Please use the following citation when referencing this work:

Maki, K. E., McGill, R. J., Conoyer, S. J., Fefer, S. A., & Ward, T. (2020). Assessing the impact of sequential data presentation on specific learning disabilities identification decisions using patterns of strengths and weaknesses methods. *Journal of Psychoeducational Assessment*. Advance online publication. https://doi.org/10.1177/0734282920983951

Assessing the Impact of Sequential Data Presentation on Specific Learning Disabilities Identification Decisions Using PSW Methods

Kathrin E. Maki[1]

Ryan J. McGill[2]

Sarah J. Conoyer[3]

Sarah J. Fefer[4]

Thomas Ward[2]

[1]University of Florida

[2]William & Mary

[3]Souther Illinois University-Edwardsville

[4]University of Massachusetts-Amherst

Author note

Correspondence concerning this article should be addressed to Kathrin E. Maki, Department of Special Education, School Psychology, and Early Childhood Studies, University of Florida, PO Box 117050, Gainesville 32611, FL, USA.Email: kathrin.maki@coe.ufl.edu

Abstract

Patterns of strengths and weaknesses (PSW) represent relatively novel methods for identifying specific learning disabilities (SLD) with proponents asserting that the incorporation of multiple sources of assessment data and professional judgment play a key role in their utility. In this study, we examined if the sequential presentation of assessment data impacted school psychologists' ratings as to whether or not hypothetical students depicted in special education evaluation vignettes should be identified with SLD. Results showed that when participants viewed vignettes that were indicative of SLD (i.e., SLD Positive), SLD likelihood ratings increased with the additional presentation of assessment data sources over time. However, when participants viewed vignettes that were indicative of a student not having SLD (i.e., SLD Negative), SLD likelihood ratings were relatively consistent over time. Moreover, participants demonstrated relatively high levels of confidence in their SLD identification decisions, and in SLD Negative vignettes, confidence increased after the fourth assessment data source was presented. Implications for SLD identification are discussed.

Keywords: specific learning disabilities, patterns of strengths and weaknesses

Assessing the Impact of Sequential Data Presentation on Specific Learning Disabilities Identification Decisions Using PSW Methods

Specific learning disabilities (SLD) identification continues to be a controversial topic because there is no agreed upon diagnostic gold standard and practitioners have multiple identification methods from which to choose when assessing children and adolescents referred for academic difficulties (Schroeder et al., 2017). Increasingly, practitioners have begun to use a class of "third-way" alternative identification methods known collectively as the *patterns of strengths and weaknesses* (PSW) method to identify students with SLD in schools (Kranzler et al., 2016). To be identified with SLD using a PSW method, a student should exhibit a cognitive weakness that is theoretically or empirically aligned with an academic achievement weakness in the presence of otherwise average or above average abilities (Fiorello et al., 2014). Although there are multiple PSW methods that have been articulated in the literature, the Dual Discrepancy/Consistency method (DD/C), a stepwise procedure that explicitly aligns with Cattell-Horn-Carroll theory (see Flanagan et al., 2018 for an in-depth description) was the most commonly reported PSW method used in a recent national survey of school psychologists' SLD identification practices (Benson et al., 2020).

Empirical Status of PSW Decision-Making

Despite its popularity, a plethora of recent research suggested that the DD/C and other related PSW methods suffer from substantive psychometric problems, which may constrain their clinical utility (Kranzler et al., 2019; McGill et al., 2018). These include, (a) poor construct validity for the scores that are the focal point of PSW analyses, (b) inconsistent long-term stability of PSW score profiles, and (c) a lack of compelling evidence for PSW identification or treatment utility. Despite these limitations, SLD identification requires decision-makers (i.e., school psychologists and multidisciplinary teams) to apply multiple gating identification criteria

to student data. Doing so involves a thoughtful decision-making process and application of professional judgment, which *may* help overcome these psychometric restrictions (Flanagan et al., 2018).

Despite the intuitive appeal of this logic, it remains unclear whether application of the DD/C method results in more accurate or consistent identification decisions given these replicated limitations. Extant empirical research examining SLD identification using PSW methods in general has resulted in inconsistent identification decisions (Maki & Adams, 2020; Schroeder et al., 2017). However, Maki and Adams presented all of the assessment data at one time to participants, which is not consistent with the iterative process by which data are actually gathered in the field. Furthermore, school psychologists frequently obtain assorted sources of information at different times for students. Thus, it is possible that the order in which certain types of information (e.g., psychoeducational test scores) are encountered in the assessment process may lead to different identification decisions depending on how the data may be differentially weighted upon receipt (McGill & Busse, 2017). However, the potential impact of order of data presentation or the amount of information encountered on the likelihood of being identified with SLD using a PSW method has yet to be investigated.

Purpose

Because PSW in general, and DD/C in particular, is a relatively novel SLD identification method, its practical implementation is still not well-understood. Eexamining the potential impact of accumulated data presentation on SLD identification decisions is important because there is significant variability in how SLD evaluations are conducted and the data that are collected as part of the evaluation process in local educational agencies (LEAs) across the nation (Benson et al., 2020; Cottrell & Barrett, 2016). Accordingly, the following research questions guided this study (1) to what extent does the amount of evaluation data and the order in which

data were presented impact SLD identification decisions using the DD/C method, (2) to what extent does the amount of evaluation data and the order in which data were presented impact practitioners' confidence in those identification decisions, and (3) what is the overall consistency of final SLD identification decisions? It is believed that the results of this investigation will further our understanding of the factors that may undergird PSW decision-making.

Method

The study methodology was developed and implemented in line with prior research examining special education identification decision-making (e.g., Maki et al, 2017; 2018; Maki & Adams, 2020; Sullivan et al., 2019). Participants were asked to review SLD identification criteria and student evaluation data, and to determine the likelihood that the hypothetical student should be identified with SLD.

Participants

Participants included 344 licensed or certified practicing school psychologists practicing in states allowing PSW methods for SLD identification. Participants were recruited from state school psychology associations (SPAs) and prominent online school psychology interest groups. Research requests were sent to SPAs in states allowing PSW for SLD identification and research with association members ($N = 30$), and the study was approved by and distributed to members of 15 SPAs. Most SPAs distributed the study at one time point, but three SPAs distributed the study to members at two time points. In conjunction with recruitment from SPAs, we posted the study information in four school psychology Facebook groups each on a Thursday afternoon to maximize recruitment (Kolowich, 2018). Participants were representative of school psychologists practicing in the U.S. (Walcott & Hyson, 2018) and were from 30 states, Washington D.C., and Puerto Rico. Three-hundred (87%) participants identified as women, and 88% identified as European American/white, 4% identified as Latinx, 3% identified as African

American/Black, 3% identified as two or more races/ethnicities, 1% identified as Asian/Pacific Islander, and < 1% participant identified as Native American/Alaskan Native. Participants reported a mean age of 39.21 years (*SD* = 11.01) with an average of 11.01 (*SD* = 8.99) years of experience. A majority of participants reported holding a specialist level degree (67%).

Evaluation Data Vignettes

Eight hypothetical student vignettes were developed, four of which represented a student identified with SLD (SLD Positive vignettes) and four of which represented a student identified not-SLD (SLD Negative vignettes). An external panel of experts examined the vignette assessment data and DD/C criteria (described below) to validate SLD Positive or SLD Negative status of the vignettes. Each of the vignettes included student background information (i.e., developmental and educational histories, teacher and parent reports), progress monitoring data (i.e., curriculum-based measures [CBM] scores and growth rate, Measures of Academic Progress [MAP] scaled scores), classroom observation data, test score data, which included subtest and composite scores from two individually administered cognitive tests and subtest and composite scores from two individually administered academic achievement tests, and behavior rating scale scores. The data sources were presented in different orders across vignettes with cognitive and achievement test score data presented together[1, 2].

Procedure

The study was conducted in Qualtrics, a secure online platform. Upon entering the study platform, participants were randomly assigned to one of eight unique study conditions if they answered affirmatively to screening questions regarding their practice status (i.e., practicing

[1] Background information was presented first for each of the 8 vignettes. Then the presentation order for the remaining data was randomized resulting in four unique orders with the same four orders applied to both SLD Positive and Negative cases to control for potential "true" status bias.
[2] Additional supplementary tables outlining the vignettes in more detail as well as analytical results not described below are provided open access through OSF (blinded for review).

school psychologist in state allowing PSW methods). After being randomly assigned to a study condition, participants provided demographic information (e.g., race/ethnicity, gender, years practicing) and indicated the state in which they practiced, which prompted Qualtrics to present their state's PSW SLD identification criteria to ensure that participants were familiar with the criteria with which they were provided. Qualtrics then presented the DD/C framework criteria, which included a table outlining the requirements for SLD identification at each DD/C level (e.g., standard score achievement deficit below 85, cognitive processing deficit below 85; see Flanagan et al., 2018 for description of all DD/C framework levels). Participants were directed to use the identification criteria they were provided to determine if the student depicted in the forthcoming data should be identified with SLD. Qualtrics then presented participants with data sources in a unique order specific to their randomly assigned condition.

Participants rated the likelihood that the student should be identified with SLD on a scale from 1 to 10 (1 = not very likely, 10 = very likely) after presentation of each assessment data source. In between each data source presentation, the participants also rated their level of confidence in their SLD identification rating on a scale from 1 to 10 (1 = not at all confident, 10 = very confident) at that time. After all data sources were presented, participants viewed all data sources again at once in the same order originally presented. Participants were then asked again to rate the extent to which the student should be identified with SLD and their confidence level in that rating. Finally, they indicated if the student should be identified with SLD with a dichotomous yes or no response.

Results

SLD Identification Likelihood

The progressive SLD likelihood ratings were compared using a repeated measures ANOVA. Results indicated that there were significant effects for vignette (F [7, 336] = 5.81, p <

.001) and the interaction between vignette and time (F [35, 1680] = 30.11, $p < .001$). Potential interactions were examined by conducting controlled, pair-wise contrasts across time within vignette and between vignettes within time (Howell, 2002). For the set of contrasts, the overall Type 1 error rate was set to .05. The post hoc contrasts indicated that the initial SLD likelihood ratings were equivalent for all conditions ($M_{Positive}$ = 5.82, $M_{Negative}$ = 5.83). For the SLD Positive vignettes, the likelihood ratings consistently increased as more data were presented. However, for the SLD Negative vignettes, the likelihood ratings remained relatively flat across the conditions until cognitive and achievement test data were presented and then immediate declines in likelihood ratings were observed in each of the vignettes (see Figure 1).

SLD Identification Decision Confidence

SLD confidence ratings were analyzed in the same manner as the likelihood ratings. Results indicated that there were significant effects for vignette (F [7, 336] = 3.19, p = .003), time (F [5, 1680] = 131.39, $p < .001$), and the interaction between vignette and time (F [35, 1680] = 7.71, $p < .001$). Potential interactions were examined by conducting controlled, pair-wise contrasts across time within vignette and between vignettes within time. The post hoc contrasts indicated that initial decision confidence ratings were equivalent for all vignettes. Whereas decision confidence improved significantly for the SLD Positive vignettes after the presentation of cognitive and achievement test score data, SLD Negative vignettes resulted in higher confidence at the presentation of the fourth source of information regardless of the type of assessment data presented. There were no significant differences in average terminal decision confidence ratings between conditions ($M_{Positive}$ = 6.53, $M_{Negative}$ = 6.34).

SLD Identification Consistency

The consistency of the final identification decisions (i.e., whether or not the student should be identified with SLD) was also examined. Table 1 presents the frequencies of final

identification decisions by vignette status (Positive or Negative). A chi-square test (χ^2 [1, n = 344] = 42.01, $p < .001$) indicated that the accuracy of final decisions was related to the vignette status. Sensitivity (true positive) was .81. However, specificity (true negative) was significantly lower at .51. Whereas the false negative rate was .18, the false positive rate was .48, resulting in a total area under the curve (AUC) of .66 which, according to Youngstrom (2014), is considered to represent low classification accuracy. The accuracy of the final identification decisions was not related to demographic variables or reported participation in professional development related to SLD.

Discussion

This study found that the interaction of data source presentation timepoint and vignette impacted participants' ratings of the likelihood that a hypothetical referred student should or should not be identified with an SLD. Such findings may be encouraging in that the different data presentations appeared to impact SLD likelihood ratings. However, having more data did not appear to serve as a "magic elixir" to ensure that appropriate identification decisions were made (Lilienfeld et al., 2007). In fact, in the SLD Positive conditions, participants' likelihood ratings consistently increased over incremental data source presentation, regardless of which data were presented. But, in the SLD Negative conditions, likelihood ratings remained relatively stable until cognitive and achievement test score data, which indicated a disconfirming PSW, were presented. Thus, these results suggested that a confirmatory PSW pattern may not function as a de facto inclusionary criterion as has been asserted in the literature (e.g., McGill & Busse, 2017). Rather, a disconfirming PSW may function as a de facto rule-out for SLD when using the PSW method regardless of the stage in the assessment process at which it is encountered. Thus, gathering a lot of assessment data may not negate the results of any particular assessment or

overcome the psychometric issues underlying cognitive and achievement test score comparisons, such as with PSW methods.

This practice may stem from the observation that practitioners appeared to overestimate the base rate of having SLD at the onset of referral as the likelihood ratings hovered around ~5.0 on a scale from 1 to 10 (chance levels) in both conditions absent any information other than a student's educational and developmental background. Put simply, school psychologist practitioners in the aggregate appeared to approach decision-making, at the onset, in a simplistic binary manner and may be suffering from base rate neglect (Meehl & Rosen, 1955) because *half* of the participants' identified students were false positives suggesting that practitioners may have difficulty identifying a referred child as not-SLD even in the presence of disconfirming information.

Finally, these findings also suggested that participants' confidence in their SLD identification likelihood ratings increased as they viewed additional sources of assessment data indicating that a larger amount of data resulted in higher levels of confidence, but not necessarily good or even adequate classification accuracy. In prior research, school psychologists similarly demonstrated high levels of confidence in their SLD identification decisions regardless of whether or not they made a consistent SLD identification decision (Maki et al., 2018). Although confidence in one's professional skills may be helpful in some aspects of school psychologists' professional roles, school psychologists and other decision-makers should ensure that high levels of confidence do not negatively interfere with making appropriate identification decisions based on assessment data.

Limitations

As with any study, the findings from the present investigation should be interpreted within the context of its limitations. First, we used two sampling methods, state school

psychology association memberships and online school psychology interest groups; thus, a response rate could not be calculated. Second, although we provided participants with the DD/C framework from Flanagan et al. (2018), we did not provide participants with access to computational software associated with the method (i.e., X-BASS). Although school psychologists have reported using DD/C without the X-BASS software in practice (Maki & Adams, 2019) and this is encouraged by proponents of this method (Flanagan et al., 2018), it is possible that using the X-BASS software could increase decision accuracy and confidence. Finally, whenever diagnostic efficiency statistics as per Kessel and Zimmerman (1993) are calculated within the context of SLD identification, it must be stipulated that no gold standard rule-in test has been developed; thus, these estimates should be regarded as our *best* estimate of how various identification methods likely operate in practice.

Conclusion

The findings furnished in this study suggested that the amount of student evaluation data and the order in which such data were presented impacted SLD identification decisions and confidence in those decisions. Such findings hold important implications for how school psychologists make defensible SLD identification decisions in practice. More importantly, the present study adds to the growing base of scientific literature calling SLD identification into question, particularly in how identification criteria are applied to specific sources of assessment data. In sum, taking an evidence-based assessment (Youngstrom et al., 2017) approach, these results indicated that practitioners applying SLD identification criteria should exercise caution given the high stakes nature of such decisions for students.

References

Benson, N. F., Maki, K. E., Floyd, R. G., Eckert, T. L., Kranzler, J. H., & Fefer, S. A. (2020). A national survey of school psychologists' practices in identifying specific learning disabilities. *School Psychology, 35*(2), 146-157. https://doi.org/10.1037/spq0000344

Cottrell, J. M., Barrett, A. B. (2016). Defining the undefinable: operationalization of methods to identify specific learning disabilities among practicing school psychologists. *Psychology in the Schools, 53*(2)*,* 143-157. https://doi.org/10.1002/pits.21892

Fiorello, C. A., Flanagan, D. P., & Hale, J. B. (2014). The utility of the pattern of strengths and weaknesses approach. *Learning Disabilities: A Multidisciplinary Journal*, *20*(1)*,* 57-61.

Flanagan, D. P., Costa, M., Palma, K., Leahy, M. A., Alfonso, V. C., & Ortiz, S. O. (2018). Cross-battery assessment, the cross-battery assessment software, and the assessment-intervention connection. In D. P. Flanagan, & E. M. McDonough (Eds.). *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 731-776). Guilford.

Howell, D. C. (2002). *Statistical Methods for Psychology*, 5th ed. Duxbury Press.

Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: Extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment, 5*(4)*,* 395-399. https://doi.org/10.1037/1040-3590.5.4.395

Kranzler, J. H., Floyd, R. G., Benson, N., Zaboski, B., & Thibodaux, L. (2016). Classification agreement analysis of cross-battery assessment in the identification of specific learning disorders in children and youth. *International Journal of School and Educational Psychology, 4*(3)*,* 124-136. https://doi.org/10.1080/21683603.2016.1155515

Kranzler, J. H., Gilbert, K., Robert, C. R., Floyd, R. G., & Benson, N. F. (2019). Further examination of a critical assumption underlying the dual-discrepancy/consistency approach to specific learning disability identification. *School Psychology Review, 48*(3)*,* 207-221. https://doi.org/10.17105/SPR-2018-0008.V48-3

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2007). Why questionable psychological tests remain popular. *Scientific Review of Alternative Medicine, 10,* 6-15.

Maki, K. E., & Adams, S. R. (2019). A current landscape of specific learning disability identification: Training, practices, and implications. *Psychology in the Schools, 56*(1)*,* 18-31. https://doi.org/10.1002/pits.22179

Maki, K. E., & Adams, S. R. (2020). Specific learning disabilities identification: Do the identification methods and data matter? *Learning Disability Quarterly, 43*(2)*,* 63-74. https://doi.org/10.1177/0731948719826296

Maki, K. E., Burns, M. K., & Sullivan, A. (2018). School psychologists' confidence in LD identification decisions. *Learning Disability Quarterly, 41*(4)*,* 243-256. https://doi.org/10.1177/073194871876925

McGill, R. J., & Busse, R. T. (2017). When theory trumps science: A critique of the PSW model for SLD identification. *Contemporary School Psychology, 21*(1)*,* 10-18. https://doi.org/10.1007/s40688-016-0094-x

McGill, R. J., Conoyer, S. J., & Fefer, S. (2018). Elaborating on the linkage between cognitive and academic weaknesses: Using diagnostic efficiency statistics to inform PSW assessment. *School Psychology Forum, 12*(4)*,* 118-132.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52*(3)*,* 194216. https://doi.org/10.1037/h0048070

Schroeder, M., Drefs, M. A., & Cormier, D. C. (2017). The messiness of LD identification: Contributions of diagnostic criteria and clinical judgment. *Canadian Psychology, 58*(3)*,* 218-227. https://doi.org/10.1037/cap0000115.

Sullivan, A. L., Sadeh, S., & Houri, A. K. (2019). Are school psychologists' special education eligibility decisions reliable and unbiased?: A multi-study experimental investigation. *Journal of School Psychology, 77,* 90-109.

Walcott, C. M., & Hyson, D. (2018). *Results from the NASP 2015 membership survey, part one: Demographics and employment conditions* [White paper]. National Association of School Psychologists.

Youngstrom, E.A. (2014). A primer on Receiver Operating Characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology, 39*(2)*,* 204-221. https://doi.org/10.1093/jpepsy/jst062

Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, Mian-Li, & Youngstrom, J. K. (2017). Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment. *Clinical Psychology: Science and Practice, 24*(4)*,* 331-363. https://doi.org/10.1111/cpsp.12207

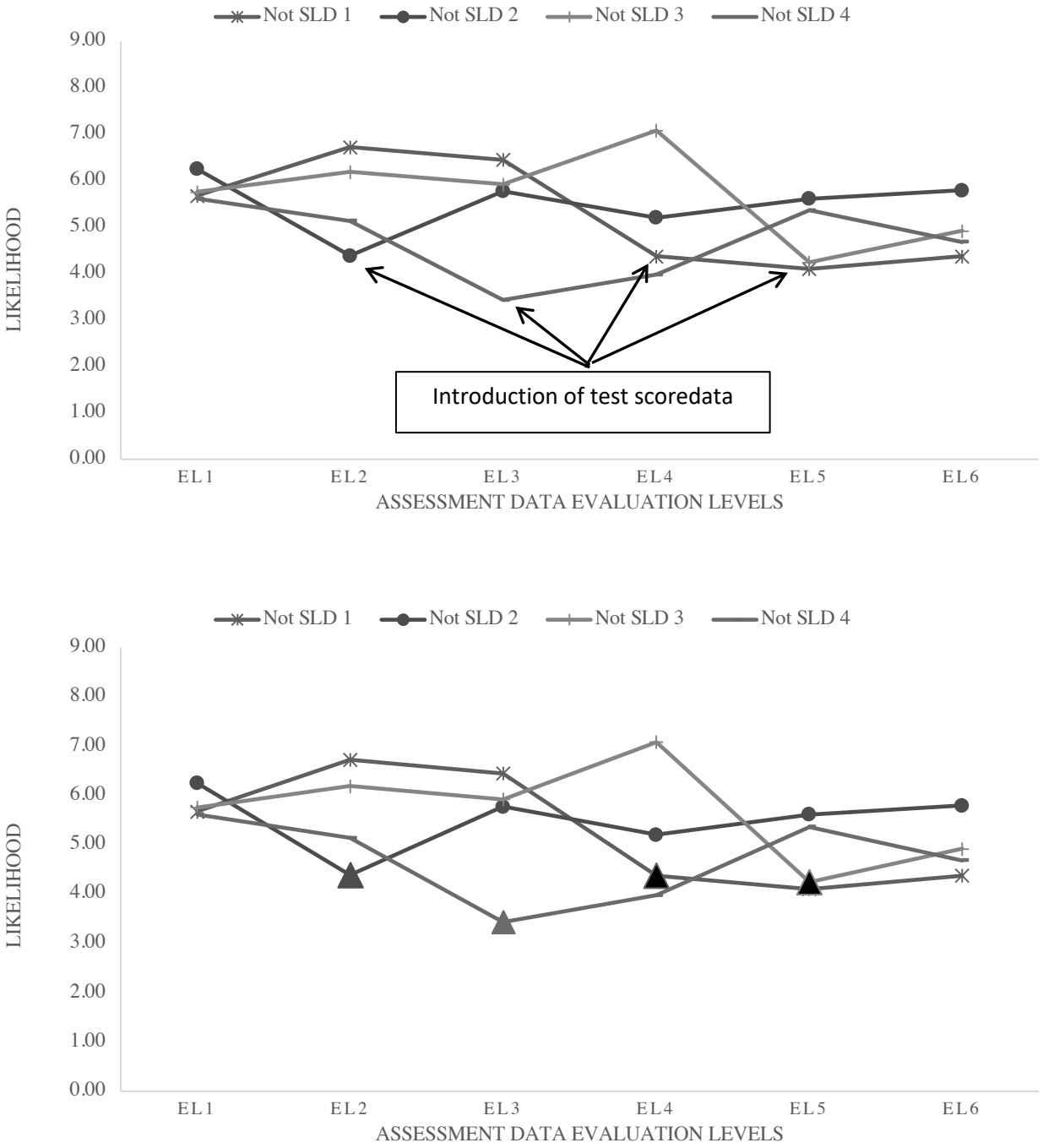
Table 1

Frequency and Percentages of Final SLD Decisions by Vignette Condition.

	Vignette Condition		
Final Decision	SLD Positive	SLD Negative	Total
SLD Positive			
Frequency	145	80	225
% within Vignette Condition	81.5%	48.2%	65.4%
SLD Negative			
Frequency	33	86	119
% within Vignette Condition	18.5%	51.8%	34.6%
Total			
Frequency	178	166	344
% within Vignette Condition	100.0%	100.0%	100.0%

Figure 1.

SLD likelihood ratings over time for SLD Negative vignettes.



Note. Triangles mark the introduction of cognitive and achievement test score data.