

# Análisis del ESLE como instrumento de evaluación

## Analysys of Spanish Proficiency Tests as an testing tool

**Marisol Andrea Méndez Cabrera**

Universidad Nacional de La Rioja

**Patricia Elena Compagnoni**

Universidad Nacional de La Rioja

**María Ana Glaccaglia**

Universidad Nacional de La Rioja

**María del Carmen Brizuela**

Universidad Nacional de La Rioja

Recibido: 27 de marzo 2023

Aceptado: 12 de mayo de 2023

---

### Resumen

El trabajo representa el resultado del análisis de los Exámenes de Suficiencia de Lengua Española (ESLE), administrados a candidatos extranjeros no hispanohablantes que aspiran a ingresar a carreras de grado en la Universidad Nacional de La Rioja (UNLaR) entre los años 2012 y 2021. Constituye la síntesis de la primera etapa del proyecto de investigación: Análisis y nuevo diseño del Examen de Español como Lengua Extranjera estandarizado de la UNLaR, cuyo objetivo es construir un diseño actualizado y con cualidades más sólidas que los exámenes hasta aquí tomados. El estudio se realizó sobre las muestras administradas durante el período indicado, en el marco de las especificaciones establecidas en el Reglamento de ESLE, aprobado por Res. N° 4035/11 HCS de la UNLaR y de las conceptualizaciones revisadas y acordadas sobre las cualidades de validez, fiabilidad, autenticidad y viabilidad, exigibles en la construcción y administración de los instrumentos de evaluación. Entre los resultados arrojados, en la muestra se observa la persistente búsqueda de la validez, sostenida principalmente en la autenticidad de las tareas y los textos, mientras que la fiabilidad y viabilidad son la respuesta a gestiones no planificadas en todas sus etapas. Para cerrar, se detallan los nuevos desafíos que implican la creación de un nuevo instrumento.

**Palabras clave:** evaluación, proficiencia, español lengua extranjera

### Abstract

This work shows the results of the analysis carried out on the Spanish Proficiency Tests taken to non-Spanish speaking foreign candidates aiming at entering the National University of La Rioja (UNLaR) in the period running from 2012 to 2022. This is a summary of the first stage of the investigation project entitled "UNLaR Standard Spanish as a Foreign Language Test Analysis and New Design" for the purpose of producing an updated test design of higher quality than the current ones.

The survey was made on the samples provided during the above-mentioned period of time under the specifications given by the ESLE Regulations approved by Resolution N° 2849/11 HCS of the UNLaR as well as under the concepts of validity, reliability, authenticity and viability required in the design and administration of evaluative tools. Among the results obtained, the samples show a permanent search for

validity mainly sustained on tasks and texts authenticity; while reliability and viability are the answers to non-planned handlings in all the stages. To sum up, the new challenges about the creation of a new testing tool are mentioned.

**Keywords:** evaluation, proficiency, Spanish foreign language

---

### Introducción

Toda evaluación supone un entramado de factores que involucran a los candidatos, interesados en ser calificados; a los examinadores, que diseñan el instrumento evaluador a partir de un conjunto de descriptores lingüísticos y gestionan el proceso en su complejidad; y a la propia institución, que certifica el dominio para realizar ciertas tareas demandadas por la comunidad. En los últimos quince años, la Universidad Nacional de La Rioja (UNLaR) ha recibido a numerosos contingentes de estudiantes extranjeros no hispanohablantes, interesados en su oferta académica de grado. Algunos de ellos formaron parte de proyectos de intercambio, como parte de las políticas de internacionalización de la universidad; otros fueron el resultado de la movilidad migratoria regional que espontáneamente buscaban oportunidades en las universidades argentinas. En este contexto, la UNLaR se vio en la necesidad de asumir políticas lingüísticas que contuvieran la demanda social.

Es entonces que en el 2009, de acuerdo a lo estipulado en la Resolución N° 2849 del HCS, se establece como requisito de ingreso para los estudiantes extranjeros, cuya lengua materna no sea el español, rendir el Examen de Suficiencia de Lengua Española (ESLE). Con este instrumento se propone evaluar la competencia lingüística comunicativa oral y escrita en español en situaciones cotidianas,

familiares, sociales, laborales o académicas de hablantes, sin tomar en consideración sus lenguas maternas ni el tipo de adquisición o aprendizaje.

En el transcurso de estos años, se dictaron cursos de español como lengua extranjera (ELE) destinados a los ingresantes, se formaron equipos docentes, se creó material auténtico con rasgos del habla local y se elaboró e implementó un examen estandarizado como el ESLE. El presente desafío es cuestionar y examinar parte de este conjunto de acciones institucionales. En este punto, fue imperioso dar inicio a un proyecto de investigación que analizara específicamente el instrumento evaluador y las actuaciones en torno al mismo poniendo en valor los logros alcanzados, en pos del objetivo y el contexto de uso. A partir de los resultados que arroje esta primera parte, rediseñar el modelo, actualizado por las nuevas necesidades institucionales y los avances teóricos en el área.

Como parte de este proyecto, el presente trabajo expresa los resultados de la primera etapa del proceso de investigación: el análisis de las muestras de ESLE, tomadas hasta el 2021. Para ello, se partirá de algunas consideraciones que definen los principios teóricos y metodológicos que acompañaron el proceso, las cuestiones principales en torno a las especificaciones del ESLE, como instrumento evaluador, y los resultados obtenidos de la observación de

cada una de las áreas que componen el examen. Por último, en la discusión, se arribará a los lineamientos a partir de los cuales se rediseñará el instrumento.

### **Marco teórico**

#### **¿Qué es evaluar en lengua extranjera?**

La evaluación de una lengua extranjera puede ser útil para diferentes propósitos. Por razones didácticas, otorga información acerca del proceso de enseñanza y aprendizaje para evidenciar si se alcanzaron los objetivos propuestos; en el marco de investigación en Lingüística Aplicada, para analizar la interlengua y los errores emergentes; y por razones operativas, para averiguar el nivel de competencia del examinado. La información que arroja la evaluación ayuda a tomar decisiones.

El instrumento más usado para recuperar la evidencia suficiente, que luego sustentará la decisión tomada, es el *examen*. Esta herramienta compleja responde a diferentes demandas en función de las necesidades e intereses de cada actor. Para los estudiantes da prueba de sus habilidades; para los docentes, la evidencia que le permitirá tomar decisiones; para las instituciones, certifica el nivel de dominio.

Un examen es un procedimiento para reunir evidencia de habilidades lingüísticas generales o específicas, a partir de la actuación en tareas diseñadas, con el objeto de proporcionar las bases para predicciones acerca del uso de esas habilidades por un individuo

en un contexto de vida real. (McNamara, citado por Prati, 2007, p.12)

Por lo tanto, de esta definición se puede considerar que *examinar* implica interpretar una serie de datos obtenidos a partir de cierta actuación del evaluado, y que estos son la *evidencia* de su conocimiento y capacidades que le permitirán desenvolverse en el mundo. ¿Cuáles son las garantías de un instrumento de tal complejidad? Bordón (2004) sostiene que un *buen examen* es posible a partir de su validez y su fiabilidad, cualidades que afectan al instrumento, y su viabilidad, dependiente de factores contextuales. Si bien no incluye en esta nómima la noción de autenticidad, no deja de ser una cualidad más a considerar, no como requisito principal, pero sí como característica relacionada con la validez. Es decir, si se desea plantar pruebas del tipo comunicativa, esta cualidad equipara la herramienta a usos reales de la lengua.

#### **Validez**

En torno a esta cualidad, los examinadores se preguntan ¿el examen mide lo que se supone que debe medir? Es decir, será considerado válido en la medida en que su diseño está orientado a obtener cierta información y no sesgado por otra. En este punto, Prati (2007), vuelve más precisa esta noción cuando la relaciona con el concepto de constructo, propuesto por Llosa (2006), y afirma que “un examen es válido en la medida en que podemos interpretar sus resultados como indicadores de las habilidades o del constructo que queremos

evaluar” (p. 15). De esta manera, la herramienta será coherente con una descripción específica de las habilidades, a la luz de principios teóricos, necesaria para interpretar la actuación del candidato.

Hay distintos tipos de validez. La *validez de contenido* es evidente en la medida en que el examen es una muestra representativa de las habilidades evaluadas. La *validez concurrente* se comprueba a partir de la concordancia de los resultados de varios individuos en exámenes diferentes, dados en condiciones semejantes. La *validez predictiva* es observable si el instrumento evaluador tiene la capacidad de predecir el desempeño futuro de los examinados. La *validez de puntaje* se comprueba cuando el puntaje se logra de la habilidad que se propone evaluar y no de otra (Prati, 2007. Pastor, 2009).

### **Fiabilidad**

Llosa (2006) sostiene que la validez es una cualidad de la interpretación que se hace de los resultados, mientras que la confiabilidad es una cualidad de los resultados propiamente dichos. En este punto, el equipo evaluador se pregunta ¿confío en los resultados obtenidos? ¿Son fiables? Para esto, se crean escenarios que permiten ensayar los alcances de esta cualidad. La medición de una habilidad o constructo será consistente o estable dados los diferentes momentos en que se administre el examen a los candidatos. Dicha comprobación arroja medidas de fiabilidad en los resultados.

Bachman (1990, citado en Pastor 2009), distingue la *fiabilidad interevaluadora*, probada cuando dos evaluadores corrigen el mismo examen y coinciden con el resultado; la *fiabilidad intraevaluadora*, demostrada cuando un mismo evaluador corrige un examen en diferentes ocasiones y da la misma valoración; y la *fiabilidad extrínseca*, evidenciada cuando las condiciones administrativas o circunstanciales no influyen en los resultados.

### **Autenticidad**

Como se ha dicho, parte de la validez de un instrumento implica que las tareas sean auténticas. Según Prati (2007), pueden serlo según “el grado de nivel o correspondencia entre las características de una determinada actividades del examen y las características de una actividad en el campo donde se usa la lengua meta” (p.17). Bachman (1996, citado en Pastor, 2009) reconoce como *autenticidad situacional*, cuando la tarea evaluadora se parece a alguna actividad lingüística que ocurre de forma natural en la vida cotidiana. A esto se suma la necesidad de incluir textos realistas que conserven las características de la lengua hablada, recuperados de instancias reales o desarrollados a partir de un guión con fines didácticos. Por otra parte, la *autenticidad interaccional* es dada cuando la tarea “requiere del alumno el conocimiento, la destreza y habilidad que necesitaría una persona para actuar en una situación de uso de la lengua” (Pastor, 2009, p.11). Es decir,

que la actividad funciona como una réplica de los procesos realizados en el mundo real.

### **Viabilidad**

Las etapas para la elaboración de un instrumento, como el ESLE, son complejas y requieren de una coherencia entre lo necesario y lo disponible. Un examen estandarizado implica formular el proyecto, diseñar el examen desde sus especificaciones hasta las herramientas para evaluar su calidad, elaborar el modelo para cada toma, incluyendo la selección del material, escribir los ítems, establecer criterios de puntuación y corrección, realizar pretesteos para luego editar el examen, entre otras acciones. Para todo este trabajo, no es menor planificar los recursos humanos, materiales y de tiempo con los que se cuenta para llevar adelante el proyecto: a esto se le llama viabilidad.

### **Metodología**

Desde los principios teóricos señalados, se abordará el análisis de las tomas del ESLE hasta el 2021. En el marco del mencionado proyecto de investigación, el universo de estudio observado consiste en el reglamento del examen y nueve modelos de exámenes con su resolución, escrita y oral. Desde su gestión original, nunca se previó un protocolo de archivo para estos exámenes por ello se contó con un corpus irregular y variable. Al momento del análisis, los resultados no pueden ser categóricos en la descripción del objeto pero, al mismo tiempo, estas

inconsistencias evidencian condiciones de fiabilidad o viabilidad.

De un total de nueve tomas se registraron 83 exámenes. Cuatro tomas corresponden al 2012 y cinco se realizaron entre el 2019 y el 2020. Debido a las condiciones impuestas por la pandemia, las dos últimas son una reformulación del instrumento en modalidad virtual. Se analizaron todas las actividades que se propusieron en cada ocasión aunque no están completas en todos los casos. En la CA, no se recuperaron los audios de las tomas de 2012; y en la EO, se conservan las láminas de trabajo y la actuación grabada de los candidatos sólo de las últimas cinco tomas.

En relación a la resolución de los mismos, se conocen las listas completas de candidatos inscritos y examinados a partir de la sexta toma, 6 de diciembre de 2019. Esto permitió verificar variables como la fiabilidad y la validez entre las consignas propuestas y las respuestas obtenidas. Sin embargo, la observación de estos rasgos se volvió difícil de comprobar en tomas previas porque se conservan muy pocos ejemplares -de algunas sólo uno y del 5 de abril de 2019, ningún examen resuelto- y se desconoce las listas de los candidatos.

### **Resultados**

Para la organización de nuestro análisis presentaremos, en primer lugar, los lineamientos generales del reglamento y luego los resultados obtenidos de la observación de los cuatro rasgos esenciales de un examen, evidenciados en el corpus.

Para abordar la validez, se detallarán los logros y limitaciones en cada uno de los apartados del ESLE. A continuación, se considerará la fiabilidad, la autenticidad y la viabilidad en la implementación del examen como una unidad.

### **El reglamento del ESLE**

Los lineamientos principales de un examen estandarizado se comunican a los interesados por medio de un documento oficial: las especificaciones, una expresión oficial acerca de lo que el examen evalúa y cómo lo evalúa. En el caso del ESLE, los puntos principales los define en su reglamento. Tal como figura en la Res. N° 4035/2001 del HCS, su propósito u objetivo es acreditar el nivel intermedio de la lengua española, considerándolo como requisito de ingreso a cualquier carrera de grado o posgrado. Es decir, el candidato que logre aprobar el examen evidenciará cierto desarrollo de su competencia comunicativa en español, suficiente para desenvolverse en la vida académica que se propone iniciar. Los destinatarios o candidatos son todos los ingresantes extranjeros, cuya lengua materna no sea el español y deseen ingresar a una carrera de grado o posgrado de esta universidad, sin tomar en consideración sus lenguas maternas ni el tipo de adquisición o aprendizaje que haya realizado.

Estos dos aspectos definen al ESLE como un examen de dominio que solo evalúa la competencia lingüística y el momento de aplicación lo establece la institución. Acredita un nivel intermedio en términos generales,

cuyos descriptores se asocian al B (usuario independiente) 1(umbral) del MCER (2002). Como examen de proficiencia, prevé una situación de uso del español en la que el usuario pueda desenvolverse con cierta fluidez y naturalidad en situaciones familiares, sociales y de servicios. Pueda iniciar estudios universitarios aunque pueda tener dificultades para enfrentar contextos desconocidos o ante la necesidad de matizar o precisar sus enunciados. Sin embargo, cabe destacar que en el documento no se discrimina niveles inferiores o superiores.

Para lograr su propósito, el ESLE se organiza en cinco áreas que proponen un conjunto de actividades orales y escritas, creadas a partir de fuentes auténticas, que suponen una generalización de las operaciones lingüísticas en la lengua meta. Se consideran las cuatro destrezas - comprensión auditiva (CA), comprensión lectora (CL), expresión escrita (EE) y expresión oral (EO)- y uso de la lengua. Cada una es definida en términos generales sin especificar lo que puede o no hacer, es decir, no cuenta con un constructo de contenidos que pudiera describir y validar los resultados de los candidatos. Se evalúan por separado, obteniendo un puntaje por cada área. Además incluye un apartado exclusivo para conocimientos gramaticales y léxicos sin asociarlo a ninguna de las demás destrezas.

Si bien el reglamento del ESLE esboza algunas cualidades de su metodología de evaluación, más específicas en unas destrezas que en otras, no se detalla otro

tipo de información necesaria para quien elabora el examen, como qué cantidad de información se dará en las consignas, en qué registro y nivel de lengua se escribirán y si incluirán ejemplos y criterios de corrección en cada una, entre otras.

Tal como se declara en el documento, los criterios de corrección se limitan a considerar la adecuación discursiva, gramatical, léxica, de contenido y registro, y agrega para la oralidad, fluidez y pronunciación /entonación. No se propone elaborar una descripción de cada uno por tarea, no se detallan el uso de escalas o bandas. Aparentemente, la interpretación de los resultados se sustenta únicamente de lo que evaluadores y correctores pudieran interpretar acerca de los descriptores del nivel intermedio, definido en el MCER (2002). El puntaje se establece en una escala del 1 al 100% por cada área, sin determinar qué cantidad de porcentaje se le asignaría a cada tarea. Para aprobar, el candidato deberá alcanzar el 60 % en cada una.

Como conclusión, podemos considerar que el reglamento del ESLE incluye gran parte de los aspectos que toda especificación de un examen de dominio estandarizado debería definir. Sin embargo, escapan a él otros muy importantes, como el constructo. Esta descripción queda solo en términos generales y el reconocimiento de los indicadores de las habilidades a evaluar descansan en la experiencia de los evaluadores. Por otra parte, hay un esmero en enunciar el tipo de ítems pero no el tipo tarea; tampoco se describen los elementos

lingüísticos, como aspectos gramaticales, semánticos contextuales y discursivos. Conservar estos criterios de trabajo descansa en la decisión de los evaluadores.

### **Validez**

En este punto del análisis de las muestras, cabe preguntarse si las propuestas del ESLE evalúan lo que pretenden observar en las actuaciones de los candidatos. Para responder, se observará las destrezas evaluadas.

### **Comprensión auditiva**

Como se mencionó, el reglamento ofrece breves definiciones e insuficientes para determinar un constructo explícito y con ello resulta difícil mensurar la validez del constructo en las muestras. Sin embargo, a partir de los descriptores de nivel, los evaluadores de ESLE han diseñado tareas e interpretado los resultados de cada toma. Las actividades propuestas para la CA están orientadas a recuperar las ideas principales de las escuchas, los textos fueron extraídos de medios masivos, principalmente programas de radio, y conservan los rasgos lingüísticos señalados en el documento. Este rasgo logra cierto sustento empírico sobre otros tipos de validez.

Atendiendo a un corpus irregular, los exámenes de CA se dividen en dos etapas distintas que agrupan en dos tipos de pruebas diseñadas. En el 2012, las tareas son de opción múltiple, siguiendo los parámetros declarados en el reglamento.

Como ventajas, ofrece un alto grado de fiabilidad porque las respuestas son cerradas, rápidas, objetivas y prácticas para corregir convocatorias numerosas. Como desventaja, asumen la posibilidad de ser respondidas por el azar. Sin embargo, esta dificultad se salvó en ocasiones en las que se diseñaron cuatro o más opciones. Por otra parte, cabe decir, que este tipo de pruebas requieren mucho tiempo para su elaboración y sólo se disponía de dos personas para todas las tareas de diseño, administración y corrección. Por eso, en términos de viabilidad operativa, resultaron inviables y se decidió cambiar el tipo de prueba.

Desde el 2019, se propuso un cuestionario que demandaban respuestas breves en torno a dos textos. En todas ellas, el examinado asume el rol de miembro de un público general, sin mayor participación más que la interpretación del sentido global y algunos aspectos puntuales de interés. Puede reconocer que se espera de él la transferencia de cierta información. La presencia de cierto fraseo textual en las preguntas y el hecho de que su orden sigue la progresión temática del texto hace más fácil la tarea. Las respuestas son semi abiertas, conducen a una producción breve - una palabra, una o dos oraciones- y bastante guiada. Como ventajas, el azar influye menos, lo que las vuelve más fiables, y son adecuadas para recuperar información explícita. Como requieren más tiempo de elaboración por parte del candidato, se redujo la cantidad de ítems y textos a escuchar, de cuatro a dos. Por otra parte, en

términos de viabilidad, para los evaluadores son más fáciles de elaborar. Como desventajas, el tiempo que se ahorró en la elaboración de los ítems, se ocupó en la corrección. Aunque las respuestas esperadas, en su mayoría son explícitas, no textuales, no se formalizaron los criterios para la redacción de las mismas. Esto pudo afectar en la validez de su interpretación y en consecuencia en la fiabilidad de sus resultados. Por eso, ante casos fronterizos, las evaluadoras consultaban entre ellas una única valoración para que el resultado no pierda consistencia.

En este punto, el ESLE conserva cierta validez de contenido porque no se aleja de los descriptores de la CA para el B1 como objetivo. Sin embargo, cambiar de tipo de tarea, de respuesta cerrada a abierta y breve, y no formalizar los criterios de interpretación abren un margen de subjetividad, difícil de regular. En relación a la validez concurrente, comprobarla resultaría prácticamente accidental, porque la administración del ESLE no planificó un análisis de esta envergadura. La observación recaería en los resultados de candidatos que debieron rendir más de una vez, hasta aprobar cada parte. Por otra parte, el alcance de la validez predictiva en los modelos del ESLE perfila un oyente, miembro del público de medios masivos, característico del nivel que acredita, pero no particulariza en los intereses del estudiante universitario. Por último, la validez de puntaje se verifica a partir de una distribución equitativa entre los ítems, sin destacar una respuesta o un texto



por sobre otro, es decir, toda la información recuperada era igualmente valiosa. Este criterio busca mitigar cualquier intento de especulación por parte del candidato. En conclusión, la evaluación de la CA en el ESLE es válida en la medida que el contenido de su diseño no se aleja de su objetivo ni su parámetro de interpretación, aunque estos sean precarios.

### **Comprensión lectora**

Como se mencionó anteriormente, el reglamento del ESLE especifica como finalidad del examen acreditar el nivel intermedio de la lengua, lo que equivale a decir, el de un usuario que puede desenvolverse con cierta fluidez y naturalidad en situaciones familiares, sociales y de servicios. A los fines del examen, se describe la comprensión lectora como la capacidad del candidato de leer textos auténticos en su mayor parte informativos y de interés general extraídos de instrucciones, folletos, periódicos, etc. El candidato contesta, seleccionando la respuesta adecuada a 12 preguntas de opción múltiple sobre cuatro textos. La extensión de estos textos no será superior a 450 palabras (Res. 4035, 2011).

En primer lugar, se evidencia que los descriptores que ofrece el marco normativo que regula el ESLE, resultan insuficientes e imprecisos y, al no describir las destrezas vinculadas con el nivel de dominio que se pretende certificar. Se torna difícil analizar la validez de constructo en las muestras de los exámenes administrados a los candidatos entre el 2012 y el 2021.

Por reglamento, las muestras analizadas atienden a los descriptores correspondientes al nivel B1 de la lengua que detalla el MCER (2002): “el candidato será capaz de comprender textos redactados en una lengua de uso habitual y cotidiano o relacionada con el trabajo, que describa acontecimientos, sentimientos y deseos en cartas personales (p.30)”. Sobre la base de estos descriptores y en relación a la validez de contenido, se observa que los exámenes administrados en los meses de abril y agosto de 2012 ofrecen textos auténticos y las actividades están centradas en su lectura y son, en ambos casos, de respuestas cerradas: selección verdadero o falso y selección múltiple. En las primeras muestras, se trata de tres textos informativos, vinculados con el área de la salud. El examen del mes de agosto incluye cuatro textos, tres de ellos informativos y uno literario. Las consignas planteadas parecen adecuadas a lo que se pretende evaluar por cuanto apelan a la capacidad de realizar inferencias y no a la identificación de respuestas literales en los textos y además están elaboradas atendiendo a las ideas centrales. En este sentido, puede decirse que los exámenes estarían mínimamente garantizando una validez de contenido. Si bien los textos elegidos son relativamente sencillos, algunos de ellos contienen muchos términos propios del mundo de la medicina o empresarial que, aunque no son palabras que puedan afectar la comprensión en el nivel que se busca acreditar, la dificultan.

La muestra tomada en el mes de noviembre de 2012 ofrece un único texto con una

actividad cerrada de selección verdadero o falso sobre diez afirmaciones que sobre él se hacen. Más allá de las desventajas o riesgos propios de este tipo de instrumentos, por ejemplo el azar o la copia, el tipo de texto elegido no parece apto para evaluar la comprensión lectora en el nivel pretendido. La particular narrativa de Eduardo Galeano, en este caso, en *El derecho al delirio*, tiene muchas expresiones, metáforas, términos y alusiones que tienen referencias socioculturales y suponen conocimientos del mundo que lo hacen un texto difícil de comprender, siempre atendiendo al nivel de CL que se ha descrito más arriba. Las afirmaciones que se ofrecen al candidato para que confirme como verdaderas o falsas, están bien construidas, ya que, igual que con las muestras anteriores, todas ellas apelan a la capacidad de hacer inferencia y no a la identificación literal en el texto. En este caso, el texto seleccionado, por su grado de complejidad pone en riesgo la validez de contenido de la muestra. Se observa además que esta muestra ofrece una única lectura y una única tarea, lo que también afecta la fiabilidad del instrumento.

Los instrumentos administrados en los años 2019 y 2021 presentan modificaciones con respecto a los anteriores, en tanto ofrecen variedad de textos y de tareas. El primer ítem de la muestra del 2019 está compuesto por cinco reseñas de libros. La tarea es, luego de leer las reseñas, sugerir qué comprar a cada lector, considerando las preferencias que aparecen expresadas al finalizar el texto. Así, el candidato debe identificar el tema de cada

libro y sus características, para poder asociar cada uno con las preferencias de los lectores. Los textos de las reseñas son cortos y sencillos y el examen presenta trece preferencias sobre las que el aspirante debe trabajar. El segundo texto de la misma muestra presenta un texto que narra una leyenda sobre el origen del mate. Si bien es sencillo tiene referencias socioculturales que es posible dificulten la lectura, afectando esto la validez de contenido del examen. El candidato debe resolver una tarea de selección múltiple. El tercer texto es un texto de opinión del área de la medicina. Se propone al aspirante que, luego de su lectura responda en forma escrita a cinco preguntas. Tres de ellas requieren que en su resolución se reponga información que no está en el texto, por ejemplo, debe fundamentar con ideas propias afirmaciones que aparecen en el texto, debe dar ejemplos propios y tiene que expresar y argumentar su acuerdo o no con la opinión del autor. Estas tareas requieren destrezas diferentes a las solicitadas en las muestras anteriores. Así, en esta prueba hay diversidad de tareas con exigencias diferentes que dotan de validez de contenido al instrumento. En su resolución se ven implicadas más destrezas de la CL y no solamente la de inferir.

En el instrumento se asigna puntaje a cada una de las tres partes otorgando la mitad a la resolución de las tareas del tercer texto. El otro 50 % lo distribuye entre las del primero (24%) y el segundo (26%). Puede afirmarse que en la construcción de esta muestra se consolida la validez de contenido y de

puntaje del instrumento. No obstante ello, se advierte una posible disminución de la fiabilidad del examen si se considera la cantidad y diversidad de textos y de tareas y el tiempo otorgado para su resolución. El instrumento no explicita los criterios para valorar las respuestas a las preguntas abiertas que propone el texto 3, lo que puede dificultar su corrección y dar lugar a subjetividades por parte de los evaluadores, debilitando la fiabilidad de la muestra.

El examen administrado en el año 2020, cumple con la validez de contenido, por cuanto propone dos textos auténticos y una diversidad de consignas que exigen poner en juego destrezas de la CL tales como: inferir, identificar información, resumir información. Las consignas que se presentan, a diferencia de la muestra del año 2019, son preguntas cuyas respuestas están en los mismos textos o se infieren. El puntaje se distribuye por partes iguales entre los dos ítems de la muestra.

### ***Expresión escrita***

Las actividades del examen para evaluar la expresión escrita (EE) son tareas integrales que están orientadas a medir el dominio de los múltiples aspectos de la competencia discursiva: género textual, coherencia, gramática, ortografía, léxico, incorporación del contexto. Según el reglamento del ESLE, “El candidato deberá ser capaz de escribir textos no especializados de naturaleza personal con informaciones relativas a la vida cotidiana y a aspectos no muy

especializados del ámbito público” (Resolución N° 4035, 2011, p.4).

La sección del examen correspondiente a la EE comprende dos tareas con dos opciones cada una, a desarrollarse en 60 minutos; el candidato debe elegir, para resolver, una opción para cada tarea. El reglamento establece que la primera tarea consiste en escribir una carta y para la segunda tarea no establece género textual aunque sugiere tipologías “composición de naturaleza narrativa, descriptiva o discursiva” (Resolución N° 4035, 2011, p.4). En ambos casos, los textos deben tener una extensión entre 150 y 200 palabras.

En las primeras ediciones del examen, las consignas de las actividades pautaban en qué debía hacer foco el candidato para responder al propósito de la actividad, lo que garantizaba la validez de contenido ya que quedaba explícito lo que se pretendía medir, pero se detallaba exhaustivamente la estructura que debían tener los textos, por lo cual la actividad se acercaba más a un ejercicio de práctica áulica que a una situación de comunicación real, a lo que debe ajustarse el examen según se especifica en el reglamento del ESLE. En este caso se observó que la validez podía ser objeto de cuestionamientos.

A partir de la edición de 2019, también se solicita la producción de un género textual determinado pero debe producirse a partir de un texto input que pretende recrear la situación de comunicación en la que estaría inserto el candidato y a diferencia de la

versión 2012 no se explicita la superestructura del texto solicitado.

En la versión 2021, en contexto de pandemia, el examen se tomó bajo la modalidad virtual, se solicitó solo una tarea de expresión escrita, con dos opciones, se redujo el tiempo de 60 minutos a 50 con la misma cantidad de palabras. En este caso, una de las opciones contextualiza la actividad vinculándola con el texto de escucha de la actividad de comprensión auditiva. Por las características que reúne esta actividad, en cualquiera de sus versiones, podemos garantizar la validez de contenido de la tarea. En cuanto a la validez concurrente, aunque el reglamento del ESLE no prevé mecanismos de valoración de los resultados para comprobarla, se pudo observar que esta sección del examen arroja resultados relativamente homogéneos. En cuanto a la validez predictiva, las actividades presuponen un escritor de textos cotidianos sin especificidad académica, a pesar de que el examen es requisito para el ingreso a los estudios universitarios. Al tratarse de un examen de destrezas integradas, en el caso de la expresión escrita, para constatar la validez de puntaje se requiere de descriptores que permitan una distribución exhaustiva de los valores para cada uno de los planos que intervienen en la escritura.

Y si bien un examen en las que se intente reproducir situaciones de comunicación cotidianas nunca será una situación realmente auténtica, las actividades de expresión escrita reúnen las características de autenticidad situacional y de autenticidad

interaccional ya que requieren del estudiante activar las competencias que le permitan usar la lengua en un acto comunicativo real.

### **Expresión oral**

La destreza de expresión oral (EO) en el ESLE cuenta con un reglamento que se acerca a las especificaciones propias de un examen de dominio. Como objetivo, se declara certificar el nivel de usuario independiente/ umbral, descrito en el MCER (2002). Se reconocen en el candidato los siguientes descriptores, para la interacción:

Sé desenvolverme en casi todas las situaciones que se me presentan cuando viajo donde se habla esa lengua. Puedo participar espontáneamente en una conversación que trate temas cotidianos de interés personal o que sean pertinentes para la vida diaria (por ejemplo, familia, aficiones, trabajo, viajes y acontecimientos actuales). (p.30)

Y para la expresión individual:

Sé enlazar frases de forma sencilla con el fin de describir experiencias y hechos, mis sueños, esperanzas y ambiciones. Puedo explicar y justificar brevemente mis opiniones y proyectos. Sé narrar una historia o relato, la trama de un libro o película, y puedo describir mis reacciones. (p. 30)

A partir de estos descriptores, los evaluadores de ESLE han diseñado tareas e interpretado los resultados de cada toma. Es importante recalcar que el examen lleva entre 15 y 20 minutos por candidato y consta de 3 partes. La primera está dedicada al conocimiento del candidato, quien se presenta, habla de sí mismo y contesta preguntas del entrevistador relacionadas con

su aprendizaje de la lengua, sus estudios, trabajo e intereses, en general, de acuerdo al diálogo que se va construyendo entre ambos. En la segunda parte se le dan al candidato dos láminas. Deberá elegir una de ellas y preparar en alrededor de dos minutos, un comentario sobre el tema de la lámina, describir y explicar lo que vea en ella y desarrollando en todo lo posible su opinión sobre el tema. Puede agregar comparaciones, ejemplos, citas, anécdotas, etc. para presentar, desplegar, justificar su opinión o mostrar su desacuerdo con otras opiniones. Después de escucharlo, entablará un diálogo breve con el examinador.

En la última actividad, encontrará en el reverso de la lámina la propuesta para representar una situación de habla con un tema relacionado al discutido previamente. Se trata de tomar un papel y conversar de manera espontánea, para demostrar su capacidad de interactuar en situaciones cotidianas.

El ESLE conserva cierta validez de contenido porque no se aleja de los descriptores de la expresión oral para el B1 como objetivo. Sin embargo, lo establecido en el reglamento se lleva a cabo en el 80% de los exámenes orales, la última actividad no se observa en los exámenes del 2019/2020, la cual consiste en representar una situación de habla simulada, para demostrar su capacidad de interactuar en situaciones cotidianas, espontaneidad. Otro elemento presente es la validez concurrente: la concordancia de los resultados observados en varios individuos con los

resultados de otros exámenes a los mismos individuos y en tiempos semejantes (Prati, 2007). Sin embargo, lo observado es de carácter accidental. En cuanto a la validez predictiva, los resultados de candidatos que debieron rendir más de una vez, hasta aprobar cada parte, nos permite predecir el desempeño futuro de los candidatos en la universidad u otros ámbitos.

Finalmente, por la validez de puntaje resulta complejo reconocerla objetivamente, no se cuenta con escalas de medición exacta, sino más bien con los parámetros antes mencionados con un margen de subjetividad. Es real que esta interpretación se basa en las observaciones de los exámenes orales realizados entre los años 2019 y 2021, y un registro escrito parcial de las mismas, que consiste en una grilla de trabajo consensuada entre los evaluadores. Toda la información recuperada es igual de valiosa. En conclusión, el ESLE es válido en la medida que el contenido de su diseño no se aleja de su objetivo ni sus parámetros de interpretación, aunque estos sean precarios. En la misma medida, busca la autenticidad y en ello, cierta validez predictiva.

### **Fiabilidad**

Analizar los resultados del ESLE propiamente dichos implica contar con un archivo sólido, previsto en las especificaciones del examen. Sin embargo el reglamento no consideró uno para posteriores investigaciones sobre las actuaciones de los candidatos. En lo que refiere a las prácticas metodológicas, ya se

describieron las condiciones del corpus con el que se cuenta.

Sin embargo, en el orden de lo práctico, el equipo evaluador debió acordar un protocolo para garantizar la interpretación más fiable de la actuación del candidato. Una vez terminada la parte escrita del examen, que incluye las secciones dedicadas a la CA, CL y EE, se procede a la parte oral. Se reconoce cierta fiabilidad interevaluadora en estas entrevistas individuales. Los dos evaluadores presencian el examen asumiendo diferentes roles en forma alternada: uno es el interlocutor, que guía la conversación con el candidato, y el otro observa para lograr una interpretación holística de la actuación, a partir de una grilla creada para la ocasión. Al finalizar, se toman unos minutos para acordar la valoración, transferida a un puntaje sobre 100%. Al término de cada toma, los evaluadores distribuyen las partes escritas y la grilla de la entrevista oral. Cada una corrige todas las áreas de un mismo examen. De esta manera, con una mirada horizontal de las actuaciones, se pretende encontrar cierta coherencia que permita otorgar o no el nivel de lengua.

Por otra parte, en un sentido vertical, los evaluadores registran las respuestas de una misma sección a medida que se puntan los ítems: cuáles son definitivamente correctas y cuáles no o están en la frontera. Ante la duda, se consulta con el otro evaluador para acordar la respuesta esperada. ¿Cómo está interactuando el examinado con la tarea? ¿Qué está midiendo la pregunta? Con esta información se definen los casos fronterizos

en equipo. Incluso se puede volver atrás y puntar una vez más. Este no es un procedimiento ideal pero sí posible cuando se cuenta con recursos limitados. Con estos acuerdos se intentó dar consistencia a la fiabilidad de los resultados.

Considerando las garantías de una fiabilidad extrínseca, siempre se previó las mismas condiciones en la administración del examen. Siguiendo el reglamento, el equipo evaluador procedió con los recursos que contaba para garantizar la fiabilidad del examen, desde un contexto seguro, un equipo de audio que proyectó un sonido de calidad, un procedimiento administrativo claro, instrucciones precisas para procedimientos regulares en cada sección, la presentación de ejemplares legibles, textos originales para cada toma, balance entre el puntaje y la actividad, procesos de corrección ecuanimes, hasta el cumplimiento de plazos antes, durante y después de la toma. Aun se puede hallar errores que escapan a la capacidad del equipo pero se considerarán en términos de viabilidad para el nuevo diseño del ESLE.

### **Autenticidad**

Si se considera la noción de autenticidad, puede ser obvio que textos auténticos motiven operaciones auténticas sin embargo esto resulta relativo para el candidato. En el caso de la CA, la escucha atenta en un examen nunca será exactamente igual a las condiciones de un medio natural. Resolver un examen con solo dos oportunidades para comprender y responder a una serie de ejercicios, demandando una audición

intensiva, que requiere de toda su atención, está lejos de parecerse a escuchar la radio en casa por entretenimiento. El candidato escucha con nerviosismo o temor por los resultados y esto puede afectar su actuación. Sin embargo, en pos de lograr cierta autenticidad, el examen ESLE evidencia rasgo que la sostiene. Entre sus tareas para destrezas receptivas, pasó de seleccionar textos pertenecientes a libros de idiomas, representativos del español peninsular, a priorizar textos de medios masivos argentinos, sin editar. Este cambio permitió que el instrumento se volviera más auténtico porque afectó la posición de oyente/lector que asume el examinado al momento de realizar la tarea: pasó de ser un oyente/lector accidental a ser miembro del público general, destinatario del producto. Todos son coherentes con los descriptores MCER (2002) que requieren discurso en lengua estándar. Por otra parte, la autenticidad en las destrezas productivas se observa en el grado de correspondencia entre las características de una determinada actividad del examen y las características de una actividad en el ámbito donde se usa la lengua meta, en este caso se crean situaciones concretas de comunicación interpersonal.

En relación a los criterios de evaluación, la interpretación de las evaluadoras está basada en el desempeño de los candidatos y se generaliza más allá del contexto de examen. Si bien es cierto que no existe una autenticidad total, como sostiene Prati (2007), las actuaciones son más o menos

auténticas al ser contextualizadas en posibles situaciones de la cotidianidad. Sin embargo, todos ellos tienen la estructura de una conversación, casual, un símil de situación concreta de comunicación y coherente con los descriptores MCER (2002) que requieren discurso en lengua estándar. Tal es el caso de la toma del año 2020 donde se tomó como referencia la situación de pandemia suscitada en ese año.

### **Viabilidad**

Un examen es viable en la medida en que es práctico y posible en la situación que se gestione. En la experiencia del ESLE, las personas que diseñaron el instrumento y su reglamento no son las mismas que lo luego lo gestionaron. Por lo tanto, la tarea fue asignada sin una planificación de los recursos en cuestión. El equipo de trabajo fue cambiando a lo largo de los años pero nunca superó los dos miembros, que cumplieron la función de elaboradores, escritores, correctores, administradores y comunicadores del instrumento evaluador.

En relación al tiempo de trabajo, y considerando al ESLE como requisito de ingreso, siempre se programaron fechas cercanas al término del período de inscripción definitiva de los ingresantes. Lo que exigió al equipo tiempos muy acotados para todos los compromisos. Se quiso programar fechas diferentes pero nunca hubo inscritos. Por otro lado, las instalaciones y equipos técnicos de la universidad hicieron posible la gestión del examen.

Sin lugar a dudas, la capacidad de un equipo reducido influyó en algunas decisiones que comprometen la validez y la fiabilidad del examen. Cambios en el tipo de prueba o acordar criterios de corrección realistas a los tiempos demandantes pudieron incluso cambiar algunas precisiones del reglamento conocido. Sin embargo, siempre se priorizó por una parte la coherencia con descriptores del nivel a certificar, objetivo principal del ESLE, y por otra parte, la comunicación clara y precisa de los procedimientos para no faltar a la confianza de los propios candidatos involucrados en el proceso.

### **Discusión**

El ESLE fue creado con la intención de garantizar el dominio del español en los ingresantes a la UNLaR, que no fueran hispanohablantes, porque de esta manera también se garantiza un proceso de aprendizaje en el entorno académico. El examen evalúa la competencia comunicativa, en forma oral y escrita, sin considerar sus lenguas maternas ni el tipo de adquisición o aprendizaje que hayan realizado.

A pesar de todas sus falencias en su formulación y diseño, el contenido del examen y la interpretación de sus resultados cumplieron con su objetivo: certificar el nivel intermedio, en coherencia con lo que se entiende por este nivel desde el MCER (2002). La autenticidad de las tareas que lo integran respalda su validez. Una viabilidad no planificada dio respuesta ante las exigencias de resultados fiables. Por eso el equipo evaluador procedió con los recursos

que contaba para garantizar la fiabilidad del examen. Estas cualidades son el resultado de un análisis sujeto a muestras parciales, que en el mejor de los casos, evidencian la continuidad de un proyecto sostenido con un gran esfuerzo.

En miras al nuevo formato del ESLE, ¿qué se debería recuperar, mantener, cambiar del modelo vigente? En el marco del proyecto de investigación que motiva este trabajo, son muchas las tareas que quedan por hacer: establecer las especificaciones acordadas, crear un banco de materiales y un archivo de exámenes resueltos, planificar herramientas de evaluación de la validez, fiabilidad y viabilidad, establecer criterios de puntuación y corrección formales, capacitar al equipo de evaluadores en tareas diferentes, entre otras. Para la próxima etapa, el equipo docente del área se ha consolidado en torno a un desafío: crear un instrumento útil y que arroje resultados consistentes, capaz de dar respuestas con la suficiente responsabilidad que la comunidad requiere, como afirma Bordón (2004) “tal vez el reto que la evaluación de la lengua tiene por delante sea el de elaborar exámenes, o tal vez otros procedimientos alternativos, cuyo objetivo sea el de ser justos” (p. 26).

### **Referencias**

- Bordón M., T. (2004). Panorama histórico del desarrollo de algunas de las cuestiones fundamentales en la evaluación de SL. *Carabela*, 55, 5-29.
- Bordón M., T. (2004) La evaluación de la expresión oral y de la comprensión



auditiva. En J. Sánchez Lobato e I. Santos Gargallo (Dires.), *Vademécum para la formación de profesores* (pp. 983-1003). Madrid: SGEL.

Bordón M., T. (2000). La evaluación de la expresión oral en el aula de E/ELE. *Carabela*, 47, 151-174.

Consejo de Europa (2002). *Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza y evaluación*. Instituto Cervantes para la traducción en español.

<http://cvc.cervantes.es/obref/marco>

Llosa, L. (2006). *Estado actual de las investigaciones sobre validación de exámenes de lengua extranjera*. Conferencia dictada en el Segundo Coloquio CELU, Universidad Nacional de Mar del Plata. <http://www.celu.edu.ar/archivos/pp.pps>.

Pastor Villalba, C. (2009). La evaluación de la comprensión oral en el aula de ELE. *marcoELE. Revista de Didáctica Español Lengua Extranjera*, 9, 1-27.

Prati, S. (2007). *La evaluación en español lengua extranjera. Elaboración de exámenes*, Buenos Aires: Libros de la Araucaria.

Resolución N° 4035 de 2011. Consejo Superior de la Universidad Nacional de La Rioja. Por la cual se establece el Reglamento de Examen de Suficiencia de Lengua Española para Estudiantes Extranjeros. 24 de octubre de 2011.

**Marisol Andrea Méndez Cabrera.**  
Licenciada en Letras y Profesora de

enseñanza media y superior en Letras. Magíster en Enseñanza de Español como Lengua Extranjera (ELE). Docente de la Universidad Nacional de La Rioja (UNLaR), dedicada al área de la Lingüística y la Lingüística Aplicada. Representante Académico de la UNLaR en el Consorcio ELSE - CIM. Sus investigaciones se relacionan con la Lingüística Aplicada, en el área del ELE y la lectocomprensión en el nivel superior.

Correo electrónico: mendezcm@unlar.edu.ar

**Patricia Elena Compagnoni.** Profesora de Castellano, Literatura y Latín para la Enseñanza Superior. Docente de Expresión Oral y Escrita y de Lengua Española III en la Universidad Nacional de La Rioja. Autora, evaluadora y correctora del examen ESLE de la UNLaR. Correctora del examen CELU. Sus investigaciones se inscriben en el área de la Sociolingüística y de la enseñanza de Español como Lengua Extranjera.

Correo electrónico:

pcompagnoni@unlar.edu.ar

**María Ana Glaccaglia,** Profesora en Ciencias de la Educación con especialización en enseñanza de la lengua y la literatura en el nivel secundario. Docente del Campo de la práctica Profesional en Institutos de Formación Docente de la provincia de La Rioja y de la Universidad Nacional de La Rioja. Coordinadora de la carrera de Profesorado en Letras de la UNLaR. Integra

el Equipo Técnico de la Dirección General de Educación Superior de La Rioja.

Correo electrónico:

[mgiaccaglia@unlar.edu.ar](mailto:mgiaccaglia@unlar.edu.ar)

**María del Carmen Brizuela.** Profesora universitaria en Letras para nivel secundario y superior. Docente de Expresión Oral y Escrita/ Docente de ELE (Español como lengua extranjera) en Universidad Nacional de La Rioja. Docente de Lengua y Literatura en el Colegio Pre-universitario Gral. San Martín. Participante activa de proyectos de extensión pertenecientes al departamento de Humanas y Cs. de la Educación. Sus investigaciones se relacionan con literatura latinoamericana y ELE. Docente de Didáctica de lengua y literatura I y Literatura Argentina en ISFD Albino Sánchez Barros.

Correo electrónico: [mcbrizuela@unlar.edu.ar](mailto:mcbrizuela@unlar.edu.ar)