

## Radosław Wujcik

**Tytuł:** Zastosowanie teorii odpowiadania na pozycje testowe (IRT) i modeli diagnostycznych w celu poprawy trafności testów stosowanych w diagnozie psychiatrycznej dzieci i młodzieży. Na przykładzie narzędzi do diagnozy ADHD i ASD.

### Streszczenie

Niniejsza rozprawa poświęcona została problematyce stosowania testów psychologicznych w diagnozie psychiatrycznej dzieci i młodzieży. Celem przeprowadzonych analiz jest ocena korzyści, jakie może przynieść wykorzystanie probabilistycznych modeli cech ukrytych, podczas konstruowania kwestionariuszowych narzędzi diagnostycznych stosowanych w psychiatrii i psychologii klinicznej. W badaniach wykorzystano polskie adaptacje dwóch kwestionariuszy: Kwestionariusz do Diagnozy Spektrum Autyzmu ASRS i Kwestionariusz do Diagnozy ADHD *Conners 3*. Pierwszy podrozdział wstępu teoretycznego odnosi się do roli jaką odgrywają testy psychologiczne w diagnozie psychiatrycznej. Omówiono w nim rodzaje testów psychologicznych stosowanych w diagnozie a także zwrócono uwagę na podstawowe problemy metodologiczne towarzyszące opracowywaniu nowych narzędzi. Następnie szczegółowo omówiono zagadnienia związane z własnościami psychometrycznymi testów, tj. rzetelność, trafność i normalizacje. W drugim podrozdziale przedstawiono założenia najważniejszych modeli wyniku testowego wykorzystywanych w psychometrii, czyli klasycznej teorii testu (KTT) i teorii odpowiadania na pozycje testowe (IRT). Przedstawiono również podstawowe założenia i zastosowania modeli diagnostycznych (CDM), stosowanych powszechnie w pomiarze edukacyjnym. Ostatnie dwa podrozdziały wstępu teoretycznego poświęcono diagnozie ADHD i diagnozie zaburzenia ze spektrum autyzmu. Przedstawiono kryteria diagnostyczne (ICD-10 i DSM-5) dla tych jednostek chorobowych, wymieniono najważniejsze metody i narzędzia stosowane do ich diagnozy w Polsce oraz szczegółowo omówiono kwestionariusze wykorzystane w analizach przeprowadzonych w ramach tej rozprawy (ASRS i *Conners 3*). Drugi rozdział poświęcono metodzie badań, omówiono próby badane każdym z narzędzi (próba normalizacyjna i walidacyjna polskiej adaptacji). Wymieniono wykorzystane wersje kwestionariuszy oraz procedurę wg, której przeprowadzono analizy. Kolejny rozdział poświęcono przedstawieniu wyników badań. Przeprowadzone analizy nie pozwoliły na wyciągnięcie jednoznacznych wniosków. Co prawda, stwierdzono poprawę trafności *Conners 3* dzięki zastosowaniu modeli IRT, lecz nie odnotowano podobnej korzyści w przypadku ASRS. W przypadku obydwu narzędzi wykorzystanie IRT pozwala na dokładniejsze oszacowanie błędu pomiaru w przypadku obydwu narzędzi. Ponadto wykazano, że niektóre pozycje obydwu narzędzi charakteryzują się zróżnicowanym funkcjonowaniem (DIF) w zależności od płci lub wieku osób badanych. Analizy przeprowadzone za pomocą modeli wielogrupowych w, których 221

oszacowano właściwości pozycji niezależnie w wyróżnionych grupach, nie poskutkowało poprawą trafności uwzględnionych w badaniach kwestionariuszy, w porównaniu do jednogrupowych modeli IRT. Pozwoliły jednak oszacować rzetelność niezależnie w każdej z grup, dzięki czemu wielkość błędu pomiaru lepiej odpowiada rzeczywistej precyzji oszacowanego wyniku badania. Analizy przeprowadzone dla wyniku szacowanego w modelach diagnostycznych nie wykazały korzyści w postaci poprawy trafności w stosunku do wyniku liczonego tradycyjnie (wg KTT). Pokazały jednak, że modele te mogą być użyteczne w przypadku identyfikacji obecności danej konfiguracji badanych cech. Może to znaleźć zastosowanie w konstrukcji narzędzi opierających się na ocenie spełniania kryteriów diagnostycznych w diagnozie różnicowej. Analiza korelacyjna wyników oszacowanych w modelach IRT i CDM wykazała, że zachowują one trafność teoretyczną wyniku liczonego zgodnie z KTT. W części stanowiącej dyskusję uzyskanych wyników, autor wylicza zalety wykorzystania modeli probabilistycznych wyniku testowego (IRT i CDM) oraz wskazuje na ograniczenia związane z tym podejściem. Podsumowując wszystkie zebrane informacje, opowiada się za konstruowaniem testów psychologicznych opartych na metodologii IRT.

**Słowa kluczowe:** diagnoza psychiatryczna, ADHD, zaburzenie ze spektrum autyzmu, trafność testu, rzetelność testu, teoria odpowiadania na pozycje testowe (IRT), modele diagnostyczne (CDM), zróżnicowane funkcjonowanie pozycji testowej (DIF), ASRS, Conners

3 222

## Abstract

**Title:** Application of test item response theory (IRT) and diagnostic models to improve the accuracy of tests used in the psychiatric diagnosis of children and adolescents on the example of ADHD and ASD diagnosis tools.

This dissertation is devoted to the issue of using psychological tests in the psychiatric diagnosis of children and adolescents. The aim of the analyses is to assess the benefits that can be brought by the use of probabilistic models of covert features when constructing questionnaire diagnostic tools used in psychiatry and clinical psychology. The study used Polish adaptations of two questionnaires: Autism Spectrum Rating Scale *ASRS* and questionnaire for ADHD diagnosis *Conners 3*. The first section of the theoretical introduction refers to the role that psychological tests play in psychiatric diagnosis. It discusses the types of psychological tests used in diagnosis and also draws attention to the basic methodological problems associated with the development of new tools. Then, the issues related to the psychometric properties of the psychological tests, such as reliability, validity and standardization, were discussed in detail. The second section presents the assumptions of the most important test result models used in psychometrics, i.e. classical test theory (KTT) and item response theory (IRT). Basic assumptions and applications of cognitive diagnostic models (CDM), commonly used in educational measurement, are also presented. The last two sections of the theoretical introduction are devoted to the diagnosis of ADHD and the diagnosis of autism spectrum disorder. Diagnostic criteria (ICD-10 and DSM-5) for these disease entities are presented, the most important methods and tools used for their diagnosis in Poland are listed, and questionnaires used in the analyses carried out in this dissertation (*ASRS* and *Conners 3*) are discussed in detail. The second chapter is devoted to the research method, the samples tested with each tool are discussed (standardization and validation test of Polish adaptation). The versions of the questionnaires used and the procedure according to which the analyses were carried out were listed. The next chapter is devoted to the presentation of research results. The conducted analyses did not allow to draw unequivocal conclusions. It is true that the improvement in the accuracy of *Conners 3* was observed thanks to the application of IRT models, but there were no similar benefits for the *ASRS*. For both tools, the use of IRT allows a more accurate estimation of the measurement error. In addition, it has been shown that some items of both tools are characterized by differential item functioning (DIF) depending on the sex or age of the subjects. The analyses prepared for multi-group models, in which the position properties were independently estimated in the groups distinguished on this basis, did not result in improvement in the accuracy of both questionnaires, compared to single-group IRT models. However, they allowed to estimate reliability independently in each group, thanks to which the size of the measurement error corresponds better to the actual precision of the estimated test result. The analyses carried out for the result estimated in diagnostic models did not show benefits in the form of improved accuracy compared to the results calculated traditionally (according to KTT). However, they showed that these models can be useful in identifying the presence of a given configuration of the examined features. This can be used in 223

constructing tools based on the assessment of the fulfilment of diagnostic criteria for differential diagnosis. Correlation analysis of the results estimated in the IRT and CDM models showed that they retain the theoretical accuracy of the result calculated in accordance with KTT. In the discussion of the results obtained, the author enumerates the advantages of using probabilistic models of the test result (IRT and CDM) and indicates the limitations associated with this approach. Summarizing all the collected information, the author is in favour of constructing psychological tests based on IRT methodology.

**Keywords:** psychiatric diagnosis, ADHD, autism spectrum disorder, test accuracy, test reliability, item response theory (IRT), cognitive diagnostic models (CDM), differential item functioning (DIF), ASRS, Conners 3