The interrater reliability of the Psychological Processing Checklist

John A. Nonnemacher
Rowan University

Follow this and additional works at: https://rdw.rowan.edu/etd

Part of the Educational Psychology Commons

Let us know how access to this document benefits you - share your thoughts on our feedback form.

Recommended Citation

This Thesis is brought to you for free and open access by Rowan Digital Works. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Rowan Digital Works. For more information, please contact LibraryTheses@rowan.edu.
The purpose of this study was to compare thirty students’ profiles under the Psychological Processing Checklist, a tool that attempts to locate the observable, measurable behaviors evident in a student’s classroom performance that might be indicative of an emergent processing deficiency. Two individual instructors rated the same students, all chosen from two fifth grade classrooms in a Southern New Jersey public school, and the results were compared using correlation data. Significant correlations between the two profiles generated for each student, as measured by the Checklist, lent support for the Checklist as a reliable tool. It was hypothesized that the PPC would demonstrate statistically satisfactory levels of correlation between the profiles generated by the teachers participating in the study. Results of this study showed modest to strong correlations between scores, which provided support for the PPC as a reliable assessment tool.
MINI-ABSTRACT

John A. Nonnemacher
The Interrater Reliability of the Psychological Processing Checklist
2001
Dr. John Klanderman & Dr. Roberta Dihoff
School Psychology

The purpose of this study is to determine the interrater reliability of the Psychological Processing Checklist. If different teachers’ assessments of the same student correlate statistically, the PPC will have demonstrated its reliability for this sample. Results of analysis indicate strong correlations, lending support for the PPC’s reliability.
ACKNOWLEDGEMENTS

Special thanks go to the school principal and four teachers who proved ultimately instrumental to the completion of this thesis paper. Thanks also to Dr. John Klanderman, whose professional associations made it possible for me to participate in a study beyond the scope of my personal interests, and Dr. Roberta Dihoff for being a statistics guru.

Thanks also for the support of family: Mom, Dad, and Gramps, without whom this would have been inestimably more difficult. *Molto grazie* to Rhonda for the Sunday dinners, and to Haddie and John Malkovich for their love.
TABLE OF CONTENTS

Chapter I–Introduction........................................................................................................1
   Need ................................................................................................................................. 1
   Purpose............................................................................................................................ 3
   Hypothesis....................................................................................................................... 3
   Theory............................................................................................................................... 4
   Limitations & Assumptions............................................................................................. 6
   Definitions....................................................................................................................... 7
   Overview.......................................................................................................................... 10

Chapter II–Literature Review ............................................................................................ 11

Chapter III–Methods........................................................................................................ 25
   Subjects ............................................................................................................................ 25
   Measures ......................................................................................................................... 26
   Design .............................................................................................................................. 27
   Testable Hypothesis ........................................................................................................ 28
   Analysis ............................................................................................................................ 28
   Summary .......................................................................................................................... 28

Chapter IV–Results........................................................................................................... 30

Chapter V–Discussion....................................................................................................... 37

References......................................................................................................................... 40
LIST OF TABLES & FIGURES

Figure 1.0–Means Comparison .......................................................... 30
Table 1.0–Descriptive Statistics for PPC ........................................ 31
Figure 1.1–Means Comparison ......................................................... 31
Table 1.1–Summary of Subscale Correlations, Means, and Standard Deviations .... 32
Figure 1.3–Scattergram for Total Raw Scores .................................... 33
Figure 2.0–Means Comparison ......................................................... 33
Figure 2.1–Means Comparison ......................................................... 34
Table 2.0–Summary of Subscale Correlations, Means, and Standard Deviations .... 35
Figure 2.2–Scattergram for Total Raw Scores .................................... 36
Figure 3.0–Means Comparison ......................................................... 37
Figure 3.1–Means Comparison ......................................................... 37
Figure 3.2–Scattergram for Composite Group Data .............................. 38
Chapter I–Introduction

Need

Multiple exigencies exist in the current approaches to meeting the requirements of both federal and state education laws. In addition to the federal special education law, the Individuals with Disabilities Education Act reinforces the democratic precept that all school-age children are entitled to a public school education. Bundled with this legislation is the expectation that students who experience sufficient difficulty learning in the classroom—enough difficulty that they either are formally labeled learning disabled—receive individual attention to address their disabilities.

It is certainly the concern of teachers, administrators, and IEP study team members how best to address learning disabled students’ needs in the classroom. However, before such needs can be addressed, the source of these needs—the very definition of the disability—must be accounted for. Various tools have been developed to measure disabilities, and terms and processes described to operationalize them. A difficulty evident in the use of these tools is, of course, first, whether the teachers that must eventually adopt these definitions and the suggested intervention grasp the nature of the student’s problem, and second, agree that the suggested implementations are
reasonable given their current responsibilities.

The Psychological Processing Checklist, developed by Swerdlick, Kahn, and Swerdlick, is a tool designed to take into account the teacher’s perspective. It attempts to locate the observable, measurable behaviors evident in a student’s classroom performance that a teacher would notice—and find problematic, either as an impediment to learning or to classroom stability—during daily instruction. The Checklist’s items query behaviors in the following domains: auditory processing, visual processing, visual-motor, social perception, organization, and attention. The behaviors described in the checklist can identify behavioral manifestations of various learning disabilities, allowing teachers who are probably unfamiliar with clinical assessments to identify these manifestations without specific knowledge of the underlying conditions. Individual behaviors lend themselves to cross disciplinary discussion; members of IEP teams, including school psychologists, can communicate with other school staff members, including teachers and administrators, as well as parents, about behaviors in a common tongue. Also, interventions based on observable behaviors identified by the Checklist would dovetail cleanly into classroom settings; teachers, having made the observations themselves, would see the intended remediation built into the intervention.

Finally, the Checklist is an attempt at a reliable, valid assessment, of which there
are currently few. In keeping with psychological protocol, the Checklist is currently being tested for both reliability and validity. The subject to be developed forthwith, then, is to determine how the same student is rated by two different instructors to help with the collection of data, suggesting the Checklist’s interrater reliability—or lack thereof. It is to be an individual study among many.

**Purpose**

The purpose of this study is to compare at least thirty students’ profiles under the Psychological Processing Checklist. Two individual instructors will rate the same student, and the results will be compared. Should the study reveal significant correlation between the two profiles generated for each student, as measured by the Checklist, the Checklist will have proven itself reliable in this evaluation. Should the correlation be weak, the study will suggest that further refinements are needed before the test can be deemed a reliable, and hence efficacious, tool for use in the development of interventions for the learning disabled.

**Hypothesis**

The Psychological Processing Checklist will exhibit strong interrater reliability, where different raters will generate statistically similar profiles of students, based on the responses to the Checklist’s items. Initial clinical trials suggest such an outcome.


Theory

Research in the fields of neurology and neuropsychology has helped locate and describe various processing pathologies. However, the question of a cure still confounds professionals. Effective and realistic treatment modalities have, then, necessarily included some kind of compensatory strategy. Hence the role of the school psychologist, who in large measure makes suggestions about how best to treat a student *despite* his learning disability, often in the face of the attitude that such a student must be removed from the classroom (Hendrickson et al 1999).

Not merely biological, processing difficulties harbor behavioral components. As such, treatments must address an individual's behavior, as success in social and academic environments requires "proper" behavior. Indeed, a lack of appropriate behavior, even though caused in part by a defined and recognized biological learning disability, can serve to exacerbate the disorder. Thus, the need for coping skills is manifest; the interaction between brain biology and environment, as well as the child's attempt to organize his or her experience into some sensible whole, require complimentary treatments that address the individual's disability. The assumption built in to the PPC—that students can learn strategies to compensate for their disabilities—meshes neatly with this interactionist paradigm.
Despite the emphasis on medical and other professional opinions, Achenbach (as cited in Teeter & Semrud-Chikeman, 1997) points to both parents’ and teachers’ behavioral assessments when evaluating a child as the most instructive. While this approach incorporates cognitive, medical, and clinical assessments, its reliance on behavioral assessments from a child’s most pervasive social environments—school and home—echoes the purpose of the PPC’s development. Outward behavior, as observed by persons outside of clinical or medical milieus, is a rich, instructive source of data. Kindergarten to fifth-grade teachers, who often see their students throughout multiple periods per day (author’s name missing, 2000), have a wealth of time and variable—and possibly more realistic—settings in which to gather such data (although they may not think of it in those terms exactly).

Anastasi and Urbina (1997) consider ratings like the PPC to offer particular strengths when used in combination with other assessments. First, the casual and informal nature of the data collection distinguish the Checklist from approaches like naturalistic observation; also, the interpretive and judgemental nature of the response format offers a qualitative aspect. Qualitative data is highly regarded in neuropsychological circles in specific “process” approaches, as opposed to strictly quantitative data (Teeter and Semrud-Ckikeman, 1997).
Limitations & Assumptions

Several considerations deserve mention as limitations of the data to be studied forthwith. Criticisms of ratings in general must be mentioned. Concerning relevant contact, the possibility exists that the instructor rating the student may not have witnessed the behavior suspected or noticed by another teacher. Processing difficulties manifest in one classroom may not appear in another, due to differences in environment, teaching style, or the subject matter being taught.

Halo effects can occur when a teacher is requested to rate a student. The possibility that a student performed well on an assignment, or performed a previous favor for the rater, can unduly influence the rater's objectivity. Also, raters of an especially democratic mindset may not want to issue too harsh a judgement on a student, resulting in errors of central tendency.

Lastly, a concern is that the teachers will not have had sufficient contact with the students. While a student's primary classroom teacher will be able to assess her behavior with a strong degree of accuracy, a teacher with more cursory exposure to the student may generate a Checklist profile that is inaccurate and incompatible with the student's primary classroom teacher.

It is assumed that the raters will understand the nature of the Checklist as an
instrument to assess troublesome behaviors, with an eye towards remediation. That is, no punishments or written evaluations will result from the profiles generated by the Checklist during this trial. Also, the status of this study—as an examination of the Checklist’s reliability—should serve to increase impartiality. Sample size should minimize the inherent weaknesses of the rating format, especially considering that this study is one of many others being done to gauge the Checklist’s reliability.

Another assumption is that the raters will have had enough exposure to the fifth graders surveyed. At the particular school where the study’s data will be collected, fifth grade is the lowest grade level (Landis Middle School holds grades five through eight; the PPC is designed for students kindergarten through fifth). Should the school have housed kindergarten through fifth graders, for instance, the likelihood of the raters having witnessed the students’ behavior in various settings over a number of years could enhance this particular data set’s instructiveness.

**Definitions**

*Least Restrictive Environment* – Individualized special education for each eligible student with a disability. To the maximum extent appropriate to the special student’s learning needs, this education is to be offered in regular settings alongside non-disabled students unless he or she could not make satisfactory progress there.
Behavioral Assessment – For the purposes of this study, a behavioral assessment is a tool that includes characteristics that are observable and measurable, and describes characteristics in terms that are easily discussed between disciplines.

IEP (Individual Education Plan) – The process and plan for supplementary aids and services to support exceptional learners in settings with non-disabled students, that emphasize student involvement in the general curriculum.

Intervention – Strategies developed to reduce the experience of a disabling condition sufficient to allow the disabled individual to participate in required or desired activities.

Proactive discipline (FBA) – Discipline-related strategies that shift their focus away from solving discipline problems through suspension and expulsion of students and concentrate their efforts on developing alternative disciplinary strategies to address the unique behaviors of individual students. Proactive disciplinary procedures address the diverse academic and behavioral challenges characteristic of students with or at high risk for emotional and/or severe behavior problems, and eschew a "zero-tolerance" mindset.

Individuals with Disabilities Education Act (IDEA) – Successor to the Education for All Handicapped Children Act, or Part B of the Education of the Handicapped Act. IDEA '97 is legislation that stresses accessibility of general education classroom placement to accountability of academic instruction as well as accountability for positive behavioral
interventions, strategies and supports.

Under IDEA, schools must address the relationship between problem behavior and the classroom learning environment. This is to be accomplished using behavioral assessment, behavioral planning as a component of the IEP process, and the establishment of positive behavioral intervention plans and supports.

IDEA requires involvement of general educators, student participation in state- and districtwide achievement testing, measurable annual goals or outcomes that mention how the disability is going to affect the student's involvement in the general curriculum, progress reports to parents, and mandatory student involvement in the general curriculum, all with the presumption that special education services will be delivered in general education settings.

Compensatory Strategy—Interventions developed to allow a student to adapt to his or her learning disability and effectively learn classroom material, essentially bypassing or compensating for an identified processing deficit.

Learning Disability — (Torgesen, as cited in McLoughlin and Lewis 1994,) Any one of a "heterogeneous group of disorders manifested as significant difficulties in the acquisition and use of listening, speaking, writing, reasoning, or mathematical abilities. These disorders are intrinsic to the individual, presumed to be due to central nervous system
dysfunction, and may occur across the lifespan.”

**Overview**

In the following chapters, readers will review the research literature pertinent to the history and development of behavioral assessments in general and the Psychological Processing Checklist specifically. A detailed methods section will delineate the procedures carried out to collect the data to be analyzed, including the treatment of the raters and the protection of the students’ anonymity. In the results section, the specific findings will be discussed in terms of statistical procedures, indicating the strength or weakness of the correlations found between ratings. Outcomes will be discussed thereafter, including suggestions about how the PPC could be used in future settings. Weaknesses and methodological flaws will be treated as well, with an eye towards how the study could have been more effectively carried out.
Chapter II—Literature Review

The following literature review will be developed in coordination with the advantages of Psychological Processing Checklist (PPC). References to possible expansions of the Checklist’s utility, based on needs highlighted in current literature, will appear throughout. The advantages of the PPC, as listed on a fact sheet by the authors, are that it:

1. “Provides a stronger link between assessment and intervention because items are observable and measurable.” Points worthy of mention that correspond to this rubric can be found under the heading *The Link Between Assessment and Intervention*.

2. “Provides data, which are more easily communicated across disciplines and between school personnel and parents.” Points worthy of mention that correspond to this rubric can be found under the heading *Opening the Lines of Communication*.

3. “Provides information which lead to the development of interventions that are more easily implemented in the students’ classrooms.” Points worthy of mention that correspond to this rubric can be found under the heading *Easily Implemented Interventions*.

4. “Can be used as part of ‘learning styles’ assessment.” Points worthy of mention that
correspond to this rubric can be found under the heading "Learning Styles."

The Link Between Assessment and Intervention

It is difficult to discuss processing disabilities without relying on psychological or neurobiological terms. Discussions of processing disabilities are often mired in academic language (Friel-Patti, 1999). Because this is so, non-specialists are unlikely to be willing to engage in such discussions, and would likely identify themselves as unable to perform any kind of assessment designed to measure processing disabilities.

The Psychological Processing Checklist (PPC) is a bold step towards meeting this knowledge gap. The behaviors commonly manifest in concordance with processing difficulties can help to identify these disabilities. Difficulties in auditory processing, for instance, have proven to be reliable in differentiating children with ADHD alone from children with ADHD and Behavior Disorder/Multiple Complex Developmental Disorder (Lincoln et al, 1998). The PPC purports to measure several auditory processing deficits: auditory discrimination, figure ground analysis, and synthesis of auditory information, as well as auditory motor integration. Also, because auditory processing difficulties can be useful in distinguishing between ADHD and Behavioral Disorder/Multiple Complex Developmental Disorder (BD/MCDD), the utility of the PPC could be expanded to early detection of developmental disorders, and thus referral to other specialists as needed.
(BD/MCDD is suspected of being a developmental disorder, and of occurring more frequently than ADHD—ibid).

Fuchs (1999) suggests that for assessment to have any connection to instruction, teachers should understand their students’ areas of difficulty. The PPC purports to do just this, claiming to measure subsets of nine processing difficulty domains: auditory processing, as mentioned above; visual processing; visual-motor processing; social perception; organization; social perception; automaticity and speed of processing; attention; conceptualization; and monitoring. However, the PPC’s mode of presentation is unique in that it measures these complex domains by compiling checklist responses keyed to behaviors often concomitant with such processing disorders. Classroom teachers necessarily witness and respond to behaviors, which are observable and quantifiable. Internal processing malfunctions, on the other hand, are clearly invisible to teachers’ eyes, and most often not within their realm of specialization.

Behavioral variables associated with ADHD symptoms in children have been able to reliably distinguish between children with ADHD and a non ADHD control group (Pineda et al, 1999). Surprisingly, behavioral variables were more successful in distinguishing ADHD children from control subjects than strictly neuropsychiatric measures, which provided more modest correlations. It should be mentioned that the
behavioral rating scale in this study measured variables similar to the PPC criterion variables for attention.

Various "alternative assessment" methods have been developed—most famously the "portfolio"—in order to provide common ground for students who typically score lower on classic classroom testing measures—multiple-choice examinations, essays, and writing. The orchestration of such assessment methods require a great deal of the teacher’s time, and many educators are suspicious of their reliability (Kleinert et al, 1999), or consider them subjective to the point of outright inaccuracy. The PPC, being short (thirty items) and keyed to measuring processing difficulties through observable behaviors, might allow teachers to retain more of their favored teaching style and methods, through the development of interventions that aim to correct disability-related behaviors. If IEP interventions are targeted on behaviors manifest due to processing difficulties, the teacher might need no or only slight modifications to his or her teaching approach.

Lastly, Schurr (1998) differentiates "authentic" assessment from other methods of assessment. Authentic assessment entails concentration on processing of information rather than traditional recall of information. The stockpiling of knowledge and skills is necessary; however, understanding the means by which students gain skills and
knowledge is crucial, too. For Schurr, the focus should be on the developmental needs of students, rather than allowing testing to determine classroom content. If Schurr is correct, the link between style and assessment will necessarily involve processing analysis, the likes of which may already be established in the PPC.

**Opening the Lines of Communication**

Concerned and interested parents typically want to be involved in the IEP process; one would guess that the literature would bear out the hypothesis that the children of parents who involve themselves actively with their child’s education—special or otherwise—would score higher on measures of academic success and achievement when compared with those children whose parents are less involved. In order for the lines of communication between parents and teachers to be open and clear, teachers need to have an understanding of their students’ special needs and challenges—to this end, a common language must be spoken (Huefner, 2000). Within the school, of course, the same is true: IEP team members and teachers must be able to share a common tongue. The PPC’s behavior-specific language meets these needs. The checklist’s items are not “mired in academic language,” as mentioned above, and it does not require a clinician to complete the checklist.

Despite the best intentions of parents, parent assessments often yield inaccurate
diagnostic profiles. Sanford et al (1999) found this to be especially true when parents and adolescents used DSM-IV subtype measures in assessments. However, behavioral measures were found more reliable in predicting the clinical course of emergent disorders.

Teachers, by contrast, tend to rate students more accurately. Handwerk and Marshall (1998) found that teachers rated students higher for behavioral or external disturbances than parents. The authors list several reasons why this is so: Parents may be prone to being less harsh in rating their children, or may not perceive the advantages to labeling their child as having some kind of disability or difficulty. Teachers may adhere to strict guidelines to ensure that children eligible for special education services receive them. The school and home environments are different, both physically and in the demands that are put on the children. Teachers tend to be more strict and less flexible; also, with a moderate to large classroom, the possibilities for distraction are numerous. Nonetheless, the PPC is a tool designed for use in school, for screening for possible processing difficulties that could hamper academic success. Teachers, it could be assumed, would be less likely to rate a student favorable than a parent, were such favorable ratings inaccurate. In this case, the facts bear out the assumption.

Based on findings by Handwerk and Marshal (ibid), the PPC may be able to
distinguish between children with Learning Disabilities and those with Serious Emotional Disturbances (SED). Typically, SED children lack processing disorders, yet come to the attention of IEP teams when their behavior becomes suspect. Because some overlap does occur—SED students could also have comorbid processing difficulties sufficient to warrant learning disability interventions—, SED students intervention or treatment needs could be neglected is LD were ruled out due to the identification of SED, or worse, should LD or SED be ruled out in favor of either case. Why? Because the typical viewpoint, according to Hartwig and Ruesch (2000) considers the causal or noncausal relationship between aberrant behavior and a learning disability to be identifiable just after a critical behavioral incident. Indeed, the National Joint Committee on Learning Disabilities, in a position paper from 1998, indicated that “Learning disabilities may occur concomitantly with other disabilities (such as...serious emotional disturbance)...[but] they are not the result of those conditions or influences.” Lack of a learning disability with the presence of a serious emotional disorder, which is defined by the NJCLD as some kind of processing impairment, could lead to no intervention, since SED often involves behavioral disruptions as well.

Relationships between behavior and academic success continue to emerge. Taylor et al. (2000) suggest that teachers, who are clearly superior judges of student academic
achievement, are very accurate predictors of learning disability-related difficulty in its nascent states. Specifically, kindergarten teachers using cognitive, achievement, and behavioral assessments could reliably identify children with phonological processing deficits. This same study also revealed that students identified by their kindergarten teachers as substandard in specific academic achievement behaviors were receiving special learning assistance by their third year in school. Two strengths of the PPC apply to the results of this study—the strength of teacher assessment, and the predictive power of behavior in identifying present or future learning disabilities. Combining the ease of communication of behavioral data with the power of teacher assessment will be a mutually beneficial arrangement for parents, students, teachers, and IEP team members.

**Easily Implemented Interventions**

In an interview with Dr. Richard Van Acker, a nation-wide school consultant for conducting functional behavioral assessments, Bullock and Gable (1999) highlight the difficulty inherent in behavioral analysis when teachers must consider their own behavior as a possible trigger for the disabled student. In fact, Arroyo et al (1999) identify teacher behavior as a key influence on student underachievement. As mentioned above, the PPC invites teacher assessment of behavior. Because it is grounded in processing measurement, and because it directly solicits teacher evaluation, the test may serve to
disarm the defensive teacher who regards IEP team suggestions about her own triggering behavior as an insult or an evaluation of her own style. The disarming of such an attitude, while understandable in some cases, would serve to make IEP implementation easier.

To this end, IDEA '97 legislation expects that general educators participate in the IEP process. For students with behavioral and possibly emotional disorders, the most profound change in the IDEA '97 may be that stipulation that schools address the relationship between problem behavior and the classroom learning environment (Hendrickson et al, 1999). Behavioral assessments, behavioral planning as part of the IEP process, and the establishment of positive behavioral interventions, plans, and supports are designed to accomplish this goal. This means that behavior analysis will be required in many more cases due to the failure of standard disciplinary practices to eliminate problem behavior (ibid)—fertile ground for a behaviorally intensive checklist like the PPC.

The expectations of assessments have increased with this newer legislation. Not only is the identification of behaviors enough; the purpose of such behaviors are to be explicated (High School Magazine, 2000). The PPC, which references the domains in which processing difficulties reside, recognizes the interplay of behavior and disability.

Behavioral assessments are intended to do just what their name suggests—identify problem behaviors. Functionally speaking, IEP team members then use the defined
behaviors to create alternative behaviors, which are to be rewarded in the classroom (Scott et al, 2000). The PPC, which attempts to measure weaknesses in processing, can also identify those modes of processing which may represent the student’s strengths in processing. Thus, the PPC could aid in establishing the alternative behaviors that IEP behavioral assessments are designed to profile.

Nichols (2000) sees behavioral measures as often inadequate in catching inconsistent, difficult to observe behaviors—those that occur with less frequency, duration, or intensity, but require intervention nonetheless. In instances such as bringing a firearm to school (considered by Nichols to be a “unique, low-frequency behavior”), typical IDEA-based strategies of dealing with aberrant behavior are usually deemed inappropriate; special education personnel and school psychologists feel that typical behavioral plans spawned from the usual assessments are unsuitable, impracticable, and ineffective in such instances (Nelson et al, 1999). As mentioned earlier, the typical viewpoint considers that causal or noncausal relationships between aberrant behavior and a disability to be identifiable after a critical behavioral incident (Hartwig and Ruesch, 2000). However, there are multiple causes to “bad” behavior, and not in every instance will the behavior be caused by a student’s disability. This suggests expanded utility for the PPC—its measures of social perception may be useful in identifying students who
could engage is such low-frequency, low intensity behaviors.

"Learning Styles"

This point is related to the research cited above in section three on the development of interventions. The PPC, while designed to measure processing deficits, can also indicate relative areas of processing strength—at least in comparison to relative weaknesses. As such, the PPC can be used as part of a “learning styles” assessment. This is helpful for the teacher who, having found something lacking in her or his teaching style, or something lacking in the way a student is processing her directives, can gain objective information about how to adjust the style in which classroom material is presented. This could be done whether or not the student is classified as learning disabled.

Chapter Two—Summary

The authors of the Psychological Processing Checklist describe five key strengths of the assessment:

1. *Observable and measurable assessment items provide a strong link between assessment and intervention.*

Not only is the behavior specific language of the checklist items easy for raters to understand, recent literature shows that behavioral assessments are especially effective in detecting processing disabilities. The behavioral nature of the checklist items measure
what the teacher sees—student behavior—rather than heady academic constructs.

Essentially, the PPC purports to measure processing difficulties, which is crucial for teachers to understand, as they are the source of many disturbing outward student behaviors. Behavioral assessments that generate classroom interventions that merely direct, and do not elucidate the source of students’ difficulties, do not allow for either the teacher to modify his style, or in other cases, forces him to modify his style beyond the boundaries of what he considers accurate classroom assessment.

2. Checklist data is more easily communicated between school personnel and parents.

The PPC’s behavior-specific language allows teachers to evaluate their students accurately, and for IEP team members to describe the processing deficits that cause aberrant student behavior to both teacher and parents. Also, the checklist data is generated by teachers, who have shown themselves to be more accurate in identifying and predicting processing disabilities. Additionally, the utility of the PPC may be expanded in that it could help discover an overlap of serious emotionally disturbed children who have comorbid learning disabilities, who might otherwise not receive proper care should LD be ruled out due to the presence of SED (which does not involve processing deficits).

3. Classroom interventions based on PPC-generated information are more easily developed.
The PPC's emphasis on processing difficulties stresses nature, rather than nurture or self. Should a teacher need to modify his or her style due to that style including some kind of trigger for the student, this focus on the processes of the brain would solve the dilemma that a teacher might believe there to be some personal conflict. Recent literature also suggests that teacher behavior has a stronger than expected influence on student underachievement. Newer expectations of assessments outlined in the IDEA '97 legislation mandate that schools more closely examine the relationship between the classroom environment and problem behavior, and so the need for an assessment tool like the PPC is more manifest than ever. Further, the PPC addresses the expectation that the source of behavioral problems be explored, rather than immediate "quick fixes" be applied without an understanding of the functional framework beneath the problem. The PPC shows promise in helping to develop classroom interventions by not only identifying weaknesses, but by potentially revealing the individual's strengths, or more reliable modes of processing. Finally, social perception measures of the PPC may expand the checklist's utility by identifying students more likely to engage in behaviors that are not suited to standard behavior analysis and intervention.

5. The PPC can be used as part of a "learning styles" assessment.

While the PPC is designed to measure processing difficulties, it de facto suggests
areas of relative processing strength. Such information would be valuable to teachers looking to modify their approaches to reach a student who seems to be having difficulty in the classroom.
Chapter III—Method

Subjects

Thirty fifth grade students, ages 10 to 12, were randomly selected from a public middle school (housing grades fifth through eighth) in Southern New Jersey. Students were primarily caucasian, latino, or African American. Fifteen students were chosen from each of two fifth grade classrooms, both of which are team-taught; that is, the two classes alternate between two teachers and their classrooms. For the purposes of this study, the two classes will be considered as either the “Teacher A-B Group” or the “Teacher C-D Group.” So,

A-B Group (n=15)  C-D Group (n=15)

/   \                      /   \
Teacher A  Teacher B  Teacher C  Teacher D

Bilingual students were eliminated from the subject pool. Due to both cultural differences and the emotional effects of being in a milieu where one does not comfortably speak the native tongue, bilingual students were considered to be nonrepresentative of the sample. Concern over how their results might skew the data and misrepresent the reliability of the PPC in this study led to the decision to exclude them from the subject
Measures

The measure used for the study is the Psychological Processing Checklist, which has undergone preliminary testing only. In addition to collecting demographic information (which accounts for special education enrollment and primary disability), the Checklist’s scales measure six processing areas: auditory (7 items), visual (7 items), visual-motor (6 items), social perception (5 items), organization (5 items), and attention (5 items). The authors of the PPC report that the scales embedded in the Checklist show satisfactory reliability (.85–.93). Four of the original six processing scales show strong construct validity; however, three of the scales (auditory, attention, and organization) did not distinguish themselves from one another.

The Checklist queries the rater to assign a score, Likert-style, from zero (“never”) to three (“often”), indicating how frequently a student engages in a behavior. Higher scores result in a higher loading of subscale items; beyond a set threshold, the student would be considered in need of special education referral. Specifically, the Checklist uses percentile rankings; “deviant” is a score in the 90th percentile (raw scores equal 58 for girls, 74 for boys).
**Design**

All students’ initials were written on small pieces of paper, folded in half so that the initials would not be visible until the paper was unfolded, dropped into brown paper bags labeled either “Class One” (Teacher A-B Group) or “Class Two” (C-D Group), and the bag was agitated. Each of the four teachers drew names, taking alternating turns, from the bag containing the students that they taught, until fifteen students had been selected from each of the two classes. The initials were then recorded on a master list, and subject numbers were later added for the purpose of analysis. (By the analysis stage, initials had been eliminated; students were codified by number for analysis).

After a brief introduction to explain the purpose of the study, the four teachers were furnished with a copy of the master list of students selected from their respective classes (Class One or Two), a large manilla envelope, and a PPC for each student with that student’s initials printed on the top of the checklist. Teachers A and B each rated the same 15 students from Class One; Teachers C and D did the same for students from Class Two; however, they were not allowed to make comparisons or discuss the ratings assigned. Each checklist was completed individually and privately, at the teacher’s discretion, within the course of twenty days. All completed checklists were placed in the furnished envelope, and sealed and returned to the investigator at the conclusion of the
twenty day test period.

**Testable Hypothesis**

The Psychological Processing Checklist will exhibit strong interrater reliability, where Teachers A and B will generate similar subscale and overall profiles of students, and Teachers C and D will do the same. These profiles will prove to be statistically significant, based on the responses to the Checklist’s items.

**Analysis**

Checklist profiles were generated following the checklist’s scoring rubric, which compiles raw scores for each of the six subscale sections mentioned above. Spearman’s Rank Order Correlation Coefficient (Spearman’s $r$) was calculated to determine whether the raters’ responses generated significantly similar profiles between students based on subscale comparisons as well as an overall raw score (composed of the individual items within the subscales), disproving the null hypothesis, or if any similarities were generated by chance, proving the null hypothesis.

**Summary**

Fifteen students were randomly chosen from two classes each of approximately thirty students, to yield a subject pool of thirty. PPC interrater reliability was measured by having the two teachers who alternately instruct each class (for a yield of four teachers)
each rate fifteen students. Hence, fifteen students were rated by Teachers A and B, while another fifteen students were rated by Teachers C and D. Interrater reliability of the Checklist was determined using Spearman’s $r$, applied to both the subscale raw scores alone as well as the individual items that comprise the entire Checklist. Correlations were calculated for both the A-B and C-D groups individually, and for a Composite group of all thirty students, considering the four teachers’ ratings.
Chapter IV–Results

It was hypothesized that the PPC, when given to different teachers to rate a sample of students that they team-teach, would show strong interrater reliability. That is, two pairs of teachers, using the PPC, would rate their students similarly, and the correlation between the profiles would be statistically significant.

Teacher A-B Group (n=15)

For the group of students rated by teachers A and B, Spearman Rank-Order Correlation analysis showed a strong correlation for raw scores in total \( r = .806, p < .01 \).

Total raw scores were obtained by adding all subscale scores. Mean scores (See Figures 1.0 and 1.1) were 32 for Teacher A and 21.8 for Teacher B.

![Figure 1.0–Means Comparison (1 of 2)](image-url)
Similar means between groups highlight the notable congruence between the two teachers’ ratings. However, large standard deviations (23.872 and 29.619 respectively) are conspicuous. Such large standard deviations are similarly reflected in the standardization studies carried out by the authors of the PPC. Interestingly, in the authors’ study, the special education sample shows a much smaller standard deviation compared to regular education samples (see Table 1.0).

<table>
<thead>
<tr>
<th></th>
<th>Regular Education</th>
<th>Special Education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=944)</td>
<td>(n=180)</td>
</tr>
<tr>
<td>Mean</td>
<td>24.96</td>
<td>49.17</td>
</tr>
<tr>
<td>Standard Dev.</td>
<td>24.46</td>
<td>24.01</td>
</tr>
</tbody>
</table>

Table 1.0–Descriptive Statistics for PPC (as noted in authors’ standardization study)

In this correlational study, three of the 15 students in the A-B group were listed as enrolled in special education. Of these 3 subjects, only one student, a male, exceeded the
deviant threshold (90th percentile; raw score=58 for girls, 74 for boys). Both teachers rated this student as deviant.

Subscale correlations, using Spearman’s $r$ corrected for ties, showed modest to strong correlations for the A-B group. Table 1.1 presents the $r$-values corrected for ties, as well as mean and standard deviation data.

<table>
<thead>
<tr>
<th>Processing subscale</th>
<th>Rho (corrected for ties)</th>
<th>Mean (teacher a)</th>
<th>S.D.</th>
<th>Mean (teacher b)</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory</td>
<td>.728</td>
<td>7.867</td>
<td>5.139</td>
<td>5.533</td>
<td>6.151</td>
</tr>
<tr>
<td>Visual</td>
<td>.817</td>
<td>5.867</td>
<td>4.897</td>
<td>3.200</td>
<td>5.493</td>
</tr>
<tr>
<td>Visual Motor</td>
<td>.760</td>
<td>3.4</td>
<td>2.898</td>
<td>3.067</td>
<td>4.667</td>
</tr>
<tr>
<td>Social</td>
<td>.677*</td>
<td>4.267</td>
<td>4.148</td>
<td>2.267</td>
<td>4.317</td>
</tr>
<tr>
<td>Organizational</td>
<td>.779</td>
<td>5.133</td>
<td>4.086</td>
<td>3.600</td>
<td>5.082</td>
</tr>
<tr>
<td>Attention</td>
<td>.831</td>
<td>5.467</td>
<td>4.155</td>
<td>4.133</td>
<td>4.868</td>
</tr>
</tbody>
</table>

*p<.05; p<.01 for all others.

Table 1.1: Summary of Subscale Correlations, Means, and Standard Deviations

Overall, correlations between total raw score measures and subscale measures were strong, evincing a linear relationship (see Figure 1.3). Large standard deviations reemerge in this sample, as they did in the PPC’s authors’ standardization samples. However, this effect is greatly suppressed when used with Special Education populations, for which the Checklist was designed.
Figure 1.3—Scattergram for Total Raw Scores.

*Teacher C-D Group (n=15)*

Spearman’s $r$ revealed a robust correlation between the C-D group (.883 corrected for ties; $p<.01$) raw score totals. Again, total raw scores were obtained by adding all of the PPC’s subscale scores together. Mean scores for Teachers C and D (see Figures 2.0 and 2.1) were 18.8 and 14.533, respectively.

Figure 2.0—Means Comparison (1 of 2)
As with the A-B group, large standard deviations emerged for the C-D group (21.318 for Teacher A, 12.872 for Teacher B). Because this effect is substantially diminished when Special Education populations are selected for (see Table 1.0 above), it should be noted that in the C-D group, only one subject was designated as Special Education in the Checklist’s demographics section. Neither teacher rated this student above the deviant threshold according to the PPC’s criteria; the one student rated above the deviant threshold was not indicated as being enrolled in Special Education.

Interestingly, for this student, only teacher C rated her above the deviant threshold (total raw score=68; deviant for girls=58 or greater). Teacher D, conversely, rated her well below the threshold at 27 (total raw score).

Corrected for ties, Spearman’s $r$ between the subscale scores for the C-D group evinced modest to robust correlations. Table 2.0 presents these correlations, as well as mean and standard deviation data.
Table 2.0—Summary of Subscale Correlations, Means, and Standard Deviations

Overall, correlations were strong for both total raw scores and subscale scores, evincing a linear relationship (see Figure 2.2). Again, and possibly due to the preponderance of Regular Education students in the sample, large standard deviations riddled the data.

<table>
<thead>
<tr>
<th>Processing subscale</th>
<th>Rho (corrected for ties)</th>
<th>Mean (teacher c)</th>
<th>S.D.</th>
<th>Mean (teacher d)</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory</td>
<td>.852</td>
<td>4.733</td>
<td>6.692</td>
<td>6.400</td>
<td>5.054</td>
</tr>
<tr>
<td>Visual</td>
<td>.924</td>
<td>4.067</td>
<td>5.391</td>
<td>1.533</td>
<td>2.416</td>
</tr>
<tr>
<td>Visual Motor</td>
<td>.785</td>
<td>2.400</td>
<td>2.849</td>
<td>1.533</td>
<td>1.246</td>
</tr>
<tr>
<td>Social</td>
<td>.614*</td>
<td>1.800</td>
<td>2.455</td>
<td>.467</td>
<td>.640</td>
</tr>
<tr>
<td>Organizational</td>
<td>.643*</td>
<td>2.800</td>
<td>3.385</td>
<td>1.667</td>
<td>2.059</td>
</tr>
<tr>
<td>Attention</td>
<td>.738</td>
<td>2.800</td>
<td>3.406</td>
<td>3.200</td>
<td>3.052</td>
</tr>
</tbody>
</table>

*p < .05; p < .01 for all others.

*Bivariate Scattergram with Supersmoother*
Figure 2.2—Scattergram for Total Raw Scores

*Composite Group (Teachers A & C, B & D; n=30)*

Composite group scores, where all teachers’ ratings were taken together, demonstrated a strong correlation ($r=.749, p<.01$). In this group, a total raw score was compiled from each student’s PPC profile, and the students profiled by Teachers A and C were compared to those profiled by B and D to compute Spearman’s $r$. Means for the Composite group were 25.4 (A & C composite) and 18.167 (B & D composite); see figures 3.0 and 3.1 for a comparison.
For the Composite group, means data correlated strongly. As with the A-B and C-D groups considered above, the data resembled a generally linear relationship (see Figure 3.2).
Considering the significance of the correlations between the profiles generated by both Teacher A and B, and then C and D, the null hypothesis can be rejected, and the testable hypothesis of this study can be accepted. Specifically, for both groups A–B and C–D, raw score totals and subscore totals showed modest to strong positive correlations. For the Composite group, raw score totals showed a strong correlation as well.
Chapter V—Summary and Conclusions

Summary

Thirty students were selected from two fifth grade classrooms in a suburban southern New Jersey public school. Each class was conducted by two teachers at different times of the day, making it an appropriate sample for an interrater reliability study such as this. As such, two groups of two teachers were solicited to complete Psychological Processing Checklist profiles for fifteen of their students. It was hypothesized that the PPC would demonstrate statistically satisfactory levels of correlation between the profiles generated by the teachers participating in the study.

Correlational data revealed that the Psychological Processing Checklist demonstrated strong interrater reliability. For the A-B and C-D groups, as well as the Composite group, total raw scores and subscale raw scores correlated in the modest to strong range. As such, the testable hypothesis—that the PPC would demonstrate strong interrater reliability—was supported.

Discussion

Large standard deviations—sometimes larger than the means data themselves—appeared for both sample groups, in both raw score and subscale analyses. While it is not the proper topic of this study to comment on the Checklist’s ability to accurately identify students with processing deficiencies, the inclusion of these standard deviations in the results chapter warrants clarification. The Checklist itself, designed to measure instances of deviance in behavior that could indicate processing deficiencies, is likely to generate
skewed data. The authors recognize this, and hence use percentile determinations rather than conversion to a z-score, which would make use of standard deviations in the style of the intelligence test. With standard deviations as large as found in this study, designations made on such a basis would be uninformative.

Regular Education students will score low—extremely low, perhaps—on the PPC. This effect keeps the Checklist from fitting nicely into a bell curve shape, especially as the Checklist is used on Regular Education populations for standardization, reliability, and validity studies. As the strengths and weaknesses of the Checklist in these areas are made manifest, it will increasingly find itself tested in the more homogeneous populations for which it was designed.

As highlighted in the results section, the effect of large standard deviations diminishes greatly when used in Special Education populations. This study’s sample consisted of only four Special Education enrollees. A preponderance of very low- and low-scoring students is probably at the root of such large standard deviations, students for whom the checklist would not likely be used.

The heart of this study is whether or not two teachers, each rating the same group of students, would generate statistically significant similar profiles of the students. Indeed, this has been shown; however, the relatively small number of Special Education enrollees means that, for the great majority of the sample, scores should have been low—well below the deviance percentile as set forth by the Checklist’s authors. Inspection of means data verifies this, so a more accurate assessment of this study’s results might read, “Teachers similarly assess students who are not enrolled in Special Education as being
largely free from processing deficiencies, the presence of which would suggest the need for Special Education services.” In fact, within the subscale data sets, there were many zeroes. Nonetheless, the correlations are strong enough to support the authors’ use of items that distinguish between disordered and healthy processing as manifest in overt classroom behaviors. Were otherwise the case, means data would be remarkably dissimilar.

**Implications for Future Research**

Research that looks into whether the PPC can adequately identify those students who have *bona fide* processing deficiencies will be most promising for the future of the instrument, as will interrater reliability studies, such as this, that determine how closely two teachers can come to identifying an existing processing deficiency. Of course, the difficulty here is between dichotomy and gradation—one either *is* processing deficient or one *isn’t*. While most professionals in the business of making disabilities determinations recognize that this black and white view of disability is an oversimplification, all instruments that purport to either diagnose or make pre-emptive designations must submit before a threshold. The point is that studies that determine the PPC’s effectiveness within the subpopulation it was designed for will contribute most to the certification of the PPC as a valid and reliable means of special education pre-screening.

**Conclusions**

All of this is necessary technical detail to create what the authors of the PPC have envisioned: a methodologically sound, scientifically tested instrument that does what it purports to do. What, then, does the PPC purport to do? By design, it is to be a
layperson’s tool for making initial determinations as to whether a child might have a processing deficiency. It is not designed to eliminate the professional from making the appropriate diagnosis; rather, it is intended to put into the language of every day behavior the outward symptoms of a burgeoning processing deficiency.

The PPC’s immediate goal of determining whether more formal screening is necessary is a short-term, limited one. In the final analysis, what makes the PPC so worth the research effort is, first, the empowerment of the classroom teacher in the IEP process, and second, the integration of considerations of processing deficiencies that make the IEP more effective.
References


Nelson, J.; Roberts, M.; Rutherford,R. A statewide survey of special education
administrators and school psychologists regarding functional behavioral assessment.


Nichols, P. Role of cognition and affect in a functional behavioral analysis.

(Spring 2000). *Exceptional Children* 66 (3), 393-402.

National Joint Committee on Learning Disabilities. Operationalizing the NJCLD definition of learning disabilities for ongoing assessment in schools. (Summer 1998).

*Learning Disability Quarterly* 21 (3), 186-93.


(February 1998). *Schools in the Middle* 6 (7), 22-7+.
