlling the mouse functions (mouse hovering or wheel scrolling). This buffering helps control “jerky” (unintended) adjustments made while adjusting the mouse for perceptual rating purposes. The choice of 1 second buffering was arbitrary and empirical testing may produce information about the optimal averaging period for controlling random human errors. The application of buffering is an analogous procedure to the use of interval scoring for improving reliability in judgments regarding behaviors that are considered to be challenging in nature to measure. There is some literature that stuttering identification may occur more reliably if interval scoring procedures are used (see Cordes & Ingham, 1994, 1995). A recent study (Howell, 2008) also brought up the issue that the selection of the interval duration is important.

A third input version, which uses a joystick, such as commonly used in flight simulator computer games, was also developed but has not been prepared for a download. The use of a joystick initially appeared to have a unique advantage in that it is spring loaded helping the user in being aware of the relative position of the cursor, thus reducing the dependence on visually monitoring a computer screen for accuracy. Availability of joysticks, and experience using them, is not as common as using a computer mouse (for hovering the mouse cursor, or using the mouse wheel). The potential use for joysticks in clinical ratings, while clinically interesting, would require further testing, but will be supported if enough interest for it becomes evident.

In conclusion, our proposed real-time severity tracking system (RPT-CS) is a promising clinical tool for providing clinical impressions regarding cluttering severity in real time and continuously throughout a sample. The strategy is proposed as an add-on to the Cluttering Severity Instrument (Bakker and Myers, 2011), to fulfill daily clinical assessment needs that would not warrant thorough assessment such as would be possible with the CSI system. Both tools may be useful for research purposes as well as they complement each other in noted ways. We are not aware of others who have reported using our strategy for these purposes, and certainly not in the case of cluttering assessment. Empirical validation of our procedure, as well as estimates of reliability when used by clinicians, are strongly needed at this point. Our ultimate intention, following completion of the tool based on audience input during the conference, is to make the software plus a user manual available through the website of the
International Cluttering Association as a freeware public domain tool for clinicians who treat individuals who clutter on a regular basis.

References


