



Open questions in the use of new technologies in psychological assessment

Pablo Santamaría and Fernando Sánchez-Sánchez

Departamento de I+D+i de Hogrefe TEA Ediciones, España

La aplicación de las nuevas tecnologías (NNTT) a la evaluación psicológica puede ampliar enormemente la capacidad de los profesionales para evaluar de forma más precisa y exhaustiva y agilizar los procesos de recogida y análisis de información. Para poder acceder a estas ventajas hay que considerar también los costes, desafíos y amenazas potenciales derivados de la implantación de estos nuevos sistemas de evaluación. Con este objetivo, en este trabajo se abordará en primer lugar en qué grado los profesionales de la psicología, los usuarios y las organizaciones están preparados para adoptar el uso de las NNTT. A continuación, se comentará el impacto que la aplicación de la tecnología está teniendo en los propios instrumentos y en el papel que los profesionales desempeñan. Tras este punto, se cuestionará en qué grado puede darse una cierta fascinación por las NNTT que ciega el análisis de los criterios psicométricos esenciales para cualquier medida. Finalmente, se apuntarán algunos desafíos que los creadores y editores de test deben afrontar durante el desarrollo de instrumentos basados en NNTT.

Palabras clave: TIC, Tecnología, Test, Evaluación.

The application of new technologies to the field of psychological assessment can greatly enhance the ability of professionals to carry out assessments more accurately and comprehensively and to streamline the processes of collecting and analyzing information. In order to access these benefits, the costs, challenges, and potential threats associated with the implementation of new assessment systems must also be considered. To this end, this paper will first address the extent to which psychology professionals, users, and organizations are prepared to adopt the use of new technologies. Then, we will discuss the impact that the application of technology is having on the instruments themselves and on the role that professionals play during the process. After this, we will question the extent to which a certain fascination with technology can blind the analysis of the essential psychometric criteria for any measurement. Finally, we will identify some of the challenges that test developers and publishers must face during the development of instruments based on new technologies.

Key words: ICT, Technology, Test, Assessment.

The implementation of new technologies (NTs) in the field of psychological assessment opens a new horizon of possibilities, resources, and approaches that until recently were unimaginable for most professionals. The development of novel and sophisticated assessment techniques (e.g., new types of tasks, virtual reality, or interactions with bots), the possibility of collecting large amounts of data and measures (e.g., continuous records through personal electronic devices or social network analysis) and the availability of new techniques to extract, analyze, and model large amounts of information (big data, artificial intelligence) have led psychology professionals and researchers to face a new and exciting scenario.

However, it is worth asking whether emerging NTs will end up constituting a new world that breaks with the previous one (technological disruption) or whether, on the contrary, they will be just toys, illusions, or appearances that will end up deflating over time (Chamorro et al., 2016).

In the wake of a pandemic that has accelerated the implementation and use of technology, this question is even more relevant. The application of NTs to assessment is full of questions, with promising aspects, but also with negative and uncertain aspects. This monograph has addressed some of the promises, and the present article will attempt to complete the picture by showing, from an

applied perspective, some potential risks and adverse effects derived from the use of NTs. The ultimate goal is to help professionals to form a critical and complex picture of the adoption and use of NTs in psychological assessment.

PROFESSIONALS, USERS, AND ORGANIZATIONS: ARE WE PREPARED FOR THE USE OF NTS IN ASSESSMENT?

Psychological assessment requires professionals to master a set of specific techniques, as well as adequate training in the use of tests and sufficient knowledge of psychometrics. Assessment using NTs, in addition to the above, also requires adequate knowledge and appropriate use of different electronic devices and technology in general (in this monograph Elosua, 2022).

Following the classification of technology users according to their degree of mastery (Rogers, 2003) (innovators, early adopters, early majority, late majority, and laggards), surveys on the use of NTs directed at psychology professionals suggest that their predominant position is late majority or laggards (Ramos-Álvarez & Rodríguez, 2021). This indicates a relatively low adoption in the use of NTs and that the digital competence of psychology professionals could be a limiting factor for the development, implementation, and good use of the new assessment tools.

Added to this circumstance is that: a) many of the environments or contexts where psychology is practiced do not stand out for a high level of technological adoption, either in terms of the availability of equipment or in terms of the digital literacy of the individuals to be assessed (e.g., elderly care, families in disadvantaged environments, etc.); b) the literature has reported the presence of a relationship

Received: 1 November 2021 - Accepted: 21 December 2021

Correspondence: Pablo Santamaría. Departamento de I+D+i de Hogrefe TEA Ediciones. C/ Cardenal Marcelo Spínola, 50. 28016 Madrid. España. E-mail: pablo.santamaria@teaediciones.com





between worse mental health and lower socioeconomic status (e.g. Reiss, 2013; Reiss et al., 2019), which suggests a potential limitation to the access and use of NTs; and c) the availability of adequate equipment and facilities to be able to perform the assessments with guarantees (areas with Wi-Fi or mobile coverage, updated devices, etc.) represents another potential barrier, related to the financial aspects and resources of the centers.

The cost of acquiring, maintaining, and upgrading both the electronic devices (e.g., tablets, virtual reality glasses, etc.) and facilities (classrooms, consulting rooms, or offices with stable and fast connections) is high, and certainly higher than more traditional paper-and-pencil-based assessment alternatives. For example, one could mention the unfortunate reality of the practice of psychology in part of the public sector (hospitals, courts, schools, etc.), an area in which a significant portion of assessments are performed, and which is frequently underfunded and underbudgeted in psychological and technological material. To this list of potential barriers could be added those related to network access restrictions imposed by the security protocols of the centers themselves or the cost of computer processing of personal data (need for secure servers). The cost of acquisition and the rapid obsolescence of the necessary devices and software pose a major challenge to efficient resource management for institutions and professionals. This scenario of costly investments with a short service life influences the implementation of NTs applied to assessment.

This set of potential barriers significantly affects the use made of these technologies. Suffice it to cite as an example that, in 2021, 90% of the applications of the SENA (Fernández-Pinto et al., 2015) or 85% of those of the PAI (Morey, 2013) were in paper-and-pencil format, compared to 10% and 15% in online format respectively. Both tests are among the most widely used in psychological assessment (Muñiz et al., 2020).

The trend is similar in the case of optimal performance tests for cognitive assessment based on performance tasks. In the case of Matrices (Matrices, Test de Inteligencia General [General Intelligence Test]; Sánchez-Sánchez et al., 2015), in 2021 87% of its applications were performed using the paper-and-pencil format and only 13% using the CAT (computerized adaptive test) format, despite the undoubted advantages of the latter in terms of brevity, precision, and quality of measurement (Abad et al. 2020; Olea et al., 2010).

A brief analysis of the relationship between technology and psychology professionals shows a slow pace of transition to the use of NTs in assessment. As has been pointed out, this may be motivated, among other reasons, by the level of technological adoption of both professionals and many of the recipients of the assessments, as well as by economic conditioning factors and the particularities of the contexts in which professional practice is carried out. It is foreseeable that the barriers related to the technological adoption of professionals and users will be reduced with the new generations and that devices and connections will become more and more affordable, with greater capacity and applications. Professionals and institutions interested in taking advantage of the benefits of NTs applied to assessment should actively direct part of their efforts to reducing these basic limitations of access, training, and availability, as well as responding to the challenge of investment, renewal, and maintenance of the necessary equipment and connections.

PSYCHOLOGICAL TESTS: ARE THEY PREPARED FOR NTS AND FOR A WORLD WITHOUT BARRIERS?

At the same time that NTs facilitate access to a large amount of information and services related to psychological assessment, they pose potential threats in relation to the vulnerability of tests and the role of the professional in the assessment process.

RESTRICTION OF ACCESS TO PSYCHOLOGICAL TESTS TO PROFESSIONALS

Any psychological test, properly constructed and tested, is a valuable tool in the hands of professionals who know how to use and interpret it properly. An accurate assessment requires a qualified professional who is able to select the aspects to be assessed, the tests and techniques to be applied and, subsequently, to integrate the results in order to arrive at a correct diagnosis, an adequate formulation of the case, an orientation report, or a well-motivated hiring proposal.

When a test is applied, it is expected that the person being assessed will directly or indirectly obtain some benefit from the results and, conversely, its inappropriate use can lead to negative consequences; in this circumstance, the professional must ensure that the evaluation process safeguards the rights and interests of the people involved at all times (AERA et al., 2014).

For this reason, it is the same as with the prescription of drugs and radiological tests in the medical field (i.e., they are only accessible if previously prescribed by an accredited professional), access to psychological assessment instruments has traditionally been restricted to professionals in psychology or related disciplines.

One of the main reasons for restricting the sale of psychological tests is to avoid misuse that contributes to the stigmatization or discrimination of the persons assessed (e.g., applying inappropriate “diagnostic” labels, excluding a person from a selection process, etc.). The restriction of test sales based on the level of qualification (a, b, or c) required of professionals to be able to use them (AERA et al., 2014) responds to a well-thought-out control system for the appropriate use of tests as the professional tools that they are. Unfortunately, many of the creators of assessment tools based on NTs are alien to the world of psychology and the associated ethical or deontological implications, offering themselves in many cases as an alternative to the professional for obtaining an assessment.

A good example is the proliferation of psychological assessments offered directly to the end user (not to the professional) openly on websites or in application stores (apps) for use by both the general public (parents, the person being assessed, etc.) and professionals from various sectors (teachers, educators, etc.). Anyone can find and acquire tools for the evaluation of developmental aspects, cognitive skills, or psychopathological traits—and therefore they could be considered “clinical”. These tests, sometimes promoted as “diagnostic tests”, are the object of online marketing campaigns aimed specifically at mothers, fathers, or relatives, who can make use of them online without the participation of health professionals in the process and, also, outside the supervision of professional associations.

The problem with this practice is not that tools specifically designed for family members or other professionals outside the field of psychology are developed and offered, which would be positive and would involve a design specifically for that purpose, but that the same



tool is offered to the entire possible spectrum of users (family members, psychology professionals, professionals from other areas, etc.), putting a specialized tool in the hands of people who do not have the ability to interpret or manage the information it provides.

As can be seen, the absence of technological barriers to access to assessment tools for the general public not only brings advantages but also opens the door to potential malpractices that are outside the scope of supervision and professional ethics.

TEST VULNERABILITY AND CONTENT PROTECTION

Restrictions on access to tests also seek to protect assessment techniques and instruments so that they can be used effectively, since the usefulness of many tests lies in the relative novelty of the task or content. Examples include the tests of a battery of intellectual aptitudes in a personnel selection process or the stimuli of a memory test in a clinical context. If the individuals being evaluated had prior access to them, they could train themselves in the tasks of the battery or memorize the stimuli of the memory test beforehand, thus compromising the efficacy of the subsequent assessments.

Any type of public exposure of the contents of a test affects its future use, so the need to protect the materials and stimuli is a key aspect explicitly contemplated in the regulations or codes of good professional practice: "All types of strictly psychological material, both for assessment and for intervention or treatment, are reserved for the use of psychologists, who shall refrain from providing them to other non-competent persons" (Article 19 of the [Spanish] Code of Ethics of Psychology).

NTs offer effective—although expensive—solutions to protect test content, such as the use of computerized adaptive tests based on large item banks (in this monograph Abad et al., 2022) or online proctoring systems (supervision online). However, in contrast to these advantages, they also contribute to increasing the risks of exposure in an extreme way; we cite as an example the test download websites or online video platforms where students or psychology professionals upload videos showing the application and correction procedures of numerous tests. Unfortunately, these videos often show all the items and stimuli, the correct answers to each of them, and how to obtain a certain score. Although these videos are mostly made as teaching resources for other psychology professionals, the lack of restrictions on their viewing means that their content is virtually exposed to the whole world.

In the same vein, it is worth mentioning the publication of articles in open access journals or the online repositories of doctoral theses that contain critical information on the tests used in the research (items, the correction key, etc.), and even access to information on the control and validity scales used by the test to identify dissimulation, simulation, or other critical aspects.

It should not be forgotten that tests are often used for very sensitive purposes, such as in the clinical or forensic context to make legal decisions (incapacitation, imputability, etc.) or in the context of personnel selection to choose the right individual for a certain position. There are multiple interests that may exist for accessing the items and tasks with not very lawful objectives, which makes an adequate protection of the stimuli essential.

It is therefore necessary to highlight the vulnerability of tests to exposure to non-professionals through social networks and the

Internet and the negative effect this can have at the applied level. The information society and the use of NTs make it necessary to accentuate the protection of tests.

ANYTIME, ANYWHERE AVAILABILITY

NTs can contribute to blurring the geographical barrier of testing. Online assessment services and instruments can be easily used from anywhere in the world and are available 24 hours a day, which is bringing the work of psychologists closer to places and contexts that were previously difficult to access.

This circumstance can become a potential source of malpractice when certain instruments originally developed in a given language and country are offered and used in other countries and different languages without following an adequate adaptation process (Muñiz et al., 2013). That is: a) through automatic translations without an adequate process of linguistic and cultural adaptation, b) without adequate local scoring, using the original scales of the country where it was developed or an aggregate of data from different countries (usually referred to as "international scales"), and c) without providing validity evidence obtained with the new versions, referring, at most, to the evidence of the original version of the test that has not been obtained either in the country or in the language with which the assessment is carried out.

These practices seriously undermine the minimum standards of test quality and use (AERA et al., 2014; ITC, 2018), which should apply to both traditional (paper-and-pencil) assessments and ones that are online or use new technology.

SCIENTIFIC QUALITY OF ASSESSMENT: DO WE EVALUATE BETTER THROUGH THE USE OF NTS?

The questions raised in the previous sections have an important impact on the applied practice of psychological assessment using NTs and the use made of them by professionals. However, undoubtedly, the central question should be the following: Does the use of NTs applied to psychological assessment increase its scientific quality?

Several articles in this monograph show that this can clearly be the case, pointing to a promising increase in the scientific quality and practical utility of the assessments. However, this is not always the case and one can find multiple examples of NT-based psychological assessment platforms that dramatically neglect substantive aspects at the psychological and psychometric level. From a legal perspective, and given the absence of a restrictive regulatory framework, little can be done to avoid these deficient tools. This circumstance places the onus on the practitioner's ability to critically examine and screen the quality of the tools with which he or she works.

DIFFERENTIATING BETWEEN TESTS AND PSEUDO-TESTS

There are few studies on the ability of professionals to carry out this task of screening and adequately assessing the quality of psychological assessment tools that use NTs. In a study carried out with graduate students in psychology attending various specialization masters ($n = 232$) and with practicing professionals attending a refresher course on the correct use of tests ($n = 73$), participants were asked to evaluate an online platform for psychological assessment that had an excellent and attractive presentation but also had serious psychometric problems (absence of Spanish scales, absence of



validity evidence, presentation of reliability data as if they were validity data, translation and not adaptation, automatic scales based on the clinical cases assessed themselves, etc.) (Santamaría, 2020). Participants were asked to carefully study the information available on the platform and to evaluate the tool as expert assessors, after which they were asked to make a judgment as to whether they would recommend its use or advise against it. Despite the obvious shortcomings, only 29% of the psychology graduates and 42% of the professionals rejected the use of the tool due to its inadequacy; more than half (55%) of the graduates and a third of the professionals (33%) rated it positively and recommended its use, and a high percentage acknowledged not feeling competent to make this assessment (16% of the graduates and 25% of the professionals). Although preliminary and merely exploratory, these data alert us to the limited capacity of graduates and professionals to detect clearly defective tools and how the use of technology, an attractive design, and adequate marketing is persuasive in these cases, even without a minimum underlying psychometric quality.

This potential difficulty for professionals to critically assess the scientific quality of the available tests could be overcome by employing independent reviews carried out by professional institutions (e.g., the COP Test Commission or similar institutions in other countries; in this monograph Hernández et al., 2022). Unfortunately, many of the assessment platforms based on NTs are outside the scope of action of these institutions—which in the case of the COP Test Commission is limited to tests published in Spain—and, therefore, they are not reviewed.

These limitations in assessing the scientific quality of some assessment tools that use NTs are aggravated by the difficulty of accessing their technical information. As with any test for psychological assessment, it is essential that professionals have access to documentation detailing the construction process followed, the psychometric evidence of reliability and validity, the scoring samples used to calculate typical scores, as well as the application and correction procedures (AERA et al., 2014). The fact that a test has been developed with NTs does not exempt it from complying with the standards required for any psychological measurement technique, nor does it exempt the available evidence from being clearly presented for analysis by professionals.

TECHNOLOGY AND KNOWLEDGE: A NECESSARY BALANCE

The incorporation of NTs into assessment has contributed to the elimination of many interprofessional boundaries, making it necessary for engineers, programmers, clinical psychologists, psychometricians, and other specialists to collaborate in the development of new tools. This multidisciplinary work responds to the need to ensure that the measure contemplates and includes the substantive aspects of what is to be assessed and that the technical procedures for carrying out the assessment allow it to be done with guarantees.

Multidisciplinary collaboration is necessary to avoid confusion between the means (technology) and the end (psychological measurement). As defined by the RAE, technology consists of the “set of theories and techniques that allow the practical use of scientific knowledge” (RAE, 2014). Technology cannot exist without the scientific knowledge upon which it is based in order to extract its practical use. In this sense, all artificial intelligence techniques rest on

the quality of the data that are introduced, which depends on scientific knowledge for its correct selection. Unfortunately, often the starting point is just the opposite: since we have the technique (programming and artificial intelligence), any issue can be addressed, without needing the scientific knowledge to guide, orient, or frame the technology; thus, the isolated value of the technology is overestimated and what psychological assessment really consists of is ignored.

Psychology faces an important challenge in facing a reality in which more and more professionals outside psychology and linked to NTs develop psychological assessment tools with limited scientific knowledge of the variables they intend to measure and of the basic metric issues associated with them. It is essential for psychologists at the social, institutional, and professional levels to understand how to reclaim the central role of psychology in these aspects as a guarantor of the adequate quality of the tools and assessments, with the aim of ensuring a good measurement that avoids negative consequences for the people being assessed. Collaboration between the scientific knowledge accumulated in psychology on the various variables of human behavior and technology is necessary and essential in order to make the best possible use of it. Without this equal interdisciplinarity, both parties will be lost.

TECHNOLOGY FOR TECHNOLOGY’S SAKE: DISTINGUISHING WHEN AND HOW TO MAKE USE OF TECHNOLOGY

The knowledge accumulated in psychometrics on good practices in the development of measurement tools is enormous. This knowledge should also be applied to NT-based tools, from the construction of the test to its validation process or to the process of collecting samples for scoring or norming.

In the construction of a test, an intelligent use of NTs must be made, integrating the psychological knowledge of the area to be assessed. In this sense, it is important to consider, on the one hand, the incremental validity, i.e., the extent to which the introduction of technology improves the quality, usefulness, and predictive value of the assessment with respect to other measures already available; and on the other hand, the added cost, both in terms of development and use. Unfortunately, in many circumstances this evaluation is not carried out and it is common to observe the use of technological solutions simply on the basis of their novelty or sophistication: technology for technology’s sake.

We can cite as an example the assessment of attentional capacity in children and adolescents using tablets and apps whose design and presentation are particularly attractive and stimulating for these ages. Paradoxically, this playful and attractive aspect can be counterproductive for the adequate assessment of cases suspected of having an attention deficit disorder, given that the attractive nature of the tablets and programs can interfere precisely in what would be the object of assessment: the child’s ability to sustain attention in unattractive, unstimulating, monotonous, and boring tasks and environments. This example shows a clear divergence between the patterns that are usually applied in NT-based platforms (highly stimulating, changing, and attractive environments based on gamification) and the characteristics necessary for a good assessment of sustained attention based on knowledge of the area (unstimulating, unchanging, unattractive, and demanding environments).



While the future of the application of NTs in assessment through the use of tablets and gamification-based apps is very promising, technology for technology's sake may be senseless if the psychological aspects being assessed as well as their implications are not taken into account in depth.

NORMING SAMPLES AND SCORING TECHNIQUES

Another central aspect in the construction and use of tests relates to the collection of norming samples and the scoring techniques employed, this being probably one of the most laborious and expensive phases of the development studies. The facilities offered by NTs in the sample collection process are enormous compared to more traditional procedures in which a professional must apply the tests in person after having carefully selected each of the individuals to be evaluated for the norming sample. There is a large body of knowledge on the criteria to be met by norming samples in terms of representativeness. These criteria are transferable and applicable to tests based on NTs, which by their very nature may be vulnerable to this circumstance. For example, the collection of samples from forwarded messages via instant messaging (e.g., WhatsApp), email or social networks, may suffer from a lack of control over the people who respond and the conditions in which they do so (with what degree of attention, in what context, etc.), which can generate a marked selection bias. From another perspective, the selection samples collected in applications such as the iPhone App Store, for example, could be the source of selection bias due to the sociodemographic profile of the people who have access to these expensive devices. In relation to this point, it is important to take into account the bias associated with differences in the profiles of people who have ease of use and access to technology versus those who have difficulties and who are very unlikely to participate in this type of "studies" (e.g., through online forms or tests hosted on apps or websites). The fact that we are able to collect large numbers of cases in a short period of time should not make us forget the importance of quality versus quantity of the samples.

A more controversial issue would be the use of online survey companies to collect the norming samples for assessment instruments (e.g., Mturk). Although their use can make the scaling processes much cheaper and faster, they are not exempt from criticisms that question the possible biases that may be committed both due to the profile of the people who usually participate in these processes and due to possible alterations or distortions to the profiles (e.g., the use of a VPN to simulate that the respondent is in Spain, when in fact he/she is responding from a Latin American country).

Although problems related to selection bias are not new to psychometrics, it seems that this is not given due attention when it comes to sampling using NTs, perhaps due to: (a) a lack of knowledge on the part of the professionals who implement them (with little training in psychometrics), (b) a certain halo effect of efficiency and precision in relation to NTs, (c) the "black box" effect of NTs that makes the processes less evident and the possible biases committed less obvious, or (d) the tendency to prioritize large samples (size) over their representativeness (quality), trusting that large numbers will be able to compensate for the biases committed.

In summary, an adequate use of technology in assessment requires the integration of a deep knowledge of NTs—their advantages,

disadvantages, and possibilities—, the substantive psychological aspects of the areas to be measured, and the associated psychometric issues. The absence of the substantive psychological and psychometric aspects often results in a hollow measurement technology.

Whatever the technology used (traditional or new), the essential and underlying issues that are required remain largely the same: adequate systematization of the assessment, a representative sample for its scales, and solid evidence regarding its reliability (internal, temporal, and inter-rater consistency) and—of course—its validity. The validation of an instrument—that is, providing theoretical and empirical evidence about what is being measured—is not a secondary issue, rather it is the central aspect of the test. Regardless of whether we are talking about a paper-and-pencil test or an NT-based test, the essence of the test and its applied use lies in the degree to which there is empirical evidence to support the inferences drawn from its scores.

IMPLICATIONS OF NTS FOR THE TESTING INDUSTRY

One of the main agents involved in the application of NTs to psychological assessment are the test developers and publishers, who face important challenges in the implementation of NTs in assessment instruments. One of these challenges relates to the different time horizons and pace of development of NTs and the test industry.

In a certain sense, the two are opposed to each other; the time horizon of NTs has a rapid and ephemeral life cycle, which persistently seeks improvement and innovation as essential and defining aspects (better resolution, faster connection, a new feature that attracts consumers, etc.). In contrast, the framework and time horizon of the creation of psychological assessment instruments seeks systematization, standardization, and stability in measurement. The objective is to reduce the effect of extraneous variables (modifications in the instructions, in the way of responding, etc.), so that the procedure is closed and systematic and any variation that occurs has to do with the person being assessed (and his or her level in the measured variable); in this sense changes in the format can be a problem for the measurement.

To illustrate this point, let us consider computer-based assessments. Since the first proposals, the evolution of computer equipment has made it possible to offer users new ways of interacting with the devices (in this monograph Elosua, 2022). While originally responses were indicated with the keyboard or mouse, touchpads soon appeared on laptops, then touchscreens, and now speech recognition. From the point of view of NTs, change is a positive and desirable thing, and innovation and the development of novel ways of interacting with devices (e.g., eye tracking with front-facing cameras on computers or phones) is constantly sought after. Flexibility and change are the hallmarks of NTs. However, the modification of response behavior can be a potential problem for tests and their systematization. For example, response times are not the same with a touch screen, with a mouse, with a touchpad, or with the voice. And applying the test with one or another system (or with one or another screen size) affects the assessment process itself (size of stimuli, distractors, times between stimuli, etc.) and as a consequence the reference scales. This is why uniformity in the application procedures is required in tests and there is usually a restriction to one single way



of indicating the responses so that they can be compared with the scale data. In this sense, standardization—understood as the stability, uniformity, and systematicity of the measures—is the hallmark of psychological tests.

On the other hand, test construction requires an economic investment of several years to systematize, standardize, norm, and validate the measurement. Creating adequate empirical support for the inferences made from the measurements requires an extensive period of time and resources. Postulating inferences from test scores about a person's possible autistic traits, ability to carry a weapon, degree of psychopathy, or competence as a parent are complex issues that require an extensive and complex validation process of several years. This circumstance is accentuated in the case of new measures based on NTs given that, precisely because of their novel and innovative nature, they require more research and evidence to support them in view of the relative absence of previous literature on the subject.

The test industry seeks, to some extent, the opposite of the NTs industry: relatively long test life cycles that allow the accumulation of evidence about the validity of test uses in different populations and for different purposes.

This contrast between the life cycles of both NTs and tests is a continuous challenge. For example, the time required for the development, norming, and validation of a test usually ranges from 3 to 5 years. During this time, there are bound to be changes in operating systems, programming languages, and the appearance or disappearance of certain programs or devices (see the cases concerning the disappearance of Flash technology, the Kinect console, or virtual reality glasses, etc.). Often, these changes during the development of NT-based tests mean that, once completed, the project must be restarted to adapt it to the new technology and avoid problems of compatibility or the design becoming obsolete. An even more serious problem is the dependence on the suppliers of the technology used (be it a device, a programming language, etc.), and the fact that they may vary their conditions over time.

All this represents a very significant challenge at an economic level that affects the development of the tests, since it requires not only an additional investment during the construction of the test itself, but also on its maintenance, and it affects the potential service life of the test, which in some cases is uncertain.

In addition, the number of test users is infinitely smaller than the number of NT users in general, which further complicates the equation for the creation of NT-based tests: a high development cost for a potentially shorter test life (due to the rapid obsolescence of the technologies), with a reduced market of potential buyers (often self-restricted to professionals as mentioned above) and with a maintenance and updating cost much higher than the paper-and-pencil versions.

Undoubtedly, this combination of factors makes the task of test development with NTs complicated, so sometimes inappropriate solutions are sought, such as: (a) trying to indiscriminately increase the number of potential users (extending it to the general public or to all types of professionals), (b) attempting to save the costs of controlled norming and validation studies (essential for performing adequate psychometric analyses and for obtaining representative samples); or (c) in an intermediate way increasing potential customers by

“exporting” the tests to other countries without performing adequate validation or norming studies for the new languages. These practices represent one of the main threats to the correct use of NTs in assessment.

BY WAY OF SUMMARY

NTs will revolutionize the way we think about psychological assessment in the coming years. Several aspects discussed in other articles in this monograph illustrate how they will provide extremely valuable new avenues of information to the psychological world.

However, the pace of this transition towards the use of NTs in assessment still seems to be slow, undoubtedly conditioned by, among other things, the level of technological adoption of both professionals and the recipients of assessments, as well as by economic constraints and the particularities of the contexts where professional practice is exercised (schools, hospitals, courts, mental health centers, private offices, etc.).

This revolution requires an adequate integration of technology and psychology, which—in our opinion—is still far from being a reality. On the one hand, technology without adequate scientific knowledge on which to base it is senseless and wastes the great wealth of psychology and psychometrics when measuring cognitive, emotional, or behavioral aspects, omitting relevant issues and introducing avoidable biases in the collection of data and the generation of results, leading to misuse with undesirable negative consequences.

Psychology and psychometrics run the risk of being left behind in a future where NTs will cover all aspects of human life; to this end, they must accelerate the pace of technological adoption, reclaim their central role as experts in psychological assessment, and warn of malpractices involved when omitting compliance with psychometric standards in terms of what is meant by an adequate adaptation, scaling, or validation that solidly supports the inferences to be drawn. All this implies responding to very different challenges, as discussed in the article, both at the level of training and technological competence, finding the balance between transparency and protection, between stability and change, between systematization and innovation, and between the different times and horizons of technology and psychology in the proper use of assessment tools.

The success of psychological assessment in the coming years, depends on the appropriate meeting of the two fields, technology and psychology.

CONFLICT OF INTEREST

There is no conflict of interest.

REFERENCES

- Abad, F. J., Sánchez-Sánchez, F., & Santamaría, P. (2020). *MATRICES-TAI. Test Adaptativo de Inteligencia General [MATRICES-CAT. Adaptive Test of General Intelligence]*. TEA Ediciones.
- Abad, F. J., Schames, R., Sorrel, M. A., Nájera, P., García-Garzón, E., Garrido, L. E. & Jiménez, M. (2022). Construyendo Test Adaptativos de Elección Forzosa “on the fly” para la Medición de la Personalidad [Building Adaptive Forced-Choice Tests “On The Fly” For Personality Measurement]. *Papeles del Psicólogo*, 43(1), 29-35. <https://doi.org/10.23923/pap.psicol.2982>



- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Psychological Association.
- Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., & Hogan, R. (2016). New talent signals: Shiny new objects or a brave new world? *Industrial and Organizational Psychology*, 9(3), 621-640. <https://doi.org/10.1017/iop.2016.6>
- Elosua, P. (2022). Impacto de la TIC en el entorno evaluativo. Innovaciones al servicio de la mejora continua [Impact of ICT on the assessment environment. Innovations for the purpose of continuous improvement]. *Papeles del Psicólogo*, 43(1), 3-11. <https://doi.org/10.23923/pap.psicol.2985>
- Fernández-Pinto, I., Santamaría, P., Sánchez-Sánchez, F., Carrasco, M. A., & Del Barrio, V. (2015). *SENA. Sistema de Evaluación de Niños y Adolescentes [SENA. Evaluation System for Children and Adolescents]*. TEA Ediciones.
- Hernández, A., Elosua, P., Fernández-Hermida, J. R. & Muñiz, J. (2022). Comisión de Test: Veinticinco años velando por la calidad de los test [Test Commission: Twenty-five years working on test quality]. *Papeles del Psicólogo*, 43(1), 55-62. <https://doi.org/10.23923/pap.psicol.2978>
- International Test Commission (2018). Guidelines for Translating and Adapting Tests (Second Edition). *International Journal of Testing*, 18(2), 101-134. DOI:10.1080/15305058.2017.1398166
- Morey, L. C. (2013). *PAI. Inventario de Evaluación de la Personalidad [PAI. Personality Assessment Inventory]* (Ortiz-Tallo, M. Santamaría, P, Cardenal, V & Sánchez, M. P., adapters). TEA Ediciones.
- Muñiz, J., Elosua, P. & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: Segunda edición [International Test Commission Guidelines for test translation and adaptation: Second edition]. *Psicothema*, 25(2), 151-157.
- Muñiz, J., Hernández, A., & Fernández-Hermida, J. R. (2020). Utilización de los test en España: el punto de vista de los psicólogos [Test use in Spain: the psychologists' viewpoint]. *Papeles del Psicólogo*, 41(1), 1-15. <https://doi.org/10.23923/pap.psicol.2921>
- Olea, J., Abad, F. J., & Barrada, J. R. (2010). Tests informatizados y otros nuevos tipos de tests [Computerized tests and other new types of test]. *Papeles del psicólogo*, 31(1), 97-107.
- Ramos-Álvarez, R. & Rodríguez, M. (9-11 July 2021). *Evolución del uso de las tecnologías en psicología en España: 2013-2020 [Evolution of the use of technologies in psychology in Spain: 2013-2020]*. V Congreso Nacional de Psicología [V National Congress of Psychology]. Spain.
- Real Academia Española (2014). *Diccionario de la lengua española [Spanish language dictionary]* (Vol. 23). Autor.
- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Social science & medicine*, 90, 24-31. <https://doi.org/10.1016/j.socscimed.2013.04.026>
- Reiss, F., Meyrose, A. K., Otto, C., Lampert, T., Klasen, F., & Ravens-Sieberer, U. (2019). Socioeconomic status, stressful life situations and mental health problems in children and adolescents: Results of the German BELLA cohort-study. *PLoS one*, 14(3), e0213700. DOI: 10.1371/journal.pone.0213700
- Rogers, E. (2003). *Diffusion of Innovations* (6th ed.). Free Press.
- Sánchez-Sánchez, F., Santamaría, P., & Abad, F. J. (2015). *MATRICES. Test de Inteligencia General [General Intelligence Test]*. TEA Ediciones.
- Santamaría, P. (12 May 2020). ¿Es oro todo lo que reluce en evaluación psicológica con nuevas tecnologías? [Is all that glitters gold in psychological assessment with new technologies?] I Jornada de Actualización en Evaluación Psicológica. [I Conference on Updating in Psychological Assessment]. Spain.