

Assessment

SYMPOSIUM

2340

Faking of personality inventories: Questioning what we thought we knew

Chair: Georg Krammer *University College of Teacher Education Styria, Austria*

Faking of personality inventories has a long research tradition but is nevertheless still poorly understood. Therefore, this symposium aims at furthering our understanding of faking. To this end, we question key aspects of what is supposedly known about faking on personality inventories: impression management, warnings, applicant faking, and faking as a static process. The first talk scrutinizes impression management. While one would expect a positive relation of the psychopathic personality with impression management, studies show the opposite relation. This first talk addresses this counterintuitive finding and tests the underlying mechanism of identity expression affecting impression management. In doing so, this first talk tries to further our understanding of impression management. The second talk questions what we know about warnings reducing faking. Warnings have been shown to reduce faking. However, warnings are often constructed atheoretical, with the underlying mechanisms still poorly understood. This second talk proposes and tests new faking warnings among military recruits. In doing so, this second talk tries to further our understanding of how faking can be reduced. The third talk challenges models of applicant faking. First, this talk questions the systematic and rational view on applicant faking put forth by models of applicant faking. Second, this talk points to problems arising when applying models of applicant faking to different selection processes (e.g. educational vs. occupational). In doing so, this third talk tries to further our understanding of how applicants fake. The fourth talk disputes that faking is a static process. It is usually assumed that the faking process of one respondent in one situation is static. The final talk will demonstrate that faking is a changing process whereby individuals do not respond using a single answering tactic. In doing so, this fourth talk tries to further our understanding of the process of faking.

Why do psychopaths score so low on impression management scales? The role of antisocial identity

Delroy L. Paulhus *University of British Columbia, Canada*

One hallmark of the psychopathic personality is the chronic tendency to dissemble regarding

their misbehavior. That tendency is assumed to apply to questionnaire responses. Previous research, however, indicates that psychopathy scores are negatively associated with scores on impression management scales: The implication is that psychopaths dissimulate less than non-psychopaths. Calling on the work of Walters and Holden, among others, we suspected that this counterintuitive finding stems from the fact that psychopaths hold deviant beliefs about desirable behavior. The hypothesis was evaluated in both student ($N = 1420$) and community ($N = 212$) samples. Both data sources replicated the typical negative association between impression management and self-report psychopathy scores. Study 1 provided evidence that the pattern is statistically mediated by individual differences in antisocial identity. Even non-psychopaths scored lower on impression management when they identified with criminals. Study 2 controlled for antisocial identity by computing a value-adjusted index of impression management. After accounting for their value profile, psychopaths had impression management scores equivalent to those of non-psychopaths. These findings highlight the overriding importance of identity expression—even to antisocial individuals. Psychopaths, in particular, have constructed an identity that integrates their antisocial tendencies into a positive self-image that clashes dramatically with that of non-psychopathic individuals.

Applicant faking warnings in the armed forces: Are they really effective?

Justin Feeney *University of Regina, Canada*

The applicant faking literature suggests that faking warnings—brief messages that dissuade applicants from faking—can reduce faking by up to 50% (Fan et al., 2012). However, the efficacy of warnings may be limited by their atheoretical construction (Goffin & Boyd, 2009) and these threatening messages can cause applicants to feel negatively about the personality test (Converse et al., 2008). We tried to improve the efficacy of faking warnings, while minimizing negative applicant reactions, by leveraging theory from the accountability and morality literatures. We tested three new faking warnings to reduce faking among military recruits that emphasized short-term accountability, long-term accountability, and morality. We also examined how these new warnings compared to a traditional faking warning and a no-warning control group. We tested 466 basic training recruits at the Canadian Armed Forces and asked them to engage in a selection simulation. We assigned groups of recruits to the different faking warning conditions and walked classes through the simulation. We found that a faking warning emphasizing short-term accountability, which threatened to detect fakers by contacting references

and using “internal integrity checks”, reduced applicant faking. None of the other messages had any effect when compared to a no-warning control group. We also found no differences in applicant reactions to the different faking warnings.

Applicant faking in college admission processes

Georg Krammer *University College of Teacher Education Styria, Austria*

Applicants faking personality inventories in selection processes is still poorly understood. Models of applicant faking portray applicant faking as a systematic and rational process where applicants systematically elevate their otherwise honest responses. This simplicity is disputed by studies showing that applicant faking is not aimed at score maximization. Furthermore, while models of applicant faking do take situational factors into account, they do not differentiate between different selection processes (e.g. educational vs. occupational). Firstly, this talk challenges viewing applicant faking as a systematic shift from otherwise honest responses. In a within-subject design, 840 applicants' Big Five personality traits were assessed in a college admission process to teacher education and again 10 months later. Systematic item parameter shifts from honest responses were tested across these applicants and incumbents with linear partial credit models. Testing patterns derived from literature of how item parameters shift showed no pattern applicable to the item parameter shifts. We therefore conclude that applicant faking shifted the item parameters in an unsystematic way. Secondly, applying models of applicant faking to educational settings (i.e. college admission processes) may need further clarification. Models of applicant faking suggest that applicants consider items' perceived relevance when responding. But how do applicants evaluate this perceived relevance in college admission processes? Are applicants considering the relevance for the education, or for the prospective profession? Before this backdrop, this talk addresses if items' perceived relevance for the education or the prospective profession can explain how applicants shift responses. Results show that both the perceived educational and occupational relevancies affect item parameters' shift. However, the perceived occupational relevancy emerges as a more dominant predictor for item parameters' shift. We therefore conclude that models of applicant faking should consider differences across selection settings not only as separate factors but as factors influencing other determinants of applicant faking.

Faking on a self-report personality inventory: A dynamic or static process?

Ronald R. Holden *Queen's University, Canada*

In this presentation, I will demonstrate that faking on a self-report personality inventory is a changing process whereby individuals do not respond using a single answering tactic. For an initial study comprised of samples responding to a personality inventory in either a real-world setting ($n = 49$) or an instructed faking context ($n = 347$), participants provided open-ended descriptions on how they would answer to maximize or minimize particular outcomes. From this open-ended responding, raters identified 11 different tactics that became the basis of experimental faking instructions under which a subsequent sample ($n = 275$) was randomly assigned to complete the NEO-PI-R. This subsequent sample was then debriefed on the tactics actually applied, revealing that individuals used multiple tactics, and sometimes seemingly inconsistent tactics, when faking. I will then show that this identified dynamic responding is consistent with recent analyses of an additional six data sets of instructed fakers ($n = 1360$) that demonstrate that, within an individual, faking varies over time and responding is more heterogeneous than for individuals who are not faking. Taken together, findings indicate that faking is a nuanced, dynamic phenomenon and not comprised of simple, static responding.

INVITED ADDRESS

697KA

What's in a Likert response? Psychometric models for measuring traits and response styles

Thorsten Meiser *University of Mannheim, Germany*

Likert items with ordinal rating scales are in widespread use in various fields of psychology and other social sciences. Rating responses, however, not only reflect the target traits that are to be assessed, but they are confounded with method effects and response styles that may bias individual measurement, inflate score correlations and distort group comparisons. The keynote talk gives an overview of psychometric approaches to account for method effects and response styles by statistical modelling and test construction. Statistical models include multidimensional and mixture-distribution extensions of item-response theory (IRT) that allow for heterogeneity in the model parameters to accommodate differences in the use of rating categories. Furthermore, multiprocess decision trees with IRT nodes (IRTtree models) allow researchers to disentangle the target traits from other judgment processes that contribute to the observed rating responses. The statistical models are introduced in an easy-to-grasp formal and graphical way, and they are illustrated with empirical applications to data from personality research and large-scale assessments. The applications demonstrate that the extended

IRT and IRTtree approaches afford unbiased measures of the target traits together with an analysis of response styles as part of the psychological response process. Beyond statistical modelling, the keynote provides an outlook on methods of test construction with alternative response formats that may alleviate or avoid biases due to method effects and response styles. Response formats that have recently been evaluated include two-step procedures of agreement and intensity, the assignment of items to categories, and item rankings in multidimensional forced-choice designs.

INVITED SYMPOSIUM

6936

MABEL-Multi-language assessment battery of early literacy overview

Chair: Gabriela S. Málková *Charles University, Czechia*

We propose a symposium to present MABEL (Multi-Language Assessment Battery of Early Literacy; www.eldel-mabel.net), a web-based tool for the assessment of children's early literacy and related skills across several European languages. MABEL comprises measures, directly comparable across languages, and targets foundational literacy skills such as phoneme awareness, letter knowledge, and rapid naming, as well as a variety of reading and spelling measures. The MABEL tests enable a systematic screening assessment for identifying children at risk of literacy difficulties, who speak any one, or a combination, of the languages currently within the battery. For this reason, it is valuable in various bilingual, multilingual educational and clinical contexts, especially at the onset of literacy instruction. For research purposes, MABEL offers a set of directly comparable measures that can be implemented in cross-linguistic or single-language studies of reading and spelling development. Each symposium presentation will highlight a specific test bundle (e.g., tests of phoneme awareness, rapid naming, reading, spelling). For each set of tests, a brief explanation of administering procedures and demonstration of materials used within the assessment will be provided. The Discussion will focus on experience for the use of the battery and especially on the benefits and limitations of directly comparable measurement tools.

MABEL – Multi-language assessment battery of early literacy – introduction, background, discussion

Markéta Caravolas *Bangor University, United Kingdom*

Introduction and Background: Cross-linguistic studies play an important role in research

on literacy development and its cognitive underpinnings. We will present the Multi-language Assessment Battery of Early Literacy (MABEL) constructed for direct cross-linguistic comparisons of children's foundation-level skills. The tests contained in MABEL were created and validated as part of the ELDEL research programme (a Framework 7, Marie Curie ITN) focusing on cross-linguistic studies of literacy development in multiple languages, including English, Spanish, French, Czech, Slovak (Caravolas et al 2012, Caravolas et al 2013). The ELDEL research programme, which investigated, longitudinally, the language-specific and general aspects of early literacy development across five languages with alphabetic orthographies. General issues relating to the development of the MABEL measures will be considered. These include how to determine the key design parameters for directly comparable cross-linguistic tests of phonological processing, reading and spelling; how to delimit the language-general and language-specific components of each skill for statistical and clinical analysis; and how to control for (or manipulate) important variables pertaining to the linguistic/orthographic, cognitive-developmental, and educational context of the studied groups. A brief overview of the MABEL website and its key features will be made. Discussion: The MABEL tool arises from a body of cross-linguistic research into the universals and language-specifics of literacy development within the European ELDEL network. This symposium highlighted how measures constructed to assess literacy and related skills in directly comparable ways across a number of languages with alphabetic orthographies can be designed to yield good psychometric properties with good cross-language validity. As such, they offer excellent tools for cross-linguistic research investigations and for clinical and educational practice within single language communities or those serving children speaking/reading in any combination of MABEL's current languages. However, across the presentations, some limitations will have also come to the fore, regarding the potential for violations of ecological-linguistic validity of the measures, the need for background tools (such as lexical corpora) for the accurate selection of comparable items, and the possibility of excluding important language-specific measures. These issues will be discussed with a view to opening the way for extending MABEL to other languages, while also maintaining good reliability, validity and practical usefulness of the measures for any given language.

Phoneme awareness tasks in MABEL

Marina Mikulajova *Pan-European University, Slovakia*

Phonological awareness (PA) skills play a crucial role in early literacy acquisition across alphabetic orthographies. PA tasks currently included in MABEL represent cognitive instruments in English, Spanish, Czech and Slovak languages. While PA skills have been for decades deeply rooted in pre-literacy and literacy research and practice in English, and to some extent also in Spanish-speaking communities, and therefore have given rise to a variety of diagnostic tools, the opposite is true for the Czech and Slovak countries. MABEL includes three measures of PA: two that assess analytic phoneme manipulation skills – namely Phoneme Isolation and Phoneme Deletion –, and one that assesses synthetic phonemic skills – namely Phoneme Blending. MABEL offers two formats of each of the above tests: a downloadable paper and pencil format, and a computer-assisted version. In this presentation, we explain the principles of test creation across languages and demonstrate components of the computer-assisted versions, such as the administration instructions, training phase, and test items. Psychometric properties of the measures will be conveyed and proposals for adaptation into additional languages will be made. We will conclude with the benefits that come from the web-based open access display of these tests for research and practice.

Rapid automatized naming and letter knowledge tasks in MABEL

Sylvia Defior *Universidad de Granada, Spain*

Along with phoneme awareness, measures of letter knowledge (LK) and rapid automatized naming (RAN) are well established early predictors and correlates of children's reading and spelling abilities. Moreover, they seem to hold the status of cognitive universals predictors for learners of alphabetic orthographies. Accordingly, the MABEL test suite includes tests of LK, assessing knowledge of letter sounds and names, in upper and lower case, as well as four measures of RAN, which assess rapid naming respectively of colours, objects, digits and letters. This presentation will highlight the factors that were taken into account in the creation of parallel test forms across five languages, including constraints on administration and scoring. Reference will be made to the research publications confirming the validity of the measures. We will demonstrate the standard paper versions of the tests, as well as the features and functioning of their web-based analogues. Considerations for interpreting individual test results, and for adapting the LK and RAN measures to additional languages will also be presented.

Reading and spelling tasks in MABEL

Gabriela S. Málková *Charles University, Faculty of Humanities, Czechia*

This paper presents the reading and spelling tasks included in the MABEL test suite, which were designed to enable direct cross-linguistic comparisons of early reading and spelling skills among children learning alphabetic orthographies. MABEL aims to provide comprehensive reading profiles of young readers, and thus contains both aloud and silent reading tasks. The tests of reading aloud assess reading efficiency of words (One Minute Word Reading Test) and pseudowords (One Minute Nonword Reading Test), which respectively measure accurate reading under time pressure. Silent reading efficiency is assessed using the Picture-Word Matching Test, which requires the selection of a target word, from four printed alternatives, that corresponds to a picture stimulus. Spelling test basic words spelling and graded spelling will be briefly introduced along the reading tasks to provide a full example of early literacy assessment test in MABEL. Each of the tests was constructed using the same criteria across languages for stimulus selection, structure, and administration procedures. MABEL offers two formats of each of the above tests: a downloadable paper and pencil format, and a computer-assisted version. The paper will briefly introduce the test content, procedures and possible applications of the available language versions, in both formats.

6058IS

Formative assessment in education – A broad perspective

Chair: Iva Stuchlikova *University of South Bohemia in Ceske Budejovice, Czechia*

Formative assessment or assessment for learning is frequently seen as a remedy for raising students' engagement in learning and improvement of students' achievement. (Black and Wiliam 2009; OECD 2005). To really support learning, the formative assessment must be based on a well-articulated model of learning (Shepard, Penuel, Pellegrino, 2018). It is, therefore, crucial to combine various perspectives on formative assessment; its actors, processes, methods. The symposium will bring together different views on formative assessment, combining the insights achieved in different contexts. The introduction to the symposium (Iva Stuchlíková) will briefly summarize a model for classroom assessment and position of formative assessment within it. Contributions will deal with the development of teachers' self-efficacy in using formative assessment strategies (Evans, Bruune), student competencies at which teachers focus when assessing formatively students' inquiry skills (Grob), communication model of formative assessment in inquiry-based learning tasks (Samková, Rokos). Further, we will discuss the implementation of formative assessment as viewed by headmasters, teachers and pupils (Laufková) and final presentation will point on relationships between students' learning

goals and their perception of teaching context and formative assessment (Kováčová, Plasová, Nohavová, Krninský).

Identifying teacher self-efficacy development in using formative assessment strategies using similar

Robert H. Evans *University of Copenhagen, Denmark*

Self-efficacies of secondary science teachers in eight participating European countries were used over three years to track capacity beliefs for using four different kinds of formative assessment. Using a pre-post-design with control group, questionnaires showed the expected overall significant gains for participants but not for non-participants. However, a novel method of network analysis, where participants across cultures were linked based on similarity networks to tease out detailed answering patterns was used. Notably, teachers were not just seen as a non-structured cohort, in which groups of answering teachers may be found by categorizing them according to background variables. Instead, the analyses revealed groups of participants across school and national and boundaries with similar profiles of self-efficacy when using formative assessment. These unique teacher profiles are useful for targeting professional development based on strengths and deficiencies in self-efficacies relevant to formative assessment. Such refined professional development strategies may be more efficient than the generalized methods of promoting the use of formative assessment which in our study, did not fully reach these cross-cultural cohorts. Instead of being invisible in the school-based results, these methods allow teacher educators to identify and target teachers with specific needs for developing formative assessment practices across schools and countries.

Competences which science teachers focus at when formatively assessing their students in inquiry-based learning tasks

Regula Grob *University of Teacher Education UTE Lucerne, Switzerland*

In this paper, formative assessment activities in the context of inquiry-based science education (IBSE) from primary and upper secondary school teachers in Switzerland will be analyzed. The rationale is to explore what competences these teachers focus at when formatively assessing their students' inquiry skills. The cases (14 from primary school, 20 from upper secondary school) were documented with teaching plans and teaching materials. In the analysis, a deductively developed coding scheme involving categories specifying the inquiry characteristics of the cases as well as the formative assessment characteristics was used. Furthermore, individual interviews

were conducted with a sub-set of the teachers ($n = 12$) in which these teachers explained how they planned the formative assessment activities. The results of the paper focus at the comparison between inquiry competences described in the literature and the competences chosen for assessment by the teachers. Furthermore, the interview data which provide insights into the teachers' reasoning for their choice of competences will be discussed.

A communication model of formative assessment in inquiry-based learning tasks

Libuše Samková *University of South Bohemia in Ceske Budejovice, Czechia*

This paper presents a communication model combining inquiry-based education and formative assessment. Based on previous research on the use of formative assessment in science and mathematics teaching, it turns out that future teachers (university students) and in-service teachers need to consider these two approaches together in order to clearly understand their connection and common characteristics. For this contribution, we have prepared a set of communication schemes for formative assessment and inquiry that is based on the ESRU model related to interactions during on-the-fly feedback. The schemes describe the main steps of inquiry as well as of formative assessment. Visualizing the connection of these two approaches should help the teachers to understand the possibilities that formative assessment offers while dealing with inquiry-based tasks. We accompany the set of the schemes with illustrative examples perceived from the perspective of an ideally realized teaching unit. The illustrative examples come from two school subjects: Mathematics and Science.

Implementation of formative assessment from headmasters, teachers and students point of view

Veronika Laufková *Charles University, Czechia*

The contribution will summarize the results of two-year action research focusing on the implementation of formative assessment in Czech schools. It is based on the concept of formative assessment which aims to improve students' learning, to provide useful information about the student's current level of knowledge and skills, and to close the gap between what students know and what we want them to know. The method implemented is a qualitative exploratory case study of formative assessment aimed to encourage 18 teachers from 9 primary and secondary schools (ISCED 1, 2). The research question is as follows: "How do headmasters, teachers and students reflect on the implementation of formative assessment tools?". The data were obtained

by interviewing headmasters, teachers, and students, and triangulated with observing lessons and a content analysis of documents. Rewritten data were analyzed in MAXQDA. The coding was done according to the axis: intentions – realization – consequences of the formative assessment. The main contribution of this study lies in the practical experience of implementation of formative assessment and the views of the participating headmasters, teachers and students. The strategies and principles on which formative practices are founded and complications in this regard will be discussed.

The relationship between students' learning goals and their perception of teaching context and form

Klára Kováčová *University of South Bohemia in České Budějovice, Czechia*

This study focused on a relationship between students' learning goals (mastery/performance) and formative assessment in the classroom. Students' learning goals were assessed by Czech adaptation of Midgley's Patterns of Adaptive Learning Scales (PALS, Kovacova, 2017) and their report of teaching context (involvement, structure, and autonomy support) was captured by the short version of Skinner's Teacher as Social Context Questionnaire (TASC-Q short, Skinner, 1992). Perception of the formative assessment was measured by the Assessment for Learning Questionnaire for students (SAFL-Q). Data were obtained from students of the 8th grade ($N = 100$); students reported about themselves and about their math teachers. Relationships between learning goals, teaching context and perception of formative assessment are analyzed and will be discussed with respect to school motivation and ability to efficiently motivate oneself.

ORAL PRESENTATION

8352

Examining the structure of occupational interest across linguistic groups in India

Purnima Singh (1), Vijai K. Pandey (2), Mariya Jilinskaya-Pandey (3) *1. India; 2. Emango Education Pvt. Ltd, India; 3. O P Jindal Global University, India*

The hexagonal RIASEC model presented by John Holland has been the de-facto model for career guidance and occupational orientation programme across the world. Several studies have been done since the first publication of the model to examine its structural equivalence across cultures and languages. Two kinds of structures are generally proposed – the hierarchical structure and circumplex structure with varying degree of support for each in different

cultural and linguistic groups. However, very little work is done to examine the structure of Holland's model in developing countries like India. The current study is an attempt to fill this gap by examining the structure of the occupational interests in 4 different linguistic and socio-economic groups. We got 1600 participants from 4 socio-economic and linguistic backgrounds to fill the Occupational Interest Questionnaire and provide self-reported information on 9 demographic variables. We analysed the data using R to examine the hierarchical as well as circumplex structure. Results indicate good support for a hierarchical structure but poor support for circumplex structure. The study has a wider implication for career and occupational guidance as it provides answers and gives further insights into some of the long-standing questions about the universality of the RIASEC model and its practical implications for occupational guidance. We also discussed the relevance of the model in the changing world of careers.

8216

Projective techniques and its applications in the Arab world

Maan A. Qasem Imam Abdulrahman Bin Faisal University, Saudi Arabia

Background: Since the development of Rorschach test in 1921, the use of projective techniques has become a part of diagnostic clinical practice. Currently, this technique plays a distinct role in revealing the hidden world of human unconscious covered by defensive mechanisms that complicate the interpretation of ambiguous/disturbed behavior, and the consequences of reactions to traumas and mental crises. Objective: The purpose of the present paper was to pay attention to the application and effectiveness of projective techniques in clinical practice in the Arab world. Method and setting: this is a cross-sectional study recruiting psychiatric outpatient and inpatient participants (male 33% and female 67%, with average age of 28years) at King Fahd University Hospital (KFUH) in Saudi Arabia. Instruments and participants: Thematic Appreciation Test (TAT) and House-Tree Person Test (HTP) were administered to participants with a previous history of domestic violence 13%, sexual orientation identity 7%, eating disorders 7%, depression 20%, anxiety 20%, and personality disorders 33%. Result: These techniques have shown to be significant and effective in assessing persons living in the Arab environment where society is characterized by conservative customs and traditions associated with tough nature (i.e. the prevalence of repression and psychological defense mechanisms, which provides diagnostic power to these tools, comprises repression 80%, projection 86%, denial 73%, displacement 46%, and sublimation 24%). The strengths, limitations, and cross-cultural variations were discussed.

Conclusion: The projective technique revealed a significant supplementary tool that could facilitate the psychiatric diagnosis and provide future recommendations on how to address problems associated with cross-cultural factors of psychological measurement and evaluation.

8214

Psychometric properties of the prenatal psychosocial profile scale for a South African sample

Maria Florence (1), Mansoo Yu (2), Tracey-Ann Adonis (1), Noluthando Mpisane (1), Gashaye Tefera (2), Omoshola Kehinde (2) 1. University of the Western Cape, South Africa; 2. Missouri University, United States of America

This study aims to assess the validity and reliability of the Prenatal Psychosocial Profile (PPP) for a sample of pregnant women ($N=172$) from low-income communities in Cape Town, South Africa. The PPP measures psychosocial constructs (stress, support of a partner, support of others, and self-esteem) that influence pregnancy outcomes. These are particularly important in low-income contexts where high rates of negative health outcomes are reported for pregnant women and their infants. The instrument has demonstrated reliability and validity for several samples of pregnant women in terms of test-retest reliability, internal consistency and convergent validity using depressive symptoms. Studies comparing samples found different properties for some of the sub-scales across different races. The data from a larger study on substance use during pregnancy collected cross-sectionally in 2018 was used to examine the underlying factors of the PPP for this sample of women. In this study, Principal Component Analysis was used to evaluate the factor structures and convergent construct validity was evaluated to assess whether the PPP sub-scales were significantly associated with substance use. Cronbach Alphas were calculated to assess the reliability of the sub-scales for this sample. While most of the items of the PPP load together as expected for this sample, almost half of the self-esteem items load on the same factor as the support of a partner items. This is a clear indication of the complex interaction between self-esteem and partner support which could add to the impact on these women's health behaviours throughout their pregnancies. Internal consistency scores range between .879 and .987. As the PPP has proved useful for research and interventions in other contexts, assessing its reliability and validity for a South African sample can prove equally useful in future efforts towards healthy pregnancy outcomes.

8180

LSI.J – A tool to investigate receptive language in German-speaking adolescents

Carina D. Krause (1), Susanne Wagner (2), Christian W. Glück (3) 1. Germany; 2. Berufsbildungswerk Leipzig, Germany; 3. University of Leipzig, Germany

Language competence is a crucial factor for socio-emotional, educational and occupational outcomes. If language impairments remain undetected, disadvantages in these areas may be the consequence (Conti-Ramsden et al., 2013). We developed a tablet-based assessment tool that profoundly describes the receptive language competence of German-speaking adolescents and young adults. The LSI.J (Leipziger Sprach-Instrumentarium Jugend/Leipzig Language Inventory Youth) comprises 10 tests in 5 areas of language comprehension: auditory processing, lexicon, syntax, pragmatics, and auditory attention. The LSI.J bases upon a psycholinguistic model derived from an extensive literature review on psycho- and neurolinguistic models of language comprehension and language impairments in adolescents. Developing LSI.J, we focussed specifically on elaborate psychometric methods to meet the quality criteria for sound diagnostic testing. In a Germany-wide norming study we acquired data from 571 adolescents and modelled them following a semi-parametric, continuous approach (Lenhard et al., 2018). The norming data is also basis for a large range of reliability and validity measures. For instance, we developed a complex nomological network depicting constructs and operationalisations as basis for sound validity computations, that is not published yet, but ready for presentation to an expert audience. Additionally, we are establishing case study-based clinical LSI.J-profiles easing the interpretation of results and providing further insights into the similarities and differences between diagnoses. Based on the test results, tailor-made supportive actions may be selected from the LSI.J-data base, including environmental adjustments, disability compensation, remedial language instruction, or speech and language therapy.

References

Conti-Ramsden, G., Mok, P. L., Pickles, A., & Durkin, K., (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. Research in developmental disabilities, 34(11), 4161–4169.
Lenhard, A., Lenhard, W., Suggate, S., & Segerer, R. (2018). A continuous solution to the norming problem. Assessment, 25(1), 112–125.

8070

Computer-based problem-solving ability assessment: Development and application in recruitment

Baiyunji Jiang (1), Danjun Wang (1), Chongli Liang (1), Huihui Xu (1)

1. Beisen Cloud Computing Co., Ltd, China

Problem-solving ability tests used in recruiting selection are mainly in a form of choosing from presented options. The test form has limitation on assessing option-choosing skill rather than actual ability of solving problems. In the current study, we develop a new form of computer-based problem-solving ability assessment for recruiting selection. During the test, candidates need to interact with situations to identify and solve the problems. The whole interaction data and final decision are recorded for analysis. There are three advantages reflected. Firstly, the new form test assesses candidates' problem-solving ability more comprehensively. The log data analysis provides assessment for dimensions such as problem-solving strategies, flexibility and willpower. And these results could be generated personal report effectively assist corporations in recruitment screening and interviews to select high-potential talents. Secondly, it brings candidates more immersive testing experience with less time-consuming. The situation simulation presented by computer programming using multistage testing technique (MST) takes 20 minutes only. Thirdly, it predicts job performance well. In the validation study ($N=55$), the test scores can predict job performance of technical employees.

7644

Development and validation of online dark triad scale (ODT) on Taiwan social media

Cheng-Yan Wang (1), Yih-Lan Liu (1) 1. The Education Institute of National Chiao Tung University, Taiwan

Objective: Past studies have shown that Dark Triad personality traits are related to traditional bullying and cyberbullying behaviors in adolescents as well as in adults. Accurate evaluation of people's dark triad is crucial to understand the impact of psychopathy, narcissism, Machiavellianism personality on the social media. However, as no questionnaires are available to capture the personality of online dark triad, this study developed and validated the online dark triad scale (ODT) on social media of Taiwan. Methods: 644 participants (62.6% females, average age = 24) completed the online dark triad personality scale in June, 2020 (Time 1), and three months later, 225 of them (64% female) completed the scales again in September 2020 (Time 2). Exploratory and confirmatory factor analysis and measurement invariance were conducted to examine the factor structure and gender invariance of ODT scale. Results: These results suggest that ODT scale had a suitable structural validity, convergent validity, discriminant validity, criterion validity, internal reliability ($\alpha=.801-.913$), test-retest reliability ($r=.570-.650, ps<.01$). Measurement invariances across gender and time was established. Conclusion: The ODT

appears promising for future research aimed at understanding the nature of online dark triad in human on the Internet. Because ODT has acceptable time stability, it can be used as a scale for long-term studies. As far as we know, this is the only online dark triad scales have this advantage.

7617

Sensitivity and specificity indices for second language testing

Sandra Figueiredo Portugal

The present study tested the validity and the cut-off point of new Second Language (L2) tools through the Receiver Operating Characteristic analysis. These tools aim to ascertain the proficiency of young learners (7–17 years old) in Portuguese as a Second Language (L2). The dimensions used were linguistic: orthographic, semantic, syntactic, lexical; and cognitive: verbal reasoning, auditory recall, dichotic listening. This study was motivated by the gap of tools and concerning the best principles of testing individuals in L2 settings, different from English as L2 context. First, we assessed 108 non-native students attending Portuguese schools, aged 7–17 years, speakers of mother tongues other than Portuguese and residing in Portugal for a period not exceeding eight years. Then, series of univariate analyses, non-parametric tests and the calculation of percentiles enabled the subsequent classification of the subjects' proficiency levels: with or without proficiency (several levels were observed). The sensitivity and specificity indexes were calculated using the Receiver Operating Characteristic (ROC) analysis to define the cut-off points for each test. The score calculation took into account the chronological age, age of immigration, laterality (right or left handed) and gender variables. The evidence from this study suggests that all tasks are suitable to assess immigrant students, but, with the ROC analysis, this assessment differs regarding degrees of proficiency between groups, given the characteristics of the tasks and also of the subjects. The importance of doing tests with discriminatory power of the subjects' performance to be able to intervene in the linguistic and cognitive areas with the greatest deficit is examined here.

7598

Developing a scale measuring children's creativity in their storytelling

Miranda Kit-Yi Wong (1), Wing Chee So (2) 1. *City University of Hong Kong, Hong Kong*; 2. *The Chinese University of Hong Kong, Hong Kong*

This study aimed to develop a scale, which features the multimodality of creativity, measuring children's creativity in their storytelling (Study 1), and examine how speech and gestures play a role in conveying creative ideas with the newly established scale (Study 2).

In Study 1, 50 Chinese (Cantonese)-speaking neurotypical children aged six to twelve provided the narrative samples, and then the psychometric properties of the Creativity Scale for Spoken Narrative (CSSN) were investigated with a sample size of 250 observations based on the raters who evaluated these narrative samples. The results of factor analyses identified and confirmed with a three-factor (Novelty, Expressiveness, and Reasonability) model of CSSN. The scale indicated good internal consistency and interrater reliability. CSSN was validated with Alternate Uses Task (AUT; Guilford, 1967) and a verbal subtest in Wechsler Intelligence Scale for Children®, Fourth Edition (Hong Kong; WISC IV-HK). Convergent validity was demonstrated by the positive and moderate correlations among the three measures; and CSSN was able to distinguish children with different levels of verbal skills to show criterion validity. In Study 2, three groups of raters evaluated children's creativity in their storytelling on CSSN in three conditions (Speech+Gesture, Speech only, and Gesture only) to explore whether speech and gestures function differently in the expression of creativity. Children scored higher on Expressiveness in the Speech only condition and on Reasonability in the Gesture only condition. To conclude, CSSN is a reliable and valid measure to evaluate children's creativity in their storytelling. With this scale, we found that speech and gestures function differently in conveying creativity. Specifically, speech plays a role in presenting creative ideas effectively (i.e., Expressiveness), while gestures help organizing and connecting these ideas in a narration (i.e., Reasonability). Gestures even facilitate the thought processes by manipulating spatial representations in storytelling, thereby enhancing children's creativity of narration.

7408

The use of virtual reality technology in relaxation of operators of underwater rescue equipment

Yuliana Petrova (1), Tatiana Chapala (2) 1. *MGIMO-University, Russia*; 2. *Togliatti State University, Russia*

Art therapy is especially efficient when used in combination with traditional counselling. It is often assigned for clients with anxieties, depression, stress and emotional instability. Moreover, it is employed in a wide range of psychological counselling such as client-centred, Adlerian, cognitive behaviour and Gestalt therapy. Cutting-edge VR technology can be used as a way of art-therapy, which can help under water rescue operators working under emotional strain to relieve stress. The use of relaxation applications Google Earth VR, National Geographic, Discovery, MasterWorks Journey through history and Looking Glass for VR devices can help clients to relieve stress more effectively. Reeder's stress test and the

method of diagnosis of depressive conditions of W. K. Zung test was implemented to assess the emotional conditions of 2 groups of operators of underwater rescue equipment. 2 groups in the total number of 20 of sample operators of underwater rescue equipment whose job requires working in rescue operations under emotional strain are analysed. The research is conducted on the basis of the developer and integrator of ship radio navigation systems JSC TsNII Kurs, Moscow, Russia. Applications for relaxation Google Earth VR, National Geographic, Discovery, MasterWorks Journey through history and Looking Glass used in the device OculusGo demonstrate different scenery of calm and quiet landscapes accompanied by a relaxing music. When exposed to relaxing 3D imaging rescue team workers find themselves surrounded by a nice relaxing atmosphere, which contributes to an adjustment of their psycho emotional state. This can help them take in control their troubling symptoms and increase self-regulation. It should be highlighted that relaxing 3D images are used as part of traditional counselling. It is for efficient for a client to get insight and overcome anxiety and emotional strain. The target group showed a stronger resilience and a better emotional response to distress.

6910

Body cues of physical fitness affect perception of formidability

Vít Trebický (1), Jitka Trebická Fialová (1), David Stella (2), Zuzana Štěrbová (2), Radim Kuba (2), Radim Pavelka (2), Klára Coufalová (2), Karel Kleisner (2), Jan Havlíček (2) 1. *National Institute of Mental Health, Czechia*; 2. *Charles University, Czechia*

A growing body of evidence suggests that people are sensitive to various cues to formidability in men and use them to assess fighting ability. Interestingly, most of the previous research on the perception of formidability is based on facial stimuli. So far, only a few studies investigated the perception of bodies though physique is arguably off primary importance when it comes to formidability judgments. In the present study, we tested a possible relationship between the perception of formidability based on bodies and various measures of physical performance. For this purpose, we collected data from 44 Czech Mixed Martial Arts (MMA) fighters regarding their fighting success and physical fitness (isometric strength, anaerobic performance, lung capacity, body composition). We created standardized life-size 360° body photographs which were assessed by 71 raters (35 males) on their formidability. We found no significant association between perceived and actual fighting ability in our sample. The results of multiple linear regression analysis showed that taller fighters and those with higher isometrics strength were

judged as more successful. In line with previous investigations, our results suggest that certain physical performance-related characteristics are perceived from bodies. However, ratings of fighting success based on bodies do not relate to the actual fighting performance. Our findings thus highlight the role of the physique in formidability assessments.

5052

Examination of the brief psychiatric rating scale (BPRS) factorial structure in a sample of outpatients
Antoni Novotni (1), Silvana Markovska Simoska (2), Biljana Blazhevsk Stoilkovska (3), Nikolina Jovanovic (4)
1. University Clinic of Psychiatry Ss. Cyril and Methodius University, Macedonia; 2. Academy of science and art of the Republic of North Macedonia, Macedonia; 3. Ss. Cyril and Methodius University; 4. Queen Mary University of London, United Kingdom

Brief Psychiatric Rating Scale (BPRS) is widely used measure aimed to assess psychopathological symptoms. In this paper factorial structure of the Macedonian version of the BPRS scale (version with 24 items) was examined. The scale (version with 24 items) was administered in a sample of 82 outpatients with chronic psychosis (female = 34) during regular clinical examinations in two hospitals in Skopje, the capital of the RN Macedonia, as a part of an extensive study approved by the ethical committee. All participants provided written informed consent prior the assessment. Exploratory factor analysis (principal component method with varimax rotation for factor extraction) was performed. Bartlett's test of sphericity was statistically significant ($\chi^2 = 939.03$, $p < .001$) and Kaiser-Meyer-Olkin measure of sampling adequacy was 0.72 (which is higher than the recommended value of 0.50). Accordingly, the assumptions for applying factor analysis were met. Produced five-factor solution was considered as the most clear/interpretative, accounting for 61.90% in the variance (first factor explained 23% of the variance, while the fifth factor accounted for 6.44% in the variance). Each factor was comprised with at least three items with factor loading greater than 0.30. Factor 1 consisted of apathy symptoms. Factor 2 was comprised by items of somatic concern, anxiety and depression. Activation symptoms defined Factor 3, while items with psychotic feature denoted to Factor 4. Guilt, hostility and disorganization items loaded on Factor 5. The obtained factor structure is similar to those presented in the literature, suggesting that the BPRS (24-item form) can be used as a useful instrument for assessment of psychopathology symptoms in Macedonian healthcare context.

4814

Concept cartoons: An educational tool in teacher professional preparation

Libuse Samkova Czechia

The contribution would present an educational tool called Concept Cartoons and my 7-year experience with the tool in teacher professional preparation. I have employed the tool in a novelty role as a representation of practice, and used it for investigating content knowledge and pedagogical content knowledge of future teachers as well as for supporting an implementation of formative assessment in their future school practice. My considerations would be accompanied by results of several empirical qualitative studies which I have performed recently with future primary school teachers during their university mathematics courses. The studies used open coding and constant comparison of data materials to sought answers to the questions:

- 1 Are Concept Cartoons suitable for investigating (pedagogical) content knowledge? In what form?
- 2 What components of content knowledge and pedagogical content knowledge could be investigated through Concept Cartoons?
- 3 Which aspects related to formative assessment could appear in future teachers' written responses to Concept Cartoons and in a subsequent discussion?

The studies showed that a suitable composition of Concept Cartoons with a suitable set of accompanying questions allowed investigating both content knowledge (e.g. solution procedures, reasoning) and pedagogical content knowledge (knowledge of pupils, tasks, instruction), as well as promoting discussions on didactical topics and on topics related to school practice (e.g. dealing with assessment on method as opposed to assessment on results, dealing with missing description of the thinking process, dealing with difficulties in seeing the thinking process that led to the solution). In the contribution, I would present the particular method of use of Concept Cartoons in accordance with the results of the above mentioned studies, and introduce several specific examples.

4672

Behavioural indicators to detect faking-good response style

Paolo Roma (1), Marco Colasanti (2), Alberto Di Domenico (3), Merylin Monaro (4)
1. Sapienza University of Rome, Italy; 2. Sapienza University of Rome, Italy; 3. University "G. d'Annunzio", Italy; 4. University of Padua, Italy

Objective. Deliberate attempts to portray oneself in an unrealistic manner is an issue commonly encountered when administering personality questionnaires. Even though many measures are equipped with validity scales designed to detect response bias, faking is still possible due to the transparency of such scales' items. Our objective is to identify external indicators that could help to detect such behavior. Methods. A series of simulation studies comparing a set of variables (*T*-scores, reaction times, mouse dynamics) with two manipulated factors: with or without time pressure, honest or instructed to fake-good. Results. The first study focused on reaction time and on the effect of time pressure: results indicated that individuals instructed to fake are slower than honest test-takers; moreover, time pressure led respondents in the faking condition to significantly improve their self-presentation. A second study, on Minnesota Multiphasic Personality Inventory-2 underreporting scales, indicated that differences in reaction times between participants instructed to fake-good and honest test-takers were only significant in the time pressure condition. A third study used machine learning analysis to evaluate behavioral deception indicators and concluded that time pressure is the most reliable method to identify faking-good, however in its absence, reaction times are more accurate than validity scales. A fourth study evaluated, along with reaction times and time pressure, the ability of mouse dynamics to detect faking-good; results suggested that mouse movements are slower for participants in the faking-good condition and, furthermore, that in the speeded condition mouse trajectories were wider and less stable than those in the unspeeded condition. A last study evaluated if these indicators are usable not only for questionnaires with dichotomous items, but also for those with 4-choices alternatives. Conclusions. Behavioral indicators are a valuable addition to validity scale for detecting faking-good.

4653

Preliminary development of conforming parenting index of the parents preference test

Cristina Mazza (1), Paolo Roma (2), Maria Cristina Verrocchio (1), Stefano Ferracuti (2)
1. University "G. d'Annunzio", Italy; 2. Sapienza University of Rome, Italy

Objective: The Parents Preference Test (PPT) is a graphical test comprised of 24 easy to understand images of daily family life, which is widely used in forensic assessments of parenting skills. Nevertheless, the PPT lacks validity scales to detect participants' attitudes toward the test; this is an important oversight, as the tendency to demonstrate faking-good parenting behaviors is common in child custody litigants. Study 1 aimed at identifying the differences in PPT responses between a forensic

sample and a control group. Study 2 aimed at developing an index to detect faking-good behaviors. Methods Study 1: A cross-sectional design was implemented. A non-parametric chi-square test for goodness of fit was used to compare the distribution of answer frequencies between the forensic ($N = 99$) and normative/control groups ($N = 110$). Results Study 1: Chi-square goodness of fit tests showed significant differences in answer preferences between groups in 11 (out of 24) vignettes. Methods Study 2: On the 11 vignettes in which significant differences in answer preferences were found in Study 1, the alternatives chosen with the highest frequency by the forensic group were added to an index called the "Conforming Parenting Index" (CPI). Results Study 2: The area under the curve (AUC) of a receiver operating characteristic (ROC) curve for a sample of 58 participants who completed the PPT under both standard and faking-good instructions demonstrated good classification accuracy ($AUC = .813$). Conclusions: Parents in the forensic group tended to take initiative in their relationship with their child; present a focus on the child, a normative style, and predominantly emotional response to their child's needs. The CPI demonstrated good accuracy in classifying caregivers who had been instructed to respond to the PPT in a fake-good way, in line with a socially desirable prototype of good parenthood.

4579

Development and psychometric evaluation of a medical practitioner compassion competency questionnaire

Michelle Visser South Africa

Medical practitioner compassion has been identified as a key construct in healthcare, not only by prominent healthcare stakeholders such as the World Health Organisation, but also by patients and practitioners themselves. In an attempt to contribute to the improvement of medical practitioner compassion in the South African public healthcare sector, the current study firstly conceptualised and constitutively defined compassion as a behavioural construct, but to also operationalise the compassion construct via a Medical Practitioner Compassion Competency Questionnaire (MPCCQ). The connotative meaning that the constitutive definition of this construct needed to capture lies in the internal structure of the construct and the manner in which the construct is embedded in a larger nomological network of other related constructs. By dissecting the competency of compassion, insight was gained into the internal structure of the construct, resulting in the identification of six structurally inter-related latent compassion dimensions. The connotative meaning of the competency was finally brought to fruition in the outcome structural model that was proposed.

The research methodology utilised in operationalising the six latent compassion dimensions in terms of their behavioural denotations, consisted of qualitative critical incident technique interviews, where medical practitioners were utilised as co-researchers in understanding their mental models of compassion from a competency perspective. The research findings from the qualitative interviews enabled the researchers to write behavioural anchors that were subsequently re-written as test items for the MPCCQ, which was then completed by medical practitioners ($n = 234$) at the 21st National Family Practitioners Congress held in Cape Town during 2019, at Karl Bremer hospital, Khayelitsha hospital, Tygerberg hospital and at Worcester hospital, all situated in the Western Cape province, South Africa. The MPCCQ showed construct validity, thus allowing the utilisation of this instrument in practice. The study concludes with practical managerial implications and suggestions for further research.

4313

Toward linguistically fair IQ screening: The multilingual vocabulary test

Julian M. Siebert (1), Kevin G. Thomas (2) 1. Stanford University, Germany; 2. University of Cape Town, South Africa

Objective: This paper presents a solution to one of the most pressing challenges to neuropsychological assessment in linguistically heterogeneous populations: Because different test-related concepts may be accessible to multilingual individuals via different languages, as their lexicon is spread over two or more languages, any monolingual instrument is likely to disadvantage them. We present a developmental account of, and psychometric evidence for, an inherently multilingual English/Afrikaans/isiXhosa screening tool for intelligence, the Multilingual Vocabulary Test (MVT). Methods: Across three independent studies, the paper describes the instrument's development, from a pilot study to a psychometric analysis of the final, digitally administered version. After Study 1 describes a pilot of the MVT's original version, Study 2 ($N = 101$) and Study 3 ($N = 494$) report criterion-related validity analyses of refined versions. Results: Highlights are: (a) significant criterion-related validity values of $r = .46$ and $.52$, respectively, with two standard measures of intelligence, the Kaufman Brief Intelligence Test-Second Edition and the Shipley-2 Verbal Intelligence scale, (b) an internal consistency reliability of $\omega = .59$, and (c) particular sensitivity and informativity in the lower half of the ability spectrum. Linear regressions showed that, although criterion measures were biased toward English first-language speakers, the MVT is largely immune to test-takers' linguistic background. Conclusion: The MVT paves the way toward more fairness in cognitive

assessments, in general, and provides a promising first step toward addressing an urgent requirement of neuropsychologists operating in resource-constrained environments—that of a quick and easy-to-administer, yet linguistically fair, cognitive screening tool.

3811

Developing a questionnaire for internal job transitions: Preliminary results

Linda Berga (1), Inese Muzikante (2), Ivars Austers (1) 1. University of Latvia, Latvia; 2. Latvia

Objective: of the study is to develop a questionnaire that measures non-intellectual predictors of job performance in times of organizational change and development. Aim of the questionnaire is to help organizations fit employees for internal job transitions (redeployment). Based on literature review, authors developed a questionnaire to assess employees' personality, vocational interests, attitudes and orientation towards change and development. Method: in two samples the reliability of newly developed questionnaire for personality and vocational interests was tested. First sample consisted of 190 students from various social science fields – economy, psychology, education. Second sample consisted of 400 employees from various companies in Latvia-telecommunication, aviation and educational sectors. Results: (1) reliability of scales in student sample ranged from unacceptable to acceptable; various changes were made in questionnaire before testing the second sample; (2) test-retest method was used to determine the reliability of questionnaire in employee sample. Conclusions: newly developed personality inventory and vocational interest inventory has overall acceptable reliability although some of the scales (humbleness and openness from personality scales and practical interests from vocational interest inventory) should be revised. Results are discussed in the light of future development of questionnaire.

3432

Measuring acculturation by two instruments: How about consistency?

Ia Shekrladze (1), Nino Javakhishvili (1), Kate Tchanturia (2) 1. Ilia State University, Georgia; 2. King's College, United Kingdom

While it is recognized that acculturation can significantly impact individual's wellbeing, there is a lack of standardized measurement approach, which hinders comparability and generalizability of findings. The objective of the current study was to compare two measures of acculturation – Vancouver Index of Acculturation (VIA) and East Asian Acculturation Measure (EAAM) – in order to minimize measurement related limitations and determine

inter-measure consistency. VIA produces 2 scores on cultural orientations (home culture orientation, mainstream culture orientation), whereas EAAM generates 4 scores on acculturation strategies of assimilation, separation, integration and marginalization. Data were collected from 253 Georgians residing in USA and UK. Validation of Georgian versions of EAAM and VIA produced acceptable fit indices for EAAM: $\chi^2 = 690.09$, $df = 316$, $p = .000$, RMSEA = 0.07, CFI = 0.81, TLI = 0.79; and for VIA: $\chi^2 = 466.94$, $df = 158$, $p = .000$, RMSEA = 0.10, CFI = 0.83, TLI = 0.79. All EAAM subscale scores significantly correlated with VIA mainstream culture orientation scores, assimilation and integration having positive ($r = 0.33$, $r = 0.48$, $p < .001$), and separation and marginalization having negative correlations ($r = -0.44$, $r = -0.24$, $p < .001$). VIA home culture orientation negatively correlated with EAAM assimilation ($r = -0.38$, $p < .001$) and positively correlated with EAAM separation ($r = 0.48$, $p < .001$) scores. Two acculturation measures appeared consistent with one another; however, while VIA effectively measured assimilation and separation, it came short in assessing integration and marginalization. Nevertheless, VIA showed a potential to properly identify four acculturation strategies.

2760

Descriptives and factor structure of IASMHS: Revalidation analysis in Indian context

Rajneesh Choubisa (1), Robinder P. Bedi (2), Namita R. Ruparel (3), Jessica Ahn (2) 1. Birla Institute of Technology & Science, Iran; 2. University of British Columbia, Canada; 3. BITS Pilani, India
Inventory of Attitudes towards Seeking Mental Health Services (IASMHS) is a twenty-four-item standardized tool developed and validated in North American settings with a three faceted factor structure. The factors that constitute the inventory constitute: (a) psychological openness factor, (b) help-seeking propensity and (c) indifference to stigma (Mackenzie, Knox, Gekoski & Macaulay, 2004). In this paper, we aim to confirm the factor structure, psychometric properties and overall utility of the IASMHS in an Indian context. We collected data from two independent samples ($n = 144$, 168) and conducted descriptive and factor analytical psychometric analyses on both samples separately. Specifically, we conducted factor analysis to find a three-factor solution with the combined consolidated dataset ($n = 312$) when certain items (namely 14, 22 and 23) were deleted from the original scale. The emergent factor solution highlights minimal cross-loadings on indifference to stigma and psychological openness factors suggesting minor revision is required on the scale to work in Indian settings. We finally deduce from our interpretations that the scale can be used in Indian conditions and

propose significant recommendations from a counseling psychology standpoint.

2195

How America owes Europe for modern testing: Intersections of international history and testing

Kurt F. Geisinger University of Nebraska-Lincoln, United States of America

This presentation will relate how the birth of intelligence testing took place in Europe but was quickly moved to the United States. Wundt's work in Germany, Binet's work in France, and Galton's work in England all set the stage for advancements in the United States by Lewis Terman and James McKeen Cattell, including norms, normative scoring, and advances in classical psychometrics and item-response theory. It is as clear as the so-called fact that Europeans "discovered" the North American continent that North American psychology has its roots in European psychology. The world's first psychology laboratory was at the University of Leipzig and was led by Wilhelm Wundt. This laboratory spawned several of the earliest significant impacts on psychological testing in the United States as well as worldwide. James McKeen Cattell was one of the Americans who came to Europe to study, initially philosophy at the University of Göttingen and subsequently psychology under the mentorship of Wundt at the University of Leipzig, where he earned his doctorate in 1886 having served as Wundt's assistant for some of his graduate student years. After receipt of his doctorate he traveled to Great Britain where he worked in Galton's Anthropometric Laboratory for two years. Both Wundt and Galton had great influence on Cattell: Wundt on the use of rigorous, standardized assessment procedures and Galton on the nature of the measures that he studied, which were mostly psychophysical in nature, but which were thought to be measures of intelligence. Terman then brought Binet's intelligence test and its foci to the English-speaking world and lead to the development of normative scores. The work of Guilford and Gulliksen explicated classical test theory in a highly organized fashion. Later the work of Frederic Lord revolutionized test theory by pushing it to item response theory.

1934

Integration of handwriting analysis into legal and forensic psychological assessment

Yury Chernov (1), Vali F. Engalychev (2) 1. Institute for Handwriting Sciences, Switzerland; 2. Tsiolkovskiy Kaluga State University, Russia

Psychological assessment in legal and forensic psychology may involve different subjects depending on specific cases and legal status of involved persons – victim, criminal or witness. Often the well-established psychometric

tests cannot be used. In this case, non-verbal methods in general and handwriting psychological analysis (HPA) in particular could be especially helpful. The formal integrative HPA proposed in the current work, particularly its computerised application HSDetect, has been validated against different psychometric tests and presented in multiple studies. That ensures its objectivity, reliability and validity. Although the procedure is computerised, it presupposes the special qualification of experts that apply HPA and evaluate handwritten texts. The system includes several hundreds of handwriting signs and over three hundreds of psychological traits. Distant assessment and possibility to identify various personality traits within one procedure are among major advantages of HPA. That makes it flexible and well adaptable for specific cases in the legal and forensic assessment. In the current work, we present two examples of HPA usage. First, several suicide cases in different countries, where the handwriting analysis of suicide victims helped to reveal their psychological conditions. The analysis of handwriting samples produced in period preceding a suicide is especially useful, because shows the changes of psychological status of a person. Secondly, evaluation of specific psychological traits, in particular implicit aggression of criminals. This allows to better understanding suspects' behaviour in the crime scene.

1057

Predictors of amelioration of psychotic symptoms for children in an inpatient psychiatric setting

Kate Cingel Rockland Psychiatric Center, United States of America

Despite the fact that research has identified difficulties in cognitive and academic functioning, attentional difficulties, and thought and language disturbance associated with a persistent pattern of psychosis or Childhood Onset Schizophrenia (COS), there has been limited research in regard to the predictability of shorter-term outcome and the immediate response to acute inpatient treatment. This study explored the extent to which Rorschach and cognitive variables could differentiate between children with psychotic symptoms whose symptoms remitted versus those whose symptoms did not over the course of an inpatient psychiatric hospitalization. The sample included 200 children, ages seven to twelve, who were assigned to three groups (never psychotic, good outcome, poor outcome) based on the therapist ratings on the psychotic symptoms scales of Children's Psychiatric Symptom Rating Scale (CPSRS). Children were assessed with a battery of psychodiagnostic instruments that included the Rorschach Comprehensive System (Exner 1990, 1991, 1993, 1995; Exner & Weiner, 1994), Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2003) and Conners Continuous Performance Test (CPT

II; Conners et al., 2000). The findings indicate that Rorschach variables (Xa%, X-%, Level 2, FAB2, WSUM6, PTI, TPC, EII-2) significantly differ between children who did not display any psychotic symptoms from children who experienced psychotic symptoms. However, none of the Rorschach variables differentiated the good outcome group from the poor outcome group. Similarly, several cognitive variables (WISC-IV) were able to distinguish the never psychotic group from children with psychotic symptoms. Children initially rated as having psychotic symptoms who displayed poor outcome demonstrated lower WISC-IV Comprehension scores than children with good outcome. Findings are discussed relative to the prognosis and treatment of children with psychotic symptoms. The hope is that the findings of this study would inform the prognosis and treatment of children with psychotic symptoms.

1016

Assessing diverse spirituality for psychologist studies in an increasingly integrated world

Amy L. Ai (1), Ray Paloutzian (2) 1. *FSU, United States of America*; 2. *Westmont College, United States of America*

Over six decades, unprecedented sociocultural changes, experienced by not only American adults but others worldwide, have subverted conventional religious worldviews and ushered in new types of spirituality. Not only societies have become more pluralistic, rapid technological developments and globalization affected people's lives and their existential concerns in important and diverse ways. Especially in the developed world, they have seen a shift from an emphasis on the institutional religious dwelling to more individualized spiritual seeking, followed by a more pragmatic focus on possible beneficial effects of various meditative and mindfulness-based practices and openness to aspects of previously unknown global spiritual worldviews. However, contemporary psychology lacks a comprehensive theory-based research approach to study the increasingly diverse spirituality among modern humans, even though deep meanings rooted in various spiritual worldviews. To address the gap, we have gathered innovative work from scholars around the globe to present at least four approaches to conceptualize diverse worldviews with promise to better understand the significance and effects of diverse spirituality in human lives. This presentation will introduce some assessment tools exemplifying these approaches: The first is a multidimensional approach to conceptualize and investigate historically unaddressed phenomena such as secular spiritual worldviews in the same way that have been used to study affective, attitudinal, and behavioral variables in mainstream Christian religions. The second is a crossover approach to integrate

different ideas associated with various spiritual worldviews, for example, conceptualizing spirituality factors from the view of multiple cultures, based on shared meaning, not merely focusing on church-based mainstream religions. The third is a domain-concept approach with a focus on specified domains of experiences and/or processes commonly observed across diverse spiritual worldviews (e.g., meaning, struggle). The fourth approach centers on faith-specific concepts (e.g., non-cline, jihad) in under-investigated traditions (e.g., Buddhism, Daoism, Islam) in current psychologist research

726

Evaluating gender discrimination in cognitive assessment using the English comprehension test

Danille E. Arendse *University of Pretoria, South Africa*

Cognitive assessment has often been used as a discriminatory tool against gender, race and/or languages. This discrimination has had dire consequences for all vulnerable groups that were deemed subordinate under patriarchy and Apartheid in South Africa. Since the demise of Apartheid, legislation and research practices have been established to guide test developers and users against gender, racial and linguistic discrimination in cognitive assessments. To ensure fairness and equal opportunities for males and females, their performance was assessed within the empirically designed English Comprehension Test. This quantitative study employed a statistical inspection of the test across the gender groups using Differential Test Functioning in Winsteps. The results indicated that the majority of the test items did not present any bias, but there were five possibly biased items identified across gender in the test. These five items which were possibly biased appear to be affected by language and not gendered knowledge, but will however necessitate further investigation. This study contributes to the cross-cultural test development and research within South Africa.

225

The role of psychosocial factors as predictors of student success

Cheryl H. Blackman (1), Ometha O. Lewis-Jack (1) 1. *Bowie State University, United States of America*

Abstract this study examined the academic, behavioral, and socio-emotional health of 125 first-time freshmen and 86 graduating seniors at a HBCU. It compared the types of challenges that students face on entering college with the challenges experienced as they are about to graduate. The goal is to develop a predictive model of student success with intervention strategies that address the challenges associated with retention and graduation. It was predicted that substance use, mental health challenges and level of stress experienced by

students will impact academic success as measured by grade point average (GPA). Also, it was hypothesized that the impact of these psychosocial factors will be different for freshmen and seniors. The results of the study indicated that grade point average was negatively related to alcohol and marijuana use for a sub-sample of females only. High levels of depression and anxiety resulted in alcohol abuse for all participants. Stress had a negative impact on grade point averages with the highest impact occurring among freshmen and intrapersonal stress was a major factor in determining academic success. These results will inform educators about the role of behavioral and socio-emotional factors in students' lives and how the relationship among these variables impact retention and graduation.

496

Forty years of psychological testing in China – The case of the 80–8 assessment scale

Qinghua Zhang (1), Icy(Yunyi) Zhang (2), Binta Zhang (1), Wenying Wang (1) 1. *Soochow University, China*; 2. *University of California, United States of America*

Shortly after the end of Cultural Revolution in China, the authors originally designed China's first psychological measurement scale, "80–8 Measurement Scale." The scale was developed based on theory at that time, in response to the Chinese scholar's and government's wish to originate a scientific method to measure and evaluate human brain function that is suitable for China. The scale is designed according to the degree of development and functional characteristics of human's cerebral cortex as the inhibitory process in the cerebral cortex is an important marker of brain developmental levels. The scale measures the speed, accuracy, flexibility and stability of cerebral cortex's inhibition by the principle of conditioning of the visual dynamic system. The scale adopts non-text symbol and suitable for individuals and groups over 6 years old. In its method, the scale adopts a combined test of 3 levels, with the difficulty increasing in each level, to observe the function of the brain and its corresponding functional characteristics under different mental workloads. In its assessment standard, it has obtained valid test of over 100,000 people in 27 provinces (cities) and established a norm for people over 7 years old. In its type-categorization, the scale proposes a scientific division of human type into 16 categorizations. Over the past 40 years, the 80–8 assessment scale has been promoted to and applied in hundreds of units and departments in various domains across China, with its participants growing to over 100,000 people. In China, scale has become an important method for selecting people for their ability, especially in the selection and configuration of special positions, enterprise management personnel,

athletes and astronauts, and in the selection and education of children with extraordinary intelligence. Through the paper, the authors tell their experience of using 80–8 measurement scale to test and select people in China.

4

Development of a receptive vocabulary assessment tool on tablet: A pilot study in 5–12 y-o children

Pauline Auphanv (1), Jean Ecalle (2), Christophe Gomes (3), Annie Magnan (2) 1. Lab EMC (EA 3085) – Univ Lyon2; Association Agir Pour l'Ecole, France; 2. Lab EMC (EA 3085) – Univ Lyon2, France; 3. Association Agir Pour l'Ecole, France

The assessment of vocabulary is important for children and their teachers. Measures of receptive vocabulary, particularly, are often used as indicators of language ability and of school outcome. Strong vocabulary knowledge is important for success in reading comprehension (Cain & Oakhill, 2014; Currie & Cain, 2015). Using computerized tools (CT) for vocabulary assessment would permit to obtain two parameters, accuracy and speed to lexical access. To our knowledge, few CT have been developed (however see Schaefer et al., 2015). Our aim is to present a new computerized tool to measure the receptive vocabulary for French-speaking children. It is based on the same principle as EVIP (Dunn & Dunn, 1981), the well-known vocabulary test. The children are asked to indicate on a tablet which of four pictures corresponds to a spoken word (noun, verb, or adjective). In this pilot study, one thousand words are tested. Children respond on tablets while results are automatically registered and scored. Then, test information is encrypted and sent to be stored on a secure server. Experimenter has just to connect to the secure server to collect data. To examine the psychometric qualities of the test, we conduct an IRT analysis using two-parameter logistic model (2PL). The item response function of 2PL model takes into account the two parameters, difficulty and discrimination of items (Baker, 2001). This type of analysis will enable to reduce the test to keep the most discriminating items in function of the age.

POSTER PRESENTATION

8169

Intuitive scoring of job interview: Evaluation from a psychological point of view

Kazumasa Mori (1), Mika Kawai (2), Tomoko Oe (3) 1. Japan; 2. Meisei University, Japan; 3. Teikyo University, Japan
In this study, we evaluated the relationship between intuitive scoring for job applications and psychological characteristics of applicants.

More precisely, we used intuitive scoring on interview selection and the Vocational Indecision Scale to ascertain the psychological state of 21 university students in terms of job preparation. We also measured the students' Big Five personality traits using the Japanese version of the Ten-Item Personality Inventory. Additionally, we analyzed the relationship between the intuitive score, psychological state of job preparation, and the Big Five personality traits. Our results show a relationship between intuitive evaluation and psychological state of job preparation, indicating that applicants who are eligible to be hired are also in a suitable psychological state for job application. Furthermore, the personality trait of conscientiousness affects the evaluation of human resources staff in a company.

8106

The validity and reliability of the Chinese vision of the vocational identity status assessment

Shijie Song (1), Ping Wang (2), Xueying Gu (3), Bin Zuo (1) 1. Central China Normal University, China; 2. Liupanshui Normal University, China; 3. Nanjing normal University, China

To introduce and revise the Vocational identity Status Assessment (VISA), and make it applicable to college students and senior high school students in China. Through random stratified sampling, a sample of 871 college students and 760 senior high school students were tested for item analysis, factor analysis, reliability and validity analysis and measurement invariance. The results showed that: The revised VISA included 26 items was divided into six factors: in-breadth career exploration, in-depth career exploration, career commitment making, identification with career commitment, career self-doubt, and career flexibility. Confirmatory factor analysis showed the structure of the VISA was rational. The subscales of career self-doubt, in-breadth career exploration, in-depth career exploration, career commitment making and identification with career commitment were significantly correlated with CES-D-13 and SWLS, indicated that the VISA had good criterion validity. The Cronbach α coefficient of the VISA and all subscales was ranged from 0.677 to 0.863, and the test-retest reliability was 0.725–0.832. The VISA was configural invariant and metric invariant across age, and can divide vocational identity into six types: achievement, foreclosure, moratorium, searching moratorium, diffusion and carefree diffusion. Therefore, the Chinese version of VISA was demonstrated as a psychometrically sound scale with good reliability and validity, and can be used as an effective tool to measure the vocational identity status of college students and senior high school students.

8091

Development of a picture personality inventory: Reliability and validity

Li Yao (1), Xiao Y. Shen (2), Li C. Liang (1), Jun D. Wang (1) 1. Beijing Beisen Cloud Computing Co.Ltd., China; 2. Beike Zhaofang (Beijing) Technology Co., Ltd., China

Personality assessments are widely used in recruiting selection. Traditional personality tests are in form of Likert scale. There is a tendency to employ shorter but better user-experience tests in recent years. This paper reports the development of a novel picture-based personality inventory (PPI), which employs the Big Five Theory and builds a 30-trait personality framework by factorial loading. Every item presented of PPI in the form of “one picture with one phrase”. Participants read each item and choose from two given options “me” and “not me”. It takes 10 minutes to complete the 160 questions. The result is calculated by IRT scoring method. Across studies, the average reliability coefficient is 0.839. Empirical evidence showed that modest, persuasive, competitive oriented, contentiousness, self-assured, power motivation and achievement striving could significantly predict sales performance in a large scale study in a leading real estate company (Beike) in China. Moreover, PPI help the company improve onboarding conversion rates and reduce turnover rates of agent candidates.

7991

A Polish adaptation of motivational postures (Towards Taxes) questionnaire

Sabina Kołodziej Kozminski University, Poland

This poster presents a Polish adaptation of the Motivational Postures (Towards Taxes) Questionnaire (MPQ). The MPQ is based on Braithwaite's (2002) concept of five tax-related motivational postures (Commitment, Capitulation, Resistance, Disengagement and Game Playing) and consists of 29 items. Three studies validating the Polish version of the MPQ are presented. The first study was conducted with a translated version of the original questionnaire and aimed to verify the factorial validity of this version using confirmatory factor analysis (CFA). Since the factor structure described by Braithwaite (ibid.) was not reproduced, exploratory factor analysis (EFA) was conducted. Study 2 used CFA to confirm the new structure of the modified version of the questionnaire evident from the Study 1 EFA, and also estimated the reliability and internal validity of the modified version. This resulted in a questionnaire consisting of 20 items and five scales (Moral Duty, Capitulation, Active Resistance, Disengagement and Pleasant Games).

The third study tested the questionnaire's construct validity. A theoretical interpretation of the scale is provided.

7716

Features of the assessment of educational innovations and conducting a pedagogical experiment

Galiya Abayeva (1), Diana Koroleva (2), Aigul Akhmetova (3), Assel Akpayeva (3) 1. Kazakhstan; 2. National Research University Higher School of Economics; 3. Abai Kazakh National Pedagogical University, Kazakhstan

Abai Kazakh National Pedagogical University in the Republic of Kazakhstan carried out a research project on cross-country research on the problem of innovation in education and a pedagogical experiment together with the National research university higher school of economics in the Russian Federation. The relevance of this project research is determined by the development of tools for identifying the effectiveness of selection, the search for innovative projects in education and pedagogical experiments. Methodological approaches to the selection and search for innovative projects in education and the implementation of the pedagogical experiment are system-activity, structural, statistical, technological. As a result of the study, we have developed a training program for workers in the education system, formed a database of educational innovations in Kazakhstan as a sub-stage of the international competition for innovations in education, carried out an experimental study of the implementation of innovative educational technologies in Kazakhstan. On the basis of the selected methodological approaches, criteria and indicators for the selection and assessment of educational innovations were developed, and the coefficient of the effectiveness of the pedagogical experiment was calculated using the methods of mathematical statistics. The module efficiency coefficient is equal to 1 and corresponds to a high level of connection between the variables that determine the quality of educational innovations in Kazakhstan. Recommendations on the use of the research results were introduced into the educational process of Kazakhstani universities, as well as into the system of advanced training of educational workers in the Republic of Kazakhstan.

7380

Caregiver needs and resources assessment: Development and validation

C. Y Lam (1), K.-K. Li (2), K.H.E. Chung (2), D.Y. Yeung (2), M.Y.L. Chiu (2), T.W. Lo (2) 1. Hong Kong; 2. City University of Hong Kong, Hong Kong

Background: Caregivers of older adults are essential in fostering aging-in-place. Caregiving can be challenging and uncertain.

Most supporting efforts mainly satisfied caregivers needs, their strengths are overlooked in individualized planning. This study aimed to develop the Caregiver Needs and Resources Assessment (CNRA) incorporates insights from dual process model (Stroebe & Schut, 1999) & Pearlin et al. (1990) caregiving stress model which can aid social workers in case management of caregivers. Methods: Items were generated based on theories, existing measures, and practitioner consultations. 351 caregivers recruited through four NGOs in HK, completed a survey including initial items in CNRA, measures of related constructs, and well-being outcomes (e.g., quality of life, burden, positive aspects of caregiving). Results: Items on needs and resources were analyzed separately, considering the differences in their nature and utilization in care management. Using EFA, five factors on needs (physiological needs, role conflict, care recipient's needs, psychological needs, and social needs; 15 items) and seven factors on resources (spirituality, self-efficacy, responsibility, community supports, family supports, closeness with the care recipient, and health awareness; 21 items) were extracted. The model fit indexes in the CFA were also satisfactory (Needs: 2 (80) = 204.97, $p < .001$; CFI = .95; TLI = .94; RMSEA = .07; SRMR = .06. Resources: 2 (168) = 373.82, $p < .001$; CFI = .94; TLI = .93; RMSEA = .06; SRMR = .05). The evidence for convergent validity was demonstrated by the moderate-to-high correlations between the factors in CNRA and the related constructs/outcomes. Discussion: Supporting evidence was found on the validity of CNRA. The factors on needs covered most areas of other existing measures. Personal agency (e.g., self-efficacy) and social agency (e.g., family supports) are two major groups of resources. Further studies are needed to examine the usefulness of CNRA in case management by social workers.

7174

Sociocultural Conditions of a Family and Its Relevance to Teacher Assessment of Adaptive Behavior

Julie Melicharová (1), Lucie Chadimová (2), Jakub Jelínek (2), Matěj Seifert (2) 1. Czechia; 2. National Institute for Education, Czechia

Adaptive behavior is defined as social, conceptual and practical abilities that enable to handle demands of daily life. Internationally there are several instruments available to assess child's adaptive behavior, however, no psychometrically sound norm-referenced instruments are available in Czech Republic. In our project we develop adaptive behavior rating scale that would (1) reflect life in Czech Republic, (2) apply to assessment of ID and (3) provide data for qualitative analysis of child's special needs. Rating scale will evaluate adaptive behavior

of children from 5 to 15 years in both family (parental form) and educational (teacher form) context. Along with cultural fairness we evaluate impact of sociocultural conditions on adaptive behavior assessment. Construct is measured by self-reported teacher scale (18 items), based on the Project of Early Care of Children from Culturally Disadvantaging Surroundings in the Field of Education (Ministry of Education, Youth and Sports, 2005). Based on analysis of 452 responses we described 4 groups with zero to high sociocultural disadvantage. We found negative relation between reported adaptive behavior and reported sociocultural disadvantage ($\omega_2 = 0.179$), membership in the lowest group predicted lower adaptive behavior ($\beta = -0.21$). Relation was stronger in self-regulation and social adaptation items ($\omega_2 = 0.26$, $\omega_2 = 0.2$, respectively) compared to practical and academical items ($\omega_2 = 0.13$, $\omega_2 = 0.05$, respectively). We found significant interaction with age, lower social disadvantage in younger children predicted lower teacher reported adaptive behavior.

7170

Conceptualising the Malaysian personality: A common factors approach

Damien Cheow (1), Yan Shan Tai (2), Sabrena G. Arosh (3), Maria F. Marnuag (4) 1. DISTED College, Malaysia; 2. University of Wollongong Malaysia KDU Penang University College, Malaysia; 3. University of Malaya, Malaysia; 4. HELP University, Malaysia

Traditional approaches to building a career narrative typically include the use of assessments of vocational interest, personality, relevant skills, and abilities. The development of the HELP CaRES (HELP Career Readiness Evaluation System) aligns with this approach in addition to grounding each assessment within the Malaysian cultural context. Previous attempts to construct a personality instrument based on the Five-Factor Model proved unsuccessful as several factors failed to emerge during factor analyses. As such, the current study based the theoretical foundation of the personality instrument on previous personality conceptualizations done in Asia such as the Chinese Personality Assessment Inventory (Cheung et al., 1996, 2001) and studies of the indigenous Malaysian personality (Mastor et al., 2015). Given the unique Malaysian cultural context with majority Malay, Chinese, and Indian populations, the current instrument encompassed constructs identified from these cited studies with the aim of capturing common constructs across the three major ethnic groups. The 287-item instrument covering 45 constructs was administered to 293 Malaysian participants. During each iteration of exploratory factor analysis (EFA) using principle axis factoring and varimax rotation, items with communalities below 0.40 and items with

cross-loadings were excluded from the questionnaire. A total of nine such iterations converged on a 23-item four-factor model of Leadership, Even-Temperament, Extraversion, and Self-Confidence, explaining 54.28% of the variance. Reliability tests revealed high Cronbach alphas for all four subscales, with an alpha reliability of 0.92 for Leadership, 0.85 for Even-Temperament, 0.83 for Extraversion, and 0.70 for Self-Confidence. Future studies aim to validate the four-factor structure with a larger sample as well as validate scale translations to be used within the career counselling context of the HELP CaRES.

7035

Education and cognitive functioning effects on the trail making test in elderly Puerto Ricans

Karen Rosado-Mercado (1), Manuel González-González (1), Giovanni Tirado-Santiago (1) 1. University of Puerto Rico, Puerto Rico

The Trail Making Test assesses speed performance (TMT-A), task switching (TMT-B) and executive control (TMT B-A), which are associated with frailty onset, cognitive decline, and visual processing in older adults. This study investigated how TMT scores are affected by education and cognitive functioning in 86 (60 ≥ year) community sample in Puerto Rico. Cognitive functioning was assessed by the Mini Mental State Examination (MMSE) and Mini-Cog. We did linear regression analyses on TMT-A, TMT-B and TMT B-A using Mini-Cog, MMSE, and education as predictors. Education contrast were CONTRAST-1 (≤12 vs. undergraduate and graduate studies) and CONTRAST-2 (undergraduate vs graduate studies). Since MMSE is affected by education this interaction was added as a predictor. Regression on TMT-A ($F(6, 73) = 22.37, p < .001$) showed that the Mini-Cog ($p = .024$), MMSE ($p = .014$), CONTRAST-1 ($p = .039$), CONTRAST-2 ($p = .001$), MMSE and CONTRAST-2 interaction ($p < .001$) were significant. Regression on TMT-B ($F(6, 73) = 9.63, p < .001$) showed that MMSE ($p = .009$) and CONTRAST-1 ($p = .001$) were significant. Regression on TMT B-A ($F(6, 73) = 7.12, p < .001$) showed that MMSE ($p = .021$), and CONTRAST-1 ($p = .001$) were significant. Although norms for Puerto Rican adults suggest education does not affect TMT-A, results show that general cognitive functioning and education impact all TMT scores. The Mini-Cog only predicted TMT-A but not TMT-B or TMT B-A, however the MMSE did. This implies task switching and executive control do not have a significant role in the Mini-Cog.

6903

Studying identity conflicts with the repertory grid technique: Method and review of findings

Guillem Feixas (1), David A. Winter (2), Víctor Suárez-Aragón (1) 1. Universitat de Barcelona, Spain; 2. University of Hertfordshire, United Kingdom

As individuals, we strive to construct a personal identity that enables a sense of psychological continuity and distinctiveness. Developmental transitions, adversity, social exclusion, and relational conflicts can be experienced as identity threats that may be unsettling to the extent of triggering substantial psychological distress. These connections between identity and symptoms or distress can be progressively revealed during a psychotherapeutic process. From the personal and social constructivist points of view, recommended changes in problematic behaviors or thoughts might involve some degree of transformation in the patient's sense of identity. These challenges to continuity in personal identity have been suggested to sometimes foster internal conflict, ambivalence towards change, and hinder clinical improvement. Indeed, the pursuit of change requires a delicate balance between the need for change and the need for psychological continuity, as it occurs in successful psychotherapy or resilient processes. Personal construct researchers and practitioners have developed the repertory grid, a semi-structured interview for the elicitation of the constructs a person uses to construe self and others using a rating scale. The resulting data matrix permits the computation of self-discrepancy and cognitive structure measures. Implicative dilemmas are detected whenever the desirable pole of a discrepant construct correlates with the undesired pole of a congruent construct. Then, a desirable change in one construct implies an undesirable change in the other. Results on a series of studies have been conducted comparing samples with different conditions (e.g., depression, bulimia, fibromyalgia) with control groups. A meta-analysis of the first set of these studies yielded a medium effect size ($OR = 3.43$; 95% CI: 2.50–4.69), indicating that participants in these conditions were more than three times more likely to present with (at least) one implicative dilemma in their grids compared to controls. A case study will be presented to illustrate the method employed and its clinical implications.

6889

Comprehensive assessment battery of language development in Czech Republic

Gabriela S. Málková Charles University, Czechia

This poster aims to introduce the theoretical and methodological background of the first comprehensive assessment battery of language development in Czech Republic. At the time of the abstract submission, the battery is about to be standardized on a sample of approx. 1000 children of early school-age children (age 3 to 8 years). As a part of the poster presentation,

the overall structure a language assessment battery and the logic of the battery construction will be presented. The special emphasis will be placed on methodological issues related to the standardizing study and to the position of a proposed battery within the Czech system of psychological assessment within the counselling and clinical praxis.

6734

Key literacy skills of Czech primary school pupils: Preliminary analysis of new diagnostic

Klára Špačková (1), Anna Kucharská (2), Gabriela Seidlová Málková (2), Hana Sotáková (2), Pavla Presslerová (2) 1. Czechia; 2. Charles University, Czechia

Though assessment of reading literacy has a long tradition in the Czech Republic, the presence of standardized reading measures is very limited, especially when it comes to the reading comprehension. The available tools either measure comprehension on the sentence level or are designed to assess comprehension of a narrative text based on retelling a story. The ongoing investigation of the TA ČR (Technology Agency of the Czech Republic) project aims to respond to the limits of current psychodiagnostic testing and develop a new diagnostic battery of the literacy skills. We build on current trends in the foreign literature and bring new approaches to evaluate different forms of comprehension which contribute to the development of functional literacy from the developmental perspective. The new comprehensive battery therefore includes variety of tools measuring decoding and reading related skills as well as text comprehension skills. Comprehension is assessed on different levels (explicit, implicit, and interpretative) and under several conditions: filling in missing words to the text, answering questions after oral/silent reading or listening to narrative/exploratory texts. The battery also includes questionnaires for children, parents and teachers to enable the examiner to map the socio-environmental and motivational factors. In the poster we present preliminary data of standardized study of the battery administered to 1000 monolingual, Czech-speaking primary school pupils (1st–5th grade) in the beginning and at the end of the academic year 2018/2019 and 2019/2020. We will present descriptive analysis of the performance of the representative sample and psychometric properties of the measures.

6708

From intention to evaluation: The instrumental practice goal realization inventory

Zofia Mazur (1), Mariola Łaguna (2) 1. Poland; 2. The John Paul II Catholic University of Lublin, Poland

Playing musical instrument requires much effort and continuous practice from musicians.

In order to understand how instrumental students, motivate themselves for long hours of individual practice, we need appropriate tools to track the process of music learning, which starts from setting intention and plans, proceeds into performing an action and evaluating its effects. To tackle this process, we developed the Instrumental Practice Inventory Goal Realization (IPGRI). This tool is based on the mindset theory of action phases and on the theory of deliberate practice in instrumental learning. Two studies: Study 1 on 171 piano students and longitudinal Study 2 on 235 students learning various instruments were performed to develop the measure and to test its psychometric properties. Their results allow identifying three higher order factors of the IPGRI and its seven subscales, three items each. Confirmatory Factor Analysis provided evidence for that the measure structure fit well with the data obtained from two studies with adolescents and adults attending to music schools. High reliability (internal consistency and test-retest stability) of scales was obtained. The new tool will be beneficial in basic and applied research in psychology of music and related disciplines.

6426

Examining the impact of online survey attention checks on the physician-patient relationship quality

Ju-Chun Chien (1), Aaron Kuo (2), Matthew S. Lee (2), Sharon Lin (2)
 1. Oriental Institute of Technology, Taiwan; 2. Morrison Academy, Taiwan

Objectives: The main purposes of this study were to examine: (1) the effects of using attention check questions on Taiwanese people's perception of physician-patient relationship (PPR) quality, physicians' patient-centered (PPC) communication style, and health behavior self-efficacy (HBSE), respectively; (2) whether there were any significant demographic differences between the attentive and inattentive online participants. Methods: This online research used a reversed question to deal with inattentiveness. Participants who showed inconsistent responses on the two duplicate questions were classified as inattentive online samples. Data collection was started on November 30 and was completed on December 19, 2019. A total of 1093 surveys were completed, about 67% were classified as attentive participants ($n = 735$); the invalid rate was 32.75 percent ($n = 358$). After conducting the EFA and the reliability analysis with 735 participants, the self-developed questionnaire using the 7-point Likert scale responses was determined to be a reliable and valid instrument. Descriptive statistics, independent sample t test, and contingency table analysis were used in this study. Results: The attentive participants revealed a significantly higher degree of PPR quality, PPC communication style, and

HBSE than the inattentive participants. In comparison to the attentive participants, the inattentive respondents tended to be older, less educated, had a neutral attitude towards the medical care process, no healthcare-related experience, and had undergone general examinations instead of specific examinations. Conclusions: Attention checks were regarded as a useful mean to facilitate data quality in this online study. However, the research topic and/or the population attributes should be taken into account when conducting online surveys.

5855

Look at the card 6 – differential power of the Rorschach popular responses

Munehika Ito (1), Hisae Tsuge (2), Shintaro Hattori (2), Kaori Sakamoto (2), Haruna Azuma (2)
 1. Gifu University, Japan; 2. Gifu Hospital, Japan

Objective: The Rorschach popular responses (P) are known as the high frequent responses and its contents or percepts are defined for each cards. We have found that the P may be useful for differential diagnosis in our study (e.g. ITO, 2019). Then, the aim of this study was to go through the differential power of the P for each cards. Methods: This study was retrospective. 26 schizophrenic and 19 ASD, 27 personality disorder patients (almost BPD) were administered the Rorschach as the daily clinical works in a hospital. The administration and the coding were conducted by the Kataguchi system (most popular in Japan). Also, the codes were translated to those of the Comprehensive System (CS) for the statistical analysis. Dealing with the frequency of the P on each card (by both of Kataguchi and CS), they were calculated statistically. The P was dependent variable. Results: Using ANOVA for groups (schizophrenia vs. ASD vs. personality disorder [PS]), the main effects were found on card 6 and total numbers of P for Kataguchi system, and on card 4 & 6, total numbers of P for CS. the analysis of HSD showed (1) $PS > S$, $PS > ASD$ on card 6 and $ASD > S$ for sum P [Kataguchi], (2) $PS > S$, $PS > ASD$ on card 6 [CS]. Conclusions: Only the P on card 6 (fur rug), common to both systems, partially differentiated groups (personality disorder patients vs. schizophrenics and personality disorder patients vs. ASDs). So it is suggested that the patients' sensitivity for shading is important. And the Rorschach cards with shading (4, 6, and 7) may have the possibility of the differential power. Focusing on shading of the cards, the further capability of the P was discussed.

5811

Collaborative feedback session using KABC-II with an advice report form for high school students

Takashi Kumagami Wako university, Japan

Objective: Feedback sessions are an important part of psychological testing for clients, parents, school teachers, and their supporters. Finn (2007) recommended a collaborative assessment and provided a protocol for feedback sessions. We consider that collaborative feedback sessions are effective for cognitive and achievement testing. Methods: We conducted five collaborative feedback sessions for students, using the Kaufman assessment battery for children, second version (KABC-II) in a public high school in Japan. We adopted an advice report form to better understand the results of the psychological testing and feedback protocol. We made a new advice report form including scores for cognitive scale, academic achievement, and cognitive characteristics, as well as our recommendations. We analyzed the oral data for the feedback session. Results: The students, parents, and teachers reported that the collaborative feedback session using the advice report form for their cognitive and academic achievement score and characteristics could easily be understood and was useful, illustrating the students' levels of cognitive and academic achievement and style, as well as their sense of security and motivation for school life. Conclusions: Some high school students had difficulties in cognitive and academic achievement. Collaborative feedback sessions using advice reports, including their characteristics and recommendations, were very useful for the students, parents, and teachers, and they promoted a team support system of special needs education for students with academic and psychological difficulties.

5547

Application of multidimensional item response theory and multidimensional computerized adaptive test

Po-Hsi Chen National Taiwan Normal University, Taiwan

The goal of this research is to apply the multidimensional item response theory (MIRT) and multidimensional computerized adaptive testing (MCAT) technique on the development of multiple intelligences test. The computerized multimedia items for assessing the multiple intelligences of grades 4–12 was developed since 2006. These items were developed based on Gardner's theory of multiple intelligence (1983). Since the bodily-kinesthetic intelligence can't be assessed with multimedia computerized test, it was excluded and replaced by the emotional intelligence items (Goleman, 1995; Salovey & Mayer, 1990). All 405 items for eight intelligences were developed, reviewed, and modified. These items were assembled into different booklets and pretested to 3,722 students of grades 4–12 with computerized testing system. Results of analyzing the pretest data indicated that

only 252 items fit the multidimensional Rasch model (Wang, Wilson, & Adams, 1997). The information function, which correspondent with the reliability, showed that only six items for each dimensions were needed to reach the acceptable precision of ability estimation with MCAT technique. The models comparison showed that eight dimensional model fit the data better than the other models, which demonstrated good construct validity. Students major in mathematics, linguistic, and design art perform better in their superior intelligence, which demonstrated good criterion-related validity. Suggestions and applications of developing MCAT had been addressed in this research.

5516

Blurred boundaries: Autism spectrum and co-morbid disorders

Melanie Johnson *United States of America*

Individuals with Autism Spectrum Disorders often present with comorbid emotional and neurodevelopmental issues. Ascertaining if we are treating Autism, early stages of serious mental illness, or both can be challenging. This session will provide tools to assess these individuals appropriately, and lead them toward more tailored treatment plans.

5506

Psychometric properties of the Spanish version of the children's inventory of anger

Julia Gallegos-Guajardo (1), Norma Ruvalcaba-Romero (2), Ania Chavez-Amavizca (1), Cecilia Martinez-Torteya (1), Michael W. Nelson III (3)
1. *Universidad de Monterrey, Mexico*; 2. *Universidad de Guadalajara, Mexico*; 3. *Xavier University, United States of America*

The Children's Inventory of Anger (ChIA) is a self-report measure that assesses the level of anger experienced by children and youth in four domains: frustration, physical aggression, peer relationships, and authority relationships. The ChIA is a 39-item measure originally based on Novaco's model of anger and Schachter and Singer's theory of emotion, which more broadly aligns with the Rational-Emotional/Cognitive-Behavioral model. The purpose of this study was to validate the ChIA for the Mexican population (ChIA-M). The participants were 638 students, 7 to 13 year old ($M = 9.8$ years ($SD = 1.51$)) attending public schools located in the cities of Guadalajara (2 schools, 32% of student sample) and Monterrey (6 schools, 68% of student sample) in Mexico. The exploratory factor analysis supported the four-factor structure with a shared variance of 47.58%, however a different distribution of items was found and three items were deleted. The confirmatory factor analysis found the new four

factor structure to be a good fit with the data [χ^2 (318) = 1109, $p < .000$, $GFI = .889$; $RMSEA = .047$], and good internal consistency was obtained for the subscales and the total score. Study findings suggest that the ChIA-M is a valid instrument for assessing anger in Mexican children and youth. Further studies are highly encouraged to continue exploring this measure. Limitations and implications for practice and further research will be discussed.

4988

Psychometric properties of the WAIS-IV: An item response theory analysis

Gomaa S. Abdelhamid (1), Juana Gómez-Benito (1), Marwa G. Bassiouni (2)
1. *Institute of Neurosciences, University of Barcelona, Spain*; 2. *Fayoum University, Egypt*

Objective: The Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) has been adapted to 27 different cultures and there has been considerable interest in examining its structure through exploratory and confirmatory factor analysis. This study examines item and scale properties of the WAIS-IV using item response theory (IRT) analysis. **Methods:** The WAIS-IV used in the present study was the version adapted for Arabic speakers, which consists of 10 core subtests and five supplemental subtests. A volunteer sample of 250 Egyptian adults was tested between 2015 and 2016. Participants ranged in age from 18 to 24 years ($M = 20.65$ years, $SD 1.71$), and 62% were female. All the IRT analyses were performed using IRTPRO 4. In addition, factor analysis was conducted with polychoric and tetrachoric correlations to investigate the dimensionality of each subtest. **Results:** The unidimensionality assumption was confirmed by our results, indicating that each WAIS-IV subtest measures one latent trait, although there were a few items that showed poor loading on the first factor. The two-parameter logistic model provided the best fit for the six dichotomous subtests of the WAIS-IV, while the graded response model fitted the polytomous data. The IRT results also indicate, first, that the WAIS-IV subtests are able to differentiate between adults with different levels of the cognitive variable being measured, and second, that they gather information especially for adults whose intellectual ability is greater than average. **Conclusion:** Overall, the IRT analysis indicated that the WAIS-IV is an adequate tool for assessing adult intelligence in another culture (in this case, the Egyptian population), with only minor modifications on some items. The current study provides new evidence regarding the utility and measurement precision of the WAIS-IV, as adapted for Arabic speakers.

3816

Psychometric characteristics of the Slovak version of drinking motives questionnaire

Marta Vavrová (1), Zuzana Rojková (1)
1. *University of Ss. Cyril and Methodius in Trnava, Slovakia*

Aims: The study deals with the verification of psychometric characteristics of Slovak Drinking Motives Questionnaire (DMQ, developed by Cooper, 1994) version to measure the four drinking motives implied by crossing two dimensions: enhancement (internal, positive; e.g., to have fun), coping (internal, negative; e.g., to forget problems), social (external, positive; e.g., to be sociable), and conformity (external, negative; e.g., to fit in with a group). The goal of study was to verify criterion and construct validity, as well as reliability such as internal consistency of the Slovak version DMQ (Vavrová, Rojková, 2019). **Methods:** The sample included 800 adolescents aged 15–17 years attending various types of secondary grammar schools. The questionnaire battery consists of original Slovak version DMQ (Vavrová, Rojková, 2019), RAPI – Rutgers Alcohol Problem Index (White and Labouvie, 1989) and SMF – Questionnaire on Situation-motivational Alcohol Behavior Factors (Rojková, Vavrová, 2017). In statistical processing the factor analysis, comparative, correlation and internal consistant reliability analysis have been applied. **Results:** The results confirmed the psychometric properties – validity and reliability of the Slovak Drinking Motives Questionnaire version. **Discussion and conclusions:** Examination of drinking motives through the validated questionnaire could be used in many areas and connections within field of Psychology and interdisciplinary as well. We can find application in the investigation of self-harm, where heavy drinking is both a form and an accompanying phenomenon and the motives of both risky behavior forms in adolescence overlap.

2950

The psychometrics properties of two versions of ANS test

Yulia Kuzmina (1), Julia Marakshina (1), Marina Lobaskova (1), Maria Sitnikova (2)
1. *Psychological Institute of Russian Academy of Education, Russia*; 2. *Belgorod National Research University, Russia*

The Approximate Number Sense (ANS) is the ability to perceive and process quantitative information without counting and using symbols. There are variety of tests that used to measure the ANS. One of the most popular test is the nonsymbolic comparison test (e.g. "blue-yellow dot test"). In this test individuals have to compare two sets of dots and choose which set contains more dots. There is evidence that during execution of this test individuals

are affected by visual properties of sets such as cumulative area or surface area. Some studies also demonstrated that indicators of acuity of ANS significantly varied across different types of nonsymbolic comparison tests depending on ways of controlling visual properties and different formats of dots presentations (e.g. mixed or separated formats). At the same time there is lack of studies where psychometric properties of different types of nonsymbolic comparison test were assessed via Item Response Theory (IRT) framework. In current study we aimed to assess psychometric properties of two versions of nonsymbolic comparison test. The first version contained stimulus which were presented in mixed format and one visual parameter (cumulative area) was controlled for. All stimulus were congruent (set which contained more dots had the larger cumulative area). The second version of the test contained congruent and incongruent trials and two visual properties were controlled for (cumulative area and surface area). The stimulus in the second version were presented in mixed or separated formats and in heterogeneous or homogenous conditions. The reliability, measurement bias and dimensionality of each test and items difficulty were assessed and compared.

2934

Social comparison as moderator of the relationship between implicit and explicit wellbeing measure

Miriam Bajo (1), Darío Díaz (1), María Stavraki (1) 1. *Universidad de Castilla la Mancha, Spain*

The study of indirect measures of psychological constructs has received increased attention by social researchers. In the study of well-being, however, traditional measures tried to capture people's wellbeing in an explicit and direct way. Recently, a new implicit life satisfaction measure using partially structured stimuli (PSWBS) has been developed. This new measure implies that people use their own circumstances as a standard of comparison to evaluate ambiguous scenarios about the circumstances of a third person's life. Therefore, the measure is very sensitive to social comparison processes. Also, consistent with other fields (e.g. attitudes domain), previous research suggests a weak relationship between direct and indirect well-being measures. As a consequence, the main goal of this research was to examine the relationship between direct (SWLS) and indirect (PSWBS) well-being measures considering the role of social comparison tendency. As expected, social comparison tendency (INCOM measure) moderated the relationship between implicit well-being (PSWBS) and life satisfaction (SWLS scale). Specifically, the interaction between PSWBS and INCOM revealed that PSWBS showed a stronger correlation with SWLS when people had a high (vs. low) social comparison tendency.

2678

The Portuguese version of the "Échelle de Fatigue de Pichot": A 3-faced construct validation method

Pedro Joel M. Rosa (1), Luís Brito (1) 1. *Universidade Lusófona de Humanidades e Tecnologias de Lisboa, Portugal*

Fatigue is a complex phenomenon for which healthy people are regularly complaining and can be found in several medical, neurologic, and psychiatric disorders. This phenomenon is very subjective and greatly varies between individuals. Several evaluation questionnaires like Pichot, or Fatigue Severity or Fatigue assessment scales were developed to quantify fatigue more objectively. Despite the common use among French psychologists and psychiatrists, the Pichot fatigue scale was not yet translated and validated for the Portuguese population. A sample of 843 Portuguese participants was randomly split into three parts, 20% for Exploratory Factor Analysis (EFA; $n = 171$), 40% for an exploratory Confirmatory Factor Analysis (CFA $n = 337$) and 40% for a cross-validating CFA ($n = 337$) as a protection against overfitting to safeguard model replicability. Measurement invariance (configural, metric and scalar) was examined over the entire sample ($n = 843$). Results on EFA and CFAs supported the unidimensionality and reliability of the European Portuguese version. Findings also suggested the model was invariant (configural, full metric, and full scalar) across gender. Overall, our study showed evidence that the European Portuguese version "Échelle de Fatigue de Pichot" is a valid and reliable scale and that a global score can be obtained for fatigue assessment in Portuguese adults.

2415

Construct validity of the WISC-V UK with Irish clinical sample

Gary L. Canivez (1), Ryan J. McGill (2); Kate James (3), Rebecca Good (3) 1. *Eastern Illinois University, United States of America*; 2. *William & Mary, United States*; 3. *Éirim: The National Assessment Agency, Ltd., Ireland*

Objective The WISC-V-UK Administration and Scoring Manual presents results from confirmatory factor analyses (CFA) of the UK standardization sample that used the same problematic methods applied to the US WISC-V (Beaujean, 2016; Canivez, Watkins, & Dombrowski, 2016a, 2016b; Canivez & Watkins, 2016). Detailed WISC-V-UK CFA procedures are lacking (Boomsma, 2000) as are examinations of bifactor structure as an alternative higher-order models promoted by the publisher. Canivez, Watkins, and McGill (2018) reported the best representation of WISC-V-UK measurement in the standardization sample was a bifactor model with four

group factors (VC, PR, WM, and PS) supported by both EFA and CFA. The present study reports preliminary results from independent WISC-V-UK EFA and CFA with Irish children referred for learning difficulties evaluation to determine if results replicate with a clinical sample. **Methods** A sample of 354 children with complete WISC-V-UK primary subtest scores served to examine the latent structure through best EFA and CFA practices. Hierarchical exploratory factor analyses with omega estimates and higher-order and bifactor models in confirmatory factor analyses were used. Omega-hierarchical and omega-hierarchical subscale coefficients helped assess unique contributions to true score variance. Results EFA results failed to support presence of a Fluid Reasoning factor and only four factors could be located (factor 5 had no salient pattern coefficients). The best structure appeared to be a four-factor model with a dominant general intelligence factor that accounted for the majority of test variance. CFA supported for only four group factors, of which Fluid Reasoning was not one, and dominance of general intelligence. Model based reliability estimates (Omega-hierarchical) showed adequate portions of true score variance captured by general intelligence but inadequate unique true score variance among the four group factors except possibly Processing Speed; thus the WISC-V-UK is primarily a measure of g .

1611

Three methods of alcohol consumption in women of childbearing potential

Ashleigh L. Chiaf (1), Andrea Blair (1), Erica K. Crockett (1), Julie M. Croff (1) 1. *Oklahoma State University Center for Health Sciences, United States of America*

Introduction: Ethyl glucuronide (EtG) concentration in hair samples have become a common tool for alcohol use detection in pre-employment screening as well as alcohol and drug treatment monitoring. The purpose of this manuscript is to evaluate the validity of hair EtG concentrations compared to transdermal alcohol concentration (TAC) and self-reported alcohol use. **Methods:** This trial included 25 adolescent and young adult females who reported at least one heavy drinking episode (≥ 4 drinks) within the past two weeks. All participants were asked to wear a Giner WrisTAS-7 ($n = 21$) or Giner WrisTAS-9 ($n = 4$) alcohol biosensor over a one-month prospective study. Participants visited the lab weekly to complete a detailed self-report of behaviors, including day of drinking events, amounts and types of alcohol use, and length of drinking events. Estimates of blood alcohol concentration (eBAC) were computed from self-report data using National Highway and Transportation Safety Administration (NHTSA) equation. TAC and eBAC data

were categorized into at risk and high-risk drinking events (BAC of >0.05 and >0.08 grams alcohol/milliliter blood, respectively). Hair EtG concentration, total number, moderate and high level of TAC and eBAC drinking events were analyzed with Spearman rank correlation test for validity comparisons. Results: No significant correlations were found between Hair EtG values and total number, moderate or high risk for detected drinking events by eBAC or TAC. Total number of drinking events detected and number of drinking events >0.08 using eBAC and TAC methods were significantly correlated with each other (respectively, $r = .33, p < .05$; $r = .42, p < 0.05$). Discussion: Our findings indicate that due to the number of false negatives hair EtG concentrations are not a valid tool for monitoring alcohol use. Results of our study confirm other studies findings that a negative result does not necessarily exclude alcohol consumption.

1257

On the distinction between cognitive “g” and academic “g” in adulthood

Ryan J. McGill (1), Gary L. Canivez (2)
 1. William & Mary School of Education, United States of America; 2. Eastern Illinois University, United States of America

Objective: Recent factor analytic investigations of achievement tests have found evidence for an academic “g” factor (ACH-g), similar in form yet, distinct from the general factor of intelligence (COG-g). However, the correlation between those two dimensions approached unity raising concern about the discriminant validity of ACH-g. Additionally, these effects have only been examined in childhood. Accordingly, the present study sought to determine whether ACH-g can be located in adulthood and whether that dimension represents a separate and distinct general factor apart from COG-g with a set of cognitive-achievement measures from the Woodcock-Johnson IV (WJ IV). Method: Participants included 1,146 adults ages 40–90 from the WJ IV standardization sample, which

were reportedly representative of the U.S. population. Mplus version 8.0 was used to conduct CFA using maximum likelihood estimation. Consistent with best practice, rival models were evaluated based on global (CFI, SRMR, AIC, BIC) and local fit. Results: Several models (omnibus oblique factors, dual hierarchical, omnibus hierarchical) were tenable for these data. However, local fit problems were encountered in all but the omnibus hierarchical model. In the dual hierarchical model, the correlation between ACH-g and COG-g was .98, suggesting those dimensions are likely isomorphic. As a result, the omnibus hierarchical model with a single COG-g factor was determined to best explain the data. Conclusion: The present study found that a single hierarchical general factor (COG-g) best explains performance in cognitive-achievement tests throughout adulthood. As a result, ACH-g dimension may not represent a viable psychological dimension in adulthood.

354

Use of psychological assessments feedback

Yoko Sugishima Hida Clinic, Japan

Objective: When clients are asked if they can remember their psychological test results, the answer often given is “No”. This study provides information on how clients can use their test results effectively. Methods 1. Psychological tests (Rorschach Test, Wechsler Adult Intelligence Scale 3rd Edition, Drawing Test, etc.) were conducted to clients who were seeking for job opportunity, for the purpose of self-understanding under doctor’s discretion. 2. Feedback session was conducted using a handout which showed client strength and weakness along with solutions to overcome their weakness. The test results were also reported to medical staff close to clients, with client consent. 3. During the two years after the feedback, five clients were interviewed on how they used information from the test. Results Most clients answered that they had reviewed the report. One client had claimed reasonable accommodation at his workplace as a result

of the handout, two had asked advice from coworkers when unconfident, and another two said that the test results had brought a clear understanding of self-concepts to find a way to deal with their weaknesses. Conclusions Handouts on test results allow clients to review clearly understand self-concepts when feedback matches their existing own preconceptions. For more effective use of the test results, it is essential they be shared and discussed with client and clinical staff who meet them daily.

221

Receiver operating characteristics of the L–H social anxiety scale (LH-SAS)

Duane A. Lundervold (1), Lauren S. Hunt (2)
 1. Program in Psychological Science, United States of America; 2. University of Missouri-Kansas City, United States of America

Essential tremor (ET), the most common movement disorder, is characterized by a rhythmic oscillation of the hand or head (no-no; yes-yes). The neuropathogenesis of ET is complex such that ET has been described as a family of diseases. ET results in significant motor impairment and psychosocial disability. Psychiatric morbidity, especially social anxiety, is common and ranges from 30%–33% in community samples. Unfortunately, social anxiety remains largely undetected among ET patients, in part due to a failure to assess and use instruments that are sensitive in detecting social anxiety related to ET. 73 adults completed measures of social anxiety including the LH-Social Anxiety Scale (LH-SAS), an idiographic-based behavioral assessment instrument. The LH-SAS is comprised of items patients with ET report eliciting fear, distorted cognition, escape and avoidance behavior. Statistical analysis of the stability, reliability, and receiver operator characteristics (ROC) are excellent. Based on ROC analysis, a cutting score of 38 resulted in a highly sensitive (91%) and specific (26%) measure of social anxiety among patients with ET. Further research and clinical use of the LH-SAS needed.