

Measuring Elementary Students' Expectancies of Success in School Science: Psychometric Evaluation of the SUCCESS Instrument

Radu Bogdan Toma ^{1*}

¹ University of Burgos, Department of Specific Didactics, SPAIN

Received 30 November 2018 • Revised 24 January 2019 • Accepted 26 February 2019

ABSTRACT

Background: The importance of valid and reliable instruments for the assessment of factors affecting students' interest in science encouraged the development and validation of a brief Spanish instrument for the measurement of expectancies of success in school science, named SUCCESS. In this study, the psychometric properties of the SUCCESS instrument are further evaluated using different psychometric tests and a different sample than the one included in the original validation study.

Material and methods: A sample of 313 Spanish elementary school students enrolled in 4th to 6th grade was drawn by means of convenience sampling techniques. Responses were analyzed in terms of construct and criterion validity, and two reliability indices.

Results: Results from confirmatory factor analysis established the unidimensional structure of the instrument, with great model fit indices. Correlation coefficients between the SUCCESS and external measures (i.e. intentions to enroll, enjoyableness, difficulty, auto-efficacy, utility and relevance of school science) provided evidence of criterion validity. Cronbach α and item-total correlation indices supported the internal consistency reliability of the instrument.

Conclusions: Taken together, this study further provide evidence to consider the SUCCESS as a valid and reliable tool for the measurement of Spanish elementary school students expectancies of success in school science.

Keywords: elementary education, expectancy of success, psychometric properties, school science education

INTRODUCTION

Science education worldwide continues to face a steady decline in students interested in scientific-related careers (Archer et al., 2010; Kennedy, Lyons, & Quinn, 2014; Lyons & Quinn, 2010, 2015; Osborne, Simon, & Collins, 2003; Potvin & Hasni, 2014). At the same time, the existence of valid and reliable instruments that would help understand the factors affecting students lack of engagement in science continue to be scarce in the science education literature (Blalock et al., 2008; Gardner, 1975; Munby, 1983, 1997).

This problem is even more pronounced in the Spanish context. Studies conducted in Spain about the factors influencing students' engagement and interest in science mainly focused on attitudes toward science construct (e.g. Marbá-Tallada & Márquez Bargalló, 2010; Pérez Manzano & de Pro Bueno, 2018; Vázquez-Alonso & Manassero-Mas, 2011). Although these studies contributed to the diagnosis of students' unfavorable attitudes towards science, the instruments used have severe limitations when faced with modern psychometric standards, with a great absence of validity and reliability evidence, or with practices in the development of the instrument not consistent with current standards (Toma & Meneses Villagrà, 2019a). For example, studies whose sample includes students from elementary school have mostly used the ROSE (e.g. Marbá-Tallada & Márquez Bargalló, 2010; Vázquez-

Contribution of this paper to the literature

- Further evaluation of the psychometric properties of a promising expectancy of success instrument.
- The SUCCESS instrument has strong content, construct and criterion validity, and internal consistency reliability.
- This study paves the way for future research related to the measurement of expectancy of success in school science.

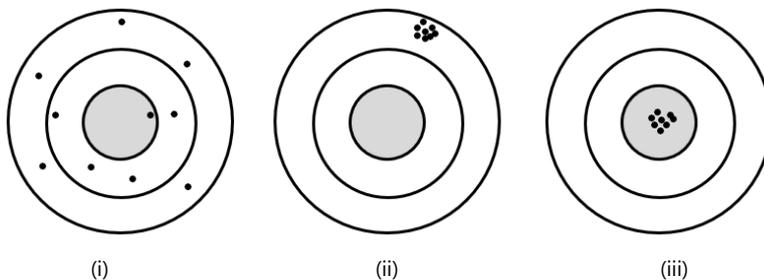


Figure 1. An illustration of validity and reliability, based on Linn and Gronlund (1995). (i) Neither valid nor reliable; (ii) Reliable but not valid; (iii) Both valid and reliable

Alonso & Manassero-Mas, 2011; Vázquez-Alonso & Manassero-Mas, 2008) or the PANA questionnaire (e.g. de Pro & Pérez Manzano, 2014; Pérez Manzano & de Pro Bueno, 2018) without providing information on their psychometric properties (Toma & Meneses Villagr a, 2019a). When validity evidence was provided, there were critical methodological misapplications, like retaining items with cross loadings between factors (i.e. Vázquez-Alonso & Manassero-Mas, 2004), that undermined the quality of the results reported.

In response to this lack of valid instruments, and because persistence has been linked with self-efficacy and expectancies of success (Ceci & Williams, 2007; Chachashvili-Bolotin, Milner-Bolotin, & Lissitsa, 2016; Guo, Parker, Marsh, & Morin, 2015; Sellami, El-Kassem, Al-Qassass, & Al-Rakeb, 2017), Toma and Meneses Villagr a (*in press*) have developed a short instrument to measure elementary school Spanish students' expectancies of success in school science (SUCCESS). Given Blalock et al. (2008) concerns about the lack of validity and reliability of measurement instruments used in science education research, who noted that among weak psychometric properties and methodological issues in the development and validation of instruments there was a proliferation of instruments that were only used in a single study with no follow-up psychometric re-analysis, in this study further psychometric analysis of the SUCCESS instrument is provided. More specifically, this study addressed one of the concerns raised by Blalock et al. (2008) about the need to subject measurement instruments to several psychometric evaluations before they can be presented as valid and reliable.

Therefore, in order to establish a cohesive and robust instrument, we follow Blalock et al. (2008) call related to the need for replication studies to generalize findings and refine instruments by submitting them to both exploratory (EFA) and confirmatory factor analysis (CFA). This practice is consistent with the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014) which states that the validation of an instrument is an ongoing process that does not end with the original validation study. Consequently, while Toma and Meneses Villagr a (*in press*) evaluated the psychometric properties of the SUCCESS instrument through different reliability and validity tests, the instrument would benefit of further analysis to establish its validity and reliability beyond the original sample included in the validation study.

BRIEF REVIEW OF PSYCHOMETRIC THEORY

Psychometric theory is concerned with the properties of measurements instruments used in social sciences and involves the development and refinement of approaches to measure latent traits (Nunnally & Bernstein, 1994). The two aspects most relevant to psychometric assessment are validity and reliability (Figure 1). Validity relates to whether an instrument measures the trait it should be measuring, and reliability focused on the extent to which an instrument is free of measurement error (Mokkink et al., 2010).

The validity of an instrument is determined upon three types of validity indices (Mokkink et al., 2010a; Mokkink et al., 2010a, 2010b). Firstly, *content and face validity* examine if the construct is adequately represented by the included items in the instrument, and whether these items are correctly interpreted and understood by the population under study. Secondly, *construct validity* determines if the instrument is consistent with theoretical expectations and with initial hypothesis. In other words, an instrument would demonstrate construct validity if it

Table 1. The SUCCESS instrument

Spanish	English
1. Las clases de Ciencias de la Naturaleza me resultan muy difíciles para mí (r)	1. School Science is very hard for me (r)
2. No soy un estudiante al que se le dé bien Ciencias de la Naturaleza (r)	2. I am not a student who does well in school science (r)
3. Soy muy bueno en Ciencias de la Naturaleza	3. I am very good at school science
4. Soy capaz de obtener buenas notas en Ciencias de la Naturaleza	4. I can get good grades in school science
5. Aprendo cosas muy rápidamente en Ciencias de la Naturaleza	5. I learn things quickly in school science
6. Puedo hacer bien las tareas de Ciencias de la Naturaleza	6. I can do my science schoolwork well

(r) refers to items that must be reverse scored

captures the dimensionality of the trait being studied, if it does not measure aspects that should not be measuring, and if it can discriminate between groups that have been hypothesized to be different regarding the focal trait (Polit, 2015). Lastly, *criterion validity* measures if the proposed new instrument is consistent with existing instruments that can be considered “gold standards” in the measurement of that particular trait under study (Polit & Yang, 2016). However, in the absence of a gold standard, it should be examined to what extent the new instrument is correlated to external measures of constructs with which conceptual convergence is expected (Polit, 2015).

The reliability domain measures mainly two aspects related to the internal consistency and the measurement error of change scores. Internal consistency examines the degree to which “(...) the items on a scale are measuring the same underlying construct” (Polit, 2015, p. 1747). The measurement error measures the variation in a set of scores to examine the extent to which “(...) scores are stable and reproducible” (Polit, 2015, p. 1748), with the assumption that a reliable instrument should be able to discriminate between “(...) random temporal fluctuations (unreliability) and true change on an attribute” (Polit, 2015, p. 1749). Likewise, a reliable instrument would correctly measure a trait that has not changed over time.

THE SUCCESS INSTRUMENT

More than two decades ago, Bandura (1977, 1997) proposed in his social cognitive theory that self-efficacy is one of the key motivational aspect that affects individual’s choice, achievement, and persistence (Bandura, Barbaranelli, Caprara, & Pastorelli, 2001). Bandura defined self-efficacy as individuals’ beliefs in their own ability to effectively perform a given task or problem. He distinguished between “outcomes expectations”, which are individuals’ beliefs about whether a specific behavior or task will lead to a certain result or outcome (e.g. studying daily will improve science achievement), and “efficacy expectations”, which relates to individual’s beliefs about whether he/she will be able to perform that specific behavior or task (e.g. being able to study long enough to improve science achievement).

Closely related to Bandura’s (1997) theory is Eccles and colleague’s expectancy-value model of achievement motivation (EVT, Eccles et al., 1983; Eccles, Wigfield, Harold & Blumenfeld, 1993; Eccles & Wigfield, 1995; Wigfield & Cambria, 2010; Wigfield & Eccles, 1992, 2000, 2002). The EVT model posits that choice, motivation and persistence in a given task are directly influenced by individuals’ expectancies of success and the values related to that specific task. The first key construct of the theory is the “expectancies of success”, defined by Eccles et al. (1983) as an individual’s beliefs about the success he will have in a given task or activity, whether in the immediate future or in the long term. Although conceptually the expectancies of success refers to individual’s internal beliefs related to both the ability to perform a task in the present and in the future, this differentiation has not been yet empirically distinguished (Eccles et al., 1993; Wigfield & Eccles, 2000).

The SUCCESS instrument is rooted in the Expectancy-Value (EVT) model of achievement motivation postulated by Eccles and colleagues (Eccles et al., 1983, 1993). More specifically, this instrument is composed by a total of six Likert type items that focuses on the first key construct of the EVT theory, called “expectancies of success”. Consistent with the EVT model, it is postulated that individuals will tend to become involved and persist in those tasks or activities in which they feel highly competent. Therefore, high levels of expectancies of success among young students may lead to the development of positive beliefs about their ability to succeed in future science courses or subjects, which may promote the selection of additional science subjects in secondary education or extracurricular activities in science, and eventually the enrolment in a science-related career (Andersen & Ward, 2014). **Table 1** displays SUCCESS items in both English and Spanish language.

Previous Psychometric Evaluation of the SUCCESS Instrument

The SUCCESS instrument has been validated by addressing concerns that have been highlighted as compromising the validity and reliability of many existing instruments in science education (Aydeniz & Kotowski, 2014; Blalock et al., 2008). Thus, the items included in this instrument have been selected from an initial pool that was examined by a panel of experts composed of science educators and psychology university professors and

elementary education teachers, and have been tested through think-aloud interviews (Beatty & Willis, 2007) with elementary education students, which provides strong evidence of content and face validity.

Subsequently, the instrument was administered to 418 Spanish elementary school students and the construct validity and reliability of the instrument were examined. The results of exploratory factor analyses reported an unidimensional structure consistent with the theoretical expectations postulated by Eccles et al. (1983). Factorial loadings were higher than .40 and with satisfactory reliability indices (Cronbach $\alpha = .761$), given the young nature of the sample studied (Newman & McNeil, 1998). To date, this is the first instrument in Spanish that has been validated following modern psychometric recommendation for the specific measurement of elementary school Spanish students' expectancies of success in school science.

Therefore, while the original validation study provided evidence for face, content and construct validity, as well as for the reliability of the SUCCESS instrument, in this study, the psychometric properties of the SUCCESS instrument are further examined by submitting it to construct validity through CFA, criterion validity using external measures, and various reliability indices.

METHOD

Sample

The sample was drawn from 15 elementary schools situated in the province of Burgos, located in the north of Spain. The initial sample was composed of 313 students, however, since goodness of fit statistics are known to be sensitive to abnormal data, Mahalanobis distance was used to detect multivariate outliers. After excluding 32 questionnaires identified as multivariate outliers, a final sample size of 281 elementary school students enrolled in 4th ($n = 66$), 5th ($n = 96$) and 6th grades ($n = 119$) was established. Half of the sample size were girls (50.2%), and the mean age of the participants was 10.35 years old ($SD = 1$, range 8-12 years).

The sample size reached approximately a 47:1 distribution (47 valid cases for each item included in the instrument). Therefore, the sample size is well above the minimum criterion set at 5:1, 10:1 and 20:1 in the literature for factor analysis (de Winter & Dodou, 2012).

Measures

The Spanish-School Science Attitude Survey (S-SSAS) adapted by Toma and Meneses Villagr  (2019a) from the Kennedy, Quinn, & Taylor (2016) instrument was used as an external measure. The S-SSAS is a short instrument composed of 7 Likert-type items and 3 semantic differential items that measures the six most common attitudes toward school science constructs explored in the literature. Four constructs (intention to enrol in future school science, enjoyableness of school science, perceived difficulty of school science, and perception of self-efficacy in school science) are examined through single-item measures, and the remaining two constructs (usefulness of school science, and relevance of school science) are measured through two and four items, respectively. **Table 2** displays S-SSAS instrument items in both English and Spanish language.

In the original validation study (Toma & Meneses Villagr , 2019a) the S-SSAS reported a Cronbach α of .704, with item-total correlations ranging from .243 to .560. In subsequent studies (Toma & Meneses Villagr , 2019b), the S-SSAS instrument reported a Cronbach α of .763, with item-total correlation ranging from .313 to .552. For the present study, the Cronbach α was .758 and item-total correlation ranged from .341 - .596. Based on these results, it can be concluded that the S-SSAS is a reliable instrument that can be used as an external measure for computing the criterion validity of the SUCCESS instrument. It is worth noting that both S-SSAS and SUCCESS instruments are concerned with the school science domain and not with science in general, which is consistent with the literature suggesting that school science may be a better predictor of students' interest and future enrolment in science studies than science in general (Osborne et al., 2003).

Table 2. The Spanish School Science Attitude Scale (S-SSAS)

(I) Intention to enroll in further science	
1. I am very likely to enroll on a science course in Year 11	1. Es muy probable que me apunte a Ciencias de la Naturaleza en la ESO
(E) Enjoyableness of school science	
2. I think science is (boring – fun)	2. Pienso que Ciencias de la Naturaleza es (aburrida – divertida)
(D) Perceived difficulty of school science	
3. I struggle with completing the assignments for science class	3. Me cuesta terminar las tareas para la clase de Ciencias de la Naturaleza
(S) Perception of self-efficacy in school science	
4. I think I am very good at science	4. Pienso que soy muy bueno en Ciencias de la Naturaleza
(U) Usefulness of science to careers	
5. A job as a scientist would be interesting	5. Un trabajo como científico sería interesante
6. For my planned career, knowledge of school science will be (SD worthless – required)	6. Para mis futuros estudios, el conocimiento de las clases de Ciencias de la Naturaleza es (inútil – útil)
(R) Relevance of school science	
7. Science helps to make life better	7. La ciencia ayuda a mejorar la vida
8. I want to learn about plants in my area	8. Quiero aprender sobre las plantas de mi entorno
9. For my everyday life, I think school science is (irrelevant – relevant)	9. Para mi vida diaria, creo que Ciencias de la Naturaleza es (poco importante – muy importante)
10. I want to learn about electricity and how it is used in the home	10. Quiero aprender sobre la electricidad y saber cómo se usa en una casa

Data Analysis

Construct validity

Construct validity was examined using CFA, which is a structural equation modelling technique (Brown, 2006). Exploratory factor analysis (EFA) aims to find a model that fits the data, so different alternative models are specified until establishing a final model that fits the data and that has theoretical support (Schumacker & Lomax, 2010). However, the CFA approaches seek to statistically verify the significance of the factorial model that has been previously established through EFA, thus providing more evidence of construct validity (Schumacker & Lomax, 2010). Therefore, CFA is appropriate when the structure of the latent construct is known based on theory, empirical research, or both (Byrne, 2010), just as in the case of the SUCCESS instrument.

For the CFA, the maximum likelihood estimation method (ML) was used through the AMOS v.23 software (Arbuckle, 2014). Following consensual recommendations in the literature (Brown, 2006; Hair, Black, Babin, Anderson, & Tatham, 2010; Kline, 2005), the following six fit indices were examined to establish model fit:

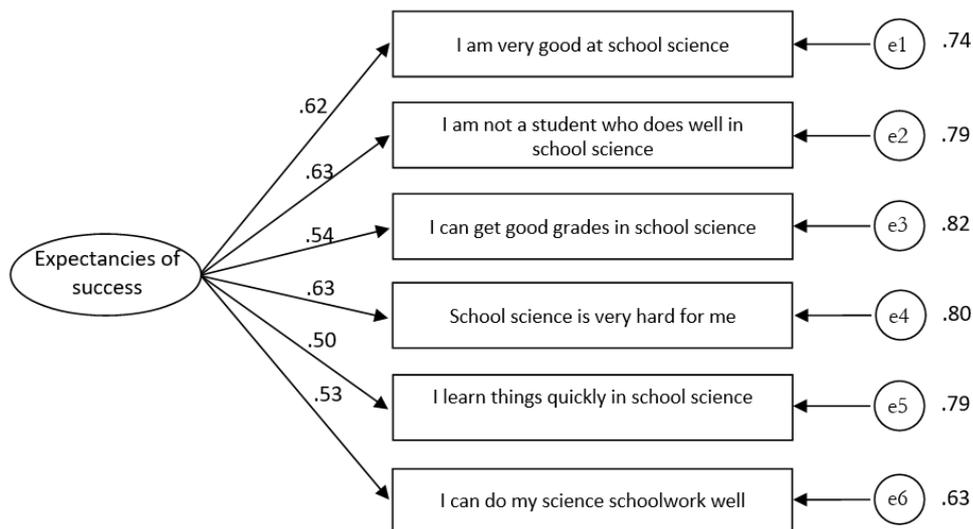
- Chi-square (χ^2). The χ^2 measures the extent to which the variance and covariance matrices differ between the observed and the implicit model. A non-significant p value is desired, however, χ^2 is sensitive to sample size, so a sample greater than 200 cases drastically increases the probability of obtaining a significant p value.
- Comparative Fit Index (CFI). The CFI indicates the degree of fit of the proposed model compared to a null model, based on the sample size. CFI values of $\geq .90$ indicate an appropriate fit.
- Goodness of Fit Index (GFI). The GFI measures the relative amount of variance and covariance that is explained by the proposed model. GFI values of $\geq .90$ indicate an appropriate fit.
- Adjusted Goodness of Fit Index (AGFI). The AGFI is similar to the GFI, however it measures the relative amount of variance and covariance adjusted to the degrees of freedom of the proposed model. AGFI values of $\geq .90$ indicate an appropriate fit.
- Root Mean Square Residual Index (RMR). The RMR index represents the average value of all standardised residuals. RMR values of $\leq .05$ indicate an appropriate fit.
- Root Mean Square Error of Approximation (RMSEA). The RMSEA examines the degree of error in the proposed model considering the sampling approximation error. RMSEA values of $\leq .08$ indicate an appropriate fit.

Criterion validity

The Pearson correlation coefficient was used to examine the correlation between the SUCCESS instrument and six other external measures with which conceptual convergence is expected, using the statistical package SPSS v. 24 (IBM, 2016).

Table 3. Model fit indices

Fit indices	Recommended values	1 factor model	Evaluation
χ^2	$p > .05$	$\chi^2 = 15.274, p = .054$	Adequate
CFI	$\geq .90$.980	Optimal
GFI	$\geq .90$.984	Optimal
AGFI	$\geq .90$.958	Optimal
RMR	$\leq .05$.034	Optimal
RMSEA	$\leq .08$.054	Optimal

**Figure 2.** Confirmatory factor analysis results

Reliability

The reliability of the SUCCESS instrument was examined through the Cronbach alpha coefficient with a minimum criterion of $\alpha > .70$ (Nunnally & Bernstein, 1994), and through item-total correlation with a minimum criterion of $r > .30$ using the statistical package SPSS v. 24 (IBM, 2016).

FINDINGS

Construct Validity

CFA results revealed that the hypothesized single-factor model obtained an optimal fit in five of the six model fit indices. The index obtaining only adequate fit was χ^2 with a value of $p = .054$, although this is a parameter that is affected by the sample size, as indicated previously.

The degree of fit of the proposed model was adequate, as shown by the CFI value greater than .90. The relative amount of variance explained by the model was also adequate, obtaining values greater than .90 in both GFI and AFGI indices. The fit of the model was also appropriate when considering the overall amount of error, with the RMR reaching a value lower than .05 and the RMSEA a value lower than .08. These results are reported in [Table 3](#).

All items loadings were statistically significant. As can be seen in [Figure 2](#), the lowest value was obtained for the item "I learn things quickly in school science", with a value of $\lambda = .50$. The strongest loadings were obtained for both reversed items "I am not a student who does well in school science" and "School science is very hard for me", with a value of $\lambda = .63$. These results confirm the initial hypothesis that the SUCCESS instrument consists of a unidimensional scale that measures the expectancies of success in school science.

Criterion Validity

Pearson correlation coefficient revealed that the SUCCESS instrument was significantly correlated to six external attitudes toward school science measures. The correlation ranged from $r = .193$ to $r = .740$. More specifically, a significant, small and positive correlation was found between students' expectancies of success and their

Table 4. Pearson correlation coefficient between the SUCCESS instrument and external measures

External measures	SUCCESS instrument
Intentions to enrol	.193 ¹
Enjoyableness	.353
Difficulty	-.485
Auto-efficacy	.740
Perceived utility	.201
Perceived relevance	.322

¹All correlation were significant at $p < .01$ level

Table 5. Reliability results

SUCCESS items	Item-total r	Cronbach α after removing items
I am very good at school science	.602	.721
I am not a student who does well in school science (r)	.598	.720
I can get good grades in school science	.458	.757
School science is very hard for me (r)	.493	.748
I learn things quickly in school science	.446	.759
I can do my science schoolwork well	.532	.739

(r) refers to items that must be reverse scored

intentions to enroll in further science courses. A statistically significant, large and positive correlation between students' expectancies of success and self-efficacy and a significant, moderate and negative correlation between students' expectancies of success and their perceived difficulty of school science was identified. Finally, a significant, moderate and positive correlation was found between students' expectancies of success and their perceived relevance of school science, and perceived usefulness of school science. These results are reported in [Table 4](#).

Reliability

Cronbach's α was .775 for the entire instrument, with ranges between .720 and .759 if items were deleted. Since the Cronbach α does not improve considerably after deleting items, it means that all items included in the instrument are relevant and measure the same underlying construct. These results, which are above the minimum .70 criteria required for exploratory studies (Nunnally & Bernstein, 1994), coincide with the original validation study.

In addition, the items included in the SUCCESS instrument are highly correlated with each other, with item-total correlation ranging from .446 to .602. These results also satisfy the minimum criterion of $r > .30$, and likewise coincide with those obtained in the original validation study. Taken together, these results provide additional evidence of the internal consistency of the SUCCESS instrument, confirming that it is a instrument with satisfactory levels of reliability. [Table 5](#) shows the internal consistency results of the SUCCESS instrument.

DISCUSSION

This study provides further evidence on the psychometric properties of the SUCCESS instrument. In relation to reliability, the internal consistency results are in line with the results of the original validation study and are above minimum recommended values for exploratory studies (Nunnally & Bernstein, 1994). Regarding validity evidences, the structural validity of the SUCCESS instrument was confirmed with fit model indices results above the recommended values in the CFA literature (Brown, 2006; Hair et al., 2010; Kline, 2005). Likewise, evidence of criterion validity has been provided using external measures that have already been validated in Spanish and postulated as robust instruments. In this regard, the extent to which the SUCCESS instrument correlates with a total of six external measures with which conceptual convergence was expected was assessed.

Firstly, given that expectancies of success greatly influence individuals' decisions regarding career choice and persistence (Lent & Brown, 1994; Lent, Brown, & Hackett, 2002), it was hypothesized that the SUCCESS instrument would be positively related to students' intentions to enroll in future school science courses, assumptions that have been confirmed in this study.

Secondly, since expectancies of success are defined as individuals' beliefs about the success of performing a given task or activity (Eccles et al. 1983), it was hypothesized that the SUCCESS instrument will be negatively related to students perceived difficulty of school science and positively related to students' self-efficacy in studying school science. The results of this study are also in line with these theoretical assumptions.

Finally, since participation in tasks or activities that require knowledge and skills application in demanding situations enhances students' self-efficacy (Van Dinther, Dochy, & Segers, 2011), and since enjoyment, interest and satisfaction have been identified as factors influencing student self-efficacy (Hutchison, Follman, Sumpter, & Bodner, 2006), it was hypothesized that the SUCCESS instrument will be positively related to students enjoyment, perceived usefulness, and perceived relevance of school science. These initial theoretical expectations have also been confirmed in this study.

Taken together, these results confirm that the SUCCESS is a valid and reliable instrument with great construct and criterion validity, as well as with structural and content validity, as assessed and reported in the original validation study (Toma & Meneses Villagr , *in press*).

IMPLICATIONS AND AVENUES FOR FUTURE RESEARCH

Several implications are derived from this study. To the best of the author's knowledge, the original study (Toma & Meneses Villagr , *in press*) and this follow-up study represent the first efforts to develop and validate an instrument that specifically tackle Spanish-speaking elementary school students' expectancies of success in school science. Since both studies show that the SUCCESS instrument is valid and reliable according to modern psychometric standards, it can be used by science teachers and researchers to examine whether different activities and teaching methodologies are successful at improving students' expectancies of success in school science. Therefore, the SUCCESS instrument could be used in studies aimed at the development of learning strategies for the improvement of students' expectancies of success in school science, which could tackle students decline in interest toward school science from early stages where scientific vocations can be more easily fostered (Lindahl, 2007).

Another implication derived is related to the need to develop large-scale studies to identify at what stage of elementary education students experience a decline in their expectancies of success. Also, given the short administration time of the proposed instrument, it could be used in conjunction with other relevant measures to identify those variables that may influence students' expectancies of success.

Additionally, although expectancies of success have been linked to persistence (Chachashvili-Bolotin et al., 2016; Guo et al., 2015), it is not clear to what extent high levels of expectancies of success may be related to an actual increase in students desire and interest in enrolling in science courses at later stages of the educational system, aspect that requires future empirical research. In addition, given that gender has been identified as a critical variable in the development of interest and attitudes toward science (Osborne et al., 2003; Sellami et al., 2017; Toma & Greca, 2018), it would be interesting to test in future validation studies the measurement invariance of the SUCCESS instrument and to compare its latent factor mean in girls and boys enrolled in elementary education. This would guarantee the integrity of the results by reducing the possibility of obtaining contradictory results derived from different factorial structures of the SUCCESS instrument in girls and boys.

Furthermore, the test-retest reliability of the SUCCESS should be examined in future research, which will allow the use of this instrument in longitudinal studies that requires several data collection phases. Finally, future studies should focus on validating this instrument for middle and high school students, which would allow cross-sectional and longitudinal studies focusing on those stages of the Spanish educational system in which a greater decrease in students interest in science has been identified (MECD, 2016).

CONCLUSIONS

The results of this second psychometric evaluation of the SUCCESS instrument reveal that the proposed instrument is conceptually consistent with the "Expectancies of success" construct first introduced by Eccles et al. (1983), and methodologically robust in terms of modern validity and reliability psychometric evidences.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME, 2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Andersen, L., & Ward, T. J. (2014). Expectancy-value models for the STEM persistence plans of ninth-grade, high-ability students: A comparison between black, hispanic, and white students. *Science Education, 98*(2), 216–242. <https://doi.org/10.1002/sce.21092>
- Arbuckle, J. L. (2014). Amos (Version 23.0). Chicago: IBM SPSS.

- Archer, L., Dewitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2010). "Doing" science versus "being" a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education*, 94(4), 617–639. <https://doi.org/10.1002/sce.20399>
- Aydeniz, M., & Kotowski, M. R. (2014). Conceptual and methodological issues in the measurement of attitudes towards science. *Electronic Journal of Science Education Electronic Journal of Science Education*, 18(3), 1–24.
- Bandura A, Barbaranelli C, Caprara G.V., & Pastorelli C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, 72, 187–206. <https://doi.org/10.1111/1467-8624.00273>
- Bandura A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287–311. <https://doi.org/10.1093/poq/nfm006>
- Blalock, C. L., Lichtenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. A. (2008). In pursuit of validity: A comprehensive review of science attitude instruments 1935–2005. *International Journal of Science Education*, 30(7), 961–977. <https://doi.org/10.1080/09500690701344578>
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (Vol. 2). New York: Routledge Taylor & Francis Group. <https://doi.org/10.4324/9781410600219>
- Ceci, S. J., & Williams, W. M. (2007). *Why aren't more women in science. Top researchers debate the evidence*. Washington, DC: American Psychological Association. <https://doi.org/10.1037/11546-000>
- Chachashvili-Bolotin, S., Milner-Bolotin, M., & Lissitsa, S. (2016). Examination of factors predicting secondary students' interest in tertiary STEM education. *International Journal of Science Education*, 38(3), 366–390. <https://doi.org/10.1080/09500693.2016.1143137>
- de Pro, A., & Pérez Manzano, A. (2014). Actitudes de los alumnos de Primaria y Secundaria ante la visión dicotómica de la Ciencia. *Enseñanza de Las Ciencias*, 32(3), 111–132. <https://doi.org/10.5565/rev/ensciencias.1015>
- de Winter, J. C. F., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, 39(4), 695–710. <https://doi.org/10.1080/02664763.2011.610445>
- Eccles, J. S., & Wigfield, A. (1995). In the mind of the achiever: the structure of adolescents' academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin*, 21, 215–225. <https://doi.org/10.1177/0146167295213003>
- Eccles, J. S., Wigfield, A., Harold, R. B., & Blumenfeld, P. B. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 64, 830–847. <https://doi.org/10.2307/1131221>
- Eccles, J., Adler, T., Futterman, R., Goff, S., Kaczala, C., Meece, J., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: W. H. Freeman.
- Gardner, P. L. (1975). Attitudes to science: A review. *Studies in Science Education*, 2, 1–41. <https://doi.org/10.1080/03057267508559818>
- Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, 51(8), 1163–1176. <https://doi.org/10.1037/a0039440>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis (Vol 7.)*. Upper Saddle River: Pearson Prentice Hall.
- Hutchison, M. A., Follman, D. K., Sumpter, M., & Bodner, G. M. (2006). Factors influencing the self-efficacy beliefs of first-year engineering students. *Journal of Engineering Education*, 95(1), 39–47. <https://doi.org/10.1002/j.2168-9830.2006.tb00876.x>
- IBM, C. (2016). IBM SPSS Statistics for Windows. Version 24.0. Armonk, NY: IBM Corp.
- Kennedy, J. P., Lyons, T., & Quinn, F. (2014). The continuing decline of science and mathematics enrolments in Australian high schools. *Teaching Science*, 60(2), 34–46.
- Kennedy, J., Quinn, F., & Taylor, N. (2016). The school science attitude survey: A new instrument for measuring attitudes towards school science. *International Journal of Research & Method in Education*, 39(4), 422–445. <https://doi.org/10.1080/1743727X.2016.1160046>

- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Lent, R. W., & Brown, S. D. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45, 79–122. <https://doi.org/10.1006/jvbe.1994.1027>
- Lent, R. W., Brown, S. D., & Hackett, G. (2002). Social cognitive career theory. In D. Brown, L. Brooks, & Associates (Eds.), *Career choice and development* (4th ed., pp. 255–311). San Francisco, CA: Jossey-Bass.
- Lindahl, B. (2007). Longitudinal study of students' attitudes towards science and choice of career. Paper presented at the 80th session of the International Conference of the National Association for Research in Science Teaching. New Orleans.
- Linn, R. L., & Gronlund, N. E. (1995). *Measurement and assessment in teaching*. New Jersey: Prentice-Hall Inc.
- Lyons, T., & Quinn, F. (2010). *Choosing science. Understanding the declines in senior high school science enrolments*. Armidale: University of New England.
- Lyons, T., & Quinn, F. (2015). Understanding declining science participation in Australia: A systemic perspective. In E. K. Henriksen, J. Dillon, & J. Ryder (Eds.), *Understanding student participation and choice in science and technology education* (pp. 153–168). Springer Netherlands. https://doi.org/10.1007/978-94-007-7793-4_10
- Marbá-Tallada, A., & Márquez Bargalló, C. (2010). ¿Qué opinan los estudiantes de las clases de ciencias? Un estudio transversal de sexto de Primaria a cuarto de ESO. *Enseñanza de Las Ciencias*, 28(1), 19–30.
- MECD. (2016). *Datos y cifras del sistema universitario español. Curso 2015/2016*.
- Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., ... de Vet, H. C. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*, 10(22), 1–8. <https://doi.org/10.1186/1471-2288-10-22>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... de Vet, H. C. W. (2010a). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... de Vet, H. C. W. (2010b). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, 19(4), 539–549. <https://doi.org/10.1007/s11136-010-9606-8>
- Munby, H. (1983). Thirty studies involving 'Scientific Attitude Inventory': What confidence can we have in this instrument? *Journal of Research in Science Teaching*, 20, 141–162. <https://doi.org/10.1002/tea.3660200206>
- Munby, H. (1997). Issues of validity in science attitude measurement. *Journal of Research in Science Teaching*, 34(4), 337–341. [https://doi.org/10.1002/\(SICI\)1098-2736\(199704\)34:4<337::AID-TEA4>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1098-2736(199704)34:4<337::AID-TEA4>3.0.CO;2-S)
- Newman, I., & McNeil, K. (1998). *Conducting survey research in the social sciences*. New York: University Press of America.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049–1079. <https://doi.org/10.1080/0950069032000032199>
- Pérez Manzano, A., & de Pro Bueno, A. (2018). Algunos datos sobre la visión de los niños y de las niñas sobre las ciencias y del trabajo científico. *IQual. Revista de Género e Igualdad*, 0(1), 18–31. <https://doi.org/10.6018/iQual.306091>
- Polit, D. F. (2015). Assessing measurement in health: Beyond reliability and validity. *International Journal of Nursing Studies*, 52(11), 1746–1753. <https://doi.org/10.1016/j.ijnurstu.2015.07.002>
- Polit, D. F., & Yang, F. (2016). *Measurement and the measurement of change: A primer for health professionals*. Philadelphia: Lippincott Williams & Wilkins.
- Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, 50(1), 85–129. <https://doi.org/10.1080/03057267.2014.881626>
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling*. Routledge (Vol. 47). New York: Routledge Taylor & Francis Group.
- Sellami, A., El-Kassem, R. C., Al-Qassass, H. B., & Al-Rakeb, N. A. (2017). A path analysis of student interest in STEM, with specific reference to Qatari students. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(9), 6045–6067. <https://doi.org/10.12973/eurasia.2017.00999a>

- Toma, R. B., & Greca, I. M. (2018). The effect of integrative STEM instruction on elementary students' attitudes toward science. *EURASIA Journal of Mathematics, Science & Technology Education*, 14(4), 1383–1395. <https://doi.org/10.29333/ejmste/83676>
- Toma, R. B., & Meneses Villagr a, J. A. (2019a). Validation of the single-items spanish-School science attitude Survey (S-SSAS) for elementary education. *PLoS ONE* 14(1), e0209027. <https://doi.org/10.1371/journal.pone.0209027>
- Toma, R. B., & Meneses Villagr a, J. A. (2019b). Preferencia por contenidos cient ficos de f sica o de biolog a en Educaci n Primaria: un an lisis cluster, *Revista Eureka sobre Ense anza y Divulgaci n de las Ciencias*, 1(1), 1104. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2019.v16.i1.1104
- Toma, R. B., & Meneses Villagr a, J. A. (in press). Development and validation of the SUCCESS instrument: towards a valid and reliable measure of expectancies of success in School Science.
- Van Dinther, M., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research Review*, 6(2), 95–108. <https://doi.org/10.1016/j.edurev.2010.10.003>
- V zquez-Alonso,  ., & Manassero-Mas, M. A. (2008). El declive de las actitudes hacia la ciencia de los estudiantes: un indicador inquietante para la educaci n cient fica. *Revista Eureka sobre Ense anza y Divulgaci n de Las Ciencias*, 5(3), 274–292. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2008.v5.i3.03
- V zquez-Alonso,  ., & Manassero-Mas, M. A. (2004). Imagen de la ciencia y la tecnolog a al final de la educaci n obligatoria. *Cultura y Educaci n*, 16, 385–398. <https://doi.org/10.1174/1135640042802473>
- V zquez-Alonso,  ., & Manassero-Mas, M. A. (2011). El descenso de las actitudes hacia la ciencia de chicos y chicas en la Educaci n Obligatoria. *Ci ncia & Educa o*, 17(2), 249–268. <https://doi.org/10.1590/S1516-73132011000200001>
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: definitions, development, and relations to achievement outcomes. *Developmental Review*, 30(1), 1–35. <https://doi.org/10.1016/j.dr.2009.12.001>
- Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: a theoretical analysis. *Developmental Review*, 12, 265–310. [https://doi.org/10.1016/0273-2297\(92\)90011-P](https://doi.org/10.1016/0273-2297(92)90011-P)
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield y J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91–120). San Diego: Academic Press. <https://doi.org/10.1016/B978-012750053-9/50006-1>

<http://www.ejmste.com>