Best Practices and Practical Strategies for Assessment and Diagnosis of Autism

Adrienne Perry, Rosemary A. Condillac and Nancy L. Freeman

Abstract

This paper describes a current "best practices" approach to the assessment and diagnosis of individuals with autism, including those who may be very difficult to test because of behavioural issues or their degree of cognitive limitation or both. The aim is to share our assessment process which is tied to the "best practices" empirical literature and is consistent with relevant ethical and professional regulations and guidelines. Our approach is also heavily influenced by our clinical experience and the practical strategies we find effective and is, further, grounded in the ultimate purpose of making the information obtained useful and helpful to the individual, family, and staff working with the person with autism.

The Purposes and Uses of Assessment for People with Autism

We believe that good assessment is crucial to understanding and assisting the children, adolescents, and adults we serve, as well as their families and others who support them. Assessment is a process of gathering useful information about a person from relevant sources for a particular purpose or purposes. Assessment is a much broader concept than "testing" (though one or more standardized tests may be part of an assessment) and the results of assessment are much more than test scores or labels. Assessments of people with autism may be done for a variety of reasons at various times in their lives.

The main purposes for assessment

There are four main purposes for carrying out an assessment. These are:

1. *To help understand the person.* The primary clinical reason for any kind of assessment is to provide useful information about the person which will help us understand them better and guide us to design, provide, or refer them for

appropriate educational and therapeutic intervention. It is important to incorporate the principle of individualization into program planning by taking into account each person's developmental level, strengths and needs, the degree of scatter among skills, their likes and dislikes, behavioural difficulties, effective teaching strategies and motivational systems. In this way we can maximize the effectiveness of intervention and, perhaps, minimize behavioural disruptions which may result from the individual’s environment not being well-matched for them. In addition, it is sometimes important to have a baseline against which to measure change (either cognitive improvement or decline), particularly in older individuals.

2. *To obtain or clarify an initial diagnosis.* For parents of young children with autism, the initial diagnosis is inevitably a very difficult process. However, it is frequently more time consuming, frustrating, and confusing than necessary because of a lack of community expertise regarding autism, some professionals’ own difficulty in delivering difficult news (Abrans & Goodman, 1998), and some practitioners’ assumption that the child is too young to undergo a diagnosis. The literature is now quite clear in supporting early screening and referral programs to ensure that parents’ concerns are taken seriously right away and that children are referred for full diagnostic assessments sooner (Filipek et al., 1999; New York State Department of Health [NYSDH], 1999). Research has also demonstrated that it is possible to diagnose autism reliably and specifically by age 2 or 3 certainly, and some would argue even at 18 months (see Baron-Cohen et al., 1996; Lord, 1995; Stone et al., 1999). We are frequently called upon to confirm, clarify, or further specify and explain a vague diagnosis (e.g., delayed with autistic features) or suspicion ("query Pervasive Development Disorder, PDD") or clarify what may initially appear to be contradictory diagnoses (e.g., PDD by one practitioner; global delay and autism by another).

3. *To document diagnostic status necessary for access to services or funding.* At various stages later in the child or adult’s life, a diagnostic reassessment may be required for the purpose of documenting a particular diagnosis, such as autism, or a certain classification, such as developmental disability, for the purpose of determining school placement, or accessing funding supports or services, in any of the social services, health care, and education systems. It is typically necessary to document that a person qualifies for a particular government-funded service targeting a particular group or to have an accurate diagnostic description to assist in advocating for higher service allocations for a client with high needs. While this reason for assessment is sometimes dismissed as administrative "gatekeeping", when examined from the larger system perspective it can be seen as a necessary and rational approach to service allocation. Although there can be risks and

disadvantages to the inappropriate use of "labelling", this is one of the appropriate and helpful uses. (See Sattler, 2001 for a good discussion of the pros and cons of labelling.) Further, a reassessment conducted ostensibly for this purpose frequently has the beneficial byproduct of providing to caregivers a greater understanding of the person’s abilities, resulting in better tailored interventions and environments.

4. *To obtain information for program evaluation or research purposes.* Sometimes assessment information is used to document the progress of an individual by comparing test results over time which can assist in evaluating and planning appropriate services for that person. When combined with other individuals in a program or classroom, this kind of information can be very helpful for program evaluation purposes, even though it was initially collected for clinical purposes. It can help us determine whether interventions are effective, for whom they work best, how client characteristics impact on effectiveness, and a variety of other factors. In some situations, additional assessments may be desirable for research purposes.

Assessment and Diagnosis of Autism: "Best Practices"

A "best practices" approach to assessment means that the specific measures used and the whole assessment process should be carefully tied to the most recent professional literature based on evidence-based evaluation of the measures and procedures. Recently, there has been a trend, in many areas of health and human services, towards multidisciplinary expert panels reviewing all the literature and pooling their clinical experience and publishing a Clinical Practice Guideline or a set of consensus panel recommendations.

In the area of autism assessment, there are two such documents which now set the standard for "best practices": **1)** Filipek et al. (1999) of the American Academy of Neurology, chaired a panel including representatives of associations from neurology, audiology, child and adolescent psychiatry, pediatrics, occupational therapy, psychology, speech and language pathology, child neurology, and developmental pediatrics, which made recommendations regarding early screening as well as comprehensive assessment and diagnosis; **2)** The New York State Department of Health (NYSDH,1999) also commissioned a multidisciplinary panel to produce a Clinical Practice Guideline regarding assessment/diagnosis and intervention for autism, though this document focusses only on young children.

Other general and comprehensive references which may be helpful to the reader and which have influenced our thinking (or which parallel it), include the following: Marcus & Stone, 1993; National Research Council [NRC], 2001; Sattler, 2002;

Schopler & Mesibov, 1988; and Sparrow, 1997. Clinical practice should also be consistent with relevant legislation, regulatory college regulations, standards, ethics, agency policies, and other relevant guidelines. In our case, we practice under the College of Psychologists of Ontario (CPO) (which has clear standards around assessment and communicating a diagnosis; see CPO, 1995) and in accordance with the American Educational Research Association [AERA] Standards for Educational and Psychological Testing (1999). The remainder of this paper describes the process we use for our psychological assessments, which draws heavily upon these sources, with the intention of being ethically and professionally appropriate, as well as consistent with a "best practices" approach. Other components of a comprehensive assessment (Filipek et al., 1999) are provided by other disciplines in our settings.

Standardized measures for autism

The DSM-IV criteria for the diagnosis of Autistic Disorder (American Psychiatric Association [APA], 1994) are the widely agreed "gold standard" and we use them systematically, rating each criterion relative to developmental level (measured as described below), based on a combination of parent report and observation, preferably over more than one occasion and in more than one setting (i.e., with parents, with other adults, alone, during free toy play, during structured demands). In very young and/or very low functioning individuals, there are several criteria that cannot be rated because they are beyond the child’s developmental level (e.g., pretend play), thus reducing the denominator for the diagnostic formula. Diagnostic assessment should also include history by parent report (and videos if available) and a review of previous assessments and reports. Difficult diagnoses or ambiguous distinctions require more observation, additional testing in some cases, and often collegial consultation.

Best practice (e.g., Filipek et al., 1999; NRC, 2001) requires the use of a standardized autism observation measure, with the Autism Diagnostic Observation Schedule (ADOS-G; Lord et al., 2000) and the Childhood Autism Rating Scale (CARS; Schopler, Reichler & Renner, 1988) being the two most prominent measures. We use the CARS, which includes primarily observation, but also some structured interaction sequences (e.g., to assess imitation and nonverbal communication), and some parent and/or staff report information. We have found it to have excellent reliability (internal consistency, inter-rater agreement, and test-retest stability) and validity (i.e., correlates well with independent clinical diagnosis) in our sample of older, lower functioning individuals (Perry & Freeman, 1996) as well as in our multi-site, diagnostically heterogeneous preschool aged sample (Perry, Freeman, Condillac, Dunn Geier & Belair, 1999; Freeman, Perry, Condillac, Dunn Geier & Belair, 2000). We have developed our own training procedure for the CARS, with empirical criteria for establishing reliability, and it should be acknowledged that other clinicians may not use the CARS as systematically. In addition, there are a number of parent interview measures and questionnaires designed for different purposes and with varying psychometric properties, which we use occasionally (reviewed by Lord, 1997; Lord et al., 2000; NYSDH, 1999; NRC, 2001; Parks, 1988).

Standardized measures for developmental disability

There are a number of popular beliefs and professional biases we often hear articulated, which include:

- intelligence is irrelevant in the presence of autism
- cognitive abilities can't be measured in these individuals
- many people with autism are unmotivated during testing but are actually quite intelligent
- people with autism are frequently "untestable" on standardized tests
- it's not fair to use verbally-laden measures with nonverbal people, and
- one can never know their true intellectual capabilities.

We regard the above concerns as little more than myths, unsubstantiated by or frankly inconsistent with the data and with best practices. In fact, the vast majority of people with autism (approximately 75-80%) also meet the criteria for a diagnosis of mental retardation or MR (APA, 1994; NRC, 2001; Sattler, 2002). In older individuals with autism seen in treatment centres, this proportion is likely even higher. Many of the learning problems and behavioural difficulties experienced by people with autism are also seen in developmental disabilities (DD) and may be more related to the person’s "level of functioning" than to their autism. Thus, it is crucial that people be evaluated for their level of cognitive and adaptive functioning according to professional guidelines in the DD field (American Association on Mental Retardation [AAMR], 1992; APA, 1994; Jacobson & Mulick, 1996).

Best practice necessitates administering a standardized, reliable, and valid measure of verbal and nonverbal intellectual functioning appropriate to the person’s age and functioning level. The test(s) used should be recently published and have adequate psychometric properties (reliability and validity). In many cases, it may be necessary to use a test intended for younger children, which may mean that standard scores and percentiles cannot be calculated because the individual is older than the normative group or has a cognitive level lower than the lowest in the normative sample. As a result, we sometimes must resort to age equivalents (e.g., Todd scored at a 4 year, 8 month level) or ranges (e.g., Martha’s skills ranged from a 2- to 3-year level or descriptive analysis (which may be more meaningful anyway (e.g., Hassan could complete simple inset puzzles, but was unable to copy geometric designs with blocks). We have to do this for many, if not most, individuals we assess. In such situations, it is important that there be a sufficient range of items in the person’s developmental range regardless of their age, not just two or thee of the lower test items (Filipek et al., 1999; Sattler, 2002). For readers less familiar with principles of psychological assessment, please see Sattler (2002), which is an excellent resource for both theory and practice.

The lack of verbal expressive skills should not, in our view, preclude scoring items requiring receptive or expressive communication, since these tasks form part of the construct called intelligence we purport to be measuring and, thus, should be measured in the same way we would for other people (see AERA, 1999 regarding rationale for not adapting tests when the disability is directly related to the construct of interest). However, separate scores in verbal and nonverbal domains may be clinically useful and in the case of extreme verbal-performance discrepancies, further nonverbal assessment methods may be warranted (but this occurs in a limited proportion of people we see). There is a typical profile of cognitive strengths and weaknesses frequently seen in individuals who are high functioning enough to complete a WISC-III (or similar test). This includes strengths in visual-spatial abilities, eye-hand coordination and rote memory and weaknesses in social comprehension, language comprehension and expression, and abstract reasoning (Sattler, 2002). In a person with a very extreme profile of strengths and weaknesses and/or an extreme verbal-performance discrepancy, the full scale score cannot be interpreted. Also, it is important not to overgeneralize from a particular strength or "splinter skill" (e.g., amazing memory for all radio station call letters) and assume it reflects high intelligence overall. Usually it does not (Sattler, 2002; Sparrow, 1997).

We use a variety of tests including, most often, the Stanford-Binet (4th ed.; Thorndike, Hagen & Sattler, 1986), the Bayley Scales of Infant Development (2nd ed.; Bayley, 1993), and the Wechsler Intelligence Scale for Children (WISC-III; Wechsler, 1991). We would normally begin with a test designed for the person’s age and then, if necessary, supplement with a test intended for younger children. Other cognitive, memory, or neuropsychological testing may also be useful on occasion, as well as academic achievement testing (not discussed in this article).

In addition to the cognitive measure, in order to make the diagnosis of DD (AAMD, 1992; APA, 1994; Jacobson & Mulick, 1996), there must be a measure of the person’s adaptive behaviour in everyday situations, based on information reported by parents or staff who know the person well (it is inappropriate and, indeed, unethical, to diagnose MR/DD based solely on an IQ test). We frequently use the well-known Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla & Cichetti, 1984) as a parent interview as part of initial diagnosis, and have the Classroom Edition completed by staff or teachers for repeated assessments as well as for program evaluation. Although the VABS is frequently recommended for use in autism (e.g., Filipek et al., 1999) and we have found it to correspond well with the Adaptive Behavior Scale (Lambert & Windmiller, 1981; Perry & Factor, 1989), there are a number of difficulties with the VABS, in our experience. In particular, it has rather poor reliability with quite large standard errors (sometimes almost one standard deviation) which often span two classifications.

In addition, the standard scores and age equivalents on the VABS frequently do not correspond well. This may give the impression of strengths and weaknesses when, using the standard scores, (statistically speaking, only standard scores should be compared across scales or tests), there is a flat profile. For example, a 3 year, 3 month-old-child could receive virtually identical standard scores (e.g., 60, 62, 64, 63) on the four domains of the VABS, but obtain age equivalent scores which imply relative differences across domains (e.g., 15, 18, 15, and 23 months respectively). Note also that the ordering of the four domains is not consistent across the two sets of scores.

Also, our clients frequently do not obtain a proper basal on the VABS. Carter et al. (1998) have published "autism norms" for the VABS but, in our view, these are not helpful clinically (e.g., what does it mean to say a person with autism is in the average range on the social domain compared to a group of people with social deficits?) and make little sense conceptually (norms should be based on a normal distribution), although it would depend on the purpose of the assessment.

We also use a variety of curriculum-referenced measures (e.g., Brigance, 1991; Partington & Sundberg, 1988) for intervention planning. In some cases, it may be appropriate to include behavioural assessment of problematic behaviours (e.g., aggression, self-injury). This should include assessment of biological (e.g., does the child have an earache?) and environmental (e.g., is the classroom too noisy?) antecedents as well as functional analysis of the consequences or effects of problematic behaviour (e.g., does the person gain peer attention, or are unpleasant tasks terminated?). These assessment methods are utilized with a view to designing prevention strategies and/or treatment strategies based on the function of the behaviour for the person. Behavioural assessment may be done using a variety of formal assessment measures or informally depending on the clinical situation and severity of the behaviour (Carr et al., 1994; Powers, 1997) but further discussion of these methods is beyond the scope of this paper.

Shared variance

It is important to understand that the construct of severity of developmental disability and the construct of the severity of autism, though conceptually distinct, are not entirely independent empirically. There is some overlap, or "shared variance" (Perry, 2001). To use an analogy, when we say a person is "big", we could mean tall or heavy or both (since height and weight are correlated or have shared variance). As an example of this in the context of assessment, consider the following example: in testing young children with autism on a "shell game" task in which a small toy is hidden under a cup, they will usually find the toy correctly on the first trial. On the second trial, however (when the position of the toy is different) the child will frequently return to the first correct position, even though they watch the object being hidden each time. Is this a symptom of autism (perseveration) or an indication of the person's cognitive level (an "A not B error", typical of toddlers in Piaget's sensorimotor stage 4) or both?

In our research, we have consistently found a moderate negative correlation (in the -.40s to -.50s) between severity of autism, as measured by CARS score and severity of cognitive impairment That is, people who are more severely autistic tend also to be lower functioning cognitively. This finding occurs whether we use cognitive level (MA or IQ type scores) or adaptive behaviour scores (standard scores or age equivalents on the Vineland). It occurs in a large sample heterogeneous for age but homogeneous for level of functioning (Perry & Freeman, 1996) and in a large sample of young children who are heterogeneous with respect to diagnosis and cognitive level (Freeman et al., 2000). This is not simply an artifact of the CARS; we find correlations of similar magnitude and direction with other measures of autism, i.e., the DSM-IV and a variation on the Checklist for Autism in Toddlers (CHAT; Baron-Cohen et al., 1996) (Perry et al., 1999).

We believe these findings can be interpreted as "shared variance" between the two constructs. Correlations of -.40 to -.50, when squared indicate there is about 20 to 25% of shared variance between severity of DD and severity of autism. It does not mean autism and DD are the same thing (75 to 80% of variance is not shared and someone with autism can certainly have an IQ in the average or even superior range), but they are not totally independent either. This is one of the reasons why it is so important to control for cognitive level in research on autism and to ensure adequate measurement of developmental level clinically.

Practical testing strategies

Standardized test administration with this population requires a high degree of skill and clinical judgment regarding acceptable adaptations to standardization. The Standards for Educational and Psychological Testing (AERA, 1999) note that any departure from standardization may affect the validity of the test to an unknown degree. On the other hand, there is an ethical obligation, when working with special populations to make adaptations which may be clinically necessary or justifiable to "minimize the impact of test-taker attributes that are not relevant to the construct that is the primary focus of the assessment" (AERA, 1999, p. 101) and provide the most accurate possible measure of the construct(s) of interest. These include suggested modifications in presentation format, response format, timing, test setting, use of portions of tests, and test substitutions. There are few published guidelines as to what is justifiable (but see Sparrow, 1997) nor any empirical evidence as to their impact

(AERA, 1999). The following are some of the adaptations we consider appropriate and the practical strategies we use to assess clients often considered difficult because of behavioural issues and/or level of functioning.

We want the testing situation to be as comfortable as possible for the person. Depending on the person's age and where the session is taking place, we often have a parent or familiar staff person present at the table (if this is helpful and not disruptive for the person being tested), or in the room but somewhat removed. Testing could take place in the child's familiar classroom, in the person's group home or day program, in our office, on the parent's lap, on the floor, or standing up, as needed. Again, depending on the individual and the location, we may observe and/or play with the child for a time to build rapport and then move gradually to more structured activities, or move more quickly.

There are a few general departures from standardization that we routinely use which, in our clinical judgment, are valid adaptations. We sometimes start at the beginning of a particular subscale (i.e., the easiest item) rather than determining the basal in the prescribed way, based on age. This ensures that the child receives some success right away, minimizes frustration, and it helps build rapport. Similarly, we may begin with a particular type of task (e.g., form board puzzle) we know the child likes or may be successful with and administer subscales in a different order than that suggested in the manual so as to maximize cooperation (e.g., by interspersing easier and more difficult tasks), we and may repeat tasks the person enjoyed following some frustrating tasks prior to a break.

In terms of standardized administration of specific tests and tasks, we typically administer the item as per the manual, at least initially. Then we begin "testing of the limits". First, on tasks which are not inherently verbal ability tasks but which have specific and complex verbal directions, we may try again with abbreviated or slightly altered verbal instructions (e.g., "put it together"), especially if we have reason to believe certain phrases are more familiar to the child (e.g., "match" versus "find me another one just like this"). We may try a generic verbal prompt (e.g., in picture vocabulary, "What is this? Is it a _______?". Sometimes we may do the same task with different materials, especially if we think they are more interesting or motivating for the child (e.g., in a shell game task, rather than the small toy we might put a treat under the cup). We may provide a partial physical prompt to encourage the person to respond (without directing their choice). We may lightly hold their hands until we are finished giving the instruction to inhibit impulsive responding and imitative pointing or tapping of test materials. In a visual matching task, we might use one of the child's hands to point to the model as a way of drawing their attention to the stimulus and then request they "find another one" with their other hand. Often, we will try an item later in the assessment or on another day. Some items (e.g., verbal imitation tasks) we would try in a naturalistic situation (e.g., during a break playing on the floor rather that at the testing table).

In certain situations, we might try to teach the child the task by doing several repeated trials with a consistent instruction and reinforcement and then retest the item. Sometimes we will ask the parent or staff person present if they think the child should be able to complete the task and/or have them try it. Whether or not we would count such modified items as correct for scoring purposes depends on the particular test (some tests allow you to teach the item, some allow item substitutions, etc.) and the degree of departure from standardization, and these decisions require considerable clinical judgment. Usually, we would not count them in the scoring, but use the information clinically in understanding the person and making recommendations about what and how to teach them. In practice, this is less of an issue than it might appear, because, in these situations, usually the person really cannot do the task regardless of how much prompting is provided.

There are a number of legitimate reasons for making these kinds of adaptations. First, they help determine whether the person has the skill rather than demonstrating the skill under very specific conditions. Second, they ensure fairness in not permitting arbitrary or irrelevant factors to inadvertently invalidate test results for a person. Third, they provide some ecological validity or assurance to parents or other observers that the person was given every opportunity to demonstrate the skill and this often helps them accept the validity of the results if the person cannot respond correctly. Finally, it provides clinically useful information about the person’s skills and strategies, about their level of comprehension, the degree of structure they need to perform certain tasks, and how quickly they learn. Note that any adaptations should be individualized for the person being tested (AERA, 1999) and are never made solely for the convenience of the examiner.

The assessment situation can be very difficult and frustrating for clients and it is important to watch carefully for the warning signs of emotional upset and behavioural outbursts. We use standard behavioural strategies to increase cooperation and maintain motivation (Carr et al., 1999; Koegel, Koegel, & Dunlap, 1996; Luiselli & Cameron, 1999). These include reinforcing the person for coming to work, sitting, trying, pointing, and working, without specifically reinforcing correct answers. We use a variety of social, activity, and edible reinforcers as needed and appropriate to the client and the situation (giving them some choice if possible). In terms of dealing with problematic behaviour, we typically use planned ignoring for minor disruptive behaviour (when the client is "testing" us) and attend to appropriate behaviour. If the person throws or attempts to destroy testing materials, we try to ignore the behaviour, rescue and remove the materials from view, and go on to something else and present the item again later. We also use escape from task as a motivator at times (if we do two more, you can have a break) and sometimes prime the person at the beginning of the session regarding how they can ask for a break (i.e., with a break card, sign, or verbal request). We usually take two or three breaks during one cognitive testing session, but it varies depending on the person (it may require more than one session).

If the client is not being successful and compliance versus ability is an issue, we will give some easy tasks or high probability requests to bring the person’s motivation back up or, at least, to end the session on a note of success and cooperation.

In the case of tantrums and refusal to cooperate, sometimes by waiting out the tantrum or working through it (particularly with a young child) and then engaging the child positively in highly motivating activities, the session can be completed, but occasionally it may have to be rescheduled, possibly in a different location. Severe problematic behaviour can usually be avoided by using these strategies but if a client is known to become seriously self-injurious or aggressive, we would plan a strategy in advance and coordinate with the staff present to ensure the client’s safety and that of others. We cannot underestimate the importance of these motivational strategies to assist the person in demonstrating their skills. Rarely, if ever, would we consider a person impossible to test.

Finally, competent performance by the examiner also requires a high degree of familiarity with the test, and the ability to manipulate test materials, and record scores and clinical observations, all at a very quick pace, while maintaining the child’s interest and motivation. We have two people working together, if possible, particularly with certain complex tests (e.g., the Bayley).

Making Assessment Information Useful

Feedback to parents

We take very seriously our professional/ethical obligation (Canadian Psychological Association, 2000) to provide timely and appropriate feedback from assessment in a form than is understandable and helpful to the client (usually the parents in our case). The fact that, in Ontario, the "Controlled Act" of Diagnosis literally refers to "Communicating the Diagnosis" to the client or their representative (CPO, 1995) is an indication of the level of responsibility this process requires (by law, it can only be done by members of the College of Psychologists or the College of Physicians and Surgeons, or students under supervision). We frequently deal clinically with the repercussions of previous inadequate feedback sessions by other professionals. When working with young children, the initial or early diagnosis is a wonderful opportunity to prevent some of the negative consequences of inadequate or poorly done feedbacks. Space precludes a detailed description of our process for providing feedback (described elsewhere by Perry & Condillac, 2000) which, in any case, depends on the purpose of the assessment, the person’s age, and the particular clinical issues involved. However, these sessions involve a combination of providing information which is accurate, thorough, and realistic, in an atmosphere of empathy (see also Abrans & Goodman, 1998; Shea 1984). Our goals include: facilitating at least some initial processing of difficult emotional reactions; offering a sense of confidence, hope, and optimism about the future; and making the connection between assessment and intervention.

Recommendations: From Assessment to Intervention

The written report and verbal feedback resulting from an assessment should be articulated in such a way as to be useful, though what form that takes may vary considerably depending on the purpose(s) of the assessment described at the opening of this paper. For more information on report writing see Sattler (2001). It is obviously important to address referral questions (answering them if possible), to make statements or recommendations regarding eligibility for service if that is part of the purpose of the assessment, to recommend levels of supports, resources, and placements which are appropriate, based on the results. Sometimes, it is helpful to the family to write a separate 1-page letter documenting the main factors (e.g., diagnosis or classification) needed for obtaining resources or advocating for the person. Often, other referrals should be made for more specialized assessment and/or for specific services.

Ideally, there should be a close link between assessment and intervention. In our reports, we make recommendations for further learning which cover a variety of domains (e.g., academic, communication, social/play, self help, etc.) based on the person's current level and the next few steps developmentally in that domain. We will often try to use the person's strengths and interests to help address their weaknesses. If there are many needs and areas requiring intervention, we will usually prioritize those we consider primary. Particular teaching strategies may be suggested based on observations of the person during testing and based on general knowledge of instructional methods. Approaches and procedures for improving attending, cooperation, and appropriate behaviour may also be suggested. These strategies will depend upon the environment and the personnel who will be implementing the recommendations, as caregivers may have considerable experience and skill to draw on or may benefit from more concrete suggestions.In these ways, feedback, reports, and recommendations from an assessment should accomplish one or more of the purposes of assessment described earlier. The information should be useful in helping others to understand the person better so that they can teach or work with the person more effectively, improve their environment, resources, and quality of life. The information should be useful in answering or clarifying diagnostic questions, and in documenting characteristics which provide access to services, supports, and resources for the person. Sometimes, the information will be useful in helping us evaluate programs and services.

Because of the importance of assessment and diagnosis in the lives of people with autism, as well as the risk of harm and the potential benefits, we believe that we have an obligation to take a "best practices" approach to the task. This necessitates remaining current in the field, bringing to bear the most recent evidence-based procedures, evaluating our practices, and holding them together with our accumulated clinical experience and our commitment to following ethical and professional guidelines. This effort will always be a journey rather than a destination. We appreciate the opportunity to describe our journey (so far) to you, the readers of this paper, and hope it will be helpful to you.

Acknowledgements

This paper is based partially on a presentation by Adrienne Perry entitled "Severity of Autism and Developmental Disability: Measurement Issues and Shared Variance" at the Eclipse Conference, Toronto, ON, January, 2001. We are grateful to Bruce Sparks, Discussant at the Eclipse Conference, for the suggestion of a publication describing departures from standardization and practical strategies for testing difficult individuals. We also thank Kim Harris and David Factor for their helpful comments on this manuscript. The views expressed are those of the authors and do not necessarily represent the views of Surrey Place Centre, Thistletown Regional Centre or the Ministry of Community, Family and Children’s Services.

References

Abrans, E. Z., & Goodman, J. F. (1998). Diagnosing developmental problems in children: Parents and professionals negotiate bad news. *Journal of Pediatric Psychology, 23*(2), 87-98.

American Association on Mental Retardation (1992). *Mental retardation: Definition, classification, and systems of support.* Washington, DC: Author

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Educational Research Association (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., et al. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry, 168,* 158-163.

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd ed.) San Antonio, TX: Psychological Corporation.

Brigance, A. (1991). *Revised Brigance Diagnostic Inventory of Early Development.* North Billerica, MA: Curriculum Associates.

Canadian Psychological Asociation (2000). *Canadian Code of Ethics for Psychologists.* Ottawa: Author.

Carr, E. G., Horner, R. H., Turnbull, A. P., Marquis, J. G., Magito-McLaughlin, D., McAtee, M. L., et al. (1999). *Positive behavior support as an approach for dealing with problem behavior in people with developmental disabilities: A research synthesis.* American Association on Mental Retardation Monograph Series.

Carr, E. G., Levin, L., McConnachie, G., Carlson, J., Kemp, D. C., & Smith, C. E. (1994). *Communication-based intervention for problem behavior.* Baltimore, MA: Brookes.

Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J. J., Lord, C., Dawson, G., et al. (1998). The Vineland Adaptive Behavior Scales: Supplementary norms for individuals with autism. *Journal of Autism and Developmental Disorders, 28,* 287-302.

College of Psychologists of Ontario (1995). The regulations, standards of professional conduct and guidelines of the College of Psychologists of Ontario. *The Bulletin, 22*(3), suppl.

Filipek, P. A., Accardo, P. J., Barnek, G. T., Cook, E. H., Dawson, G., Gordon, B. et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29,* 439-484.

Freeman, N. L., Perry, A., Condillac, R. A., Dunn Geier, J., & Belair, J. (2000, June). *Comparison of DSM-IV diagnosis with CARS scores in four diagnostic groups.* Presentation in Symposium entitled, "Scientist-practitioner approach to autism: From early screening to clinical intervention research" at the Canadian Psychological Association, Ottawa , ON.

Jacobson, J. W., & Mulick, J. A. (1996). *Manual of diagnosis and professional practice in mental retardation.* Washington, DC: APA Books.

Koegel, L. K., Koegel, R. L., & Dunlap, G. (1996). *Positive behavioral support: Including people with difficult behavior in the community.* Baltimore, MD: Brookes.

Lambert, N. M., & Windmiller, M. (1981). *AAMD Adaptive Behavior Scale--School Edition.* Washington, DC: American Association on Mental Deficiency.

Lord. C., (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36,* 1365-1382.

Lord, C. (1997). Diagnostic instruments in autism spectrum disorders. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (pp.460-483). New York: Wiley.

Lord, C,., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C. et al. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205-233.

Luiselli, J. K., & Cameron, M. J. (1999). *Antecedent control: Innovative approaches to behavioral support.* Baltimore, MD: Brookes.

Marcus, LM., & Stone, W. L. (1993). Assessment of the young autistic child. In E. Schopler & G. B. Mesibov (Eds.), *Preschool issues in autism* (pp. 149-173). New York: Plenum Press.

National Research Council (2001). *Educating children with autism.* The Committee on Education and Interventions for Children with Autism. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

New York State Department of Health (1999). *Autism/Pervasive Developmental Disorders.*Clinical practice guideline technical report. New York: Author.

Parks, S. L. (1988). Psychometric instruments available for the assessment of autistic children. In E. Schopler & G. B. Mesibov (Eds.), *Diagnosis and assessment in autism* (pp. 123-136). New York: Plenum Press.

Partington, J.W., & Sundberg, M.L. (1988). *The assessment of basic language and learning skills.* Pleasant Hill, CA: Behavior Analysts Inc.

Perry, A. (2001, January). *Severity of autism and developmental disability: Measurement issues and shared variance.* Presented at the Eclipse Conference, Toronto, ON.

Perry, A., & Condillac, R. A., (2000, June). *Assessment and diagnosis of autism spectrum disorders in preschoolers.* Workshop presented at the Canadian Psychological Association, Ottawa , ON.

Perry, A., & Factor, D. C. (1989). Psychometric validity and clinical usefulness of the Vineland Adaptive Behavior Scales and the AAMD Adaptive Behavior Scale for an autistic sample. *Journal of Autism and Developmental Disorders, 19,* 41-55.

Perry, A., & Freeman, N. L. (1996, March). *Psychometric properties of the Childhood Autism Rating Scale (CARS).* Poster presented at the Ontario Association on Developmental Disabilities Research Special Interest Group Research Day, Toronto, ON.

Perry, A., Freeman, N. L., Condillac, R. A., Dunn Geier, J., & Belair, J. (1999, April). *A comparison of developmental and diagnostic indicators in preschoolers with and without a Pervasive Developmental Disorder.* Poster presentation at the Society for Research in Child Development Biennial Conference, Albuquerque, NM, April 14-18, 1999.

Powers, M. D. (1997). Behavioral assessment of individuals with autism. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 448-459). New York: Wiley.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Author.

Sattler, J. M. (2002). *Assessment of children: Behavioral and clinical applications* (4th ed.). San Diego, CA: Author.

Schopler, E. C., & Mesibov, G. B (Eds.). (1988). *Diagnosis and assessment in autism.* New York: Plenum Press.

Schopler, E. C., Reichler, R., & Renner, B. (1988). *The Childhood Autism Rating Scale (CARS).* Los Angeles, CA: Western Psychological Services.

Shea, V. (1984). Explaining mental retardation and autism to parents. In E. Schopler & G. B. Mesibov (Eds.), *The effects of autism on the family* (pp. 265-288) . New York: Plenum Press.

Sparrow, S. (1997). Developmentally based assessments. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 411-447). New York: Wiley.

Sparrow, S. S., Balla, D. A., & Cichetti, D. V. (1984). *Vineland Adaptive Behavior Scales.* Circle Mines, MN: American Guidance Service.

Stone, W. L., Lee, E.B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., et al. (1999). Can autism be diagnosed accurately in children under three years? *Journal of Child Psychology and Psychiatry, 40,* 219-226.

Thorndike, R L., Hagen, E.P., & Sattler, J. M. (1986). *The Stanford-Binet Intelligence Scale:* Fourth Edition. Chicago, IL: Riverside.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children (WISC-III).* San Antonio, TX: Psychological Corporation.