

Testing International

Volume 41, June 2019

Editor: Nicky Hayes



International Test Commission

PRESIDENT

Kurt Geisinger

Buros Center for Testing / University of Nebraska, USA

PRESIDENT-ELECT

Aletta Odendaal

Stellenbosch University, South Africa

SECRETARY-GENERAL

Paula Elosua, Universidad del Pais Vasco, San Sebastian, Spain

TREASURER

April Zenisky, University of Massachusetts Amherst, USA

COUNCIL MEMBERS

Elected Members

Anna Brown, University of Kent, United Kingdom

Neal Schmitt, Michigan State University, USA

Stephen G. Sireci, University of Massachusetts Amherst, USA

Solange Wechsler, Pontificia Universidad Catolica de Campinas, Sao Paulo, Brazil

Co-Opted Members

Dragos Iliescu, SNSPA Bucharest, Romania

Kadriye Ercikan, Educational Testing Service, USA

Peter Macqueen, Compass Consulting, Australia

Wayne Camra ACT Inc.

Observers

Samuel Greiff, Université du Luxembourg, Luxembourg

Gonggy Yan, Beijing Normal University

REPRESENTATIVES

IUPsyS Representative Ann Watts, IUPsyS

IAAP Representative Jacques Gregoire, Université Catholique de Louvain, Belgium

EDITORS

International Journal of Testing

Stephen Stark, University of South Florida, USA

Testing International

Nicky Hayes, United Kingdom

Testing International
is a publication of the
International Test Commission

Contents

ITC President's Message	2
Editor's Message	4
A tribute: Fons van de Vijver	5
Conference News	7
Testing News	9
Food for Thought By Norman Buckley	11
Publication and Communications Committee	13
International Journal of Testing update	14
Membership, Involvement & Marketing Committee	15
Book Review:	17
Around the World: Evaluation & Assessment in Mexico	18

The 12th Conference of the International Test Commission

Save the date

Diversity and equity in a globalized digital world:
Opportunities and challenges for assessment

Call for Papers and Symposia
15 July 2019

Deadline for Workshop Submissions
31 October 2019

Deadline for Paper and Poster Submissions
15 December 2019

Deadline for Symposia Submissions
15 December 2019

Registration starts
1 February 2020

For news updates, visit www.itc-conference.com.
For questions, contact us at itc@uni.lu

uni.lu
UNIVERSITÉ DU LUXEMBOURG

Faculty of Language and Literature, Humanities, Arts and Education

ITC President's Message



Kurt Geisinger
President,
International Test
Commission

Kurt Geisinger
Buros Centre for Testing /
University of Nebraska

I am writing shortly after members of the ITC Council met for its annual meeting in Vilnius, Lithuania. All reports are that the meeting was extremely successful. Let me share why we met in Lithuania; it will provide some information to you about the current running of the association. My graduate assistant (from Belize) at the University of Nebraska-Lincoln and I computed some estimated airline expenses for the entire Council and determined that the lowest sum of airfares that we could achieve was to meet in northern Europe. Then I surveyed some individual psychologists whom I know in northern Europe to determine an interesting place that we could visit that would also keep our prices low. I was told that one of the most cost-effective places would be in the Baltic States, perhaps Lithuania. The Lithuanian Psychological Association has long been a member of ITC, a good reason for us to consider Lithuania. I contacted the Lithuanian/Vilnius travel office, got a listing of hotels that would meet our needs and selected one in the Old City. My colleagues tell me it was a beautiful city and they especially appreciated a traditional Lithuanian dinner one night.

Now let me tell you about the leadership teamwork for the ITC organization. In preparation for the Vilnius meeting, it was truly heartwarming to see the work of our Secretary-

General Paula Elosua, our President-elect Aletta Odendaal, our inveterate office manager and general-jane-of-all trades Ananda van Tonder, and our new treasurer, April Zenisky. I would like to think that I helped too, but it was just fascinating to see highly focused people working with so much teamwork, so inter-dependently, and effectively. They did and continue to do so for the best of the organization.

Then, my crisis hit. The day before I was to leave for Europe I was in a terrible automobile accident and transported to the hospital in an ambulance. I fought with the doctors to let me continue to the Vilnius meeting until they told me I had 8 broken ribs and a broken sternum and that flying to Europe would likely lead to collapsed lungs. At that point, I contacted Aletta and let her know as President-elect she would need to be running the meeting. And it should surprise no one that she did outstandingly and that the ITC Council meeting was effective, who knows, perhaps more effective without me! Thanks to all involved!

The Council learned that the 2020 meeting in Luxembourg was moving forward wonderfully; they have set up the Scientific Committee, and made many great decisions regarding the future success of the conference. We have primarily Samuel Greiff and Lindie van der Westhuizen to thank for this planning. The Council also looked at some possibilities for future meetings, looked into options related to the local incorporation of the International Test Commission, ways to encourage sponsors to support our conferences, reviewed the ITC budget, and so on.

I have not yet seen the minutes of the meeting, but by all accounts it was effective. I also have an apology to make based upon my previous presidential column. I mentioned in that column that Professors David Bartram and Barbara Byrne received ITC fellow status at the 2018 Montreal conference and would be giving keynote presentations at the 2020 Luxembourg

conference. A corrected set of facts includes that both of them received fellow status in ITC at the wonderful 2012 ITC conference in Amsterdam, that they independently received the Tom Oakland Award at the 2018 conference, and that they would be each be providing Tom Oakland Lectures at the 2020 Luxembourg conference. That may be reason enough to plan your attendance at the Luxembourg conference! Please accept my apology, Dave and Barbara! We all look forward to your talks in Luxembourg!

I have one last idea. ITC as an organization has largely grown and spread by word of mouth and by professionals telling others about our super professional, and friendly meetings. We also have a well-developed mentorship program. Please share information about our upcoming meeting in Luxembourg with your graduate students and junior colleagues. I think that ITC is an especially valuable conference for such junior colleagues and colleagues from countries that do not yet have well-developed assessment systems. I have been to the setting at the University of Luxembourg personally and it is an exciting place to visit. The ITC conference is a great place to share one's research, to network, and to develop international partnerships and associations. We all know that the world is shrinking and connections around the world abounding. Let ITC be a catalyst to testing research and practice, and by implication, world peace. If politicians cannot bring about world peace, maybe we can! Please have a great summer!

Kurt Geisinger

Afterword

I am now doing much better. After being in what was essentially a head-on crash on April 23, people at the scene told me I had to go the hospital. I walked to the ambulance, climbed in, and walked into the hospital on my own, at which time they made me ride in a wheelchair. It turned out that I had six broken ribs on my right side, two on the left, as well as a broken

sternum and three breaks to my pelvis. They believe that all the breaks were due to my wearing a seatbelt. I had been telling the doctors in the emergency room that I needed to fly to Europe the next day for the ITC Council meeting. After a CT scan they told me I could not fly; having the breaks to my ribs, they believed that in an airplane my lungs would likely collapse and I would die. Fairly convincing. I ultimately spent 5 days in intensive care, then 6 days in rehabilitation in the hospital. Initially they were mostly concerned that I had internal bleeding near my heart due to the broken sternum. After three days that quasi-crisis passed without surgery. When I came home, I stayed home for almost two weeks. I have now been back at work for over two weeks and am making progress, walking a bit more with a cane each day. I expect to be back to 100% sometime in July. Moral of the story: wear your seatbelt every time. We cannot afford to lose ITC members involuntarily! I consider myself very lucky. I would probably have died without it.

The International Journal of Testing

The *International Journal of Testing (IJT)* is dedicated to the advancement of theory, research, and practice in the areas of testing and assessment in psychology, education, counseling, organizational behavior, human resource management, and related disciplines.

IJT publishes original articles addressing theoretical issues, methodological approaches, and empirical research, as well as integrative and interdisciplinary reviews of testing-related topics and reports of current testing practices. All papers are peer-reviewed and are of interest to an international audience. The journal is published quarterly, with each issue containing 4-5 articles.

If you have a paper you would like to be considered for publication in IJT, or if you are interested in reviewing papers for publication, please contact the editor, Stephen Stark, by email: (sestark@usf.edu).

Greetings from the Editor



Welcome to the latest issue of Testing International. We begin on a sad note, with a tribute to Fons van de Vijver and the sterling contributions he has made to psychological testing in general, and to the ITC in particular. Fons' untimely death while exercising came as a shock to all who knew him, and he will be sadly missed.

On a less sombre note, our conference organisers Lindie von der Westhuizen and Professor Samuel Greiff have provided us with a preview of next year's conference programme, and I'm sure you, like me, will be looking forward to visiting the Grand Duchy of Luxembourg in July 2020. Other Testing News includes a report on an English translation of the COTAN review system and an update on how the revision of the Guidelines for Technology-Based Assessment is progressing.

We follow that with some food for thought, as Norman Buckley deconstructs a significant testing shibboleth and challenges us all to re-think some cherished beliefs. You may have your own opinions about some issue in testing. If so, Testing International would love to hear from you! Send your contribution (no matter how small) to me at newsletter@intestcom.org.

Neal Schmitt tells us about new developments with ITC publications while Stephen Stark gives us the latest news from the International Journal of Testing and Peter Macqueen updates us on membership and involvement issues. Peter has also included his ever-helpful summary of forthcoming conferences

I find that list really useful - it's always good to know what's going on around the world, and Peter goes to a lot of trouble to compile it. So if you're aware of a forthcoming conference or congress with any relevance to testing coming up in your region, do let him know about it! You can contact him at: p.macqueen@compassconsulting.com.au

Following Peter's contribution we have a book review, and then an interesting overview of evaluation and assessment in Higher Education in Mexico. Our thanks to Agustin Tristan-Lopez for this contribution, which I hope it will be the first of many: country overviews like this are definitely helpful for anyone working at an international level, and equally definitely interesting for the rest of us! So if you feel like giving us a picture of testing in your country, please do it!

So there it is: this summer's Testing International. I hope you enjoy this issue - but remember, it's only our members who make it interesting! You can send your contribution, or anything else you'd like to see included, to me at newsletter@intestcom.org.

Happy Reading!

Nicky Hayes
Editor, Testing International

Call for Articles and Announcements: *Testing International (TI)*

Deadline for the December 2019 issue:
November 11, 2019

TI is the newsletter of the International Test Commission, and disseminates information about national / international assessment projects and initiatives, new test developments, recently published books / articles, upcoming conferences and workshops, and topical issues in the field of testing and assessment to the international community.

Please contact Dr. Nicky Hayes with your ideas, proposals, announcements, and brief papers

newsletter@intestcom.org

Fons van de Vijver, PhD

1952-2019



It was with a great deal of sadness that members of the ITC Council were recently informed of the very untimely and sudden death of Professor Fons van de Vijver, one of the longstanding leaders of the International Test Commission (ITC), in addition to a host of other professional organizations.

Given the multiplicity of important contributions Dr. van de Vijver made over the past 25 years to the ITC, we believe that all ITC members will very much enjoy knowing his overall academic credentials and accomplishments in general, and his cutting edge contributions to the ITC, in particular. We begin with a summary of his academic credentials and follow with his more specific work for the ITC.

Alphonsius (Fons) Josephus Rachel was born on October 4, 1952 in Koewacht, The Netherlands and was 66 at the time of his passing. Fons studied psychology at Tilburg University in the Netherlands, where he also earned a Ph.D., and stayed there as a faculty member for the major portion of his career. He held a chair in cross-cultural psychology as well as an extraordinary chair at North-West University, South Africa, and the University of Queensland, Australia. Over the years, Fons held a variety of important leadership positions at Tilburg University, authored or co-authored over 400 publications - mainly in the domain of cross-cultural psychology, served recently (2016-2018) as president of the International Association of Cross-Cultural Psychology, and was a former president of Division 2 (Assessment and Evaluation) of the International Association of Applied Psychology as well as a former

President of the European Association of Psychological Assessment.

The main topics in his research involved bias and equivalence, psychological acculturation and multiculturalism, cognitive similarities and differences in the cognitive domain, response styles, translations and adaptations. His work on analytic approaches and conceptualization of acculturation and multiculturalism has led to the development of an indigenous personality scale relevant to 11 South African languages. At Tilburg University, he supervised approximately 35 PhD studies and 6 post-doctoral students. Among the books he wrote include the following: *Methods and data analysis for cross-cultural research* (both individually and with Kwok Leung), *Cross-cultural survey methods* (with Janet Harkness and others), and *Cross-cultural research methods in psychology* (with David Matsumoto). Indeed, he has been one of the most frequently cited cross-cultural psychologists in Europe, if not worldwide.

Fons has presented keynotes and invited lectures at a vast number of appropriate conferences and offered workshops worldwide, often including of course, the ITC. He was a member of the following professional organizations: International Association for Cross-Cultural Psychology, International Association of Applied Psychology, European Association of Psychological Assessment, International Academy for Intercultural Research, and the ITC. He is the former editor of the *Journal of Cross-Cultural Psychology* and the *European Journal of Personality Assessment*, and has often served as a reviewer and editorial board member on dozens of journals. Of course, he was considered especially appropriate to read papers dealing with the adaptation of tests and measures and methodological issues in so doing.

Finally, Fons was the 2013 recipient of the American Psychological Association's Award for Distinguished Contributions to the International Advancement of Psychology (for his contributions to international cooperation and to the advancement of knowledge of psychology) and the 2014 recipient of the International Association of Applied Psychology Fellows Award (for contributions to applied psychology).

We turn now to Fons' critically important contributions to the ITC, where his first position on Council was as editor of what was originally called the ITC Bulletin, which at that time existed as a separate section in the *European Journal of Psychological Assessment*. Fons held this position from 1994 until 1999 at which time the ITC president, Thomas Oakland, signed a contract with Lawrence Erlbaum Associates for the ITC to publish its own journal entitled the *International Journal of Testing*.

During this time period also, Fons worked closely with Ron Hambleton in developing a detailed set of guidelines for use in the adaption of educational and psychological tests for use in other languages and/or cultural contexts. Unquestionably, Fons' cross-cultural and methodological expertise has been critically needed and thus greatly appreciated for all subsequent updates of these adaptation guidelines. Indeed, the ITC will be forever grateful for the dedicated contributions of Fons van de Vijver over a span of several years in working with Ron Hambleton and other ITC members in their development and updating of the *ITC Guidelines for Translating and Adapting Tests*.

Many of us believe that we will miss his contributions tremendously, but will miss his friendship more.

Kurt F. Geisinger and Barbara M. Byrne



The ITC International Handbook of Testing and Assessment

Editors: Frederick T. L. Leong, Dave Bartram, Fanny M. Cheung, Kurt F. Geisinger & Dragos Iliescu.
Oxford University Press ISBN 9780199356942.

The ITC International Handbook of Testing and Assessment addresses the many challenges facing the cross-cultural applications of psychological and educational globalization. It represents and showcases the concerted efforts of tackling the wide range and complexities in psychological testing; information and help guide professionals and graduate students regarding testing and assessment from an international and global perspective.



Conference News

The 12th Conference of the International Test Commission: Luxembourg 2020

By Lindie van der Westhuizen (Secretary of the Local Organising Committee) & Prof Samuel Greiff (Chair of the Local Organising Committee)

Moien! Bonjour! Guten Tag! Good day!
We are delighted to invite you to the 12th Conference of the International Test Commission which will be held in the Grand Duchy of Luxembourg from the 14th to the 17th of July, 2020.

The 2020 ITC Conference promises to be an exceptional professional and scientific experience in a unique cultural environment renowned for its vibrancy and cosmopolitan flair. Exploring the theme *Diversity and equity in a globalized, digital world: Opportunities and challenges for assessment*, the 12th ITC conference will present the latest developments and innovations in assessment.

The theme highlights cultural diversity resulting from globalisation and the rapid advancements in the field of technology-based assessment. It is further divided into five sub-themes: (1) *Test development and validation in international and multicultural environments*; (2) *Innovations and advances in psychometric theory, modelling, and technologies*; (3) *Current issues of policy, diversity, and equity*; (4) *Best practices in testing and assessment*; and (5) *Test security and privacy in a globalized digital world*.

An impressive line-up of keynote speakers has been secured that is guaranteed to excite and engage your mind:

Prof Sacha Epskamp, one of the pioneers in network modelling, will present a state-of-the-art lecture on the *Current State and Future of Network Psychometrics*.

Prof Jonathan Templin, a renowned leader in the field of diagnostic classification models, will present a state-of-the-art lecture on *Building a Diagnostic Model-Based Formative Assessment System for Personalizing Educational Progress*.

Prof Anita Hubley, recognized internationally for her expertise in test development, validity, and

psychological and health assessment, will share her expertise on best practices in testing and assessment with her talk on *Contributions of Response Processes to Test Validation and Development*.

Given the impact of recent data protection and security developments on testing, **Prof John Fremer** will discuss *Challenges Confronted and Lessons Learned: Protecting Test Content and Personal Information from Test Security Threats in International Testing Programs*.

Dr Sara Ruto's talk on *Measuring Learning for All Children: The Citizen Led Assessment Approach* will address the increasing importance of and challenges associated with inclusivity in educational testing in the developing world.

Playing devil's advocate, **Prof John O'Gorman**, will consider the implications of neuroscience for industrial organizational psychology in his talk *The Future for Organisational Neuroscience in Selection and Assessment, or Is There One?*

Representing the growing field of language testing, **Prof Lianzhen He** will use China's Standards of English framework as a case study and will share lessons learned for language learning, teaching, and assessment.

Lastly, **Prof Aletta Odendaal**, Industrial Psychologist and ITC President-Elect, will present a developing country perspective on challenges and critical issues in response styles and test use.

The stimulating scientific programme will be complemented by a packed social programme, which includes the opening ceremony, the gala dinner, sightseeing tours, and new scientific speed networking and speed mentoring sessions. We have also revitalised the poster session format to increase presenter and audience engagement. Be sure to keep an eye out for the second conference announcement and the call for papers that will be released on the 15th of July 2019. For more information and updates on the conference, you can visit the conference website at <https://www.itc-conference.com>.

We look forward to seeing you in Luxembourg in July 2020!

The 12th Conference of the International Test Commission

Diversity and equity in a globalized digital world:
Opportunities and challenges for assessment

Save the date

Call for Papers and Symposia

15 July 2019

Deadline for Workshop Submissions

31 October 2019

Deadline for Paper and Poster Submissions

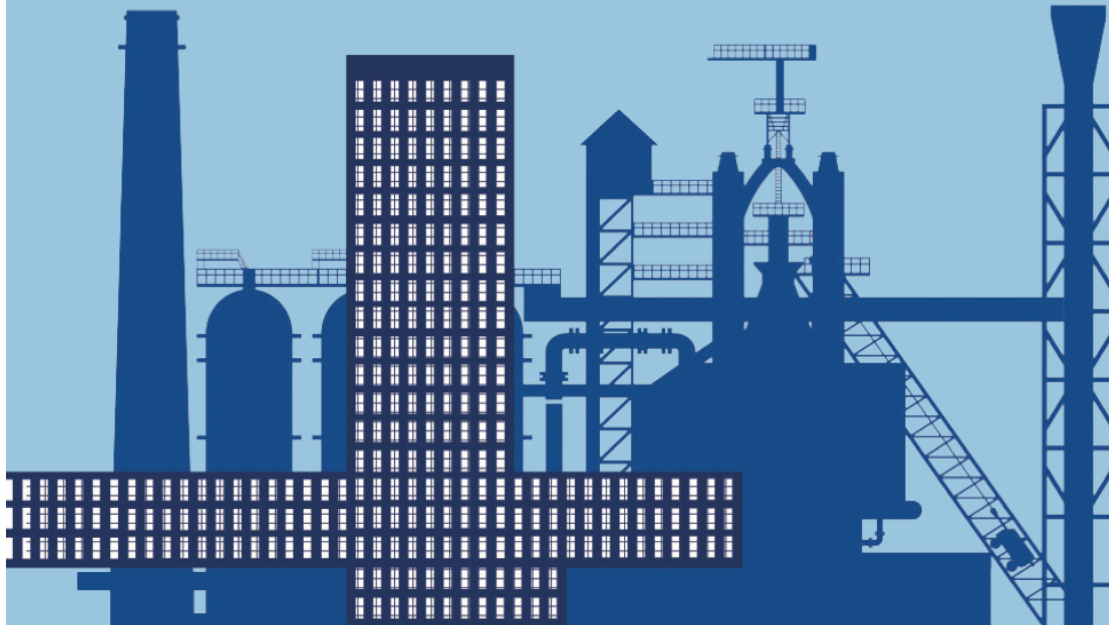
15 December 2019

Deadline for Symposia Submissions

15 December 2019

Registration starts

1 February 2020



UNIVERSITÉ DU
LUXEMBOURG



Faculty of Language
and Literature, Humanities,
Arts and Education

For news updates,
visit www.itc-conference.com.

For questions, contact us at
itc@uni.lu

Testing News

English version of the COTAN review system

The COTAN review system for evaluating test quality (2010) provides criteria for rating psychological and educational tests, scales and questionnaires used in the fields of work and organisational psychology, education, clinical psychology, developmental psychology and other contexts. Test reviews are published by the COTAN (only in Dutch via www.cotandocumentatie.nl) in order to improve tests and testing in the Netherlands. These reviews can be used by psychologists and other professionals as an aid in making a responsible choice for a test.

Recently, the Dutch Association of Psychologists (NIP) published the English translation of the COTAN review system for evaluating test quality (2010). This review system is not only used by the reviewers of the Dutch Committee on Tests and Testing (COTAN) and Dutch test publishers, but is also frequently consulted at Dutch universities in the context of psychology education. To illustrate: several Psychology courses in the Netherlands require students to systematically assess the quality of a test and its manual based on the COTAN review system for evaluating tests. However, in the Netherlands more and more psychology courses are taught in English, with Dutch-language students increasingly enrolling in the English-language specializations of these programs.

In addition, foreign researchers and psychologists increasingly asked the COTAN for an English translation of its test review system. Until now only a Dutch version of the COTAN review system has been available. The COTAN and the NIP therefore considered it important to publish an English translation of the COTAN test review system.

The pdf document can be found on this page:

<https://www.psynip.nl/en/dutch-association-psychologists/about-nip/psychological-testing-cotan/>

Karin Vermeulen
Stafmedewerker COTAN



More Testing News

Update on the Guidelines for Technology-Based Assessment

The International Test Commission has partnered with the Association of Test Publishers (ATP) to jointly develop Guidelines for Technology-Based Assessment.

The former ITC guidelines on computer-based and internet-delivered testing were released in 2005. ATP's Guidelines for Computer-Based Testing predate that document by three years. The purpose of the new guidelines is to provide information about the key factors and issues that should be considered when designing and delivering tests using digital platforms, and to provide guidance to test developers, test administrators, and test users, on how to best ensure fair and valid assessment in a digital environment.



The development of these joint guidelines is being led by Stephen Sireci for the ITC and John Weiner for ATP. They serve as the Co-Chairs of the Steering Committee, which is populated with members as follows:
ITC: Kadriye Ercikan (USA); Dragos Iliescu (Romania); April Zenisky (USA)
ATP: Alex Tong (China); Alina von Davier (USA); Linda Waters (USA)

Over the past year, the Steering Committee has held monthly conference calls, during which the focus of the committee's work has primarily focused on two areas. One critical task for the committee was to conceptualize the collaborative process to be implemented to produce these Guidelines with respect to the role of the Steering Committee as well as various contributors.

Intentionally so, the process was designed to fully incorporate many individuals who were purposefully identified and recruited to ensure that a diversity of perspectives are represented at multiple stages in the development of these Guidelines, to ensure that the final product would be representative of global technology-based testing applications.

The guidelines themselves are to be developed by individual content authors in coordination with several advisory groups that are being formed to reflect a) practice areas (Education, Certification/Licensure, Workplace, Clinical, Security, Technology) and b) regions (Europe, Asia, India, Middle East, Australia, Africa, South America). As guidelines and related content is written, it will be reviewed by the advisory groups, as well as ad hoc and legal reviewers, and once the full document is drafted, it will be opened for a period of public comment.

The second, but equally critical task, has been to outline the structure of the new Guidelines and to develop introductory text reflecting the purpose, rationale, scope, and intended audiences. The Steering Committee members discussed the positive aspects as well as limitations of the prior sets of Guidelines to reflect current and emerging practices in the aim of promoting best practices in test development, administration, and scoring to facilitate fair and valid measurement of the psychological and educational characteristics targeted by contemporary assessments.

Recently, the Co-Chairs presented a session at the 2019 ATP meeting in February, entitled "Technology-Based Assessment Guidelines: Present and Future". In addition, a meeting of the steering committee and several other practice area and regional leads was held with those present at ATP (with others calling in) to touch base on the process and discuss the work completed to date.

Submitted by April L. Zenisky, on behalf of the Steering Committee.



Food for Thought



Norman Buckley gives us his views on a widely held article of faith.

Deconstructing the shibboleth: Feedback

Shibboleth: A custom, principle, or belief distinguishing a particular class or group of people, especially a long-standing one regarded as outmoded or no longer important.

My first job as a baby psychologist was with a firm of "Industrial" Psychologists in Australia. We used the term Industrial rather than Organisational back then. Yes, I'm that old. The job consisted of taking a set of cognitive, attitudinal, preference and personality results and writing a report that pulled them all together. If you got it right, it produced a picture of the person. If you got it wrong, it was still a good picture – just not of that person. I'd like to say it was interesting and I suppose I can. I wrote over 3000 of these in three years.

But writing wasn't the issue. It was standard practice that each person received detailed feedback on the results. Not a full coaching session since the feedback was rather one way but it meant that the person knew everything in the report and where it came from. No secrets. These were the days when test results were really nothing more than a set of numbers. You needed to be trained (accredited) to interpret them and that was the role of the psychologist. Hard to believe as it is, computers were around but not very friendly. But we knew that they would make a difference. We were certain that reports could be computer generated. We were also convinced that they'd never replace the psychologist. When the first workable computer-generated reports emerged, it looked as if we were right. They were OK but we could always do better. Ian Fraser when he was at what was then SHL commented that "a computer generated OPQ report is never going to be as good as that produced by a good, trained psychologist. But," he said, "it was way better than the average report that was being written in the real world!"

become more situation specific and user friendly. But it seems that we still believe in the importance of feedback. And that it should be given by trained people. But feedback has two elements. First there are the results of whatever test, profile or inventory was used. The client needs to know this so that it is less of a black box. A respondent can look at a construct like Extraversion (Energy in Facet5) and see that, compared to other Danes, Chinese or Argentines, the score is quite high. And with accurate understanding of what an extravert is like, he/she should be able to say "Yep, sounds about right". This is the first element of feedback. Agreement as to the output.

But there's the next stage. My son's English tutor was charged with teaching him how to write essays. Not his strong point when he was 15. Her advice was to imagine a parrot on his shoulder and whenever he wrote something, his parrot would say "so what?". So when the budding essayist wrote a few struggling words he would hear "so what?" and think about the implication of what he was saying. Same for feedback. Stage two of the process is "so what?" So I'm extraverted. So what?

Feedback without the "so what" is little more than an ego trip for the practitioner (look what I found out) or a bit of fun for the client. Counsellors have figured this out. They focus on the "so what". In fact, given the enthusiasm of counsellors for tools like the MBTI, you could argue that they're not that interested in stage 1 – the accuracy of the results. Many years ago I knew a British psychologist who reckoned he got great information from a Sentence Completion Test. I suspect he was using it to open a "so what" discussion.

Now automated reports have got much better. They provide really good, often graphical results so the person can see where they sit in the psychometric universe. And frequently reports are contextualized to open the "so what" discussions. What does this mean for me in conflict? What does it say about my interaction in a team? How do I compare to Jean over there? Where are we similar? Where different? What does it mean? So I think we can have some confidence that current generation reporting can provide good feedback of content and increasingly useful "so whats".

But what about the feedback process? Does it have to be given by a psychologist? Do we need to be trained?

Tool specific vs level specific training has been a debate within the BPS. And what about the method? Is the medium the message.

Roy Childs at TeamFocus asked this question in a short paper titled *"To give or not to give"*. He had a sample of people who had received feedback in different ways. Some face-to-face. Some by video. Which was better? The short answer he got was it didn't seem to matter much. A bit embarrassing for the highly skilled, empathic practitioners but the delivery vehicles were broadly equally effective in the view of the respondents.

But most practitioners I know would prefer feedback to be face-to-face and that can be difficult. What about oil rig workers? Or a factory manager in Bangladesh? I once gave a 90 minute telephone feedback to the head of one of the US security services. It's one with three letters but I can't tell you which one. They're listening, you know. I was dreading it, (had my tinfoil hat on) but it went very well (I think). It helped that he was very, very smart and deducted, inducted and inferred at will.

I was brought up to believe that we should always offer feedback to the respondent and I have never heard anybody counter this. Some have said they can't afford to or don't really care enough but they agree somebody should give feedback. Ideally this is face-to-face but telephone or video feedback is OK. But don't just dump the report on them. My colleague Rebekah Williams reckons that if you do that, when you do sit with them or phone them, you end up spending half the time unwinding the misunderstandings from their initial reading of the report. So my thought, and I think it's quite common, is to present the report during the feedback. If it's a paper report, you have control of the page. If it's video or telephone, you can do it page by page where again you have control.

But hang on a minute. There are two other questions. First, who should give the feedback. For many years this was limited to qualified psychologists. Largely due to the work of UK commercial test publishers this has been overturned in many countries. It's now a competency-based model. But we need to be aware of a resurgence of this idea. In Brazil, Italy and some other places there is pressure for only PhD psychologists to be allowed to feedback psychometric test data. Anders Wahlberg – chairman of the Swedish Psychologists Association recently told me there is occasional pressure there. But it's not going

to happen. That horse has bolted. If official bodies try to impose, business will just find a way around it. I have seen psychometrics reclassified as "behavioural profiles". Nothing "psychometric" at all. Nothing to see here. Having said this, the imprimatur of the BPS and Norsk Veritas is still highly valued especially in Scandinavia.

Second relates to management of the data. Some countries have an opinion on how data is managed, and it has a significant impact. The NIP (*Nederlands Instituut van Psychologen*), for example, says the respondent has the right to see the results of any test first and it is up to them as to whether anybody else sees it. This is very different. We have clients who state contractually that they "own" the data and can decide what to do. But they probably don't. As far as I know this hasn't actually been challenged but under the slow awakening of companies and individuals to the impact of the GDPR (General Data Protection Regulation), it may require a change of thinking. The rights of "data subjects" (in GDPR lingo) are very tightly defined including who has it, who uses it, what it is used for and where and for how long it is stored. GDPR does not differentiate between character traits, education, political opinions or medical history. It's all personal data. Same rules apply.

We already have clients requesting that test results should be sent automatically first to the respondent who then decides whether they would like personal feedback.

So with regard to feedback it may be moving from "good practice" to law. And that will put pressure back onto test publishers to make sure that their outputs are designed to be seen first by the client and only then by somebody else of the respondent's choosing.

So what started as a clear agreement to the need for qualified people to present feedback in a controlled and precise manner via the channel of their choosing may have to change. It may no longer matter what they think. The world has moved on. That shibboleth has been deconstructed.

Note: The General Data Protection Regulation 2016/679 is a regulation in EU law on data protection and privacy for all individuals within the European Union and the European Economic Area. It also addresses the export of personal data outside the EU and EEA areas.

ITC Committee Reports

Publication and Communications Committee Report *Neal Schmitt, Chair*



The Publication Committee's activities include the International Journal of Testing, Testing International, and the book series. The first three volumes in the ITC book series published by Cambridge University are now available. The first book authored by Dragos Iliescu is titled "Adapting tests in linguistic and cultural contexts". John Scott, Dave Bartram and Douglas Reynold's book titled "Next generation technology: Enhanced assessment: Challenges and Opportunities" is also available.

The volume by William Schmidt and his colleagues titled "Measuring opportunity: Insights from international large-scale assessment" was just published.

Craig Wells continues working on "Assessing measurement invariance for applied researchers"

Maria Elena Oliveri and Cathy Wendler are editing a volume titled "Higher education admission practices: An international perspective" which will soon be available.

We have also contracted with Sumaya Laher to do a volume on the history of psychology. She is making excellent progress and has identified section editors and many chapter authors.

The contract to do the series ended on March 1, 2019 and we have signed another three-year contract to do five books. We look forward to continuing working with Cambridge as we develop a series of books that discuss the role and impact of assessment in a global context.

Continued viability of the series does depend on getting new authors. The following list includes some books I thought might be good additions to the series (some have been discussed before). I would like some reaction to this list and suggestions regarding potential authors. Additional topics that you think

would be a good match for the series are also welcome.

- **Security issues worldwide.**
- **Personality testing** (I think a volume on etic and emic approaches to personality testing is clearly within the purview of the series and should be a good seller as well. I have asked several people to do a volume on personality and no one has agreed to take it on for various reasons.
- **Principles of Testing and Assessment.** I think we could use an introductory textbook on testing for undergraduates or those in need of an introduction to psychometrics.
- **Licensing and certification tests.** This could be a compendium of how tests are used by regulatory authorities in different countries to certify one is competent in a profession. I have talked with several potential authors with interest but unable to do the work at this time.
- **Global issues in diversity and testing.** In at least in the US, Europe, and Australia, a long standing concern has been the fair and unbiased use of tests with underrepresented minorities. The groups about which there is concern, the manner in which the testing community addresses these concerns, and the manner in which governments deal with these issues varies widely across the globe. A book that details these concerns and how they are addressed would be useful to anyone working in an international context.

The Publications Committee has also considered a proposal from Larry Rudner that ITC consider publishing Practical Assessment and Research Evaluation which Larry has developed and edited since its inception over the past two decades. The journal has an excellent reputation and appears to have had a significant impact on assessment practice. The ITC Council discussed this proposal at length but decided to decline for two reasons. One was financial; we have very limited funds with which to support another journal. The second was our reluctance to take on the time commitment that would be necessary to continue this journal's success. We wish Larry the very best as he seeks a partner to assume responsibility for this publication.



International Journal of Testing

News and Updates

The *International Journal of Testing (IJT)* is dedicated to the advancement of theory, research, and practice in the areas of testing and assessment in psychology, education, counseling, organizational behavior, human resource management, and related disciplines. IJT publishes original articles addressing theoretical issues, methodological approaches, and empirical research, as well as integrative and interdisciplinary reviews of testing-related topics and reports of current testing practices. IJT is published quarterly with each issue containing 4-5 articles.

We are delighted to announce the **impending publication of our special issue** on **“Challenges and Opportunities in the Design of Next Generation Assessments of 21st Century Skills,”** guest edited by **Drs. Robert Mislevy and Maria Elena Oliveri**. We were fortunate to receive manuscripts from highly distinguished authors in psychometrics and technology-based testing, so be on the lookout for its release very soon!

This summer, we will also issue a **call for papers** broadly related to **gamified assessment and new developments in scoring outside the realm of traditional psychometrics** (e.g., machine learning, artificial intelligence). This is a rapidly growing area of interest in educational and workplace contexts, so we plan to highlight some of the technology and draw attention to the need for research on reliability and validity per the APA/AERA/NCME Testing Standards and ITC Guidelines. We will issue this call by early July with approximately 90 days to submit a paper, so if you are interested in doing so, please periodically check for journal news on the IJT and ITC websites, as well as on Linked In.

In addition to this special issue, we remain interested in publishing some “target articles” and commentaries on various topics. The idea is to submit a paper on an issue of broad interest, which has empirical underpinnings, but takes a position that is likely to generate pro and con responses. If you have a topic idea or a paper that you feel might be suitable, please contact me by email (sestark@usf.edu).

Finally, we continue to look for new reviewers and streamline our manuscript review process. Please contact me if you are willing to review for IJT, or if you would like to nominate someone (including yourself) for our Board of Consulting Editors. To be selected for the editorial board, a nominee must have an accomplished record of publication and expertise in measurement/testing methodology and practice. To expand the journal’s outreach and international impacts, we encourage nominations of individuals from underrepresented groups, cultures, and regions beyond North America.

Stephen Stark

University of South Florida

Editor of International Journal of Testing





**Membership,
Involvement and
Marketing (MIM)
Committee**
Peter Macqueen, Chair

The recent ITC Council Meeting in Vilnius, Lithuania was a valuable opportunity to discuss a range of important matters, including those relevant to the MIM Committee.

1. Membership and Fee Payment:

Thank you very much to our members (Full, Affiliate, Individual) who have paid their 2019 dues, and perhaps prior dues. Several of you have responded to our call for such, and we appreciate it. This is particularly true as we lead into our next conference in Luxembourg (July, 2020) and the need to draw upon our funds to ensure the success in staging this event.

I would like to thank Ananda, Committee Members, and Council Members for their assistance.

We still need to follow up on a few organisations, and Ananda and I will provide a friendly reminder of outstanding dues over the next few months.

2. MIM Committee Membership:

We are in the process of refreshing membership of this Committee. In particular, we are keen to add a person with **social media skills**. There is no doubting that professional bodies, increasingly, need to address this important channel in attempts to engage with members, and potential members.

If you know of any ITC member (or potential member) who would like to assist the ITC from this perspective, please contact me (at p.macqueen@compassconsulting.com.au) and the ITC Secretary-General, Paula Elosua at secretary@intestcom.org. Please note that this is a voluntary role.

3. Upcoming Testing Events:

As part of our activity in fostering or supporting relevant professional events around the globe, it would be great for the ITC to be more aware of such so that these may be listed (if not promoted) on the ITC website, or within Testing International, this newsletter.

Please send your suggestions to:

secretary@intestcom.org

Ananda van Tonder (Office Manager) or Paula Elosua (Secretary-General) can then direct your email for action.

Here are some of the offerings of which we are aware:

AFRICA

The African Journal of Psychological Assessment (AJOPA) has now launched. Published via African Online Scientific Information Systems (Pty) Ltd (AOSIS), this online journal is edited by Associate Professor Sumaya Laher.

<https://ajopa.org>

(Articles initially online then printed in an end-of-year publication.)

Association for Educational Assessment in Africa (AEAA)

5 – 9 August 2019, Abuja, Nigeria

<https://www.aeafrica.org/programme>

ASIA

22nd Japan Language Testing Association (JLTA) Conference

11 – 12 September 2019, Niigata, Japan

www.jlta2016.sakura.ne.jp

5th Annual India-ATP (I-ATP) Conference

22 November 2019, New Delhi, India

(See the main ATP website)

EUROPE

15th European Conference on Psychological Assessment (EAPA)

7 – 10 July 2019, Brussels, Belgium

<https://ecpa15.com>

European Association of Test Publishers (E-ATP) Conference

25 – 27 September 2019, Madrid Spain

<http://www.eatpconference.eu.com/>

61st International Military Testing Association (IMTA) Conference

7 – 11 October 2019, Tallinn, Estonia

http://www.imta.info/conference/conference_home.aspx

*** **12th ITC Conference** ***

14 – 17 July 2020, Luxembourg
<https://www.itc-conference.com>

See the report on page 7 of this newsletter

NORTH AMERICA

International Association of Computerized Adaptive Testing (IACAT)

10 – 13 June 2019, Minneapolis, MN

<http://www.iacat.org/2019-iacat-conference-minneapolis-usa>

National Conference on Student Assessment (NCSA)

24 – 26 June 2019, Orlando, FL

<https://ccsso.org/events/national-conference-student-assessment-ncsa>

Conference on Test Security

16 – 18 October 2019, Miami, FL

<https://conferenceontestsecurity.org/>

National Council on Measurement in Education (NCME) Annual Meeting

16 – 20 April 2020, San Francisco, CA

<https://www.ncme.org/home>

Association of Test Publishers (ATP) Conference

29 March – 1 April 2020, San Diego, CA

<http://www.innovationsintesting.org/atp2018/future-conferences.aspx>

OCEANIA

There are no test specific events scheduled although broader conferences will offer sessions related to testing and assessment.

Australian Psychological Society (APS): IOP Conference

July 2019, Adelaide, Australia

[I will be Chairing a Professional Practice Forum: Innovations and Developments in Psychological Testing and Assessment]

New Zealand Psychological Society (NZPS)

August 2019, Rotorua, New Zealand

SOUTH AMERICA

9th Congress of The Brazilian Institute of Psychological Evaluation (IBAP)

25 – 28 June 2019, Salvador, Brazil

www.ibapnet.org.br/congresso2019 (In Portuguese)

37th Congress of Interamerican Society of Psychology (SIP)

15 – 19 July 2019, La Habana, Cuba

<http://www.cipcuba2019.com/2019July> (Spanish and English websites)

[38th Congress: 2021 in Paraguay]

NB: Some of the events I have listed may have passed prior to publication of this issue of TI. However, we suggest you take note of the conference websites for possible updates on further conferences in 2020 or 2021.

Peter Macqueen

p.macqueen@compassconsulting.com.au

Brisbane, Australia



Book Review

Adapting Tests in Linguistic and Cultural Situations

Dragos Iliescu Cambridge University Press 2018
ISBN 9781107110120

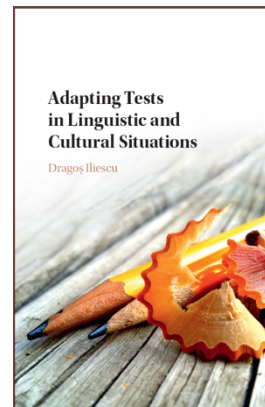
As part of the ITC series of books, Cambridge University Press published "Adapting Tests in Linguistic and Cultural Situations" by Dragos Iliescu, the ITC Past-President, in 2018.

It is challenging to write a book about test adaptation, and that for at least three reasons. First, test adaptation does not follow a clearly prescribed process. The possible variations in the process are very large, and depend not only on the focal test, but also on a number of contextual variables (e.g., the source and target culture, the objectives of the researcher, etc.). Ultimately, they depend on the choices made by the researcher during every phase in the adaptation process. Second, test adaptation has no clearly prescribed outcomes. Although a large number of projects are identified as "test adaptations", the differences between them are sometime larger than the similarities. Third, test adaptation literature is not a clear body of knowledge. The literature is very much split between reports of cross-cultural research (both small-scale and large-scale), reports of actual test adaptations, test manuals, technical details that are usually related to some form or another of statistical procedure, and a great body of advice literature.

Despite the difficulties intrinsic in this topic, the author has risen to the challenge and has provided an important contribution to this area. The book, the first of its kind, provides an in-depth treatise and guide on why and how to adapt a test to a new culture in such a way as to preserve its psychometric value. Over more than 700 pages, the author explores test adaptation, test localization, and test indigenization, an important scientific and professional activity that is now ubiquitous in the social and behavioural sciences as part of work in today's globalized world.

Adapting Tests in Linguistic and Cultural Situations discusses the principles and historical roots of test

adaptation. It acknowledges the interdisciplinary vocation of this activity, stating that it is unclear whether it "belongs" - at least from a historical point of view - more to any given science than to another. At the same time, it underlines that psychology, and specifically cross-cultural psychology, has probably the strongest claim on the domain of test adaptation, with other sciences and branches also offering strong contributions, such as educational sciences (e.g. educational testing), sociology (e.g. survey translation), medicine (e.g. epidemiological studies) and others.



The book has three parts. The first part covers some preliminary issues in test adaptation. It discusses terminology, and explains the various forms in which adaptations can be found; delves into issues related to copyright and the protection of intellectual property; and presents a number of documents which constitute the standards, guidelines, and recommendations of the domain.

The second part focuses on equivalence - arguably the most important concept in this process. It explains what equivalence is, describes its various forms, discusses in detail how to test for the various forms of equivalence and to detect bias, and how to deal with bias once it has been detected.

The third part examines the critical phases in the adaptation process. It discusses various translation designs, describes pretesting and norming, and looks at how to complete the process by finally publishing the test materials. The book concludes with a large and very applied chapter with practical examples of test adaptations.

What makes this book a particularly strong contribution is its practical stance. It provides a step-by-step approach to test adaptation, presenting it emphatically as a *mélange* between science and practice. The volume is driven by the first-hand practical experience of the author in a large number of test adaptation projects in various cultures, and is supported by the consistent scientific body of knowledge accumulated on the topic over the last several decades. It will prove invaluable for anyone involved in this aspect of psychological testing.

Around the World



Agustin Tristan-Lopez discusses evaluation in Mexico.

Evaluation and Assessment in Higher Education in Mexico

Agustin Tristan-Lopez, PhD. Director. Instituto de Evaluación e Ingeniería Avanzada, Mexico & ITC member.

Introduction

During the last four decades, there has been fluctuating interest in evaluation topics in Mexico. Nowadays, the unfavourable appreciation is mainly due to a history of informal, subjective, non-systematic evaluation performed by teachers from basic education to higher education, where the indiscriminate use of exams (and actually with the abuse of rubrics and portfolios) is usually reduced to the assignment of grades and not to inform or feedback to students or teachers. The effect that grades have on the allocation of prizes and punishments distorts the interest in evaluation in many school environments, contributing to a poor appreciation of large-scale tests that are supposed to inform and support academic and professional development.

The application of nationwide large-scale tests initiates in the 1960s, with exams produced by the Secretariat of Public Education (SEP), the so called “Knowledge Olympics” and the final (or extraordinary) tests to pass a course (Martinez-Rizo, 2001). Subsequently the SEP produced learning tests (once per semester) with scarce technical foundations, administered to student samples. At that time people believed that it affected minority or disadvantaged populations, assuming – without demonstration – a cultural bias against indigenous populations and students of low socioeconomic status. These SEP tests were officially used not only to evaluate students, but also for the ranking of teachers, institutions, state subsystems and the national system. In any case those exams had the implicit purpose of improving planning and favouring the accountability of education stakeholders. The diversity of uses of the tests involved several defects of interpretation, errors of appreciation and annoyance in some academic and social groups.

For Higher Education Level, large-scale evaluations were developed from 1960. In particular, the National Autonomous University of Mexico (UNAM) began to apply admission tests with more than one hundred multiple-choice items (MCI), graded by software provided by the optical mark reader.

These initial tests did not guarantee the equalization of test versions in the same promotion or over several years, nor did they have calibration analysis.

Interest in evaluation in Higher Education increased in the 90s when universities and external agencies started to provide test services. It is possible to classify the tests according to the moment and purpose of the evaluation:

- a) Selection tests. Designed as norm referenced tests for the admission to the institution (produced by universities and external agencies).
- b) Diagnostic tests. Designed to identify the strengths and weaknesses of students, in order to provide feedback in their competencies for success in the university career. These tests were designed by external agencies and applied at different times (on admission or at the end of the first two years of the career). Unfortunately, they are still not widely used.
- c) Degree tests. Generally produced by an external agency based on a general specification for all universities in the country. These tests are defined by technical committees from professional boards and are only an option at the end of the College or University course, because in certain universities, students could get the professional diploma by finishing all the courses with a high mean of grades or with a short thesis.
- d) Certification tests. These exams are professional academic certification of the performance of professionals who have more than 3 to 5 years of experience since finishing University. These tests are offered by Professional Boards in a given area (for instance Accountancy, Medicine, Nursing, and Civil Engineering). The Professional Board must have a contract with an external evaluation body and be authorized by the General Directorate of Professions of the Secretariat of Public Education.

All these exams are independent of those prepared by individual teachers or institutions. Generally, class exams and quizzes, as well as so-called authentic assessment tools, are informal tools showing deficiencies on validity, objectivity and reliability. Institutional exams (called departmental exams) are more structured and are prepared by a group of teachers according to the planned progress at the end of a unit or area, but they are not always administered or scored under standardized conditions. In consequence they offer only a limited advantage for quality control of learning and teaching. A small number of institutions have calibrated item banks, and a lesser number of those analyse the tests to calibrate the items or determine the measurement error. It is a common practice to grade students by combining the results of exams and portfolios with homework, participation

in class and a variety of other elements, leading to subjective evaluations and low-quality assessment processes. Currently there are some universities (mainly private) that have computer platforms to administer quizzes, quick tests or exams at the end of a topic or unit. They have tools to store some evidence and documents to integrate a portfolio and for administrative purposes. Unfortunately, these platforms in general are not focused on providing feedback to students and teachers.

Evaluating Agencies

In Mexico there are some agencies or institutions that produce tests for sale as a service, in schemes similar to the ACT or the ETS of the United States, CITO in the Netherlands or the NFER in the UK. Mexican evaluation agencies are highly professional and possess technical capabilities to provide quality products to evaluate Higher Education. As a new form of attracting projects, some university institutions offer evaluation services for specific purposes. For instance, since 1990 the Autonomous University of Baja California has offered the Knowledge and Basic Skills Test (EXCOHBA by its name in Spanish). The UNAM (Autonomous National University of Mexico), the Universidad Iberoamericana University and the Universidad Veracruzana offer some tests for professional certification since 2013. Within the evaluation agencies, the National Centre of Evaluation for Higher Education (CENEVAL by its name in Spanish), based in Mexico City since 1994, offers some products (<http://www.ceneval.edu.mx/>), such as: (a) admission tests to high school, university and postgraduate courses (called EXANI I, II and III), (b) graduation tests for some professions (called EGEL) and (c) certification tests as the External Evaluation Body for Professional Boards (such as accountants, actuaries or dentists). These criterion referenced tests (CENEVAL, 2017) define a professional profile based on academic knowledge, with the tests containing between 120 and 200 multiple choice items organized in two sessions (from 3 to 4 hours each). In the teacher's evaluation, the certification process includes a review of a video performance in class. The webpage of each evaluation project provides the candidate with a guide in PDF format. In general, this agency follows some of the standards from AERA-APA-NCME (1999).

The Institute of Evaluation and Advanced Engineering (IEIA by its name in Spanish) is a nongovernmental and non-profit civil association based in San Luis Potosí since 1989 (www.ieia.com.mx). The IEIA offers several products: (a) diagnostic assessments of competencies for admission of students at undergraduate or graduate levels (called DICOIN), (b) formative evaluation system for higher education teachers (ECPEMS) (Tristan & Ylizaliturri, 2012), (c) certification tests as External Evaluating Body for Professional Boards in the areas of health (Nursing, General Medicine, Pharmaceutical

Chemistry) and engineering (Civil, Mechanical, Electrical, Petroleum, Appraisals) and (d) software for test scoring, item banking and on-line testing. The assessment projects of the IEIA are criterion referenced, based on the definition of a profile: academic success expectations for the diagnostic tests or competencies for professional certification. Each project includes a variety of instruments: objective tests with multiple choice and forced response items, open response sections for clinical cases (in health) or practical cases (in engineering), simulators, analysis of performance using a sequence in a video; lexical analysis and automatic scoring of texts. Each project includes a formative component on a web-based platform, where the candidate and the institution can find feedback reports and a variety of resources such as white papers, links to articles, videos and activities. This agency follows the standards from AERA-APA-NCME (2014), the ICT guidelines recommendations for fair testing, test scoring and report and the IEIA proprietary standards (Tristan & Vidal, 2006).

Evaluation experiences in accordance to areas of knowledge

The most interesting areas of development in evaluation pertain to Medicine and Engineering.

Health Sciences

Due to the large number of students in Medicine, since 1973 the Area of Medical Education of the Faculty of Medicine (UNAM) has produced an objective professional test. At the beginning, it was a rigid model of 280 multiple choice items, where the first 210 were independent, with the remaining 70 items organized in 14 testlets or clinical cases, consisting of 5 items each. This model subsequently changed to encompass more clinical case testlets, on the same basis of 5 items for each case. For many years responses were registered using pre-punched cards and scored using a software programmed in FORTRAN or ALGOL by a responsible individual in the same Faculty. This software changed to other tools for PC in the 90s. The test blueprint and specifications were aligned to the curriculum and more recently to some professional competencies. A study guide is offered to the student, which has an interesting evolution, in that at first, the item bank from which each test was derived was fully disclosed, effectively requiring the student to memorize several thousand questions. This erroneous practice of preparing for the test was gradually modified and is no longer used. Graduation tests are currently divided into a theoretical phase with multiple choice items and a practical phase including Objective Structured Clinical Examinations (OSCE). Other universities in the country are using a similar model. Professional certification tests are applied when a person has more than 5 years of experience after completing the medical school. This nationwide certification has been produced for

the Mexican Board of Certification in Medicine by the IEIA since 2010.

Nursing certification is produced by three agencies (CENEVAL, IEIA and Universidad Iberoamericana) for three different Professional Boards. That from the Mexican Board of Certification in Nursing is the oldest, having been established in 2007. The exam contains from 220 to 280 multiple choice items, a portfolio, and a clinical case to develop in an open response answer sheet.

Engineering

The second profession with a solid trajectory in evaluation is engineering. Since 1976, evaluations have been prepared by the Faculty of Engineering (UNAM). It emphasises the evaluation of mathematics for six areas during the first four semesters, including diagnostic tests for admission, as mentioned in the UNESCO's report on Mathematical Evaluation (Kilpatrick, 1979). The evaluation model originally included the design of a scoring software in FORTRAN IV able to handle multiple choice, true-false, ranking, column relationship, and brief response open questions. Initially it used pre-punched response cards, but then moved to an open format needing a team of typists. A specific program was produced to score and calibrate the items, providing statistical analysis of the test and the students. Feedback reports were provided to students, teachers, academic coordinators and faculty directors. An interesting innovation was the report going to the schools of origin of the students who had entered the university. The item bank system offered innovative services with quick tests (quizzes) and informative reports to the teachers, to improve their class planning and their own evaluations. Those original systems have been updated for the Windows environment and are now commercially available on the Web by an external agency.

Other Domains

In Psychology, the focus in psychometrics is mainly on the use of commercial tests such as MMPI or Wechsler, for clinical applications and employment selection. A graduate test is produced by CENEVAL, but no other professional nationwide certification testing is actually produced.

Since 1960, several Mexican institutions have hired the SAT to the College Board of Puerto Rico. This produced some initial problems of decontextualized language for Mexico, which have been improving over time.

Evidence points to evaluation in higher education being a heterogeneous panorama. On one side, there are internationally known important agencies, producing projects, software and psychometric models using classical test theory, Rasch and IRT, with a significant emphasis on assessment and comprehensive reports.

However, in the past few years, political and social requirements are changing the focus of evaluation. For instance, the current holistic educational model is not

interested in measuring models, mastering psychometric techniques or providing technical feedback to the students, professionals and institutions. Universities aim to sell evaluation services for professional certification, despite it not being part of their core competencies. The government and unions are reluctant to participate in high stakes testing, arguing, without evidence, that they are unfair for their labour contracts. The new trends in evaluation are responding to complex and subjective political and social decisions that affect the future condition of all Mexicans. Other countries in Latin America, Africa and even Europe are using some of the approaches and developments produced in Mexico. In addition, at university level in general and in the evaluation field in particular, Mexico has worldwide recognized evaluation agencies and specialists with a very high level of expertise at the top of the development of measurement models and software design.

References

- AERA-APA-NCME (1999/2014). *Standards for educational and psychological testing*. Washington, American Educational Research Association, American Psychological Association & National Council on Measurement in Education. Washington, D.C. Available in: <https://www.apa.org/science/programs/testing/standards>
- CENEVAL (2017). *Anuario Ceneval*. México, CENEVAL. Available in: http://www.ceneval.edu.mx/documents/20182/49855/Anuario+Ceneval+2017_2/5d95c427-bd58-45df-b052-a332d626cf5c
- Martínez-Rizo, F. (2001). Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate. *Revista de la Educación Superior*. ANUIES. Vol. 120. México. Available in: <http://publicaciones.anui.es.mx/revista/120>
- Kilpatrick, J. (1979). Methods and results of evaluation with respect to mathematics education. In *"New trends in mathematics teaching"*, Vol. 4, pp. 162-179. International Commission on Mathematical Instruction. UNESCO. Paris. Available in: <https://unesdoc.unesco.org/ark:/48223/pf0000124827>
- Tristán, L.A. e Ylizarri, S.M.A. (2012). *Certification of Teaching Competencies at the High-School Level in Mexico: An Evaluation Model for a Social Challenge*. The International Journal of Educational and Psychological Assessment. 9(2) p. 139-164. Available in: <https://sites.google.com/site/tijepa2012/articles/vol-9-2>