Revista Latinoamericana de Psicología

http://revistalatinoamericanadepsicologia.konradlorenz.edu.co/

ORIGINAL

A lexical decision task to measure crystallized-verbal ability in Spanish

Graham Pluck[ab]*

[a] *Instituto de Neurociencias, Universidad San Francisco de Quito, Cumbayá, Ecuador*
[b] *The University of Sheffield, United Kingdom of Great Britain and Northern Ireland*

Received 15 May 2019; accepted 29 November 2019

KEYWORDS
Verbal ability,
crystalized intelligence,
lexical decision,
reading,
cognitive assessment

Abstract A classical distinction in cognitive science is between fluid and crystalized abilities. Fluid ability is captured by many common executive function and intelligence tests. Crystalized ability, on the other hand, can be measured quite simply via lexical decision tasks including the English-language Spot-the-Word Test. However, no similar Spanish-language test has been available up to now. This paper presents a Spanish-language Lexical Decision Task that is quick to administer and was tested on sample of 139 normal adult participants. Results indicate that the new test has good internal consistency and test-retest reliability. An analysis of the correlations between this new test and demographic variables, as well as with the subtests of the Wechsler Adult Intelligence Scale suggest that it is a valid measure of crystalized-verbal ability. It also appears to be a brief but valid assessment of intelligence in general, and its positive correlation with academic achievement establishes predictive validity. The new test has the potential to be a useful research tool to rapidly measure reading ability, crystalized-verbal ability, and intelligence in Spanish speaking adults.

Una tarea de decisión léxica para medir la capacidad verbal cristalizada en español

PALABRAS CLAVE
Habilidad verbal,
inteligencia cristalizada,
decisión léxica,
lectura,
evaluación cognitiva

Resumen Una distinción clásica en la ciencia cognitiva es entre las habilidades fluidas y cristalizadas. La habilidad fluida es medida por muchas funcionas ejecutivas y tests de inteligencia. Por otro lado, la habilidad cristalizada puede ser medida sencillamente mediante una tarea de decisión léxica, como en la versión en inglés conocida como Spot-the-Word Test. Sin embargo, hasta ahora no ha habido una versión similar de este test en español. Aquí les presento una Tarea de Decisión Léxica en Español que es de rápida aplicación. Esta fue aplicada en una muestra de 139 participantes, adultos normotípicos. Los resultados indican que este nuevo test tiene buena consistencia interna y confiabilidad test-retest. Los análisis de las correlaciones entre este nuevo test y las variables demográficas, al igual que con las sub pruebas de las Escala de Inteligencia de Wechsler para Adultos, sugiere que es una medida confiable de la habilidad verbal cristalizada.

* Autor para correspondencia.
e-mail: g.c.pluck@gmail.com

https://doi.org/10.14349/rlp.2020.v52.1

También parece ser una breve, pero válida evaluación de inteligencia en general, con validez predictiva establecida por sus correlaciones positivas con el logro académico. Este nuevo test tiene potencial para ser una herramienta útil para medir rápidamente habilidad de lectura, habilidad verbal cristalizada e inteligencia en adultos hispanohablantes.

The use of sophisticated symbolic language is, of course, recognised as a crucial feature that separates human from non-human animals (Berwick, Friederici, Chomsky, & Bolhuis, 2013). Indeed, humans are much more intelligent than any other animals (Premack, 2007), and language seems to underlie, or at least greatly influence, many other aspects of human cognition (Baldo, Bunge, Wilson, & Dronkers, 2010; Pinker, 2010).

That our mental lexicon is an essential part of our human intelligence is highlighted by the fact that vocabulary is almost always found to be the most significant part of general intelligence when compared with other cognitive function tests (Johnson, Bouchard Jr, Krueger, McGue, & Gottesman, 2004; Kan, Wicherts, Dolan, & van der Maas, 2013). Furthermore, lexical ability is a strong predictor of academic achievement at undergraduate level (Pluck, 2019a), suggesting a wide role in intelligent behaviour. Indeed, the mental lexicon of word knowledge, is one of the fundamental bases of what is frequently referred to in psychology as crystalized intelligence (Schipolowski, Wilhelm, & Schroeders, 2014). Crystalized intelligence is contrasted with fluid intelligence in Raymond Cattel's influential theory (Cattell, 1963, 1967, 1973). Supposedly, fluid ability deals with on-line novel problem solving in the present, while crystalized ability is based on learnt verbal information, which is acquired through experience, particularly education (Kan et al., 2013). This classic approach to understanding intelligence, separated into fluid and crystalized components, is supported by modern research (Hampshire, Highfield, Parkin, & Owen, 2012).

In fact, this distinction between fluid and crystalized ability has recently become important in cognitive neuroscience through the identification of two broad, but not overlapping, systems within the human brain. These have been called the 'multiple-demand system' and the 'language system' (Blank & Fedorenko, 2017; Blank, Kanwisher, & Fedorenko, 2014; Woolgar, Duncan, Manes, & Fedorenko, 2018). The distinction is important because it appears that they may be two fundamental parts of the human cognitive system, with language being 'domain-specific' (i.e., processing only linguistic information, and nothing else), and the 'multiple-demand system' being domain general (i.e., involved with multiple, attentional top-down tasks), for example, executive functions and fluid intelligence, irrespective of the type of material (Campbell & Tyler, 2018). These neuroscientific findings, therefore, closely reflect the original fluid and crystalized ability division of Cattell.

On the one hand, there are many available tests for fluid or multiple-demand system processing. Such fluid intelligence-based and executive top-down systems have been intensely studied in psychology; these include tests on planning, working memory, reasoning etc. (see e.g. Delis, Kaplan, & Kramer, 2001). This is partly because such abilities have proven to be essential in understanding many themes in contemporary psychology research including child development (e.g. Blakemore & Choudhury, 2006), educational achievement (e.g. Pluck, Ruales-Chieruzzi, Paucar-Guerra, Andrade-Guimaraes, & Trueba, 2016), and mental illness (e.g. Snyder, Miyake, & Hankin, 2015). On the other hand, crystalized ability, e.g. measurement of the extent of the mental lexicon, has been much less aggressively researched in relation to how we understand the wider psychological context. In addition, whereas many of the executive and general intelligence tests are deliberately made language-free (e.g. Pluck, 2019b), by their very nature, crystalized ability tests are limited by linguistic culture.

This is less of a problem in English speaking countries, where there are multiple, validated crystalized ability tests. However, in Spanish there are fewer options. One very widely used assessment of the language system that can be used in research or for educational or clinical purposes is the Vocabulary subtest, which features in several versions of the Wechsler intelligence tests. This is available in Spanish (Wechsler, 2012), and performance on this test is considered to be a strong measure of crystalized ability (Wechsler, 1999). However, the Wechsler tests have several drawbacks, particularly for Latin American countries; as commercial products they are expensive. Furthermore, administration and scoring are not straight forward; training is required and even then, as participant responses are open-ended, scoring is ultimately subjective, introducing unwanted inter-rater variation in scores. One method which avoids this and has been widely used in Spanish is the Word Accentuation Test, in which word pronunciation is assessed (Pluck, Almeida-Meza, Gonzalez-Lorza, Muñoz-Ycaza, & Trueba, 2017). This is simpler, as the psychologist need only listen for the correct stress pattern. In addition, word pronunciation has been shown to figure very highly on a crystalized verbal ability component with factor analysis (Crawford, Stewart, Cochrane, Parker, & Besson, 1989). However, at its core, word pronunciation is probably controlled by non-declarative procedural ability rather than declarative semantic knowledge, as it is possible to pronounce words without knowing their meaning. Furthermore, there are several problems when using reading aloud as a psychological assessment because performance can be hindered by illiteracy, articulation, or visual difficulties; group testing is not possible; and it penalizes the self-educated who have acquired knowledge through reading but may not know the correct pronunciation in languages with irregular grapheme-phoneme matching (Baddeley, Emslie, & Nimmo-Smith, 1992).

In the English language, a test that avoids all of these problems and has been widely used to measure crystalized-verbal ability is the Spot-the-Word Test (Baddeley et al., 1992; Baddeley, Emslie, & Nimmo-Smith, 1993). This

is a simple lexical decision task in which pairs of words are presented to participants, and the task is to choose which of the two is a real word (one is always a legitimate looking non-word and one is a real English word). This is usually done with visual presentation, but for illiterate participants the words can be read aloud by the psychologist. The test has been shown to have good psychometric properties, and validity was established by its high correlation (r = .861) with an established measure of vocabulary: The Mill Hill Vocabulary Scale (Baddeley et al., 1992).

Despite the widespread clinical and research use of the original English-language Spot-the-Word test, and despite its versatility and simplicity as a measure of crystalized-verbal ability, there is no published version in Spanish available for use in research. There is an experimental version, similar to the Spot-the-Word test, which has been used to measure various aspects of word recognition impairment in Spanish patients with Alzheimer's disease (Cuetos, Arce, Martinez, & Ellis, 2017). In that version, patients and controls were required to identify real words in sets of four words, in which three were distractor non-words. However, as it was an experimental measure, the psychometric properties of the test were not reported.

In this paper, I present a Spot-the-Word type task in Spanish (called the Spanish Lexical Decision Task), and detail its reliability and validity as a measure of crystalized-verbal ability. The aim is to provide a preliminary development of the test, which could be used with adults from the Spanish-speaking community.

Method

Participants

Two different samples were taken. The Development Sample contained 63 participants and was used to establish the basic psychometric properties of the Spanish Lexical Decision Task. All participants were recruited from the Universidad San Francisco de Quito (Ecuador) community, but this included friends and family members of people already recruited. The aim was to recruit a diverse sample of people so that the natural range of performance would be present in the overall data set. Although ten undergraduate students were included, the researchers also recruited one mature student, three postgraduate students, ten members of the cleaning staff, ten security guards, and five university professors. The rest were from a range of professions including psychologists, schoolteachers, maintenance staff, clerical staff, etc. The mean age of this sample was 34.13 (SD = 11.93, range = 18-63), and 32/63 (51%) were female. The mean years of education was 14.27 (SD = 4.10, range = 6-26). The majority, 54/63 (85.71%) identified as being of Mestizo ethnicity. The whole sample used Spanish as their primary language.

The second sample was called the Test Sample. The sample was primarily included in order to examine the ability of the Spanish Lexical Decision Task to predict real-world performance, e.g. academic achievement. Thus, the researchers recruited a sample of 75 undergraduate students. The mean age of this sample was 20.56 (SD = 3.43, range = 17-41), and the mean years of education was 14.47 (SD = 1.60, range = 13-22). The majority of this sample were female (56/75, 74.67%) and identified as of Mestizo ethnicity (61/75, 81.33%). All used Spanish as their primary language (studying for degrees taught in Spanish). They were studying for a range of majors; however, the most common subject was psychology (21/75 participants, 28.00%).

Assessments

The Spanish Lexical Decision Task was specifically designed for this study. The original English-language Spot-the-Word Test used 58 pairs of words in which one word is real, and one looks like an English word but is not. This allows for items to be successfully guessed about 50% of the time. Instead, the current researchers opted to use one real word together with two false words for each item, which would, therefore, reduce the guessing rate to about 33%.

The researchers first selected 45 real Spanish words. These were all nouns or adjectives and were selected from the list of modern Spanish word frequencies provided by the Real Academia Española (Real Academia Española). Words were selected to represent a wide-range of frequencies in modern Spanish, from occurrences of once per 2015 words (hierro) to once per 566577 words (hespéride). No attempt was made to choose from prespecified semantic categories. The researchers then produced 90 non-real words that conform to Spanish spelling rules, but do not actually exist. These were then checked in dictionaries and general internet searches to eliminate any words which did exist, for example as obscure words, or proper nouns. These 90 non-words were randomly used to produce 45 triplets of one real word with two non-real foils. These 45 triplets were ordered based on the frequency of the real word (from high to low). Examples from this list can be seen in Table 1. Each triplet containing one real word and two foils was presented in a single line on an A4 page, with about 12 triplets on each page (i.e. over four pages). The words were presented in lower-case sans-serif font, size 14.

To assess the relationship between ability on the Spanish Lexical Decision Task and intelligence the researchers also used the Wechsler Intelligence Scale-IV Spanish Version. This is the most commonly used intelligence test in Spanish-speaking countries. The Spanish-language version of the WAIS-IV was normed in Spain and published in 2012 (Wechsler, 2012). Factor analytic studies suggest that the WAIS-IV is a strong measure of general intelligence (Canivez & Watkins, 2010). The full administration involves ten subtests and can take around 90 minutes. However, the current researchers used a validated abbreviated version that is commonly used in research. This seven-subtest version can be administered in around 40 minutes and contains the following sub-tests, Block design, Similarities, Digit span, Arithmetic, Information, Digit symbol, and Incomplete figures. It is almost as good as the full-test in estimating full-scale IQ, with the seven-test version correlating r = .99 with the full administration (Meyers, Zellinger, Kockler, Wagner, & Miller, 2013).

Procedure

All participants were recruited and tested at a private university in Quito, Ecuador. All participants were assessed by research assistants in a private interview room. The

WAIS-IV was administered as per the standard instructions. For the Spanish Lexical Decision Task, the pages were placed in front of the participant and they were given a pen. They were instructed to circle the real word in each triplet.

The Development sample (n =63) were paid US$20 for participation. Forty-three participants in this group performed both the Spanish Lexical Decision Task and the WAIS-IV on the initial assessment. Twenty participants were involved in the test-retest arm of the project and performed only the Spanish Lexical Decision Task (all 45 items) and some other cognitive tests not reported in this paper. The test and retest sessions were separated by four weeks and each participant was assessed by the same researcher and in the same location on both occasions. However, seven of the participants from the test-retest arm were asked to return for a third visit, when they completed the WAIS-IV. Therefore, the full Development sample was composed of 63 participants, 50 of whom completed both the Spanish Lexical Decision Task and the WAIS-IV.

The Test sample were all undergraduate students. These 75 participants were all assessed with both the Spanish Lexical Decision Task and the WAIS-IV. The assessments were performed in a quiet interview room at the University. The participants were not paid for participation but were given course credits.

Research Ethics

All participants provided written informed consent to participate, in accordance with the ethics committee approved research protocols.

Statistical Analyses

Scores on the WAIS-IV were converted to age-corrected standard scores to calculate the full-scale IQ. Raw scores were used for the Spanish Lexical Decision Task. Distributions of scores were checked with Shapiro-Wilk tests. Parametric procedures (e.g. Pearson correlations, r) were used whenever the data did not differ from the normal distribution based on the Shapiro-Wilk tests. For data that was not normally distributed, non-parametric alternatives were employed (e.g. Spearman's rank correlation, r_s). For all correlational analyses, 95% confidence intervals of the r values were estimated using boot strapping with 1,000 iterations. All analyses are two-tailed with a significance threshold of 0.05. To estimate effect sizes, the researchers used the conventions proposed by Cohen (1992).

Results

These first analyses are all limited to the Development Sample (general population). The researchers first examined the internal consistency of the full 45-item version Spanish Lexical Decision Task with Cronbach's alpha. This produced an initial alpha of 0.811, which would be considered 'good' internal consistency by standard qualitative interpretations. However, four items had zero variation as they were answered correctly by all participants. Furthermore, several items had low item-total correlations. The analysis followed Clark and Watson (1995) who suggest that item-total correlations should, in general, be greater than 0.15. These were removed sequentially starting from the lowest values until all the remaining items had item-total correlations of greater than 0.15. This produced a set comprised of 33 items. However, the researchers randomly added back in three of the four items with perfect scores in order to extend the range at the lower end of performance. These 36 items were used to form the final version of the Spanish Lexical Decision Task. The Cronbach's alpha of this final version was 0.846, which would be considered 'good' internal consistency. The final version is shown in Table 1, which includes the mean scores, item:total correlations, and the resulting Cronbach's alpha values for each item if removed.

The sequence of items in the final 36-item version is the same as in the original 45-item version, except that two of the very easy items were moved to later in the sequence. This is so that people with generally poor performance may be less disheartened by the progressive difficulty of the task as they will encounter items that they will probably know. The positions of the correct responses were altered slightly so that the correct word appears equally often in each of the three columns.

Next, the researchers examined the distribution of the scores on the final version. As there are 36 items, the theoretical range of scores is 0-36; however, as each item is composed of one real word and two foils, the guessing rate is one-third, so effectively the score range would be expected to be from about 12 to 36 (a range of 24 points). In fact, the lowest score in the development sample was 13/36 (achieved by one participant, 1.59% of the sample) and the highest score was 36/36 (achieved by two participants, 3.17% of the sample). This suggests a small ceiling effect; however, the participants in this sample are probably somewhat above average in terms of education (mean 14.27 years). The mean Spanish Lexical Decision Task score was 26.92 (SD = 5.63), which was slightly higher than the median score of 26, suggesting some positive skewing within the distribution. A histogram showing the score distribution is shown in Figure 1 (upper panel). Despite the slight positive skew, the score distribution did not differ significantly from normal (Shapiro-Wilk Test (63) = .969, p = .118).

Twenty of the participants in the development sample were tested twice on the Spanish Lexical Decision Task. For the final 36-item version, the correlation between these two administrations was high, r = .890, p < .001, 95%CI of r = .752 - .960, suggesting 'good' test-retest reliability. The mean test score was 25.35 (SD = 4.58), and at retest it was 26.00 (SD = 5.05) showing only a small practice effect of 0.65 points, and this change was not significant, $t(19)$ = -1.264, p = .222. To quantify the magnitude of this practice effect, the researchers calculated Cohen's d (Cohen, 1992), which was d = 0.283. This would be considered a qualitatively 'small' practice effect.

Still on the topic of the Development Sample, the researchers next examined the association of Spanish Lexical Decision Task scores with demographic variables. There was no significant correlation between age and task performance, r_s = .103, p = .422, 95%CI of r = -.177 - .350. This is consistent with findings showing that crystallised verbal ability of adults is little effected by age and remains stable into at least the eighth decade of life (Park, 2000). This lack of correlation also suggests that the Spanish Lexical Decision Task can, in general, be used in research without

Table 1 Items included in the final version of the Spanish Lexical Decision Task, the mean scores in the Development Sample, and their relationship to Cronbach's alpha variables

Item				Mean (+variance)	Corrected Item:total Correlation	Alpha if item removed
1	**hierro**	fabrasión	lliure	1.000 (0.000)	0.000[1]	0.847
2	**zapato**	fidesión	bausitar	0.953 (0.046)	0.241	0.844
3	adalberos	anacletosión	**nicotina**	0.953 (0.046)	0.282	0.844
4	**insigne**	esnaola	solsonasión	0.921 (0.074)	0.343	0.842
5	roselición	corpor	**rodeos**	0.968 (0.031)	0.359	0.843
6	rafisos	caguan	**jovial**	0.873 (0.113)	0.238	0.844
7	estuveso	**cohecho**	viparas	0.619 (0.240)	0.315	0.843
8	**ciénaga**	gomaración	zolaso	0.714 (0.207)	0.470	0.838
9	boloñaso	querolesión	**adusto**	0.524 (0.253)	0.477	0.837
10	iruretación	alavedra	**cubículo**	0.952 (0.046)	0.419	0.842
11	**terraplén**	chocor	balbicito	0.762 (0.184)	0.467	0.838
12	mipar	**toga**	susasión	0.857 (0.124)	0.457	0.839
13	tamapal	amanciosión	**alabastro**	0.857 (0.124)	0.423	0.840
14	bacomer	**bechamel**	cemiso	0.524 (0.253)	0.508	0.836
15	jerfresa	caprileción	**acritud**	0.698 (0.214)	0.229	0.845
16	córquibo	**aforismo**	pabonesión	0.841 (0.136)	0.423	0.840
17	**lumbares**	enojioso	vorgovés	0.921 (0.074)	0.299	0.843
18	simogeso	pregosión	**saúco**	0.698 (0.214)	0.249	0.845
19	**tórrido**	gorasde	macoriso	0.953 (0.046)	0.187	0.845
20	efesación	**lumpen**	iralaso	0.365 (0.236)	0.306	0.843
21	camarasación	**monacal**	mururoso	0.587 (0.246)	0.476	0.837
22	linoges	**brillo**	citrenos	1.000 (0.000)	0.000[1]	0.846
23	roucosación	valviso	**flexor**	0.778 (0.176)	0.323	0.842
24	**fístula**	remisoción	biljanaso	0.873 (0.113)	0.397	0.841
25	sarres	cliones	**mustia**	0.730 (0.200)	0.331	0.842
26	paineso	**gamo**	puren	0.571 (0.249)	0.391	0.840
27	**incruento**	flagleresión	dincotes	0.635 (0.236)	0.264	0.844
28	poblaso	**lada**	puskas	0.460 (0.252)	0.383	0.841
29	anergiaso	cristicosión	**macarra**	0.619 (0.240)	0.211	0.846
30	**misógino**	ucedistasión	ortuelaso	0.778 (0.176)	0.216	0.845
31	bernaolaso	**sesudo**	curitibas	0.540 (0.252)	0.472	0.838
32	paduación	**adagio**	boitel	0.801 (0.157)	0.344	0.842
33	**novato**	osoritación	deglanes	1.00 (0.000)	0.000[1]	0.846
34	julenoción	larderes	**sínodo**	0.635 (0.236)	0.518	0.836
35	**ferrado**	viamonte	profepaso	0.460 (0.252)	0.196	0.847
36	delanoso	**hespéride**	fayes	0.492 (0.254)	0.450	0.838

The correct responses in each trial are shown in bold. [1]indicates items which showed zero variance in the Development Sample (all participants were correct).

Table 2. Correlations of scores on the Spanish Lexical Decision Task with WAIS-IV subtest scores

	Correlation with The Spanish Lexical Decision Task scores			
	Development Sample		Test Sample	
	n = 50		*n* = 75	
WAIS-IV Subtest	*r*	95%CI of *r*	*r*	95%CI of *r*
Block design	.506***	.239 - .670	.091	-.133 - .306
Similarities	.674[1,***]	.507 - .808	.254*	.054 - .429
Digit span	.646[1,***]	.443 - .787	.146[1]	-.087 - .353
Arithmetic	.620***	.429 - .785	.120[1]	-.138 - .349
Information	.771[1,***]	.616 - .865	.494[1,***]	.288 - .640
Digit-symbol	.508***	.330 - .696	.052	-.162 - .267
Incomplete figures	.598***	.410 - .755	.105	-.165 - .341

[1]Speraman's Rho, *p < .05, ***p < .001

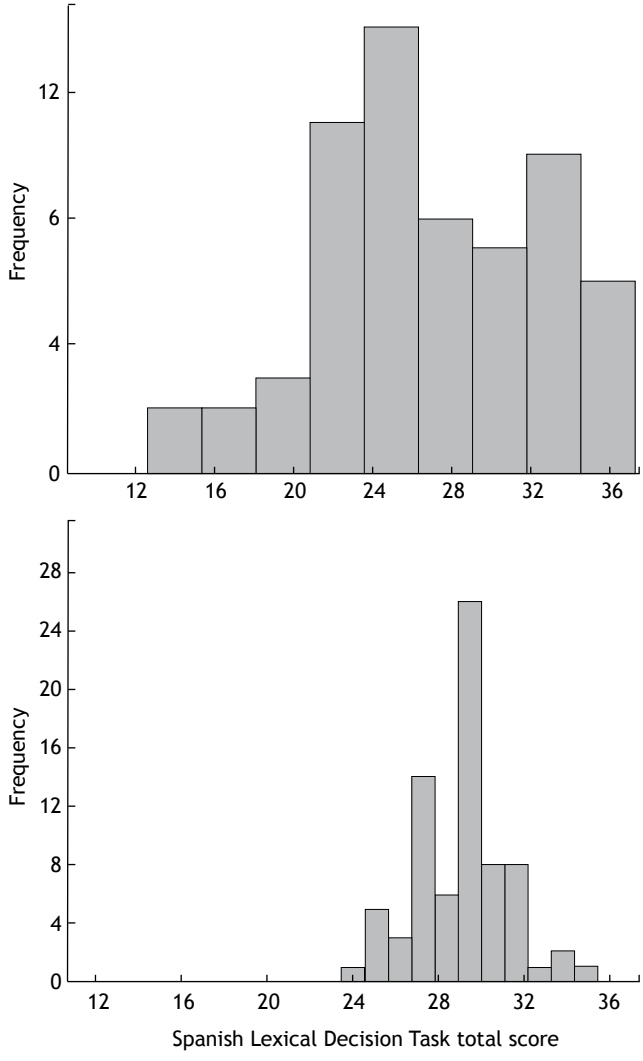

Figure 1. Score distributions of the Spanish Spot-the-Word test in the Development Sample (upper panel) and the Test Sample (lower panel)

the need for age correction of scores. On the other hand, there is a large, positive and significant correlation between test scores and years of formal education, r_s = .630, p < .001, 95%CI of r_s = .458 - 762. This suggests that the Spanish Lexical Decision Task is measuring crystallised verbal ability, which is defined as *"skilled judgment habits (that) have become crystallized . . . as a result of earlier learning"* (Cattell, 1963; pp 2-3).

As a further examination of the validity of the Spanish Lexical Decision Task, the researchers examined its correlations with scores on the WAIS-IV in the 50 participants who completed both assessments. There was a large, positive significant correlation between Spanish Lexical Decision Task scores and full IQ on the WAIS-IV, r = .715, p < .001, 95%CI of r = .597-828. The correlations with the subtest standard scores are shown in Table 2, where it can be seen that the Spanish Lexical Decision Task had significant and positive correlations with all of the subtest scores. All of these would be considered qualitatively 'large' correlations, the largest being with the Information subtest (r_s = .771, p < .001).

Next, the researchers examined the statistical performance of the final 36-item Spanish Lexical Decision Task in the Test Sample, which was composed of 75 undergraduate students. This sample scored a mean score of 29.05 (SD = 2.36) and a median score of 29 (range 24-35). Thus, in this sample there is no evidence of skewing nor of a ceiling effect; however, there is also considerably less variation in the data. This can be seen in Figure 1 (lower panel). The distribution of scores did not differ significantly from the normal distribution (Shapiro-Wilk Test (75) = .974, p = .126).

The correlation between the 36-item Spanish Lexical Decision Task and the full WAIS IQ score in the Test Sample was r = .411, p < .001, 95%CI of r = .197 - .588. This r value is somewhat lower than that observed in the Development Sample, but this is likely caused by the restriction of range in the data of the Test Sample. The correlations between the Spanish Lexical Decision Task and the WAIS subtests in the Test Sample are also shown in Table 2. Again, the correlation r values are much attenuated in this restricted-range sample; however, the overall pattern is similar to that observed in the Development Sample, in which the two highest correlations are for the mainly verbal assessments of Similarities and Information. They are particularly strong for the latter.

There was also a significant and positive correlation between Spanish Lexical Decision Task scores and Grade Point Average scores, r_s = .318, p = .006, 95%CI of r = .070 - .534. This essentially provides predictive validity of the test as a measure of intelligence. However, there was no significant correlation of the Spanish Lexical Decision Task with the number of years spent in education, r_s = .191, p = .100, 95%CI of r = -.063 - .421. However, this, again, may be attenuated by the limited range within the data.

Discussion

In this paper, the researchers present preliminary validity of a simple lexical decision task, similar to the English-language Spot-the-Word Test. The test has good psychometric properties. It was found that both internal consistency and test-retest reliability would be considered qualitatively 'good' by the normal interpretations of the values. This newly-developed test can be constructed from the items shown in Table 1 and the instructions in the Method section. Alternately, a ready to use version can be downloaded for free from http://www.gpluck.co.uk. This test could be used whenever a simple and fast assessment of language competence is required for research purposes.

More specifically, this Spanish Lexical Decision Task can be considered as a valid measure of crystallised-verbal ability (as opposed to other forms of non-crystallised linguistic ability such as verbal fluency). The fact it measures crystallised ability is demonstrated by its pattern of correlations with demographic and cognitive variables. Firstly, a measure of crystallised ability would be expected to be unrelated to age, which is exactly what was found. This is because while fluid abilities are known to decline in older adulthood, a hallmark of crystallised abilities is that they remain stable throughout adulthood and into old age (Craik & Bialystok, 2006; Park, 2000). Secondly, a measure of crystallised ability would be expected to be closely associated with years of education, which, again, is what was found. This is because crystallised ability is defined by it being based on acquired information (Cattell, 1963). This leads us to the third supporting point. Crystallised intelligence is thought to be based on two different forms of acquired information: verbal-lexical information and general knowledge (Schipolowski et al., 2014). Of the seven WAIS subtests that were examined, the only test that is clearly a measure of crystallised ability (in the form of general knowledge) is the Information subtest, and this is the test that has the highest correlations with the Spanish Lexical Decision Task.

The correlations between the WAIS subtest scores and the Spanish Lexical Decision Task were all qualitatively 'large' within the Development Sample. However, the correlations in the Test Sample comprised of university students, were much smaller, and the only two that were statistically significant were Similarities and Information: both assessments of verbal skills. This again, is consistent with the suggestion that the new test reported here is a measure of crystalized-verbal ability. The reason for the discrepancy in general between the large correlations observed in the Development Sample and the weaker correlations observed in the Test Sample is likely due to the variation of abilities within the two samples. As observed in Figure 1, there was much more variation in task performance within the Development Sample, compared to the Test Sample. It is known that correlation r values are attenuated when variance within the variables is restricted (Howell, 1992).

Therefore, this Spanish-language test, equivalent to the English-language Spot-the-Word Test, can be considered as a reliable and valid measure of crystallised-verbal ability. It should also be noted that the test had impressive correlations with full-scale IQ scores in both samples studied, and it also correlated significantly with academic achievement as measured by undergraduate GPA scores. It could also, therefore, be used as a simple measure of intelligence for research purposes. As a self-completion scale with only 36 items, it has significant advantages over other intelligence measures, which are slow to administer and usually require one-to-one interviews.

A further potential application could be as a measure of premorbid ability, which is particularly important in neuropsychology (see e.g. García-Torres, Vergara-Mor-

agues, Piñón-Blanco, & Pérez-García, 2015; Pluck et al., 2012). Measuring premorbid ability is also important in psychopathology research in which scales that are likely not affected by mental-illness-related cognitive impairments are very useful, for example when matching patients to control participants (see e.g. Pluck, Barajas, Hernandez-Rodriguez, & Martínez, 2020). Initial reports with the English-language Spot-the-Word Test suggested that it was resistant to age-related cognitive declines that were evident with other neuropsychological tests (Yuspeh & Vanderploeg, 2000). Furthermore, one study found that even demented Alzheimer's disease patients could perform the Spot-the-Word Test normally (Law & O'Carroll, 1998). This resistance to cognitive impairment is very useful to psychologists as it means that the patient's normal level of cognitive ability can be estimated. However, other studies have suggested that the Spot-the-Word Test is not as resistant as sometimes claimed, and that dementia does in fact have a substantial effect on performance (Beardsall & Huppert, 1997). Nevertheless, further research and modifications of lexical decision tasks such as the Spanish Lexical Decision Task may be able to improve this feature. One particularly promising aspect may be that studies with lexical decision tasks in Spanish have shown that age of acquisition, imageability, and frequency of words within the language are an important factor in resilience to cognitive impairment. Words that are highly imageable, high frequency, and early acquired are generally quite resistant to dementia and, therefore, may be useful in lexical decision tasks that aim to estimate premorbid ability (Cuetos et al., 2017; Cuetos, Herrera, & Ellis, 2010).

Despite its potential use in neuropsychology, it must be noted that this test has not been validated for clinical or educational decision making and should not be used for such purposes. However, the test may be useful in clinical and educational research contexts where a rapid assessment of crystallised-verbal knowledge, intelligence, or simply reading ability is required. It is notable that the Spanish Lexical Decision Task seems to be stable over ageing, as shown by the lack of correlation with age. This is as expected for a test primarily measuring crystalized-verbal knowledge, as such skills are known to be stable well into old age (Craik & Bialystok, 2006; Park, 2000). A useful consequence of this is that there is no need for age adjustment. Raw scores from groups can be used for between-subject comparisons, or for correlational/regression analysis with other variables. As the task is designed for research, and not applied assessments, conversion to normative scores is not necessary.

The test was administered via reading in the current research. However, as with the English Spot-the-Word test, it can potentially be administered via speech with the psychologist reading the words aloud. Such an administration format has not yet been specifically assessed, and so it should be used cautiously. It would be useful for future research to establish the psychometric parameters of that alternative administration format, which would allow the test to be used with illiterate or visually impaired participants or simply to test auditory word recognition.

There are some limitations to the research that should be addressed. The Test Sample was comprised of 75 undergraduate students. While this did allow the predictive validity of the Spanish Lexical Decision Task to be examined via the correlations with GPA, there are some problems with the homogeneity of the sample. Student samples from specific sites in general tend to lack variety because they are drawn from limited demographic backgrounds. In the current research, that meant that there was very limited range of ages and educational experience. In addition, as they were drawn from a private university population, they tended to be from a relatively high socioeconomic background. This was not the case for the Development Sample, however, in which there was a deliberate attempt to recruit a wide range of participants. Although socioeconomic status (SES) was not measured directly, education level is considered a strong proxy measure of SES (Galobardes, Shaw, Lawlor, Lynch, & Davey Smith, 2006), and, in our sample, the range was from 6 -26 years of formal education, indicating heterogeneity within our Development Sample. A further limitation is that although crystalized intelligence is thought to be comprised of both lexical word and general knowledge (Schipolowski et al., 2014), the researchers only correlated scores with the latter aspect, as measured by the Information subtest of the WAIS-IV. This was because the abbreviated form of the intelligence that we used does not include the Vocabulary subtest, which would be the obvious measure of stored language. This omission also meant that the researchers could not calculate a measure of verbal IQ, which should also be seen as a limitation. Similarly, the research did not include any measures of linguistic variables such as reading level, or semantic fluency. Such measures would have allowed a more detailed analysis of the validity of the Spanish Lexical Decision Task, and their absence should be considered a limitation.

In summary, the Spanish Lexical Decision Task is a reliable and valid measure of crystallised-verbal ability, reading, and also intelligence. The test is fast to administer and as a self-report scale does not have any issues associated with inter-rater reliability. In addition, it can be used either in one-to-one interview sessions or also potentially with appropriate control, in group-testing contexts. The Spanish Lexical Decision Task contributes to the available tools for use in cognitive testing of Spanish speaking research participants.

Acknowledgments

This work was supported financially by the Chancellor's Grant Scheme of Universidad San Francisco de Quito. I would also like to thank Sarahi Ponton Junes for help with data collection and processing.

References

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language-Processing Test*. Bury St Edmunds: Thames Valley Test Company.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: a robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology, 32*(1), 55-65. https://doi.org/10.1111/j.2044-8260.1993.tb01027.x

Baldo, J. V., Bunge, S. A., Wilson, S. M., & Dronkers, N. F. (2010). Is relational reasoning dependent on language? A voxel-based lesion symptom mapping study. *Brain and Language, 113*(2), 59-64. https://doi.org/10.1016/j.bandl.2010.01.004

Beardsall, L., & Huppert, F. (1997). Short NART, CCRT and Spot-the-Word: comparisons in older and demented persons. *British Journal of Clinical Psychology, 36*(4), 619-622. https://doi.org/10.1111/j.2044-8260.1997.tb01266.x

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences, 17*(2), 89-98. https://doi.org/10.1016/j.tics.2012.12.002

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3-4), 296-312. https://doi.org/10.1111/j.1469-7610.2006.01611.x

Blank, I., & Fedorenko, E. (2017). Domain-general brain regions do not track linguistic input as closely as language-selective regions. *Journal of Neuroscience, 37*, 9999-10011. https://doi.org/10.1523/JNEUROSCI.3642-16.2017

Blank, I., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology, 112*, 1105-1118. https://doi.org/10.1152/jn.00884.2013

Campbell, K. L., & Tyler, L. K. (2018). Language-related domain-specific and domain-general systems in the human brain. *Current Opinion in Behavioral Sciences, 21*, 132-137. https://doi.org/10.1016/j.cobeha.2018.04.008

Canivez, G. L., & Watkins, M. W. (2010). Investigation of the factor structure of the Wechsler Adult Intelligence Scale--Fourth Edition (WAIS-IV): exploratory and higher order factor analyses. *Psycholological Assessment, 22*(4), 827-836. https://doi.org/10.1037/a0020429

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1-22. https://doi.org/10.1037/h0046743

Cattell, R. B. (1967). The theory of fluid and crystallized general intelligence checked at the 5-6 year-old level. *British Journal of Educational Psychology, 37*(2), 209-224. https://doi.org/10.1111/j.2044-8279.1967.tb01930.x

Cattell, R. B. (1973). *Cattell Culture Fair Intelligence Tests*. Champaign, IL.: Institute for Personality and Ability Testing.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319. https://doi.org/10.1037/1040-3590.7.3.309

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends in Cognitive Sciences, 10*(3), 131-138. https://doi.org/10.1016/j.tics.2006.01.007

Crawford, J. R., Stewart, L. E., Cochrane, R. H. B., Parker, D. M., & Besson, J. A. O. (1989). Construct validity of the National Adult Reading Test: a factor analytic study. *Personality and Individual Differences, 10*(5), 585-587. https://doi.org/10.1016/0191-8869(89)90043-3

Cuetos, F., Arce, N., Martinez, C., & Ellis, A. W. (2017). Word recognition in Alzheimer's disease: Effects of semantic degeneration. *Journal of Neuropsychology, 11*(1), 26-39. https://doi.org/10.1111/jnp.12077

Cuetos, F., Herrera, E., & Ellis, A. W. (2010). Impaired word recognition in Alzheimer's disease: the role of age of acquisition. *Neuropsychologia, 48*(11), 3329-3334. https://doi.org/10.1016/j.neuropsychologia.2010.07.017

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System: Technical manual*. San Antonio, TX: Psychological Corporation.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Davey Smith, G. (2006). Indicators of socioeconomic position (part 1). *Journal of Epidemiology and Community Health, 60*(1), 7-12. https://doi.org/10.1136/jech.2004.023531

García-Torres, A., Vergara-Moragues, E., Piñón-Blanco, A., & Pérez-García, M. (2015). Alteraciones neuropsicológicas en pacientes con VIH e historia previa de consumo de sustancias. *Revista Latinoamericana de Psicología, 47*(3), 213-221. https://doi.org/10.1016/j.rlp.2015.06.001

Hampshire, A., Highfield, R. R., Parkin, B. L., & Owen, A. M. (2012). Fractionating human intelligence. *Neuron, 76*(6), 1225-1237. https://doi.org/10.1016/j.neuron.2012.06.022

Howell, D. C. (1992). *Statistical methods for psychology*. Belmont (CA): Duxbury Press.

Johnson, W., Bouchard Jr, T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: consistent results from three test batteries. *Intelligence, 32*(1), 95-107. https://doi.org/10.1016/S0160-2896(03)00062-X

Kan, K. J., Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. (2013). On the nature and nurture of intelligence and specific cognitive abilities: the more heritable, the more culture dependent. *Psychological Science, 24*(12), 2420-2428. https://doi.org/10.1177/0956797613493292

Law, R., & O'Carroll, R. E. (1998). A comparison of three measures of estimating premorbid intellectual level in dementia of the Alzheimer type. *International Journal of Geriatric Psychiatry, 13*(10), 727-730. https://doi.org/10.1002/(sici)1099-1166(1998100)13:10<727::aid-gps851>3.0.co;2-2

Meyers, J. E., Zellinger, M. M., Kockler, T., Wagner, M., & Miller, R. M. (2013). A validated seven-subtest short form for the WAIS-IV. *Applied Neuropsychology: Adult, 20*(4), 249-256. https://doi.org/10.1080/09084282.2012.710180

Park, D. C. (2000). The basic mechanisms accounting for age related decline in cognitive function. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 1-21). Philadelphia, PA: Psychology Press.

Pinker, S. (2010). Colloquium paper: the cognitive niche: coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences of the United States of America, 107*(Supplement 2), 8993-8999. https://doi.org/10.1073/pnas.0914630107

Pluck, G. (2019a). Lexical reading ability predicts academic achievement at university level. *Cognition, Brain, Behavior, 22*(3), 175-196. https://doi.org/10.24193/cbb.2018.22.12

Pluck, G. (2019b). Preliminary validation of a free-to-use, brief assessment of adult intelligence for research purposes: The Matrix Matching Test. *Psychological Reports, 122*, 709-730. https://doi.org/10.1177/0033294118762589

Pluck, G., Almeida-Meza, P., Gonzalez-Lorza, A., Muñoz-Ycaza, R. A., & Trueba, A. F. (2017). Estimación de la función cognitiva premórbida con el Test de Acentuación de Palabras. *Revista Ecuatoriana de Neurología, 26*(3), 226-234.

Pluck, G., Barajas, B. M., Hernandez-Rodriguez, J. L., & Martínez, M. A. (2020). Language ability and adult homelessness. *International Journal of Language & Communication Disorders*. Advance online publication. https://doi.org/10.1111/1460-6984.12521

Pluck, G., Lee, K. H., Rele, R., Spence, S. A., Sarkar, S., Lagundoye, O., & Parks, R. W. (2012). Premorbid and current neuropsychological function in opiate abusers receiving treatment. *Drug and Alcohol Dependence, 124*(1-2), 181-184. https://doi.org/10.1016/j.drugalcdep.2012.01.001

Pluck, G., Ruales-Chieruzzi, C. B., Paucar-Guerra, E. J., Andrade-Guimaraes, M. V., & Trueba, A. F. (2016). Separate contributions of general intelligence and right prefrontal neurocognitive functions to academic achievement at university level. *Trends in Neuroscience and Education, 5*(4), 178-185. https://doi.org/10.1016/j.tine.2016.07.002

Premack, D. (2007). Human and animal cognition: continuity and discontinuity. *Proceedings of the National Academy of Sciences of the United States of America, 104*(35), 13861-13867. https://doi.org/10.1073/pnas.0706147104

Real Academia Española. Banco de datos (CREA) [on line]. Corpus de referencia del español actual. Retrieved 19 September, 2014, http://www.rae.es

Schipolowski, S., Wilhelm, O., & Schroeders, U. (2014). On the nature of crystallized intelligence: The relationship between verbal ability and factual knowledge. *Intelligence, 46*, 156-158. https://doi.org/10.1016/j.intell.2014.05.014

Snyder, H. R., Miyake, A., & Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. *Frontiers in Psychology, 6*, 328. https://doi.org/10.3389/fpsyg.2015.00328

Wechsler, D. (1999). *The Wechsler Abbreviated Scale of Intelligence professional manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2012). *Escala de Inteligencia de Wechsler Para Adultos-IV*. Madrid, Spain: Pearson.

Woolgar, A., Duncan, J., Manes, F., & Fedorenko, E. (2018). Fluid intelligence is supported by the multiple-demand system not the language system. *Nature Human Behaviour, 2*(3), 200-204. https://doi.org/10.1038/s41562-017-0282-3

Yuspeh, R. L., & Vanderploeg, R. D. (2000). Spot-the-Word: a measure for estimating premorbid intellectual functioning. *Archives of Clinical Neuropsychology, 15*(4), 319-326. https://doi.org/10.1016/S0887-6177(99)00020-7