



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

DISEÑO Y DESARROLLO DE HERRAMIENTA PARA EL DESARROLLO DE HABILIDADES DE PENSAMIENTO CRÍTICO Y SU IMPLEMENTACIÓN EN UN MODELO *FLIPPED CLASSROOM*

VICENTE REINALDO MARTIN TARUD

Tesis para optar al grado de
Magister en Ciencias de la Ingeniería

Profesor Supervisor:
MIGUEL NUSSBAUM

Santiago de Chile, (Enero, 2019)

© 2019, Vicente Martin



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

DISEÑO Y DESARROLLO DE HERRAMIENTA PARA EL DESARROLLO DE HABILIDADES DE PENSAMIENTO CRÍTICO Y SU IMPLEMENTACIÓN EN UN MODELO *FLIPPED* *CLASSROOM*

VICENTE REINALDO MARTIN TARUD

Tesis presentada a la Comisión integrada por los profesores:

MIGUEL NUSSBAUM

VALERIA HERSKOVIC

ALEJANDRA MENESES

DIEGO LÓPEZ-GARCÍA

Para completar las exigencias del grado de
Magister en Ciencias de la Ingeniería

Santiago de Chile, (Enero, 2019)

A mi familia y amigos

AGRADECIMIENTOS

Agradezco a todos los que fueron parte de este proyecto, y también a quienes aportaron y colaboraron con ideas buena disposición y apoyo moral, en particular a Miguel Nussbaum por invitarme a ser parte de este proyecto y guiar el proceso de investigación. Por sobre todo agradezco a mis padres quienes me dieron la posibilidad de estudiar, quienes junto a Patricia, mis hermanos y mis amigos estuvieron a mi lado en todo momento, escuchando, apoyando y aguantando mis momentos de mal genio.

INDICE GENERAL

	Pág.
DEDICATORIA	ii
AGRADECIMIENTOS	iii
INDICE DE TABLAS	viii
INDICE DE FIGURAS.....	xiii
RESUMEN.....	xiv
ABSTRACT	xv
 1. MOTIVACIÓN, MARCO TEÓRICO Y SOLUCIÓN PROPUESTA.....	 1
1.1 Habilidades y Competencias del Siglo XXI.....	3
1.2 <i>Flipped Classroom</i>	5
1.3 Contexto, Funcionamiento, Diseño y Desarrollo de la Solución	6
1.3.1 Contexto.....	6
1.3.2 Plataforma web	9
1.4 Contribución de la solución	29
1.5 Conclusiones	30
 2. ARTÍCULO ENVIADO: TECHNOLOGY-SCAFFOLDED PEER ASSESSMENT FOR DEVELOPING CRITICAL THINKING IN PRE-SERVICE TEACHER TRAINING	 33
2.1. Introducción	34
2.2. Metodología	36
2.2.1. Estrategia utilizada: modelo de andamiaje para la evaluación de pares	 36
2.2.2. Contexto y participantes	41
2.2.3. Instrumentos de evaluación	43
2.2.4. Aplicación de rúbricas	47
2.2.5. Análisis cuantitativo	48
2.2.6. Análisis cualitativo	53

2.2.7. Consideraciones éticas.....	53
2.3. Resultados	54
2.3.1. Cuantitativos	54
2.3.2. Cualitativos	59
2.3.2.1. Evaluación de pares y propiedades.....	60
2.4. Discusión	62
2.4.1. Desarrollo de las habilidades de orden superior	62
2.4.2. Influencia de la evaluación de pares en el desarrollo de habilidades superiores.....	63
2.4.3. Modelo de andamiaje para profesores en formación.....	65
2.4.4. Uso de tecnología.....	68
2.5. Conclusiones.....	69
2.6. Apéndices.....	71
A. Detalle de actividades realizadas y habilidades de pensamiento crítico involucradas.....	71
B. Síntesis de habilidades de pensamiento crítico, indicadores y descriptores de rúbricas R y NR	72
C. Instrumentos de evaluación: rúbrica de R y NR, F e integración de F en NR	73
D. Cohen Kappa de instrumento de evaluación.....	79
E. Tabulación de Reponse, New Response y Feedback para cada actividad.....	86
F. Descripción mediante ejemplos y cálculo de variables consideradas en el modelo de regresión lineal	96
G. Proceso de selección de variables modelo de regresión lineal de <i>Avance</i>	99
H. Datos utilizados en el modelo de regresión lineal.	100
I. Distribución del <i>Avance</i>	101
J. Ejemplo de aplicación del modelo de andamiaje.....	102
K. Tablas de correlaciones de variables eliminadas del modelo de regresión.....	105
L. Doble codificación de <i>Response</i> (R), <i>New Response</i> (NR) e Incorporación del <i>Feedback</i> en la <i>New Response</i> (NR(FR)).....	107
M. Guía para la Retroalimentación	119

BiBliografía.....	121
Capítulo 1:	121
Capítulo 2:	123
A N E X O S.....	134
A- ARTÍCULO TRADUCIDO: TECHNOLOGY-SCAFFOLDED PEER ASSESSMENT FOR DEVELOPING CRITICAL THINKING IN PRE-SERVICE TEACHER TRAINING	135
1. Introduction	136
2. Methodology	137
2.1. Chosen strategy: Scaffolded peer assessment.....	137
2.2. Context and participants	141
2.3. Assessment instruments.....	143
2.4. Applying the rubrics	145
2.5. Quantitative analysis.....	146
2.6. Qualitative analysis.....	150
2.7. Ethical considerations	150
3 Results	151
3.1 Quantitative.....	151
3.2 Qualitative.....	155
4 Discussion	158
4.1 Developing higher-order thinking skills	158
4.2 Influence of peer assessment on the development of higher order thinking skills.....	159
4.3 Scaffolding model for pre-service teachers	160
4.4 Use of technology	162
5 Conclusions	163
6 References	165
7. Apendix	172
A. Description of the activities and the critical thinking skills they involve	172
B. Summary of the critical thinking skills, indicators and descriptors for the rubrics applied to R and NR.....	173

C.	Assessment instruments: rubric for assessing R, NR, F and incorporation of F into NR.....	174
D.	Cohen's kappa for the assessment instrument	181
E.	Scores for Response, New Response and Feedback for each activity	188
F.	Description, examples and calculation of the variables included in the linear regression model	198
G.	Process for selecting variables for the linear regression model for <i>Improvement</i>	201
H.	Data used in the linear regression model	202
I.	Distribution of Improvement	203
J.	Sample use of the scaffolding model.....	204
K.	Tables with the correlations of the variables discarded from the regression model	207
L.	Double coding of Response (R), New Response (NR) and Incorporation of Feedback into New Response (NR(FR))	209
M.	Set of the guidelines for giving feedback	221
B-	Carta de recepción de artículo	223

INDICE DE TABLAS

	Pág.
Capítulo 1:	
Tabla 1-1: Requisito Funcional 01.....	17
Tabla 1-2: Requisito Funcional 02.....	18
Tabla 1-3: Requisito Funcional 03.....	18
Tabla 1-4: Requisito Funcional 04.....	19
Tabla 1-5: Requisito Funcional 05.....	19
Tabla 1-6: Requisito Funcional 06.....	20
Tabla 1-7: Requisito Funcional 07.....	20
Tabla 1-8: Requisito Funcional 08.....	21
Tabla 1-9: Requisito Funcional 09.....	21
Tabla 1-10: Requisito Funcional 10.....	22
Tabla 1-11: Requisito Funcional 11.....	22
Tabla 1-12: Requisito Funcional 12.....	23
Tabla 1-13: Requisito Funcional 13.....	23
Tabla 1-14: Requisito Funcional 14.....	24
Tabla 1-15: Requisito Funcional 15.....	24
Tabla 1-16: Requisito No Funcional 01.....	25
Tabla 1-17: Requisito No Funcional 02.....	25
Tabla 1-18: Requisito No Funcional 03.....	26
Capítulo 2:	
Tabla 2-1: Resumen de las actividades realizadas	42
Tabla 2-2: Variables consideradas para construir el modelo de regresión lineal.....	50
Tabla 2-3: Resumen análisis de regresión lineal de variables que predicen el avance de las habilidades de pensamiento crítico (N = 35)	55
Tabla 2-4: Estadísticos descriptivos de Response (R), New Response (NR) y Avance (A) y t-tests de significancia estadística del Avance (A), en cada una de las actividades..	56

Tabla 2-5: Estadísticos descriptivos de Response (R), New Response (NR) y Avance(A) y t-tests de significancia estadística del Avance (A), del promedio de las habilidades involucradas en cada actividad y del promedio de las tres actividades.57	57
Tabla 2-6: Estadísticos descriptivos Feedback Received (FR), Feedback Delivered (FD) y de la Integración de FR en NR (NR(FR)).58	58
Tabla 2-7: Distribución resultados de aplicación de rúbrica de R y NR según criterios de evaluación de la rúbrica59	59
Tabla 2-8: Categoría clave y categorías60	60
Tabla 2-9: Categorías y citas60	60
Tabla D1: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR). Dimensión: Habilidad de Análisis.79	79
Tabla D2: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR). Dimensión: Habilidad de Explicación.79	79
Tabla D3: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR). Dimensión: Habilidad de Interpretación.79	79
Tabla D4: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR). Dimensión: Habilidad de Evaluación.80	80
Tabla D5: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR). Dimensión: Habilidad de Inferencia.80	80
Tabla D6: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Criterios de Evaluación.80	80
Tabla D7: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Claridad y Pertinencia de la Explicación.....81	81
Tabla D8: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Intención de la Argumentación.81	81
Tabla D9: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Formulación de la Retroalimentación.81	81
Tabla D10: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Solicitud de Evidencia.82	82

Tabla D11: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).Dimensión: Relectura de un Argumento o Presentación de un Nuevo Argumento y/o Contraargumento.	82
Tabla D12: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Solicitud de Reformulación.	82
Tabla D13: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F). Dimensión: Llamado de Atención.	83
Tabla D14: Matriz de nivel de acuerdo entre correctores en Rúbrica de integración de Feedback (F) en New Response (NR).	83
Tabla D15: Frecuencias esperadas de Cohen Kappa de cada criterio de evaluación: Rúbrica de Response (R) y New Response (NR), Rúbrica de Feedback (F) y Rúbrica de integración de Feedback (F) en New Response (NR) según dimensiones correspondientes.	84
Tabla D16: Factor Kappa de cada dimensión evaluada por Rúbrica de R y NR, Rúbrica de F y Rúbrica de integración de F en NR.	85
Tabla E1: Tabulación de Response (R) y New Response (NR) de Actividad 1.	86
Tabla E2: Tabulación de Response (R) y New Response (NR) de Actividad 2.	87
Tabla E3: Tabulación de Response (R) y New Response (NR) de Actividad 3.	88
Tabla E4: Tabulación del Avance entre R y NR.	89
Tabla E5: Tabulación de Feedback Received (FR) y Feedback Delivered (FD) de Actividad 1.	90
Tabla E6: Tabulación de Feedback Received (FR) y Feedback Delivered (FD) de Actividad 2.	91
Tabla E7: Tabulación de Feedback Received (FR) y Feedback Delivered (FD) de Actividad 3.	92
Tabla E8: Tabulación de Integración de Feedback Received en New Response NR(FR). .	93
Tabla E9: Tabulación de promedios de actividades en Response (R), New Response (NR), Avance, Feedback Received (FR), Feedback Delivered (FD) e Integración de FR en NR (NR(FR)).	94

Tabla E10: Tabulación de Test 1 (T1) y Test 2 (T2).	95
Tabla G1: Modelos engenerados con stepwise selection method en funcion de variable dependiente Avance	99
Tabla G2: Forward stepwise selection method para modelo de Avance (N=35)	99
Tabla H: Datos utilizados en STATA v. 12.0. para elaboración de la regresión lineal que predice el Avance de las habilidades de pensamiento crítico.	100
Tabla I1: Distribución del avance entre Response (R) y New Response (NR) en cada actividad, según habilidad de Pensamiento Crítico.....	101
Tabla K1: Correlación de <i>R</i> con variables descartadas del modelo por redundancia de información con respecto a <i>R</i>	105
Tabla K2: Correlación de <i>NR</i> con variables descartadas del modelo por redundancia de información con respecto a <i>NR</i>	105
Tabla K3: Correlación de <i>F</i> con variables descartadas del modelo por redundancia de información con respecto a <i>F</i>	106
Tabla L1: Puntajes de la doble codificación de la rúbrica aplicada a <i>Response</i> (R) en la Actividad 1.	107
Tabla L2: Puntajes de la doble codificación de la rúbrica aplicada a <i>Response</i> (R) en la Actividad 2.	108
Tabla L3: Puntajes de la doble codificación de la rúbrica aplicada a <i>Response</i> (R) en la Actividad 3.	109
Tabla L4: Puntajes de la doble codificación de la rúbrica aplicada a <i>New Response</i> (R) en la Actividad 1.	110
Tabla L5: Puntajes de la doble codificación de la rúbrica aplicada a <i>New Response</i> (R) en la Actividad 2.	111
Tabla L6: Puntajes de la doble codificación de la rúbrica aplicada a <i>New Response</i> (R) en la Actividad 3.	112
Tabla L7: Puntajes de la doble codificación de la rúbrica aplicada a <i>Feedback</i> (F) en la Actividad 1.	113

Tabla L8: Puntajes de la doble codificación de la rúbrica aplicada a <i>Feedback</i> (F) en la Actividad 2.	114
Tabla L9: Puntajes de la doble codificación de la rúbrica aplicada a <i>Feedback</i> (F) en la Actividad 3.	115
Tabla L10: Puntajes de la doble codificación de la rúbrica aplicada a la <i>Incorporación del Feedback en la New Response</i> (NR(FR)) en la Actividad 1.	116
Tabla L11: Puntajes de la doble codificación de la rúbrica aplicada a la <i>Incorporación del Feedback en la New Response</i> (NR(FR)) en la Actividad 2.	117
Tabla L12: Puntajes de la doble codificación de la rúbrica aplicada a la <i>Incorporación del Feedback en la New Response</i> (NR(FR)) en la Actividad 3.	118

INDICE DE FIGURAS

	Pág.
Figura 1-1: Marco de referencia para las habilidades del siglo XXI. Fuente: http://www.p21.org/our-work/p21-framework	4
Figura 1-2: Pirámide de Bloom. Fuente: http://marian81fc.blogspot.co.uk/2016/11/taxonomia-de-bloom.html	6
Figura 1-3: Estructura semanal aplicación modelo <i>Flipped Classroom</i> . Fuente: elaboración propia	9
Figura 1-4: Ejemplo de vista de alumno para <i>responder</i> . Fuente: elaboración propia	12
Figura 1-5: Ejemplo de vista de alumno para retroalimentar. Fuente: elaboración propia ..	13
Figura 1-6: Ejemplo de vista de alumno para rehacer. Fuente: elaboración propia	14
Figura 1-7: Ejemplo de vista de profesor para responder. Fuente: elaboración propia	15
Figura 1-8: Patrón Modelo-Vista-Controlador. Fuente: https://helloacm.com/model-view-controller-explained-in-c/	16
Figura 1-9: Simplificación de diagrama Entidad-Relación original. Fuente: Elaboración propia	28
Figura 2-1: Actividad de construcción de respuesta y evaluación de pares a través de dispositivos en red	41
Figura 2-2 Modelo de fases discursivas en actividades que desarrollan las habilidades del pensamiento crítico de interpretación, análisis, evaluación, inferencia, y explicación	67

RESUMEN

En los últimos 30 años, el mundo en que vivimos ha sufrido grandes transformaciones. La era digital invade cada vez con más fuerza distintos aspectos de nuestra vida diaria. Aún así, aspectos como la educación, han permanecido estáticos y no han sufrido grandes cambios en ese sentido. Este proyecto de investigación intenta generar un aporte de innovación y cambio para transformar la enseñanza. Para esto se desarrolló una herramienta digital que propone una metodología de enseñanza especializada. La herramienta fue utilizada en cursos los de Didáctica de la Matemática y Didáctica de la Lectura impartidos por la Facultad de Educación de la Pontificia Universidad Católica de Chile, mientras en éstos se implementaba una metodología de enseñanza *Flipped Classroom*. Con el uso de la plataforma se promueve el desarrollo de habilidades superiores, entregando un rol activo a los estudiantes. La metodología de la plataforma permite que los estudiantes realicen actividades de interacción pasiva, entregando y recibiendo retroalimentación de sus compañeros, construyendo aprendizaje en función de ese trabajo.

En el primer capítulo se introduce y analiza el problema que presenta el modelo educativo actual y luego se entregan detalles de la solución propuesta a dicho problema. El segundo capítulo es el artículo científico con título “*Technology-Scaffolded Peer Assessment For Developing Critical Thinking in Pre-Service Teacher Training*.” En él se presenta un estudio del uso de la herramienta desarrollada en el curso de Didáctica de la Lectura, y cómo su uso promueve el desarrollo de las habilidades de pensamiento crítico. Como principal resultado de este proyecto, se observa que con el uso de la herramienta desarrollada efectivamente los estudiantes desarrollan habilidades de pensamiento crítico, una de las Habilidades del Siglo XXI.

Este trabajo contó con el apoyo del proyecto FONDECYT-CONICYT 1150045 y 1180024.

Palabras Claves: Pensamiento Crítico, Flipped Classroom, Formación de Profesores, Habilidades del Siglo XXI.

ABSTRACT

In the last 30 years the world has changed. The digital age pervasively invades every aspect of our daily life. Yet, it could be said that education has remained static in relation to other transformations in society. This project is meant to contribute and innovate in the teaching activities, by developing a digital tool with a specific teaching methodology. A web platform was developed to promote superior thinking abilities, giving the students an active learning role when using the platform. The platform methodology allows the students to participate in activities with passive interaction, learning by the action of giving and receiving feedback from their classmates. The platform was used in the courses of Mathematics Methods and Reading Methods, for trainee teachers from the Pontificia Universidad Católica de Chile, while a Flipped Classroom model was implemented in both courses.

The first chapter of this document is an introduction and analysis of the problem that is present in the current education model, then, details of the proposed solution are given. The second chapter is the paper “Technology-Scaffolded Peer Assessment For Developing Critical Thinking in Teacher Professional Development”. The paper studies how the use of the tool in the Reading Course helped the students promote critical thinking skills. It was found that with the use of the platform students developed critical thinking skills, considered one of the 21st Century Skills.

This investigation was supported by the FONDECYT-CONICYT project numbers 1150045 and 1180024.

Keywords: Critical Thinking, Flipped Classroom, Teacher Training, 21st Century Skills.

1. MOTIVACIÓN, MARCO TEÓRICO Y SOLUCIÓN PROPUESTA

Según el sitio *The Statistics Portal* (Statista, 2018), para el año 2020 se estima una población mundial de 7.800 millones de personas. El mismo sitio, en otro estudio estima que en el año 2017 3.580 millones de personas tenían conexión a internet. Es decir, hoy cerca de la mitad de la población mundial tiene acceso a internet. Internet es una de las principales fuentes de acceso a la información y al conocimiento. En ese sentido, la creciente cobertura y alcance de la población conectada a internet, se traduce en condiciones de fácil y rápido acceso que se tiene al conocimiento y la información. Producto de la creciente cobertura de la conexión a internet, para los profesores es cada vez más difícil mantener a los estudiantes motivados y comprometidos con su aprendizaje (Brabazon, 2016). El *overflow* de información y los dispositivos tecnológicos son una creciente distracción para los jóvenes, los que no sienten una necesidad de dependencia de los profesores. Una forma pragmática de analizar esta situación, es que los estudiantes sienten que internet puede suplir en gran medida lo que puede entregarles un profesor con clases presenciales, y así lo evidencia la creciente demanda en cursos online (Vioreanu, 2018). Hoy se tienen alumnos que no asisten a clases, que están distraídos en clase y que sienten que lo que puede enseñarles un profesor, es posible encontrarlo y reemplazarlo fácilmente con contenido de internet. Este fenómeno presenta grandes desafíos en cuanto a la formación y la enseñanza, desafíos que están presentes tanto en la educación primaria y la educación secundaria, como en la educación superior.

Por otro lado, la automatización de tareas, de la mano con las nuevas tecnologías como la robótica, la inteligencia artificial, *machine learning* y en general todas las tecnologías emergentes y los avances tecnológicos de los últimos veinte años, han cambiado y seguirán cambiando el mundo y la sociedad en la que vivimos. De acuerdo a un estudio realizado por la consultora McKinsey (Manyika J. et. al., 2017), entre 400 y 800 millones de empleos podrían desaparecer en el año 2030 y al mismo tiempo se crearán un número similar de nuevos puestos de trabajo. En otro artículo, la misma consultora

establece que “...la mitad de las actividades laborales de la actualidad pudieran automatizarse para 2055.” (Manyika J. et. al., 2017, p. 4).

Con estos antecedentes, si se tiene en cuenta el desafío que hoy implica el mantener la atención y motivación de los estudiantes, el “¿Cómo se está enseñando a los estudiantes?”, adquiere una gran importancia. Pues como nos dicen García F., Portillo J., et. al. (García F., Portillo J., et. al., 2007), “...la docencia y los procesos de aprendizaje deben adaptarse permanentemente a las características de los individuos que en cada momento la componen.” En la actualidad se habla de individuos o estudiantes nativos digitales, quienes están capturados por como la tecnología es parte de sus vidas e invade cada espacio de su entorno (Prensky, 2001). Sin embargo, los procesos de aprendizaje no han tenido grandes transformaciones en relación a los grandes cambios de la sociedad, y en cierta medida la falta de motivación de los estudiantes podría ser explicada por un modelo de enseñanza que aún no logra adaptarse del todo a las nuevas necesidades. El dominio de modelos de enseñanza centrados en el profesor, con clases frontales y expositivas, son evidencia empírica que demuestra la problemática recién mencionada, metodologías de enseñanza que no se adaptan a las nuevas necesidades.

Sin embargo el “¿Qué se está enseñando?” no deja de ser importante, más aún si se consideran los cambios que se esperan en el mundo laboral en los próximos 20 años. En ese sentido, es lógico cuestionar si lo que se está enseñando a niños y jóvenes, en las escuelas y universidades, les entrega las herramientas suficientes para enfrentarse y adaptarse a estos cambios. Hajnal (2018) presenta la influencia crucial que ha tenido en el incremento y desarrollo del trabajo de las organizaciones civiles, la eficiencia y las habilidades en el uso de las tecnologías de información y comunicación. Desde esa perspectiva Asselin (2016) identifica los objetivos de la enseñanza como la necesidad de desarrollar aprendedores de por vida (*lifelong learners*), y establece la importancia de preparar a los estudiantes para la literatura de las tecnologías de información y comunicación, ya que estas son centrales en la adquisición de conocimiento.

Las dos problemáticas anteriormente descritas, evidentemente no tienen una única solución. Este proyecto de investigación toma las Habilidades y Competencias del Siglo XXI como una solución al “qué”, y el *Flipped Classroom* como una solución al “cómo” y muestra su implementación. En otras palabras, el objetivo principal de este trabajo es desarrollar las Habilidades y Competencias del Siglo XXI, a través de una metodología de *Flipped Classroom*. En función de ese marco teórico, se diseñó y desarrolló una herramienta computacional con la que se desarrollan actividades al interior del aula. A continuación, se explica por qué las Habilidades del Siglo XXI y la metodología *Flipped* son una solución que se ajusta a las necesidades de la sociedad. Luego se expone la herramienta, su diseño, desarrollo y funcionamiento, como parte de la implementación de un modelo *Flipped Classroom* que busca desarrollar habilidades de pensamiento crítico, parte de las Habilidades del Siglo XXI.

1.1 Habilidades y Competencias del Siglo XXI

La OCDE ha enfrentado la problemática de el “¿Qué se está enseñando?”. Desde el año 2009 ha puesto el énfasis de la enseñanza en el desarrollo de las ya mencionadas Habilidades y Competencias del Siglo XXI (Ananiadou K. & Claro M., 2009), las que se piensa son las habilidades necesarias para adaptarse a los cambios que trae la llamada “4ta Revolución Industrial”. Tal como su nombre lo denota, las habilidades del siglo XXI se definen como las habilidades necesarias que deben desarrollar los estudiantes para tener éxito en el trabajo, en la vida y como miembros de una sociedad civil.

La Figura 1-1 muestra como el *Partnership for 21st Century Learning* ha separado las Habilidades y Competencias del Siglo XXI en tres áreas, las habilidades de aprendizaje e innovación; las habilidades digitales; y las habilidades de desarrollo personal y profesional. Este esquema no pretende dejar de lado el aprendizaje del

contenido, sino más bien propone llevar el aprendizaje de las materias como idiomas, arte, matemáticas, economía, ciencia, geografía, historia y educación cívica a un nivel más elevado de abstracción (Partnership for 21st Century Learning, 2015).

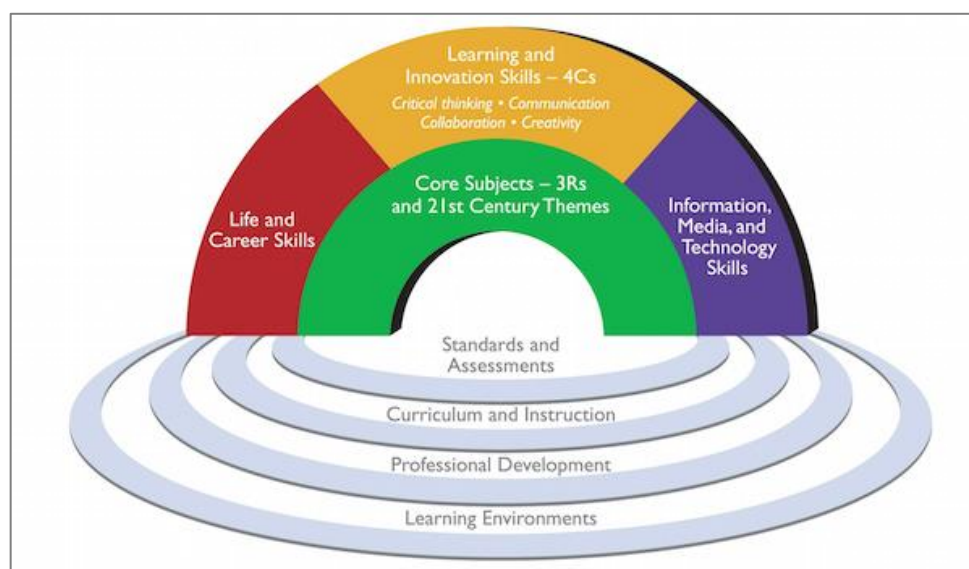


Figura 1-1: Marco de referencia para las habilidades del siglo XXI.

Fuente: <http://www.p21.org/our-work/p21-framework>

La herramienta desarrollada para la solución se centra en el desarrollo de las habilidades de aprendizaje e innovación. De acuerdo al mismo *Partnership for 21st Century Learning*, éstas son las habilidades que separan a los estudiantes que se encuentran preparados para el cada vez más complejo ambiente del siglo XXI en el mundo laboral y en la vida, y aquellos que no lo están. Las habilidades de aprendizaje e innovación, agrupan habilidades como la creatividad, la innovación, el pensamiento crítico, la solución de problemas, la comunicación y la colaboración. En el *paper* del Capítulo 2 se profundiza en cómo a través de la

herramienta y del modelo que ésta aplica, se desarrollan en particular las habilidades de pensamiento crítico.

1.2 *Flipped Classroom*

El *Flipped Classroom*, o aula invertida, es una metodología de enseñanza centrada en el estudiante. Varios estudios han demostrado que las metodologías de aprendizaje activo como el *flipped* hacen que los alumnos se sientan más comprometidos y motivados con el aprendizaje (Gilboy M. B., Heinrichs S. & Pazzaglia G., 2015) y (Anna Therese & Njå, 2017). Por tanto, el *Flipped Classroom* es una metodología de enseñanza que ha demostrado ser efectiva en términos de lo que este trabajo intenta resolver, estudiantes que le encuentren sentido a lo que un profesor puede entregarles.

Se le llama aula invertida, pues en términos de la taxonomía de Bloom (2001) los estudiantes deben realizar las actividades de menor nivel cognitivo, recordar y comprender el contenido, fuera de clases (lo que corresponde a los dos primeros niveles de la pirámide de Bloom mostrada en la Figura 1-2). Mientras que dentro de clases, se trabajan los niveles más altos de la pirámide (niveles 3 a 6 de la misma figura), junto con el apoyo de compañeros y tutores (Brame C., 2013). En la práctica, esto se traduce en que la primera aproximación de los estudiantes a un nuevo contenido se realiza de manera individual, mediante material audiovisual o de lectura facilitado por el profesor. Luego en clases, los estudiantes pueden trabajar en tareas más complejas como el análisis, la aplicación de los contenidos, la resolución de problemas y construcción de nuevas ideas. Esta secuencia en el proceso de aprendizaje es todo lo contrario a lo que plantea una clase tradicional, donde las clases son principalmente expositivas, con alumnos pasivos y dejando que estos desarrollen las actividades cognitivas más complejas en sus casas y sin el apoyo de compañeros y tutores.



Figura 1-2: Pirámide de Bloom. Fuente:

<http://marian81fc.blogspot.co.uk/2016/11/taxonomia-de-bloom.html>

1.3 Contexto, Funcionamiento, Diseño y Desarrollo de la Solución

La herramienta creada en este proyecto, es la implementación práctica de la solución para el desarrollo de las Habilidades y Competencias del Siglo XXI, a través de una metodología de *Flipped Classroom*. En concreto, se desarrolló una plataforma web para ser usada en cursos de la Facultad de Educación de la Pontificia Universidad Católica. La aplicación se encuentra montada en los servidores de la Facultad de Educación de la Pontificia Universidad Católica, y es posible acceder a ella a través del enlace <http://flipped.educacion.uc.cl/>.

1.3.1 Contexto

Para entender el contexto bajo el cual se desarrolla este proyecto, es necesario mencionar que este trabajo corresponde a una de las aristas del proyecto doctoral “*Desarrollo de un modelo de aprendizaje basado en el aula invertida en*

educación superior”. Este proyecto doctoral, a su vez forma parte de los proyectos FONDECYT 1150045 y 1180024.

El proyecto doctoral mencionado, ha estado trabajando en conjunto con la Facultad de Educación de la Pontificia Universidad Católica en la implementación de un modelo de *Flipped Classroom*, en los cursos de Didáctica de la Lectura y Didáctica de la Matemática. Durante el primer semestre del año 2017, la metodología de trabajo aplicada constó de tres etapas: (1) entrega del contenido por medio de videos, (2) una clase interactiva para la profundización de los contenidos y (3) el desarrollo de actividades de aplicación mediante el uso de la plataforma web desarrollada en este proyecto. Esta metodología se estructuró de manera semanal, sumando un total de 14 semanas en el caso de Didáctica de la Matemática y 11 semanas en Didáctica de la Lectura. Cada una de las etapas mencionadas, se desarrolló de la siguiente manera:

a) Entrega de contenido por medio de videos (1):

El día viernes eran publicados los videos que debían ser vistos por los estudiantes para el miércoles de la siguiente semana. Así los estudiantes tenían tiempo para revisar el material durante el fin de semana, o durante los días lunes y martes. El número de videos por semana variaba entre 3 y 5, y cada uno de estos tenían una duración de entre 5 y 8 minutos. Al finalizar cada video los estudiantes eran sometidos a una evaluación formativa de dos preguntas, con lo que se evaluaba la comprensión por parte de los estudiantes de los contenidos tratados en cada video. Esta evaluación luego era utilizada como un insumo por parte del profesor. Esta le permitía tener una noción de aquellos contenidos que los estudiantes manejaban mejor o peor, y así poder dar más énfasis en la clase de profundización a aquellos contenidos que requerían de mayor atención.

b) Clase interactiva para la profundización de los contenidos (2):

El día miércoles era desarrollada una clase de profundización de los contenidos tratados en los videos vistos por los estudiantes. Para la organización y planificación de esta clase, se utilizaba como insumo el resultado de la evaluación formativa realizada por los alumnos al finalizar cada video. Esto permitió a los profesores preparar las clases de acuerdo a las necesidades del grupo, pues se tenía una noción a priori de la comprensión de los contenidos por parte de los estudiantes. Además, para el desarrollo de estas lecciones se utilizó el sistema de respuesta en aula Nearpod (<https://nearpod.com/>), el cual favorece y potencia la participación e involucramiento de los estudiantes.

c) Desarrollo de actividades de aplicación (3):

La última etapa de esta estructura semanal, consta de la realización de actividades centradas en el desarrollo de habilidades superiores. Estas actividades eran realizadas los días viernes y se trabajaba durante dos módulos de clases (2 horas y 40 minutos), con el objetivo de aplicar los contenidos estudiados durante esa semana. El medio utilizado para el desarrollo de estas actividades fue la plataforma

web desarrollada en este proyecto, la cual aplica un patrón de trabajo basado en *responder-retroalimentar-rehacer*.

La figura a continuación muestra un esquema de la estructura semanal, de la metodología explicada anteriormente:

	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Semana N							
Semana N + 1							
Publicación de videos Revisión de material de videos Clase de profundización Clase de aplicación							

Figura 1-3: Estructura semanal aplicación modelo *Flipped Classroom*.

Fuente: elaboración propia

1.3.2 Plataforma web

La herramienta desarrollada corresponde al *software* utilizado para el desarrollo de las actividades de aplicación de la tercera etapa descrita anteriormente. Tanto el desarrollo, la programación y el diseño de la plataforma, como la metodología en particular que se sigue en cada actividad (*responder-retroalimentar-rehacer*), son parte del trabajo realizado en este proyecto.

a) Funcionamiento:

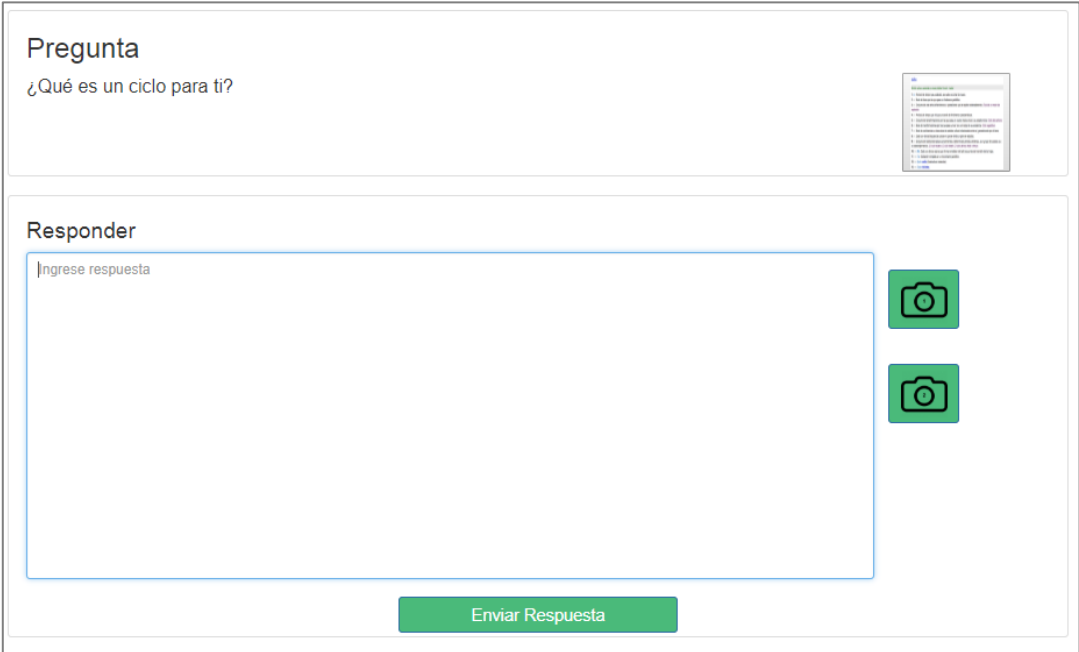
Cada actividad realizada con la plataforma involucra un proceso que sigue un patrón *responder-retroalimentar-rehacer*. La ejecución de este proceso puede

realizarse mediante la escritura de texto, la organización de esquemas, el despliegue de material visual, o la combinación de estos medios de información utilizando un dispositivo móvil (celular, *tablet* o laptop). El inicio de la actividad se gatilla mediante un estímulo entregado por el profesor, a través de la plataforma. Cada estudiante recibe el mismo estímulo. Es importante que éste se centre en el desarrollo de habilidades superiores y no en la reproducción del contenido estudiado esa semana, para lo que se requiere de la planificación y preparación adecuada de la actividad. Al recibir el estímulo, cada estudiante debe entregar una respuesta utilizando los medios mencionados. Luego la plataforma permite generar un intercambio aleatorio de las respuestas entre los estudiantes, es decir, cada estudiante envía su respuesta a un compañero y al mismo tiempo recibe la respuesta de otro. Una vez que cada estudiante recibe la respuesta de un compañero, este debe retroalimentar y complementar la respuesta recibida. Todas las respuestas son intercambiadas al mismo tiempo, de manera automática cuando el profesor lo estima conveniente y en anonimato, es decir, los estudiantes no saben de quien es la respuesta que reciben. Por último, el sistema permite nuevamente retornar la respuesta original a cada estudiante, junto con la retroalimentación que se elaboró para esa respuesta. Esto al igual que el intercambio de respuestas, se hace cuando el profesor lo estima conveniente y de manera automática para todas respuestas. A partir de eso, cada estudiante debe construir una nueva respuesta al estímulo inicial entregado por el profesor, considerando su respuesta inicial y la retroalimentación recibida para su respuesta inicial. El *paper* del Capítulo 2, entrega más detalle del análisis cuantitativo y análisis cualitativo de los factores que más afectan el desarrollo de este proceso. Las figuras 1-4, 1-5 y 1-6, muestran el proceso recientemente descrito desde el punto de vista de un estudiante.

En línea con la metodología *Flipped*, el rol del profesor durante este proceso no termina con la entrega del estímulo, sino que debe mantener un papel activo

durante el desarrollo de toda la actividad. El docente es quien debe monitorear, guiar, apoyar y dar las directrices para cada una de las etapas. En ese sentido la plataforma también es una herramienta de apoyo para la realización de dichas actividades, la Figura 1-7 muestra el desarrollo de la etapa responder desde el punto de vista del profesor y explica las funcionalidades que ayudan al desarrollo de la actividad. Esta misma interfaz está disponible para el profesor en las dos etapas siguientes (argumentar y rehacer).

La Figura 1-4 muestra la vista de un alumno en la primera etapa, donde se puede apreciar el estímulo entregado por el profesor (pregunta). Cada estudiante debe responder a este estímulo, ya sea en forma de texto escribiendo a través del dispositivo móvil que esté utilizando, o con la posibilidad de adjuntar material visual.



Pregunta

¿Qué es un ciclo para ti?

Responder

Ingrese respuesta

Enviar Respuesta

Figura 1-4: Ejemplo de vista de alumno para *responder*.

Fuente: elaboración propia

La figura 1-5 muestra la vista de un alumno en la segunda etapa, donde se puede apreciar la respuesta que recibe el estudiante (texto a continuación de Responder), a la cuál debe entregar retroalimentación. La retroalimentación, también puede ser entregada en forma de texto o utilizando material visual.

Responder

Lorem Ipsum es simplemente el texto de relleno de las imprentas y archivos de texto. Lorem Ipsum ha sido el texto de relleno estándar de las industrias desde el año 1500, cuando un impresor (N. del T. persona que se dedica a la imprenta) desconocido usó una galería de textos y los mezcló de tal manera que logró hacer un libro de textos especimen. No sólo sobrevivió 500 años, sino que también ingresó como texto de relleno en documentos electrónicos, quedando esencialmente igual al original. Fue popularizado en los 60s con la creación de las hojas "Letraset", las cuales contenían pasajes de Lorem Ipsum, y más recientemente con software de autoedición, como por ejemplo Aldus PageMaker, el cual incluye versiones de Lorem Ipsum.




Argumentar




Enviar Respuesta

Figura 1-5: Ejemplo de vista de alumno para retroalimentar.

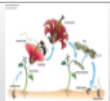
Fuente: elaboración propia

La figura 1-6 muestra la vista de un alumno en la tercera etapa, donde se puede apreciar la respuesta inicial del estudiante y la retroalimentación recibida (texto a continuación de Responder y Argumentar respectivamente). En esta etapa, el estudiante debe reconstruir una respuesta al estímulo inicial entregado por el profesor. El *paper* del Capítulo 2, centra gran parte de su atención en el análisis de esta última etapa, y la comparación de ésta con la primera. A través de esta comparación, se intentó determinar cómo cada estudiante logra, o no, un avance en la aplicación de habilidades superiores.

Responder

Al contrario del pensamiento popular, el texto de Lorem Ipsum no es simplemente texto aleatorio. Tiene sus raíces en una pieza clásica de la literatura del Latín, que data del año 45 antes de Cristo, haciendo que este adquiera mas de 2000 años de antigüedad. Richard McClintock, un profesor de Latín de la Universidad de Hampden-Sydney en Virginia, encontró una de las palabras más oscuras de la lengua del latín, "consectetur", en un pasaje de Lorem Ipsum, y al seguir leyendo distintos textos del latín, descubrió la fuente indudable. Lorem Ipsum viene de las secciones 1.10.32 y 1.10.33 de "de Finibus Bonorum et Malorum" (Los Extremos del Bien y El Mal) por Cicero, escrito en el año 45 antes de Cristo. Este libro es un tratado de teoría de éticas, muy popular durante el Renacimiento. La primera línea del Lorem Ipsum, "Lorem ipsum dolor sit amet...", viene de una línea en la sección 1.10.32






Argumentar

Lorem Ipsum es simplemente el texto de relleno de las imprentas y archivos de texto. Lorem Ipsum ha sido el texto de relleno estándar de las industrias desde el año 1500, cuando un impresor (N. del T. persona que se dedica a la imprenta) desconocido usó una galería de textos y los mezcló de tal manera que logró hacer un libro de textos especímen. No sólo sobrevivió 500 años, sino que también ingresó como texto de relleno en documentos electrónicos, quedando esencialmente igual al original. Fue popularizado en los 60s con la creación de las hojas "Letraset", las cuales contenían pasajes de Lorem Ipsum, y más recientemente con software de autoedición, como por ejemplo Aldus PageMaker, el cual incluye versiones de Lorem Ipsum.

Rehacer

Ingrese respuesta




Enviar Respuesta

Figura 1-6: Ejemplo de vista de alumno para rehacer.

Fuente: elaboración propia

La Figura 1-7 muestra la vista de un profesor en la primera etapa. En ella se aprecia el estímulo entregado a los estudiantes (pregunta), el avance del curso al saber cuántos alumnos ha entregado una respuesta y también es posible ver las respuestas que han elaborado los estudiantes.

Act 1 U1

Etapas: Responder

Han respondido: 22 / 32

Pregunta

Observa el video que está en el link (https://youtu.be/HO9ax7Ocr_U). Verás a un estudiante leyendo el libro "Pequeños, pequeños" de Jean Maubille. Fijate y toma nota sobre el modo en que lee el texto. Para tomar notas adecuadas y pertinentes, ten presente los conocimientos y habilidades involucrados en la comprensión lectora.

Analiza el desempeño lector del estudiante, refiriéndote a los aspectos logrados y aquellos no logrados. Además, determina la dimensión en la que según tú se necesita una intervención pedagógica para que pueda alcanzar un desarrollo lector óptimo. Justifica brevemente tu elección. Reporta tu análisis a través de un texto escrito apoyándote en las ideas propuestas por Perfetti et al. (2005) así como en las de McKenna y Stahl (2009) y justifica usando las notas de la observación del video.

[Volver a Actividades](#) [Refrescar](#) [Cerrar Actividad](#)

[Retroalimentar](#)

Botón para ver respuesta

Apellido	Nombre	Respuesta	Ver
Apellido Estudiante 1	Nombre Estudiante 1	✓	Ver Respuesta
Apellido Estudiante 2	Nombre Estudiante 2	✓	Ver Respuesta
Apellido Estudiante 3	Nombre Estudiante 3	✓	Ver Respuesta
Apellido Estudiante 4	Nombre Estudiante 4		
Apellido Estudiante 5	Nombre Estudiante 5		Ver Respuesta
Apellido Estudiante 6	Nombre Estudiante 6		
Apellido Estudiante 7	Nombre Estudiante 7		
Apellido Estudiante 8	Nombre Estudiante 8	✓	Ver Respuesta

Estudiante aún no entrega respuesta

Figura 1-7: Ejemplo de vista de profesor para responder.

Fuente: elaboración propia

b) Diseño y desarrollo:

La aplicación web fue desarrollada en Ruby on Rails 4.2 con un sistema de gestión de base de datos PostgreSQL. El patrón arquitectónico de la aplicación es Modelo-

Vista-Controlador (Leff & Rayfield, 2001), en la Figura 1-8 se muestra un esquema explicativo de la arquitectura. Esta tecnología permite el desarrollo de aplicaciones multiplataforma de manera rápida y de bajo costo. El hecho de que sea una aplicación multiplataforma, quiere decir que se puede acceder a ella desde cualquier tipo de dispositivo mediante un navegador. Por otro lado, dados los recursos disponibles para el desarrollo, una aplicación nativa para cada sistema operativo, hubiese requerido de un mayor costo en tiempo a cambio de un mejor desempeño y funcionalidades que no eran requeridas por el proyecto.

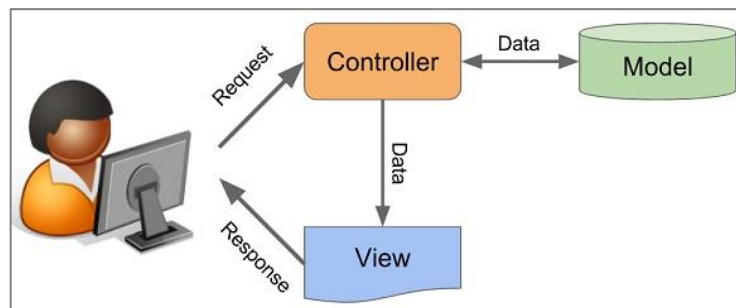


Figura 1-8: Patrón Modelo-Vista-Controlador.

Fuente: <https://helloacm.com/model-view-controller-explained-in-c/>

La metodología de desarrollo de software utilizada se basó en los principios de desarrollo ágil (Collier, 2012). Un equipo multidisciplinario, con un usuario final (docente usuario del producto) involucrado de principio a fin en el desarrollo, y un proceso iterativo e incremental de las distintas funcionalidades. Esta forma de trabajo reduce el riesgo de que el usuario final reciba un producto que no se acomode a sus necesidades, y entrega valor al producto de manera rápida, ya que se obtiene una versión con las principales funcionalidades en poco tiempo.

Las pruebas de usabilidad con los usuarios finales son de gran importancia en el desarrollo de software. Con ellas es posible determinar si el funcionamiento del

producto es confiable y si su interfaz es fácil de usar (Nielsen, 2003). El desarrollo de la aplicación consideró múltiples pruebas de usuarios. Estas fueron realizadas con el equipo de trabajo de los profesores de los cursos de Didáctica de la Matemática y Didáctica de la Lectura (profesores y grupo de ayudantes), profesores que luego utilizarían el software en sus clases durante el año 2017. Además se realizó una prueba piloto en los mismos cursos, donde se aplicó la metodología de trabajo descrita en la sección 1.3.1. Este piloto incluyó el desarrollo del contenido de una unidad completa de cada curso durante el mes de octubre del año 2016.

En función de los requisitos establecidos por el problema que se busca solucionar con el uso de esta plataforma, de las reuniones con el equipo docente de la facultad de educación, de las pruebas realizadas al sistema, de las conversaciones con el equipo de desarrollo y de las pruebas de usabilidad mencionadas anteriormente, surgieron los requisitos funcionales, requisitos no funcionales y el diagrama entidad relación que se muestra a continuación:

i) Requisitos funcionales:

Tabla 1-1: Requisito Funcional 01.

Identificador	RF01
Nombre	Crear nuevo usuario
Descripción	El sistema debe permitir crear un nuevo usuario
Información de Entrada	• Correo electrónico • Contraseña • Confirmación de contraseña
Usuarios	Estudiante y Profesor

Fuente: elaboración propia.

Tabla 1-2: Requisito Funcional 02.

Identificador	RF02
Nombre	<i>Log-in</i>
Descripción	El sistema debe ser capaz de que ingresen los distintos tipos de usuarios e identificar si es profesor o estudiante. Si el correo o la contraseña no son correctas, no se debe dejar ingresar.
Información de Entrada	• Correo electrónico • Contraseña
Usuarios	Estudiante y Profesor

Fuente: elaboración propia.

Tabla 1-3: Requisito Funcional 03.

Identificador	RF03
Nombre	Crear nuevo curso
Descripción	El sistema debe permitir que los usuarios de tipo Profesor puedan crear un nuevo curso. El curso al ser creado debe agregarse a los cursos disponibles para el profesor, pero no para los demás usuarios del sistema.
Información de Entrada	• Nombre del curso • Sigla del curso
Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-4: Requisito Funcional 04.

Identificador	RF04
Nombre	Creación de código de curso
Descripción	El sistema debe crear un código único de identificación del curso cuando un profesor crea un curso nuevo y éste debe estar disponible para el profesor.
Información de Entrada	• Nombre del curso • Sigla del curso
Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-5: Requisito Funcional 05.

Identificador	RF05
Nombre	Inscribir un curso
Descripción	El sistema debe permitir que un usuario de tipo Estudiante pueda agregar un curso mediante el código único de identificación disponible para el profesor de dicho curso. Al ser agregado, el curso debe aparecer en el listado de cursos disponibles del usuario, pero no en el de los demás usuarios del sistema.
Información de Entrada	• Código del curso
Listado de Usuarios	Estudiante

Fuente: elaboración propia.

Tabla 1-6: Requisito Funcional 06.

Identificador	RF06
Nombre	Eliminar un estudiante de un curso
Descripción	El profesor como administrador de su curso, debe poder eliminar estudiantes de su curso. Al hacerlo, el curso debe dejar de estar en el listado de cursos disponibles del usuario eliminado.
Información de Entrada	• Confirmación del usuario Profesor
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-7: Requisito Funcional 07.

Identificador	RF07
Nombre	Crear actividad
Descripción	El sistema debe permitir que el profesor de un curso pueda crear actividades para ese curso. La actividad al ser creada, debe aparecer en el listado de actividades del curso.
Información de Entrada	• Nombre de la actividad • Pregunta de la actividad • Imagen o multimedia asociada (opcional)
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-8: Requisito Funcional 08.

Identificador	RF08
Nombre	Editar actividad.
Descripción	El sistema debe permitir que el profesor del curso pueda editar los atributos de una actividad creada (nombre, pregunta y multimedia).
Información de Entrada	• Seleccionar actividad a editar • Nombre de la actividad (opcional) • Pregunta de la actividad (opcional) • Imagen o multimedia asociada (opcional)
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-9: Requisito Funcional 09.

Identificador	RF09
Nombre	Eliminar actividad
Descripción	El sistema debe permitir que el profesor pueda eliminar las actividades creadas.
Información de Entrada	• Seleccionar una actividad creada.
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-10: Requisito Funcional 10.

Identificador	RF10
Nombre	Publicar actividad
Descripción	Poner actividad seleccionada disponible para que los estudiantes la realicen.
Información de Entrada	• Seleccionar una actividad creada.
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-11: Requisito Funcional 11.

Identificador	RF11
Nombre	Cambiar etapa de actividad
Descripción	El profesor debe poder pasar a la etapa siguiente o anterior de la actividad publicada
Información de Entrada	• Seleccionar etapa anterior o siguiente.
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-12: Requisito Funcional 12.

Identificador	RF12
Nombre	Desarrollo de la actividad: Crear Respuesta
Descripción	Los estudiantes al ingresar a un curso con una actividad disponible (publicada), deben poder subir la respuesta a la etapa que corresponda (Responder, Argumentar o Rehacer).
Información de Entrada	• Seleccionar un curso disponible • Respuesta en forma de texto • Respuesta multimedia
Listado de Usuarios	Estudiante

Fuente: elaboración propia.

Tabla 1-13: Requisito Funcional 13.

Identificador	RF13
Nombre	Desarrollo de la actividad: Editar Respuesta
Descripción	Los estudiantes deben poder editar las respuestas subidas en la etapa que corresponda (Responder, Argumentar o Rehacer).
Información de Entrada	• Seleccionar un curso disponible • Respuesta en forma de texto • Respuesta multimedia
Listado de Usuarios	Estudiante

Fuente: elaboración propia.

Tabla 1-14: Requisito Funcional 14.

Identificador	RF14
Nombre	Ver respuesta alumno
Descripción	El sistema debe permitir que el profesor pueda ver las respuestas que van entregando los estudiantes en cada una de las etapas.
Información de Entrada	• Seleccionar estudiante
Listado de Usuarios	Profesor

Fuente: elaboración propia.

Tabla 1-15: Requisito Funcional 15.

Identificador	RF15
Nombre	Ver todas las respuestas de todos los estudiantes
Descripción	El sistema debe permitir al profesor ver las respuestas de todos los estudiantes en cada una de las etapas, de una actividad en particular.
Información de Entrada	• Seleccionar la actividad
Listado de Usuarios	Profesor

Fuente: elaboración propia.

ii) Requisitos no funcionales:

Tabla 1-16: Requisito No Funcional 01.

Atributo	Descripción
Identificador	RNF01
Nombre	Ambiente operacional: multiplataforma
Descripción	El funcionamiento del sistema debe ser en un entorno multiplataforma.
Tipo	Interfaz

Fuente: elaboración propia.

Tabla 1-17: Requisito No Funcional 02.

Atributo	Descripción
Identificador	RNF02
Nombre	Almacenamiento de resultados
Descripción	El sistema debe permitir que los datos generados en el uso de la plataforma y las actividades, sean de fácil acceso y manejo, para poder estudiarlos y analizarlos.
Tipo	Gestión de datos

Fuente: elaboración propia.

Tabla 1-18: Requisito No Funcional 03.

Atributo	Descripción
Identificador	RNF03
Nombre	Continuidad operacional
Descripción	Producto del uso que la plataforma tendrá en los distintos cursos y secciones, esta debe asegurar la continuidad operacional y soportar un tráfico de usuarios adecuado.
Tipo	Servicio

Fuente: elaboración propia.

Tabla 1-19: Requisito No Funcional 04.

Atributo	Descripción
Identificador	RNF04
Nombre	Soporte de red
Descripción	La red wifi de la universidad en la facultad de educación, debe tener alta cobertura en las salas determinadas para los cursos que harán uso de la plataforma y debe soportar el tráfico y conectividad de dispositivos adecuada.
Tipo	Servicio

Fuente: elaboración propia.

iii) Modelo entidad-relación: Las clases del modelo entidad-relación son *Course*, *User*, *Homework* y *Answer*.

- *Course*: corresponde a los cursos que puede crear un profesor, donde los estudiantes pueden inscribirse mediante el código único de identificación. Las actividades (*Homework*) son propias de cada curso.
- *User*: los usuarios pueden ser de tipo profesor o estudiante (*role*). Estos pueden crear cursos, suscribirse a ellos, crear actividades y completarlas.
- *Homework*: corresponde a las actividades que el sistema permite desarrollar. Se identifican por un nombre (*name*), la pregunta o contenido (*content*) y según la fase en que se encuentre (*actual_phase*) los estudiantes responderán la etapa que corresponda (responder, argumentar o rehacer).
- *Answer*: son las respuestas que los estudiantes dan a las actividades. Una única respuesta guarda cada una de las distintas etapas (responder, argumentar y rehacer).

A continuación se muestra el diagrama con las clases mencionadas:

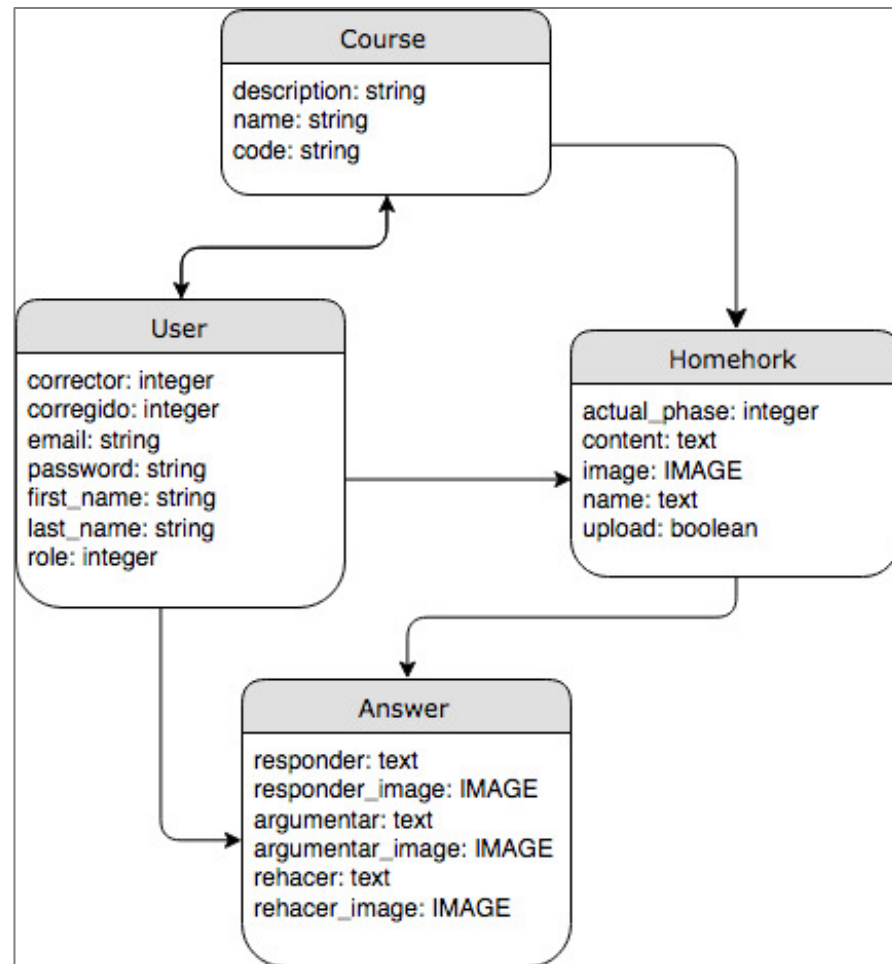


Figura 1-9: Simplificación de diagrama Entidad-Relación original.

Fuente: Elaboración propia.

1.4 Contribución de la solución

Como ya se ha mencionado anteriormente, la plataforma web es una contribución directa a la implementación del modelo *Flipped Classroom* en los cursos de Didáctica de la Matemática y Didáctica de la Lectura de la Facultad de Educación de la Pontificia Universidad Católica. Pero más allá de ser una herramienta particular para el desarrollo de esta metodología, la plataforma es en sí misma es un aporte a la enseñanza y la educación. Así lo demuestra la aceptación y adopción que ha tenido en más cursos de la misma facultad, y también en cursos de la Facultad de Ingeniería de la misma universidad. El primer semestre del año 2018, cinco cursos distintos hicieron uso de la plataforma, lo que equivale a trece secciones y más de diez profesores. A pesar de ser un software bastante básico, la plataforma es una herramienta de apoyo para los profesores y de aprendizaje para los estudiantes.

El tipo de actividades que la plataforma permite realizar en conjunto con el modelo que aplica, potencia y fomenta el desarrollo de habilidades superiores por parte de los estudiantes, y así lo demuestra el *paper* del Capítulo 2. Las actividades buscan que los estudiantes vayan más allá de la simple comprensión del contenido, es decir, desafía las capacidades de los estudiantes instándolos a aplicar lo estudiado, crear nuevas ideas y compartir el conocimiento con sus compañeros.

Por otro lado, al mismo tiempo que la herramienta es un aporte para la enseñanza desde el punto de vista de los estudiantes, también contribuye a la educación y es una herramienta de apoyo para los profesores. La plataforma facilita el trabajo de los profesores y los ayuda principalmente en dos aspectos, el *collaborative overhead* y el orden en el aula. El *collaborative overhead* se define como el esfuerzo adicional necesario para realizar una actividad de aprendizaje de forma grupal, sin que ese esfuerzo se refleje necesariamente en una contribución esencial en post de alcanzar el objetivo principal de aprendizaje. Tareas como la formación

de grupos, la necesidad de los integrantes de un grupo de sociabilizar y el establecer la forma de trabajo de cada grupo, son algunos ejemplos de este fenómeno. La comunicación e interacción social entre los participantes, es necesaria para contribuir a un ambiente que favorezca el desarrollo y la relación entre los participantes de un grupo (MacMillan, Entin & Serfaty, 2004). Pero por otro lado las interacciones sociales postergan el desarrollo de las tareas asignadas, y por consecuencia se posterga también el logro del objetivo de aprendizaje (Kaufman, 2010). En cuanto al orden en la sala de clases cuando se realizan trabajos grupales, se debe mencionar que si las interacciones sociales para realizar este tipo de trabajo se desarrollan en espacios reducidos, el ambiente generado, el ruido y el desorden son perjudiciales para el aprendizaje de los estudiantes (Bavelas, 1950 en Barkoczi & Galesic, 2016). En ese sentido la plataforma soluciona estos problemas al generar grupos de trabajo anónimos y de manera automatizada, donde las interacciones sociales son reducidas y la forma de trabajo está predeterminada (responder, argumentar y rehacer). Por último, también existen otros beneficios como el orden, la persistencia, el respaldo y el almacenamiento de los datos de trabajo al desprenderse del lápiz y del papel, y el monitoreo centralizado por parte del profesor para apoyar a los estudiantes que más lo necesiten.

1.5 Conclusiones

Este proyecto de investigación desarrolló una herramienta tecnológica y la implementó en un modelo de *Flipped Classroom*. Se estudió como el modelo mencionado y la forma de trabajo con la herramienta, fomentan el desarrollo de las Habilidades del Siglo XXI y ayudan a comprometer a los estudiantes con su aprendizaje. Además la herramienta puede extender su uso más allá del modelo *Flipped*, y su uso es adaptable a distintas necesidades que pueden surgir de parte de los profesores. El hecho de que sea un software básico y con funcionalidades

simples, implica que no se requiere invertir mucho tiempo en capacitaciones para que los estudiantes utilicen el sistema, por tanto existe una mayor probabilidad de que se adopte su uso. Sin embargo, para uso masivo se recomienda un sistema más sofisticado con desarrollo de aplicaciones especializadas para web y móvil. En particular se sugiere el desarrollo de una aplicación web mediante Angular, *framework* de desarrollo de aplicaciones web, y el desarrollo de una aplicación móvil híbrida mediante Ionic, kit de desarrollo de *software* para aplicaciones móviles. Ambos desarrollos ayudarían a mejorar y simplificar aún más el uso por parte de profesores y estudiantes, obteniendo una mejor performance en función de los requisitos no funcionales y mejorando la usabilidad a través de una interfaz más amigable y con mejores funcionalidades.

El modelo *Flipped* presenta limitaciones si se piensa en la gran inversión en tiempo y dedicación inicial que se requiere, para poner en práctica la metodología mencionada en el punto 1.3.1. La grabación de los videos con los contenidos y la preparación de las actividades desafiantes a la hora de verdaderamente fomentar el desarrollo de las Habilidades del Siglo XXI para la plataforma, requieren de un equipo de trabajo especializado, si se intenta replicar la metodología.

Por otro lado, las creaciones de los estudiantes en la plataforma, es decir, la respuesta inicial, la retroalimentación y la respuesta final, no pasan por un juicio experto que las califique en función del contenido. En otras palabras, no existe un profesor que corrija y revise exhaustivamente las respuestas, para que luego los estudiantes pudiesen saber si lo expresado en sus escritos e imágenes es correcto en términos del contenido disciplinar expresado.

Es importante mencionar la curva de aprendizaje y adaptación que tienen tanto la metodología *Flipped* como la herramienta en particular. La metodología es de gran exigencia para los estudiantes, y requiere de más de una iteración para que el

seguimiento se dé en forma natural. Es decir, los estudiantes demoran en adaptarse y recordar cada una de las etapas sin que los profesores deban recordarles lo que deben hacer. En ese sentido, también es importante que los estudiantes se hagan cargo de su aprendizaje y se responsabilicen por su nivel de participación. Al ser una metodología de aprendizaje activo, cada quién se vuelve más responsable de lo que aprende, pues el profesor pasa a ser guía y colaborador de su aprendizaje. En ese sentido, la plataforma es una herramienta de fácil uso, en las pruebas realizadas la inducción inicial era suficiente para que los estudiantes pudiesen trabajar enfocados en los objetivos de aprendizaje.

Por último, el desarrollo de este trabajo de investigación generó aprendizaje en ámbitos como la ingeniería de software, trabajo estadístico y análisis cuantitativo. En cuanto a la ingeniería de software, la plataforma fue la primera aplicación web desarrollada por el investigador para usuarios reales, lo que implicó desarrollar conocimiento en relación a tecnologías de desarrollo y nuevos lenguajes de programación. En cuanto al trabajo estadístico y el análisis cuantitativo, se hizo un extenso trabajo para elaborar modelos de regresión, test de significancia estadística y análisis de datos. En particular, para la elaboración del inciso 2.5 y 3.1 del Capítulo 2, los cuales resumen el trabajo estadístico y de análisis cuantitativo realizado.

2. ARTÍCULO ENVIADO: TECHNOLOGY-SCAFFOLDED PEER ASSESSMENT FOR DEVELOPING CRITICAL THINKING IN PRE-SERVICE TEACHER TRAINING

Abstract

El pensamiento crítico es una habilidad que escasamente se promueve de forma sistemática en las aulas universitarias. En formación de profesores esta situación no es distinta. A pesar del esfuerzo de diversas instituciones de promover prácticas de enseñanza constructivistas, éstas siguen siendo centradas en el profesor. Considerando la creciente importancia que ha tomado el desarrollo del pensamiento crítico en la preparación de futuros profesores, por un lado, y la relevancia de la retroalimentación que se entrega y la que se recibe en el proceso de aprendizaje, este artículo responde a la pregunta sobre ¿cómo actividades con un modelo de andamiaje para la revisión de pares pueden desarrollar pensamiento crítico en formación de profesores? A través de un modelo de aula invertida, en un curso de didáctica de la lectura de una facultad de educación, se realizaron actividades de aplicación e integración mediante una herramienta tecnológica para aplicar el modelo de andamiaje. Treinta estudiantes participaron en tres actividades a lo largo del semestre. Las actividades trabajadas por los alumnos consideran tanto el aprendizaje del conocimiento pedagógico disciplinar y práctico, como el desarrollo de una o más habilidades del pensamiento crítico. El modelo de andamiaje, con la ayuda de la herramienta tecnológica minimiza los problemas logísticos al interior de la sala de clases, y al mismo tiempo soluciona los problemas interpersonales para la evaluación de pares, ya que la asignación de las distintas tareas se realiza de manera aleatoria. Un metodología de análisis mixta, mostró que en formación inicial docente, la evaluación de pares basado en la retroalimentación anónima producto del andamiaje y la tecnología, permite un avance en el desarrollo de las habilidades de pensamiento crítico, y que tanto la retroalimentación recibida como la retroalimentación entregada son predictores de ese avance. Este trabajo aporta con un

modelo para el desarrollo del pensamiento crítico en el aula, más las orientaciones al docente de cómo desarrollar estas prácticas.

2.1. Introducción

El pensamiento crítico, considerado una de las habilidades del siglo XXI, es entendido como “un juicio autorregulado que considera las habilidades de interpretación, análisis, evaluación, inferencia, explicación y autorregulación” (Facione, 1990, p.4). Las actividades que promueven el pensamiento crítico sitúan al alumno en el centro del aprendizaje (DeWaele, 2015), y al docente como un mediador del aprendizaje puesto que otorga andamiajes que permiten el desarrollo del pensamiento crítico (Lajoie et al., 2001, en Kim, 2013).

El pensamiento crítico, requisito fundamental para el ejercicio profesional (Betancourt, 2014), se considera uno de los resultados más deseables de la educación superior (Dunne, 2015). Aunque se ha intentado que sea un foco del aprendizaje en la educación superior (Stupple, 2016), es una habilidad que escasamente se promueve de forma sistemática en las aulas universitarias (Hammersley, 2016), cuestión clave si es que se considera que a los estudiantes se les dificulta pensar críticamente (Ten Dam, et al, 2004).

Una de las razones sobre el escaso desarrollo de las habilidades de pensamiento crítico se vincula con la instrucción puesto que generalmente se enseña usando un modelo tradicional, centrado en el docente, quien usa el tiempo de clase para desarrollar habilidades de orden inferior (conocer y reconocer), dejando las habilidades de orden superior (application, analysis, and synthesis) para las actividades fuera de la sala de clases (Bormann, 2014). Las aulas universitarias escasamente generan ambientes donde los estudiantes se involucren activamente (Mehring, 2016), con aprendizajes que

permitan la participación por parte de los estudiantes (DeWaelche, 2015). Al tomar el alumno un rol pasivo (Mehring, 2015), el profesor espera que los alumnos practiquen la materia de manera independiente (Unruh et al., 2016).

En formación de profesores, debiese orientarse aún más a motivar a los estudiantes a desarrollar habilidades pedagógicas y de pensamiento crítico (Kostiainen et al 2018). Al combinar el pensamiento crítico con el contenido disciplinar, ambos tienen un resultado positivo en el desarrollo de habilidades por parte de los estudiantes (Ennis, 2018). A pesar de décadas de esfuerzo de diversas instituciones de promover prácticas de enseñanza constructivistas en la formación de profesores, éstas siguen siendo centradas en el profesor, donde los estudiantes son meros receptores del contenido (Vaughan, 2014). El escaso desarrollo de las habilidades de pensamiento crítico de futuros profesores se debe a una variedad de factores, tales como la ausencia de experiencias de aprendizaje de pensamiento crítico que han vivido tanto en su trayectoria escolar como en su formación universitaria (Arsal, 2017). Los futuros docentes son escasamente alentados a buscar diversos procedimientos ante la resolución de un problema, y tienden a memorizar paso a paso la solución sin comprender el cómo y porqué del proceso (Brown, 2016). Para que los futuros docentes puedan aplicar estrategias que promuevan el desarrollo del pensamiento en su práctica profesional, es fundamental que adquieran y practiquen estas habilidades (Akgun, et al, 2016).

Uno de los mecanismos centrales para el aprendizaje de los futuros docentes es el *feedback*. El *feedback* ha sido ampliamente investigado y actualmente es uno de los focos en educación (Boud & Molloy, 2013a; Evans, 2013; Nicol, 2010, en Winstone, 2017), ya que es considerado uno de los factores que más influye en el aprendizaje de los estudiantes (Hattie, 2007). De hecho, el constructivismo social ha considerado tanto el entregar como el recibir *feedback* uno de sus principios (Barak, 2017). El rol del alumno se torna protagónico, dado que al dar y recibir *feedback* el estudiante se involucra activamente en su proceso de aprendizaje (Harland, et al, 2017), y permite que el estudiante construya su conocimiento (Lin, 2018).

Proporcionar *feedback* entre compañeros es una actividad valiosa para el aprendizaje (Harland et al., 2017), pues el estudiante puede comparar su actividad con el producto generado por un compañero, otorgando sugerencias sobre su propia experiencia (Baker, 2016). Así, los alumnos aprenden tanto de la actividad como de sus compañeros (van Gennip et al., 2010), al proporcionar orientaciones sobre las habilidades y comprensiones que tuvieron sus pares (Winstone, et al, 2017). Considerando que los estudiantes indican no sentirse preparados y con las habilidades adecuadas para desarrollar evaluación de pares (Mulder et al., 2014), para lograr un *feedback* de alta calidad es necesario contar con una orientación sobre cómo otorgar *feedback* (Baker, 2016). Para eso, los profesores juegan un rol fundamental en la mediación para la implementación de actividades de evaluación de pares, indicando a los estudiantes que es relevante y determinando rúbricas de evaluación (Panadero & Brown, 2017).

Considerando la creciente importancia que ha tomado el desarrollo del pensamiento crítico en la preparación de futuros profesores, por un lado, y la relevancia de entregar y recibir *feedback* en el proceso de aprendizaje por otro, este artículo busca responder a la pregunta sobre ¿cómo actividades con un modelo de andamiaje para la revisión de pares pueden desarrollar pensamiento crítico en formación de profesores?

2.2. Metodología

2.2.1. Estrategia utilizada: modelo de andamiaje para la evaluación de pares

La estrategia de evaluación de pares analizada en este artículo se encontraba dentro de un modelo de *Flipped Classroom*. La estrategia de *Flipped Classroom* se utilizó para transformar el aprendizaje de un aula pasiva a una activa (Johnson et al., 2015), donde el tiempo presencial, es decir, la interacción profesor-alumno en clases, se maximiza (Apedoe y Col, 2017). Se reconocen las etapas de aprendizaje: conexión, comprensión y expresión, en el aprendizaje individual, las actividades fuera del aula, dónde videos y

evaluaciones formativas fueron usadas con el objetivo de construir un punto de vista en los estudiantes; y, cuestionar, construir y reflejar, en el trabajo en clases, para analizar y evaluar un tópico con la retroalimentación del docente (Jong, 2017).

a) Actividades fuera del aula:

Se consideraron videos y evaluaciones formativas para la activación de los contenidos (Lo & Hew, 2017). Entre 3 y 5 videos de entre 5 y 8 minutos cada uno eran vistos cada semana, completando un total de 16 semanas de actividad. Los estudiantes podían ver, pausar y rebobinar el los videos todo lo que necesitasen. Considerando que la retroalimentación de sistemas de información acerca del trabajo de los estudiantes hacia el docente, ha demostrado ser menos efectivo que el *feedback* que los sistemas pueden entregar a los estudiantes (Faber et al., 2017), al finalizar cada video una evaluación formativa era entregada a los estudiantes y una activación al comienzo de las actividades dentro del aula. La evaluación de pares consistía en tres preguntas, cada una de éstas con un componente de contenido y un componente de argumentación, las que entregaban como *feedback* distintos segmentos del video. El componente argumentativo, tenía el objetivo de apoyar el proceso de enseñanza del pensamiento crítico a través del contenido que se estaba evaluando (Lund-Diaz et al., 2016).

b) Actividades al interior del aula:

Para la activación, al comienzo de cada clase en el aula, dónde el objetivo era revisar las actividades realizadas entregando *feedback*, se priorizaron los contenidos y las habilidades superiores no observadas, o con menos logro, indicadas en un reporte de la evaluación formativa desarrollada fuera del aula. Con esta información, utilizando un sistema interactivo de respuestas para el aula, el docente trabajaba con los estudiantes los tópicos requeridos. Para el proceso de aplicación e integración (Lo & Hew, 2017), al interior del aula se realizaron actividades de evaluación de pares mediante un modelo de

andamiaje provisto por una herramienta tecnológica, a través de dispositivos conectados en red (Tablet, or Laptop). La actividad se dividía en tres etapas. En la primera etapa, lectura y análisis de una pregunta y construcción de una respuesta individual, los estudiantes respondían de manera individual una pregunta formulada por el profesor; en Figura 2-1, Student 1 Response (R1), y así hacia adelante. Estas actividades -situadas en escenarios de enseñanza y aprendizaje de la lectura- promueven en los futuros profesores la capacidad para elaborar una propuesta y clarificar su significado (**interpretación**), realizar una relación analítica de conceptos (**análisis**), realizar una evaluación de propuestas o posibles soluciones (**evaluación**), clarificar y articular la información relevante que está implícita (**inferencia**) y articular razonamientos que respalden la información (**explicación**) (Facione,1990). Este artículo analiza en particular la secuencia de aprendizaje en las actividades con el modelo de andamiaje para la evaluación de pares a través de una plataforma digital.

2.2.1.1. Modelo de andamiaje de evaluación de pares

La actividad trabajada por los alumnos en forma individual considera tanto el aprendizaje del conocimiento pedagógico disciplinar y práctico, como el desarrollo de una o más habilidades del pensamiento crítico. Esta se desarrolla en tres fases:

a) Fase 1: Construcción de respuesta:

Siguiendo el *framework* de Facione, para el desarrollo del pensamiento crítico (Facione, 1990; 1998), durante la fase de construcción de la respuesta (R) se promueve que los estudiantes: (1) reconozcan, interpreten, integren y desarrollen las habilidades involucradas para resolver la tareas, (2) identifiquen y comprendan el contenido pedagógico disciplinar y especializado por aplicar, declarado en la pregunta y que, además, ha sido previamente trabajado en la clase mediante la discusión de la lectura y la aplicación a situaciones pedagógicas, 3) ajusten sus respuestas a los formatos de presentación de la información y que, además, desarrollen todos los focos y partes de la

tarea dada y 4) preparen y presenten las evidencias correspondientes, las que respaldarán y sustentarán el desarrollo de su respuesta. Para esto, en algunos casos, la pregunta suponía la lectura previa de bibliografía relacionada con el tema en estudio y el aprendizaje de los contenidos propuestos en los videos.

b) Fase 2: Otorgando *feedback*:

La respuesta del estudiante (R) es enviada a un segundo estudiante para que la evalúe y otorgue *feedback* en un segundo documento. En Figura 2-1, Student 1 entrega *feedback* a Student 3 (F3), y Student 2 entrega *feedback* a Student 1 (F1), y así. Bajo las indicaciones del profesor se espera que el alumno que escribe el *feedback* para su par: (1) elabore una retroalimentación efectiva y significativa a la respuesta inicial (R) (Shute, 2008), que presente comentarios y sugerencias, con la intención de guiar, evidenciando el uso y aplicación de criterios concretos, y (2) proponga estrategias y explicaciones de mejora a la respuesta inicial (R) entregando directrices claras, pertinentes y concretas (Kluger & De Nisi, 1996). Para guiar a los futuros docentes en este proceso, se elaboró una guía de retroalimentación para ayudar a los estudiantes a evaluar los planteamientos presentados y su credibilidad, a partir de criterios definidos; y a explicar a través de argumentaciones, los razonamientos lógicos y representaciones presentes, así como las valoraciones personales detectadas, justificando con evidencia (Algassab, et al, 2018). Considerando que la apreciación que se tenga de un peer al otorgar *feedback* incide en la retroalimentación efectuada y así en su fiabilidad (Poulos y Mahony, 2008 en Berndt, Strijbos y Fischer, 2018), se decidió que la retroalimentación fuera anónima (y aleatoria) para reducir las consideraciones socio afectivas (Li, 2017) y aumentar la actitud positiva hacia la tarea (Lin, 2018) así como mayor motivación hacia la actividad (Berndt et al., 2017). La plataforma digital diseñada permitió un proceso de anónimo y aleatorización efectivo.

c) Fase 3: Construcción de una nueva respuesta:

Este proceso (NR) engloba la experiencia del estudiante en este proceso. En Figura 2-1, Student 1, para la construcción de la nueva respuesta (NR1) considera, tanto su respuesta inicial (R1) la retroalimentación escrita por el Student 2 (F1), como la experiencia de haber realizado él una retroalimentación (F3) a otra respuesta de un tercer alumno (R3). El proceso de construcción de esta nueva respuesta (NR) requiere la articulación de las habilidades de pensamiento crítico de evaluar y analizar qué tan atingentes son los comentarios de la retroalimentación a su texto, así como interpretar e inferir información necesaria para su respuesta a partir de los comentarios recibidos en la retroalimentación, más su propia experiencia de retroalimentador (Boud, 2013). En esta etapa se espera que los estudiantes: (1) reconozcan e integren los aspectos y niveles atingentes indicados por el evaluador (F1); (2) entreguen evidencias claras y coherentes que justifiquen los posibles cambios efectuados y (3) presenten una reelaboración de R en NR, que generalice los elementos relevantes analizados por el estudiante, es decir, que no copie ni pegue información, sino que reconstruya a partir de los resultados de su análisis y evaluación de los recursos correspondientes (Gibbs & Simpson, 2004).

Como se muestra en la Figura 2-1, cada estudiante asume ambos roles durante la actividad, constructor de respuesta y evaluador, y cada estudiante se relaciona con dos estudiantes, al que debe entregar *feedback* y del que recibe *feedback*. Este proceso es administrado de manera transparente a través de la plataforma digital. Observamos pues, que la evaluación de pares es clave en nuestra propuesta, pues involucra al estudiante de manera activa con otros dos pares (Phielix et al., 2010 en Alqassab et al., 2018). Asimismo, la propuesta se inserta bajo la lógica constructivista al crear una nueva respuesta sobre el conocimiento observado por otros dos pares, siendo la evaluación de pares un proceso fundamental de ésta (Lin, 2018).

El profesor tiene un rol fundamental al no solo definir el trabajo por realizar, sino también especificar los criterios para lograr cada fase, y retroalimentar a los alumnos durante las distintas fases de resolución de la actividad (Yorke, 2003).

El modelo de andamiaje es facilitado por la tecnología. Los problemas interpersonales que pueden presentarse en la evaluación de pares, desaparece producto de la asignación aleatoria y anónima de las tareas mediante la plataforma. Al mismo tiempo, la plataforma facilita y minimiza la logística en el aula.

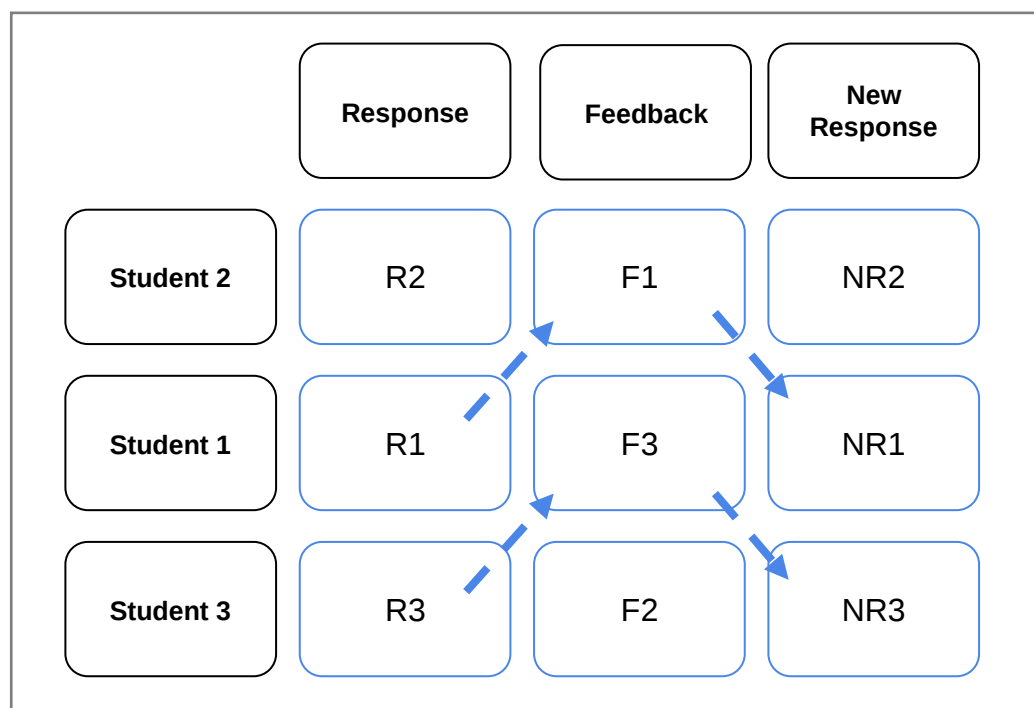


Figura 2-1: Actividad de construcción de respuesta y evaluación de pares a través de dispositivos en red

2.2.2. Contexto y participantes

La comunicación es considerada una de las habilidades clave dentro de las Habilidades del Siglo XXI (Keane et al., 2016). Estudios han demostrado la relación entre las

habilidades de pensamiento crítico y de comprensión lectora (Yousefi & Mohammadi, 2016). Por tanto se decidió enfocar el estudio en un curso de didáctica de lectura.

En este estudio, de diseño mixto, participaron 30 estudiantes (2 hombres, 28 mujeres), en un curso de didáctica de la lectura, que se dicta en el segundo año (pregrado) del programa de educación general básica de una universidad chilena que se distingue por un currículo basado en la práctica. La edad de los estudiantes varió entre los 20 y 24 años de edad. La profesora que llevó a cabo la intervención, 42 años de edad, dicta el curso hace cinco años y constantemente está buscando metodologías para optimizar el aprendizaje práctico de los estudiantes, y reportó tener conocimiento en el uso de tecnologías en el aula.

Los principales contenidos del curso fueron: conocimientos teóricos y empíricos especializados sobre el aprendizaje y desarrollo lector; evaluación y prácticas clave para el aprendizaje de habilidades del código base, evaluación y prácticas clave para el aprendizaje de habilidades de conceptos constructivistas, con énfasis en el rol del lenguaje académico y del vocabulario en la comprensión de textos complejos, el aprendizaje de cómo facilitar una discusión basada en textos académicos, y el aprendizaje de estrategias de lectura para el monitoreo de la comprensión de estudiantes entre 1° a 6° año básico. La carga del curso corresponde a 10 horas semanales incluido el tiempo de clases (4 horas y media) como de estudio autónomo (6 horas y media).

Se realizaron tres sesiones de evaluación de pares de 1:20 cada una, en los meses de marzo, mayo y junio. En el Apéndice A se detallan las tres actividades realizadas, y en Tabla 1 se resume para las tres actividades, el tipo de conocimiento implicado, las habilidades de pensamiento crítico promovidas, y la cantidad de sub-tareas involucradas.

Tabla 2-1: Resumen de las actividades realizadas

	Actividad 1	Actividad 2	Actividad 3
Conocimiento	-Conocimiento	-Conocimiento	-Conocimiento pedagógico

implicado	pedagógico disciplinar	pedagógico disciplinar	disciplinar -Conocimiento especializado sobre práctica generativa: Discusión productiva.
Habilidades de Pensamiento Crítico	Análisis y Explicación	Explicación, Interpretación e Inferencia	Inferencia y Evaluación
Cantidad de Subtareas involucradas	5	4	4

2.2.3. Instrumentos de evaluación

La efectividad de la intervención se evaluó utilizando tres rúbricas (Apéndice C) basadas en el modelo de pensamiento crítico de Facione y adaptada por los investigadores. Las rúbricas desarrolladas responden a un modelo de evaluación analítico, pues desglosan en varios indicadores y niveles de logros (o nivel de ejecución) una serie de criterios de evaluación (Airasian & Russell, 2001; Mertler, 2001). Este tipo de instrumento de evaluación permite aproximarse a un análisis detallado de competencias y habilidades y, además, posibilita a los investigadores la observación del progreso de los estudiantes y sus desarrollos y la generación de modelos de respuesta objetivos, que superen, por tanto, las subjetividades del evaluador (Saxton, Belanger y Becker, 2012).

Respecto a la construcción de los instrumentos, el proceso supuso tres momentos sustentados en tres aspectos evaluativos y pedagógicos fundamentales: *Momento 1.* Construcción de criterios de evaluación que responden a tres dimensiones: (1) práctica-textual que apunta a la formulación de la respuesta de los estudiantes; (2) habilidades de pensamiento crítico basados en el análisis y adaptación del modelo de pensamiento crítico de Facione (1990) que apunta a evaluar si la producción de los estudiantes se ajustaba al modelo de pensamiento crítico de Facione; (3) feedback, compuesta por dos sub dimensiones: deconstrucción de la habilidad de pensamiento crítico de evaluación, sumada a aspectos formales; y indicadores conceptuales dependientes de la intención del

evaluador. *Momento 2.* Construcción de niveles de desempeño abarcables que evaluarán, por lo tanto, un desarrollo progresivo, no cerrado (Mertler, 2001). Los niveles correspondientes son: logrado (4 puntos), en desarrollo (3 puntos), inicial (2 puntos), no logrado (1 punto). *Momento 3.* Construcción progresiva de descriptores de nivel. Las diferencias progresivas entre los niveles se articularon en dos polos descriptivos: un polo formal-cuantitativo, es decir, el paso de un nivel a otro supone la ausencia o no completitud de uno de los contenidos supuestos en la actividad (falta información) y un polo analítico-cualitativo, que supone un problema de análisis o no completitud del desarrollo conceptual de la respuesta.

En apéndice B se presenta, para cada una de las dimensiones del pensamiento crítico evaluadas (Interpretación, Análisis, Evaluación, Inferencia, y Explicación), los criterios definidos tanto para la definición de la habilidad como las dimensiones de la rúbrica, utilizados para determinar las rúbricas utilizadas para evaluar la Respuesta Inicial (R) y la Nueva Respuesta (NR) (Tabla C-1, Apéndice). Complementariamente, se definieron las rúbricas para evaluar el *Feedback* (F) (Tabla C-2, Apéndice) y la Integración de *Feedback* en Nueva Respuesta (F en NR). (Tabla C-3, Apéndice)

Apéndice C-1 ilustra la rúbrica para R y NR. Esta tiene por objetivo evaluar si la respuesta del estudiante, en la etapa 1, es decir, R y en la etapa 3, es decir, NR, evidencia un desarrollo adecuado de las habilidades implicadas en la pregunta. Se definieron cinco dimensiones, las que proyectan y adaptan las cinco habilidades superiores trabajadas, a saber, interpretación, análisis, evaluación, inferencia y explicación (Facione, 1990). Según lo anterior, si la pregunta del docente trabajaba las habilidades de interpretación y análisis, por ejemplo, eran estas las dimensiones que se evaluaban en la respuesta del estudiante. Para obtener el puntaje final de cada actividad, se calculó el promedio de las habilidades implicadas en la pregunta. Si la pregunta suponía habilidades de análisis e interpretación, el puntaje final correspondía al promedio entre dichas habilidades, teniendo entonces la actividad un máximo de 4 y un mínimo de 1 punto.

Apéndice C-2 ilustra la rúbrica para evaluar el *Feedback*, F. Esta se dividió en dos grandes dimensiones. La primera consistía de, criterios de evaluación (la capacidad del estudiante de proponer criterios concretos a su compañero), claridad y pertinencia de la explicación (la atingencia del comentario, i.e., que efectivamente responda al desarrollo de la tarea evaluada y no a criterios personales u opiniones no fundamentadas), intención de la argumentación (la capacidad del estudiante de orientar su *feedback* a un hacer, i.e., que provea no solo comentarios descriptivos, sino que además detalle estrategias de mejora) y formulación de la retroalimentación (la claridad de la formulación en términos discursivos). Esta dimensión se consideró de carácter obligatorio, aplicada a todas las respuestas, por la importancia a nivel evaluativo de los mismos (**Wiggins, 2012**). Lo anterior porque, como detalla **Sadler (1989)**, los componentes centrales de un *feedback* efectivo y significativo que impacte en el aprendizaje de los estudiantes comprenden tanto el desarrollo claro de un objetivo concreto como también el desarrollo de criterios evaluativos y la presentación de estrategias concretas para avanzar y superar el nivel alcanzado inicial.

La aplicación de los indicadores de la segunda dimensión era opcional, pues dependía del desarrollo de la respuesta inicial, R, o new response, NR, y las decisiones pedagógicas, evaluativas y metodológicas del retroalimentador para potenciar el avance del desarrollo de la respuesta de su par (Cowan, 1999). Estos indicadores opcionales, también se encuentran en el Apéndice C-2. Así, por ejemplo, si la respuesta del estudiante no presenta la evidencia solicitada en la actividad, el evaluador aplicará el indicador correspondiente. En suma, el evaluador deberá evaluar la pertinencia de la aplicación del indicador correspondiente, de acuerdo con el desarrollo de la respuesta que esté revisando (Brookhart, 2013).

Por último, Apéndice C-3 ilustra los criterios para evaluar el proceso de integración de los comentarios de F en la nueva respuesta, NR, que corresponde a la metacognición, según Facione (1990). La nueva respuesta, NR constituye una reformulación de la

respuesta inicial, R, como resultado de un autoexamen o autoevaluación, producto de una evaluación externa F, más la experiencia del alumno de entregar un F, a otro R que le correspondió revisar. El proceso final, de integración, por tanto, supone la articulación de todas las habilidades de pensamiento crítico. El estudiante interpreta e infiere a partir de los comentarios recibidos en F; evalúa y analiza qué tan atingente son los comentarios de F al texto R y, además, articula y explica los cambios potenciales de R a NR. Para evaluar el proceso anterior, la rúbrica de metacognición evalúa el reconocimiento e integración de los aspectos y niveles tanto positivos (o atingentes para su desarrollo) como negativos (o no atingentes para su desarrollo) indicados por el evaluador (F). Evalúa además si se entregan evidencias claras y coherentes (con los aspectos positivos y negativos mencionados) que justifiquen los posibles cambios efectuados (señalar razones de eliminación, incorporación o mantención de información). Además, la respuesta se hace cargo de las posibles deficiencias indicadas por el evaluador y expone los pasos, operaciones, nueva información, métodos o interpretaciones que permitieron superar el problema o mejorar el desarrollo (Spiller, 2012).

Las rúbricas construidas se validaron con juicio experto (Escobar-Perez, et al, 2008). Tres expertos en el área validaron de manera individual a través de un cuestionario de validación con escala de Likert para medir el grado de acuerdo o desacuerdo con el descriptor.

Finalmente, se realizaron dos evaluaciones a los alumnos. Test 1 (T1), corresponde a la calificación que obtuvieron los estudiantes en el test rendido el 28 de abril, previo al desarrollo de las actividades 2 y 3. Test 2 (T2), corresponde a la calificación que obtuvieron los estudiantes en el test rendido el 23 de junio, posterior al desarrollo de las tres actividades. Estos instrumentos consistieron de 14 y 17 preguntas respectivamente, que mezclaban preguntas abiertas, de alternativas y de selección de ítems en el texto.

2.2.4. Aplicación de rúbricas

Todas las respuestas de los alumnos (R, F, y NR), fueron evaluadas por dos evaluadores en paralelo para doble codificación (Apéndice L). En primer lugar, se aplicó la rúbrica a todas las respuestas, R, para cada una de las tres actividades (Tablas L-1, L-2 y L-3, Apéndice L); una vez completadas éstas, se evaluaron todas las nuevas respuestas, NR, para cada una de las tres actividades (Tablas L-4, L-5 y L-6, Apéndice L). Lo anterior fue para evitar que la evaluación de la primera etapa, R, afectara la objetividad de la aplicación del instrumento en la tercera etapa, NR. Respecto de la etapa de *feedback*, F, para cada una de las tres actividades (Tablas L-7, L-8 y L-9, Apéndice L), posterior a la revisión de R. Finalmente, se evaluó el proceso de integración de comentarios de *feedback*, F, en NR, para cada una de las tres actividades (Tablas L-10, L-11 y L-12, Apéndice L).

Considerando que la confiabilidad y la validez de la muestra son indicadores clave de la calidad de un instrumento (Kimberlin & Winterstein, 2008), la validación se realizó mediante Cohen Kappa (McHugh, 2012). En el caso particular de las rúbricas en estudio (Apéndice C), el nivel de acuerdo que los evaluadores logran para las respectivas respuestas corresponde a los niveles de logro alcanzados (logrado, en desarrollo, inicial y no logrado). Apéndice D ilustra el proceso detallado del cálculo de Cohen Kappa respectivo para la evaluación que realizaron las dos personas involucradas. Tablas D1 a D5, contienen las matrices de nivel de acuerdo de los evaluadores, para cada una de las cinco habilidades de pensamiento crítico evaluadas, tanto para R como NR. Tablas D6 a D13 muestran las matrices de nivel de acuerdo de la rúbrica de F, y Tabla D14 contiene la matriz de nivel de acuerdo de la rúbrica que evalúa la integración de F en NR. Para obtener la $\sum a$ se deben sumar los elementos de la diagonal de la matriz de nivel de acuerdo. Es importante mencionar que se tiene una única tabla para cada dimensión de la rúbrica de R y NR, puesto que la rúbrica que se utiliza en ambas etapas es la misma. Tabla D15 contiene la suma de las frecuencias esperadas de cada criterio, para cada dimensión de cada una de las tres rúbricas. Finalmente, Tabla D16, muestra que las

dimensiones que evalúa la rúbrica de R y NR tuvieron un Cohen Kappa de 0,83, mientras las dimensiones que evalúa la rúbrica de F, tuvieron valor de 0,76, mientras que la rúbrica que evalúa la integración de F en NR tuvo un Cohen Kappa de 0,92.

Luego de la revisión y doble codificación, se analizaron y discutieron las diferencias de aplicación y se calibraron las potenciales problemáticas de criterios.

2.2.5. Análisis cuantitativo

Para calcular los distintos estadísticos y realizar el modelo de regresión lineal, se utilizó el software estadístico (STATA v. 12.0).

2.2.5.1. Modelo de regression *Stepwise*

Para conocer cuáles de los factores estudiados en la experiencia presentada inciden en el desarrollo del pensamiento crítico (Tabla 2-2), se desarrolló un modelo de regresión lineal con *forward stepwise selection method* (Tabla 2-3). Este método permite explorar la relación de los distintos factores estudiados, y en base a esto establecer una comparación relativa entre ellos, identificando cuales son los factores más determinantes (Lewicki et al, 2006 En Luciando, et al, 2018).

El modelo se construyó en función de la variable dependiente *Avance (A)*. Esta variable mide la mejora en las habilidades de pensamiento crítico, mientras se desarrollan las actividades, i.e., refleja la diferencia entre los niveles de logro en *Response (R)* y *New Response (NR)*. Para el modelo se consideraron las variables de *Response (R)*, *Feedback (F)*, el cual para efectos de la regresión se divide en el *feedback received* de cada estudiante (*FR*) y *feedback delivered* de cada estudiante (*FD*). También se consideró la variable que mide la integración del *feedback* recibido en la *new response (NR(FR))*, y

por último el Test 1 (T1) y el Test 2 (T2) que corresponden a los tests que miden los conocimientos pedagógicos disciplinares del curso.

La Tabla 2-2 contiene todas las variables mencionadas, dependiente y predictoras, en sus distintas dimensiones, es decir, la variable dependiente A , se divide A_h , A_n y $\bar{A}$, las variables independientes tales como la variable R se divide en R_h , R_n y $\bar{R}$, la variable NR se divide en NR_h , NR_n y $\overline{NR}$, la variable F se divide en F_n y $\bar{F}$, la variable $NR(FR)$ se divide en $NR(FR)_n$ y $\overline{NR(FR)}$, y las variables $T1$ y $T2$. En la tabla se muestra el objetivo de cada variable, como se calcula cada una de las dimensiones y dónde se encuentra el detalle.

Tabla 2-2: Variables consideradas para construir el modelo de regresión lineal

Variable	Objetivo	Cálculo para obtener el valor de la variable	Detalle de resultados
A_h	Medir el avance de las habilidades de pensamiento crítico de cada actividad.	Diferencia entre NR_h y R_h de cada habilidad en cada actividad	Apéndice E Tabla E4
A_n	Medir el promedio del avance de las habilidades de pensamiento crítico de cada actividad.	Diferencia entre el promedio de las <i>new response</i> , NR_n , y el promedio de las <i>response</i> , R_n	Apéndice E Tabla E4
$\bar{A}$	Medir el avance de las habilidades de pensamiento crítico en el promedio de las tres actividades.	Diferencia entre $\overline{NR}$ y $\bar{R}$	Apéndice E Tabla E9
R_h	Medir las habilidades de pensamiento crítico en la respuesta inicial de cada actividad	Corresponde al puntaje que se obtiene al aplicar la rúbrica del Apéndice C-1 a las respuestas iniciales de cada actividad	Apéndice E Tablas E1, E2 y E3
R_n	Medir el promedio de las habilidades de pensamiento crítico en la respuesta inicial de cada actividad	Corresponde al promedio de las R_h de las habilidades involucradas en cada actividad	Apéndice E Tablas E1, E2 y E3
$\bar{R}$	Medir el promedio de las respuestas iniciales de las tres actividades	Promedio de R_1 , R_2 y R_3	Apéndice E Tabla E9
NR_h	Medir las habilidades de pensamiento crítico en la respuesta final de cada actividad	Corresponde al puntaje que se obtiene al aplicar la rúbrica del Apéndice C-1 a las respuestas finales de cada actividad	Apéndice E Tablas E1, E2 y E3
NR_n	Medir el promedio de las habilidades de pensamiento crítico en la respuesta final de cada actividad	Es el promedio de las NR_h de las habilidades involucradas en cada actividad	Apéndice E Tablas E1, E2 y E3
$\overline{NR}$	Medir el promedio de las respuestas finales de las tres actividades	Promedio de NR_1 , NR_2 y NR_3	Apéndice E Tabla E9
F_n	Medir las componentes obligatorias y opcionales del feedback en cada actividad. Se divide en FR_n y FD_n	Suma de los componentes obligatorios y opcionales que se obtienen al aplicar la rúbrica del Apéndice C-2	Apéndice E Tabla E5, E6 y E7
$\bar{F}$	Medir el promedio de las suma de las componentes obligatoria y opcionales de las tres actividades. Se divide en $\bar{FR}$ y $\bar{FD}$	Promedio de F_1 , F_2 y F_3 , por componente FR y FD	Apéndice E Tabla E9
$NR(FR)_n$	Medir como los estudiantes integran el feedback (F) en sus nuevas respuestas (NR) en cada actividad.	Corresponde al puntaje que se obtiene al aplicar la rúbrica del Apéndice C-3	Apéndice E Tabla E8
$\overline{NR(FR)}$	Medir como los estudiantes integran el feedback (F) en sus nuevas respuestas (NR) en el promedio de las tres actividades	Promedio de $NR(FR)_1$, $NR(FR)_2$ y $NR(FR)_3$	Apéndice E Tabla E9
$T1$	Medir el desempeño en la primera evaluación convencional del curso	Calificaciones en la primera evaluación formal del curso	Apéndice E Tabla E10
$T2$	Medir el desempeño en la segunda evaluación convencional del curso	Calificaciones en la segunda evaluación formal del curso	Apéndice E Tabla E10

Nota: Para X_h "h" representa una habilidad de pensamiento crítico y para X_n "n" representa el número de la actividad.
 (*) Tomando como referencia el modelo de Figura 2-1, y considerando a Student 1

Es importante mencionar, que, dado que no todos los alumnos participaron en todas las actividades, las variables que se calculan como el promedio de las tres actividades, se calcularon como el promedio de los puntajes de las actividades en las que participó cada estudiante; e.g., si el estudiante solo participó de las actividades 1 y 3, para calcular $\bar{R}$ se obtuvo el promedio entre R de la actividad 1 (R_1) y R de la actividad 3 (R_3). El Apéndice F muestra cómo se calcularon estas variables para tres posibles casos.

Las variables $\bar{R}$, $\bar{F}$ (que se divide en $\overline{FR}$ y $\overline{FD}$), $T1$ y $T2$, son las variables predictoras y la variable $\bar{A}$ es la variable dependiente del modelo. Como se menciona anteriormente, las otras dimensiones de las variables no fueron consideradas en el modelo y, por tanto, fueron descartadas. Estas variables son: las variables de NR_h y NR_n , las que son utilizadas para construir la variable $\overline{NR}$; lo mismo sucede con las variables R_h y R_n que se utilizan para calcular $\bar{R}$; las variables A_h y A_n que se utilizan para calcular $\bar{A}$; las variables F_n que se utilizan para calcular $\bar{F}$; y las variables $NR(FR)_n$ que se utilizan para calcular $\overline{NR(FR)}$. $\overline{NR}$ también es descartada, ya que se utiliza para calcular $\bar{A}$ ($\bar{A} = \overline{NR} - \bar{R}$), por lo tanto la información que entregan $\bar{R}$ y $\overline{NR}$ en conjunto, es la misma que entregan $\bar{R}$ y $\bar{A}$ en conjunto. Por último, la variable $\overline{NR(FR)}$ también fue descartada, ya que entrega información redundante con respecto a $\bar{F}$. En el Apéndice K, las Tablas K1, K2 y K3 muestran la correlación de las variables que fueron descartadas en función de las variables $\bar{R}$, $\overline{NR}$ y $\bar{F}$ respectivamente. La fuerte correlación entre las variables es evidencia empírica de la redundancia en la información (Taylor, 1990 En Olasina, 2018).

Siguiendo el *forward stepwise method*, se determinó el ingreso de las variables al modelo en el siguiente orden: $\overline{FR}$, $\overline{FD}$, $T2$, $\bar{R}$ y $T1$. El *forward stepwise method* establece que, a partir de un modelo sin variables, se añaden las variables una a una en base a un determinado criterio estadístico. En cada paso, se añade al modelo la variable que entregue mayor significancia estadística en función del criterio elegido, hasta que ninguna de las variables cumpla con dicho criterio (Lewicki et al, 2006 En Luciando, et

al, 2018). El criterio de selección utilizado para la construcción de este modelo es el de R^2 ajustado, es decir, a partir de un modelo sin variables se añadió en cada paso la variable que entregase al modelo un mayor incremento en el valor de Adj R^2 . Finalmente, el modelo que se obtiene, maximiza la variabilidad explicada. Cada paso, en donde se agrega cada una de las distintas variables, se encuentra en el Apéndice G, mientras que el modelo final con todas las variables se encuentra en la Tabla 2-3.

2.2.5.2. Estadísticos descriptivos and T-test

Con el objetivo de identificar las variaciones entre variables, entre las distintas sesiones y en el desarrollo de una actividad, se hizo un análisis de los estadísticos descriptivos de R , NR , A , F (FR y FD) y $NR(FR)$. Los estadísticos de R , NR y A se encuentran en las Tablas 2-4 y 2-5, mientras que los estadísticos de F , divididos en FR y FD , y los estadísticos de $NR(FR)$ se encuentran en la Tabla 2-6.

Por último, para conocer como los alumnos desarrollaban las distintas habilidades de pensamiento crítico, a través de las distintas actividades realizadas, se realizó un análisis de la distribución de la variable avance, el que se muestra en la Tabla 2-7. Para esto, se compara la distribución de la *response* (R), con la distribución de la *new response* (NR) para cada una de las actividades y por cada una de las habilidades de pensamiento crítico. Con esta comparación, fue posible identificar como era la mejora de los estudiantes y cómo se movían de un nivel de desempeño a otro.

Para estudiar si el desarrollo de las habilidades de pensamiento crítico, entre R y NR , era estadísticamente significativo, se evaluó la significancia estadística entre estas variables con T-Students (t-test). Estos tests, se aplicaron en cada una de las dimensiones, es decir, por cada una de las cinco habilidades del pensamiento crítico estudiadas ($A_h = NR_h - R_h$), Tabla 2-4. Además se estudió para cada una de las actividades, ($A_n = NR_n - R_n$), y para el promedio de las tres actividades ($\bar{A} = \overline{NR} - \bar{R}$), Tabla 2-5.

2.2.6. Análisis cualitativo

Con el objetivo de conocer las diversas perspectivas de los estudiantes en relación a la evaluación de pares, se realizó un análisis cualitativo. Este fue producto de un *focus group*, de aproximadamente 1 hora de duración, con 6 estudiantes elegidos aleatoriamente post implementación, el cual fue grabado y transcrito de manera literal. El estudio abordaba tanto los aspectos de las actividades al exterior del aula, como los aspectos de las actividades al interior. La pregunta abierta orientadora correspondió a ¿Cuál es la percepción que tienen los estudiantes aprendiendo bajo la lógica del *Flipped Classroom* en Didáctica de la Lectura? En este estudio nos concentraremos en los aspectos del trabajo en aula.

El análisis del proceso resultante se basó en la teoría fundamentada (Murphy, 2017; Sinclair, S et al, 2017). El análisis comenzó con la lectura de la transcripción del *focus group*. Posteriormente, por medio de una lectura línea a línea se identificaron y fragmentaron los datos, generando de esta forma las unidades de sentido. A partir de los datos procedentes de las unidades de sentido se realizó el procedimiento de codificación abierta del texto, por medio de la pregunta ¿De qué habla el texto? Se realizaron las primeras categorías, y por medio de la pregunta ¿qué dice de lo que habla? las propiedades. El proceso se realizó hasta lograr la saturación teórica, dado que los datos no otorgaban información adicional. A partir de la codificación efectuada se realizó la lista jerárquica y el primer árbol, posteriormente una etapa de revisión detallada de cada una de las categorías se reflexionó en torno a la pregunta de investigación y se articuló nuevamente el material en función de ésta, por lo que a raíz de esta reflexión se realizó una recodificación y reordenación de la lista.

2.2.7. Consideraciones éticas

Los estudiantes antes de participar de la primera actividad obtuvieron información sobre el propósito de la investigación, sobre las actividades específicas, duración de estas, así

como riesgos y beneficios de la intervención. Se hizo hincapié en la voluntariedad de participar en el estudio. Asimismo, se les solicitó firmar un consentimiento informado.

Respecto el análisis cualitativo, los participantes obtuvieron información y explicación sobre el propósito de la investigación, y la duración del *focus group*; se les explico que toda la información sería usada para fines investigativos. Así mismo, se les pidió autorización para grabar. Todos los estudiantes otorgaron su consentimiento de manera escrita.

Finalmente, el estudio contó con la aprobación ética otorgada por el comité de ética de la Universidad.

2.3. Resultados

2.3.1. Cuantitativos

2.3.1.1. Modelo de Regresión Stepwise

La Tabla 2-3 muestra el modelo de regresión lineal para la variable dependiente *Avance (A)* en función de las variables predictoras *Feedback Received (FR)*, *Feedback Delivered (FD)*, *Test 2 (T2)*, *Response (R)* y *Test 1 (T1)*. El Apéndice G detalla el proceso de selección paso a paso del método *stepwise*, con el que finalmente se construyó el modelo mostrado. La primera variable seleccionada fue el *Feedback Received ($\overline{FR}$)*, el que explica un 25% de la variabilidad ($Adj R^2 = .25$), de manera significativa $F(1, 33) = 12.44, p < .05$; la segunda variable fue el *Feedback Delivered ($\overline{FD}$)*, el que explica un 11% adicional de la variabilidad ($Adj R^2 = .36$), de manera significativa $F(2, 32) = 10.52, p < .05$; la tercera variable en incluirse fueron los resultados en el *Test 2 (T2)* ($Mean = 5.41, Std. Deviation = 1.14$), esta explicó un 5% adicional de la variabilidad ($Adj R^2 = .41$), de manera significativa $F(3, 31) = 8.73, p < .05$; la cuarta variable que se ingresó al modelo fue el *Response ($\overline{R}$)* el que explicó un 3% adicional de la variabilidad ($Adj R^2 = .44$), de manera significativa $F(4, 30) = 7.69, p < .05$; y la

quinta y última variable fueron los resultados en el *Test 1* ($T1$) ($Mean = 5.13$, $Std. Deviation = 0.53$), los que explican un 2% adicional de la variabilidad ($Adj R^2 = .46$), de manera significativa $F(5, 29) = 6.85$, $p < .05$.

Tabla 2-3: Resumen análisis de regresión lineal de variables que predicen el avance de las habilidades de pensamiento crítico (N = 35)

Variables	b	Std. Error	β	t	p-value
$\overline{FR}$	.18	0.1495	0.6177	4.13	< 0.001
$\overline{FD}$	.12	0.1429	0.4031	2.82	0.009
$T2$	.13	0.1577	0.4419	2.8	0.009
$\bar{R}$	-.10	0.1752	-0.3316	-1.89	0.068
$T1$	-.05	0.1359	-0.2024	-1.49	0.147

Nota: modelo con variables estandarizadas ($Mean = 0$ y $Std. Deviation = 1$).

Los datos utilizados para construir la regresión se encuentran en el Apéndice H.

Las variables independientes *Feedback Received* ($\overline{FR}$), *Feedback Delivered* ($\overline{FD}$), y *Test 2* ($T2$) son predictores significativos de la variable dependiente *Avance* ($\bar{A}$), siendo la primera la variable la que más influye. No son predictores y tienen coeficientes $\beta < 0$ el *Promedio Response* ($\bar{R}$) y los resultados del *Test 1* ($T1$).

2.3.1.2. Estadísticos descriptivos y T-test

En la Tabla 2-4 se puede apreciar que el *Avance* entre R y NR es estadísticamente significativo para todas las habilidades de pensamiento crítico involucradas en la Actividad 1, correspondiente a la habilidad Análisis ($p\text{-value} = 0.0008$) y para la habilidad Explicación ($p\text{-value} = 0.0003$) y para todas las habilidades involucradas en la Actividad 2 correspondientes a la habilidad explicación ($p\text{-value} = 0.0025$), la habilidad Interpretación ($p\text{-value} = 0.0119$) y para la habilidad Inferencia ($p\text{-value} = 0.0101$). En cambio, para la Actividad 3, en la habilidades involucradas correspondientes a Inferencia y ($p\text{-value} = 1.00$) y evaluación ($p\text{-value} = 0.3265$), por lo que el avance no es significativo. En esta actividad únicamente un estudiante logra un avance entre R y NR y

lo hace para ambas habilidades involucradas, además es la única actividad donde un alumno presenta un retroceso en una habilidad, el cual se manifestó en la habilidad de inferencia. En este caso, dado que el avance del alumno que mejora es de igual magnitud que el retroceso del alumno que empeora, se anulan los efectos y se tiene que la media del avance para la habilidad de inferencia en la actividad 3 es cero.

Tabla 2-4: Estadísticos descriptivos de Response (R), New Response (NR) y Avance (A) y t-tests de significancia estadística del Avance (A), en cada una de las actividades

	Actividad 1 (N = 13)			Actividad 2 (N = 33)			Actividad 3 (N = 27)		
	Mean	Std. Dev.	p-value (t-test)	Mean	Std. Dev.	p-value (t-test)	Mean	Std. Dev.	p-value (t-test)
$R_{análisis}$	2.54	0.88		—	—	—	—	—	—
$NR_{análisis}$	3.54	0.97		—	—	—	—	—	—
$A_{análisis}$	1.00	0.82	0.0008	—	—	—	—	—	—
$R_{explicación}$	2.77	0.73		2.48	0.80		—	—	—
$NR_{explicación}$	3.77	0.60		2.79	0.93		—	—	—
$A_{explicación}$	1.00	0.71	0.0003	0.30	0.53	0.0025	—	—	—
$R_{interpretación}$	—	—	—	2.58	0.83		—	—	—
$NR_{interpretación}$	—	—	—	2.76	0.90		—	—	—
$A_{interpretación}$	—	—	—	0.18	0.39	0.0119	—	—	—
$R_{inferencia}$	—	—	—	2.52	0.91		2.44	0.58	
$NR_{inferencia}$	—	—	—	2.82	0.98		2.44	0.58	
$A_{inferencia}$	—	—	—	0.30	0.64	0.0101	0.00	0.28	1.00
$R_{evaluación}$	—	—	—	—	—	—	2.93	1.14	
$NR_{evaluación}$	—	—	—	—	—	—	3.04	1.09	
$A_{evaluación}$	—	—	—	—	—	—	0.11	0.58	0.3265

Nota: estadísticos se obtienen a partir de la tabulación de resultados en Apéndice E, Tablas E1, E2, E3 y E4.

Analizando las tres sesiones, en la Tabla 2-5, se observa que el promedio del curso tuvo siempre un incremento entre R y NR para cada una de las tres sesiones, y también en las tres sesiones promediadas, y que este avance es significativo, salvo para la tercera sesión (A_3). La primera actividad es la que muestra mayor *Avance* ($Mean = 1.00$ punto), lo que es equivalente a pasar al siguiente nivel de logro: de no logrado a inicial, de inicial a en desarrollo y de en desarrollo ha logrado.

Tabla 2-5: Estadísticos descriptivos de Response (R), New Response (NR) y Avance(A) y t-tests de significancia estadística del Avance (A), del promedio de las habilidades involucradas en cada actividad y del promedio de las tres actividades.

	N	Mean	Std. Dev.	range of scale	p-value
R_1	13	2.65	0.72	1 - 4	
NR_1	13	3.65	0.77	1 - 4	
A_1	13	1.00	0.68	0 - 3	0.0002
R_2	33	2.53	0.77	1 - 4	
NR_2	33	2.78	0.88	1 - 4	
A_2	33	0.25	0.43	0 - 3	0.0017
R_3	27	2.69	0.82	1 - 4	
NR_3	27	2.74	0.79	1 - 4	
A_3	27	0.05	0.40	0 - 3	0.4773
$\bar{R}$	35	2.54	0.62	1 - 4	
$\bar{NR}$	35	2.80	0.75	1 - 4	
$\bar{A}$	35	0.26	0.05	0 - 3	< 0.0001

Nota: estadísticos se obtienen a partir de la tabulación de resultados en Apéndice E, Tablas E1, E2, E3, E4 Y E9.

Respecto al *Feedback Received*, FR , y *Feedback Delivered*, FD , Tabla 2-6, éstos debieran ser iguales tanto en la media como en desviación estándar, dado que el feedback que un estudiante recibe es el mismo que el que otro estudiante entrega. Considerando que no todos los estudiantes completaron la etapa new response, NR , y que para el análisis es necesario considerar solo aquellos alumnos que si la completaron, pues no se podría calcular el avance, A , sin NR , habrá en la muestra analizada alumnos que entregaron feedback, FD , (si considerado en Tabla 2-6), a estudiantes que no realizaron la etapa de new response, NR (con FD s que no fueron considerados en Tabla 2-6).

Tabla 2-6: Estadísticos descriptivos Feedback Received (FR), Feedback Delivered (FD) y de la Integración de FR en NR (NR(FR)).

Variable	N	Mean	Std. Dev.	range of scale
FR_1	13	18.08	6.28	4 - 32
FD_1	9	20.33	2.83	4 - 32
$NR(FR)_1$	13	2.92	1.26	1 - 4
FR_2	33	9.67	7.29	4 - 32
FD_2	21	14.52	2.82	4 - 32
$NR(FR)_2$	33	1.67	1.47	1 - 4
FR_3	27	9.85	6.23	4 - 32
FD_3	21	12.67	3.62	4 - 32
$NR(FR)_3$	27	1.44	1.34	1 - 4
$\overline{FR}$	35	10.28	6.24	4 - 32
$\overline{FD}$	35	14.45	3.51	4 - 32
$\overline{NR(FR)}$	35	1.72	1.17	1 - 4

Nota: Cálculos de tabla en función de Apéndice E, Tablas E5, E6, E7, E8 y E9.

En la Tabla 2-7, se muestra la distribución del desempeño de los estudiantes, para cada una de las actividades según las habilidades de pensamiento crítico involucradas en la actividad. En la Actividad 1 hay movilidad en todos los niveles de desempeño desde R a NR, para las dos habilidades involucradas. En la Actividad 2, a pesar de que el porcentaje de estudiantes que se encuentran en el nivel en desarrollo para las habilidades de explicación e interpretación se mantiene igual para R y NR, hay movilidad en todos los niveles con excepción del nivel no logrado. En la habilidad de inferencia en la Actividad 3, un estudiante se mueve del nivel inicial a nivel en desarrollo, mientras que otro estudiante lo hace en sentido opuesto. El mismo estudiante que mejora en la habilidad de inferencia, también pasa de nivel no logrado a nivel logrado en evaluación. El Apéndice I, detalla cómo se distribuyen las magnitudes del avance entre R y NR por habilidad, para cada una de las actividades.

Tabla 2-7: Distribución resultados de aplicación de rúbrica de R y NR según criterios de evaluación de la rúbrica

	Actividad 1 (N = 13)		Actividad 2 (N = 33)		Actividad 3 (N = 27)	
	R	NR	R	NR	R	NR
<i>Habilidad de Análisis</i>						
Nivel no logrado	15.38%	7.69%	-	-	-	-
Nivel inicial	23.08%	7.69%	-	-	-	-
Nivel en desarrollo	53.85%	7.69%	-	-	-	-
Nivel logrado	7.69%	76.92%	-	-	-	-
<i>Habilidad Explicación</i>						
Nivel no logrado	0.00%	0.00%	6.06%	6.06%	-	-
Nivel inicial	38.46%	7.69%	51.52%	36.36%	-	-
Nivel en desarrollo	46.15%	7.69%	30.30%	30.30%	-	-
Nivel logrado	15.38%	84.62%	12.12%	27.27%	-	-
<i>Habilidad de Interpretación</i>						
Nivel no logrado	-	-	6.06%	6.06%	-	-
Nivel inicial	-	-	45.45%	36.36%	-	-
Nivel en desarrollo	-	-	33.33%	33.33%	-	-
Nivel logrado	-	-	15.15%	24.24%	-	-
<i>Habilidad de Inferencia</i>						
Nivel no logrado	-	-	6.06%	6.06%	3.70%	3.70%
Nivel inicial	-	-	57.58%	39.39%	48.15%	48.15%
Nivel en desarrollo	-	-	15.15%	21.21%	48.15%	48.15%
Nivel logrado	-	-	21.21%	33.33%	0.00%	0.00%
<i>Habilidad de Evaluación</i>						
Nivel no logrado	-	-	-	-	14.8%	11.1%
Nivel inicial	-	-	-	-	22.2%	22.2%
Nivel en desarrollo	-	-	-	-	18.5%	18.5%
Nivel logrado	-	-	-	-	44.4%	48.2%

Nota: distribución de criterios de evaluación se obtienen a partir de tabulación de resultados en Apéndice E, Tablas E1, E2 y E3.

2.3.2. Cualitativos

Al desarrollar la teoría fundamentada, la categoría relacionada con el trabajo de aula es “experiencia de clase” la cual será categoría central y unificadora de los datos. Tabla 2-8, muestra la categoría clave y sus categorías correspondientes. Considerando que el objetivo central es explorar la evaluación de pares, desde el estudio cualitativo, nos centraremos en esta categoría.

Tabla 2-8: Categoría clave y categorías

Categoría Clave	Categoría	Explicación de la Categoría
Experiencia de Clase	Importancia del Docente	Esta categoría indica la acción del docente en la sala de clases, entendido como la influencia del profesor en la clase. Los alumnos indican el valor de poder realizar actividades prácticas con el profesor, en su función de guía y moderadora del aprendizaje de los alumnos
	Importancia del grupo curso	Indica cómo los estudiantes influyen en cómo se desarrollan las actividades de aula.
	Evaluación de pares	Indica cómo fue el proceso de evaluar a los pares dentro de la sala de clases.

2.3.2.1. Evaluación de pares y propiedades

Tal como muestra Tabla 2-9 que ilustra las categorías y citas de los estudiantes, las propiedades que emergen son: Interés y disposición frente a la tarea, proceso de aprendizaje frente a la evaluación de pares, proceso de adopción de la tecnología y construcción social del conocimiento por medio de la evaluación de pares y desarrollo de las habilidades superiores.

A continuación, la tabla muestra las categorías relacionadas con la evaluación de pares.

Tabla 2-9: Categorías y citas

Categorías -Explicación de la categoría	Citas
<u>Interés y disposición frente a la tarea</u> Los estudiantes indican que el interés y la disposición influyen directamente en la retroalimentación que ellos hicieron a sus compañeros	<i>Student 2: “la calidad de la retroalimentación depende del interés”.</i> <i>Student 1: “Depende de la disposición; mejoró a hacer una buena retroalimentación”.</i> <i>Student 6: “Depende del profe, del grupo, del alumno, todo depende de la disposición”.</i> <i>Student 4: “Como las tareas eran</i>

	<i>desafiantes daban ganas de retroalimentar”.</i>
<u>Proceso de aprendizaje frente a la evaluación de pares:</u> Los estudiantes indican que el proceso de retroalimentar un compañero mejoraba en el tiempo.	<i>Student 5: “Cuando comenzamos el curso yo no sabía cómo hacer una retroalimentación y tuve que aprenderlo, fue difícil al comienzo, pero después pude hacerlo”.</i> <i>Student 4: “Al principio yo hacía muy malas retroalimentaciones porque no sabía cómo hacerlo”.</i> <i>Student 1: “Fue mejorando con el tiempo”.</i> <i>Student 6: “Lo novedoso y fácil de utilizar, la retroalimentación entre compañeros”.</i>
<u>Proceso de Adopción de la tecnología</u> Los estudiantes indican que es necesario conocimiento de la herramienta y tecnología.	<i>Student 3: “Me costaba al principio, conectarme a internet, abrir la página (...) no sabía si había enviado mi trabajo”.</i> <i>Student 2: “(...) Me demoraba porque no estaba acostumbrada a trabajar con computador en las clases”.</i> <i>Student 1: “Cuando empezamos no entendía como tenía que evaluar a mi compañero, me costaba comprender si había escrito todo bien en la plataforma”.</i>
<u>Construcción social del conocimiento por medio de la evaluación de pares.</u> Los estudiantes indican cómo se construyó el conocimiento por medio de las actividades de evaluación de pares.	<i>Student 2: “Facilita la construcción del conocimiento, ya que hay otro que te da consejos de cómo puedes mejorar tu respuesta”.</i> <i>Student 6: “Todo estaba armado y tenía un sentido. Las actividades (peer review) eran aterrizadas y te permitía aplicar lo que habías aprendido de manera más profunda”.</i> <i>Student 1: “La actividad se armaba entre todos” (...) y vas construyendo mucho.</i> <i>Student 3: “Aplicar todo lo que sabes para que la otra persona pueda mejorar bien”.</i> <i>Student 2: “Es difícil retroalimentar cuando uno no tiene clara la respuesta o no se ha estudiado bien el contenido”.</i> <i>Student 5: “Las actividades te ponían en el</i>

	<i>lugar del profesor”.</i> <i>Student 1: “Ver otra respuesta te permite mejorar tu respuesta” o al revés “Cuando te llegaba (una) muy mala respuesta para revisar, no te servía a ti, porque le faltaba todo... no te ayuda a ti”.</i> <i>Student 6: “El sistema permite ver desde otra perspectiva”.</i> <i>Student 4: “Nos permite mejorar respuesta, dar y recibir consejos”.</i> <i>Student 5: “Había muchos aspectos en la retroalimentación, ver desde otra perspectiva (...) que bueno los consejos”.</i>
--	--

2.4. Discusión

2.4.1. Desarrollo de las habilidades de orden superior

Se observa una mejora en el desarrollo de las habilidades superiores medido a través de las rubricas utilizadas para medir interpretación, análisis, evaluación, inferencia y explicación, tanto en la respuesta inicial R, como en la respuesta después del feedback, NR. Los resultados, Tabla 4, indica un avance entre la respuesta **R** y **NR**, en dos de las tres actividades realizadas. El no avance de la tercera actividad, se explica por que la profesora indicó a los alumnos que se concentraran en argumentar mas que en retroalimentar. Esto queda en evidencia en Tabla 7, donde se observa que en ambas habilidades (Inferencia y Evaluación) presentes en Actividad 3, practicamente todos los alumnos se mantienen en sus habilidades iniciales. Otros de los posibles factores involucrados, es el tiempo dedicado a cada fase, y las discusiones y andamiajes dados por el docente entre fases. Es importante destacar la importancia de la estructura de la clase. La orquestación de los recursos involucrados (Nussbaum, et al, 2013; Dillenbourg, 2013), que en este caso son tanto los de fuera de la clase como los de dentro, son relevante de considerar al planificar la clase. Esto se reflejó en el comentario de un

alumno *“Student 6: Todo estaba armado y tenía un sentido. Las actividades (peer review) eran aterrizadas y te permitía aplicar lo que habías aprendido de manera más profunda”*.

Tal como se muestra en Tabla 2-4, es relevante indicar que, salvo la habilidad de Evaluación, que solo se realizó en la tercera sesión, todas las otras habilidades (Análisis, Explicación, Interpretación e Inferencia), mostraron mejoras en su evolución en la actividad gracias al proceso de *feedback*. Es necesario trabajo futuro para estudiar si es posible también desarrollar esta habilidad con el modelo de andamiaje propuesto, o el desarrollo de la habilidad de evaluación es más complejo que las otras habilidades (Zaidi, et al, 2018)

2.4.2. Influencia de la evaluación de pares en el desarrollo de habilidades superiores

El modelo de regresión lineal muestra que la retroalimentación recibida tiene el impacto mayor en el avance de las habilidades de pensamiento crítico (Tabla 2-3, $\overline{FR}$ ($\beta = 0.62$), ($p\text{-value} = <0.001$). Esto es consistente con lo que dicen los alumnos, por ejemplo, *Student 3: “Aplicar todo lo que sabes para que la otra persona pueda mejorar bien”*. Una buena retroalimentación muestra un camino de cómo mejorar una respuesta, y la evaluación de pares puede brindar un apoyo para que el estudiante mejore (Vásquez-Colina M, et al 2017).

Por otro lado, consistente con la bibliografía, la evaluación de pares es un proceso donde el estudiante tiene un papel activo en su propia evaluación y en la de otros (Planas-Lladó et al, 2018). Un alumno, *Student 2*, destacó que la evaluación de pares *“Facilita la construcción del conocimiento, ya que hay otro que te da consejos de cómo puedes mejorar tu respuesta”*. Finalmente, la retroalimentación no solo beneficia la motivación si no que depende de ella (Lai y Hwang, 2015; Tseng & Tsai, 2007; Hwang, Hung, y Chen, 2013), lo que destaca *Student 1: “Depende de la disposición mejoró a hacer una buena retroalimentación”*.

El modelo de regresión lineal también muestra que el *feedback* entregado es predictor significativo en el avance de las habilidades de pensamiento crítico (Tabla 2-3, $\overline{FD}$ ($\beta = 0.40$), (p -value = <0.01). El hecho de que la retroalimentación entregada tenga un alto impacto en el avance se explica por el hecho que los estudiantes comparan su propio escrito con el que observan y de esta manera incorporan este conocimiento (Huisman, et al, 2018). Esto se ejemplifica a través de *Student 1*, de otorgar otro punto de vista: *“ver otra respuesta te permite mejorar tu respuesta” o al revés, cuando te llegaba (una) muy mala respuesta para revisar, no te servía a ti, porque le faltaba todo ...no te ayuda a ti”*.

Concordamos con Newman et al. (2018) que para entregar *feedback* de calidad es necesario entregar directrices a los alumnos tales como, mirar más allá del caso propio, y no solo destacar los problemas. Retroalimentar a un compañero es una habilidad compleja que necesita ser aprendido bajo oportunidades de enseñanza (Sluijsmans et al. 2004, En Alqassab M et al 2017). Así, es relevante una adecuada instrucción previa para que los alumnos tengan más seguridad y habilidad entregar *feedback* (De la Cruz, et al, 2015). En el *focus group* queda en evidencia que la evaluación de pares no es una habilidad instalada en los futuros profesores y es necesaria de desarrollar, considerando que permite apoyar la colaboración entre los estudiantes (Alqassab, et al, 2018). Por ejemplo, *Student 4*: *“al principio yo hacía muy malas retroalimentaciones porque no sabía cómo hacerlo”*, *Student 5*: *“cuando comenzamos el curso yo no sabía cómo hacer una retroalimentación y tuve que aprenderlo, fue difícil al comienzo, pero después pude hacerlo”*, y *Student 1*: *“Fue mejorando con el tiempo”*. Asimismo, trabajar en actividad de evaluación de pares permite que los futuros profesores mejoren sus habilidades de evaluación (Sluijsmans et al., 2004 En Alqassab, et al, 2018).

Considerando los resultados de $\overline{FR}$ y $\overline{FD}$ en la regresión, Tabla 2-3, concluimos la relevancia de la construcción social del conocimiento donde los alumnos construyen su conocimiento a través de dar y recibir *feedback* (Lin, G, 2018). Esto también se refleja

en los comentarios *Student 1*: “la actividad se armaba entre todos”, o *Student 2*: “Facilita la construcción del conocimiento, ya que hay otro que te da consejos de cómo puedes mejorar tu respuesta”.

Finalmente, el hecho de que el Promedio de la variable Response ($\bar{R}$ en Tabla 2-3) tenga un $\beta = -0.33$, puede ser interpretado de forma que aquellos alumnos que tienen un estado inicial más alto, tienden a tener menos espacio para mejorar, pues están más cerca del nivel más alto (Yang et al, 2013).

2.4.3. Modelo de andamiaje para profesores en formación

Es sabido que para que las evaluaciones de pares sean efectiva, los estudiantes requieren entrenamiento en el modelo específico (Hanrahan & Isaacs, 2001). A partir de un análisis de las respuestas de los estudiantes que alcanzaron el nivel más alto, correspondiente al nivel logrado, se creó un modelo de andamiaje para profesores en formación, que considera directrices para desarrollar las habilidades de pensamiento crítico que permite a los estudiantes estructurar sus razonamientos y enriquecer sus reflexiones pedagógicas.

El propósito de este modelo es a) orientar al docente formador de profesores sobre cómo trabajar indicadores concretos y progresivos que permitan entregar herramientas al estudiante para construir su respuesta y que permitan evaluar además la complejidad de la tarea (las fases y contenido a trabajar orientan al formador de profesores sobre las metas de aprendizaje que espera que los estudiantes logren en el tiempo asignado); b) reconocer cuáles son las etapas o procesos comunes a desarrollar en todas las habilidades, es decir, y como muestra la Figura 2-2, existen fases discursivas que son comunes a todas las habilidades, por lo que deberán ser dominadas por los estudiantes para lograr un adecuado desarrollo de la habilidad correspondiente (en este sentido, el modelo propuesto permite al estudiante generar una estructura previa de respuesta que permita generalizar ciertas fases; por ejemplo, todas las habilidades suponen una fase de

situación y de *instancia pedagógica*, por lo que estos elementos deberán ser trabajados y guiados previamente en las discusiones del aula por el docente); c) diferenciar cuáles son tareas específicas de cada habilidad de pensamiento crítico (lo que potencia la guía de escritura de la respuesta y la orientación que debe tener el *feedback* realizado. Sobre esto último, si se trata, por ejemplo, de una actividad que trabaje la explicación, la fase de “conceptualización”, por ser exclusiva de esta habilidad, será sumamente necesaria para el desarrollo adecuado de la actividad).

La Figura 2-2 sintetiza las fases o etapas de escritura que debiese tener una respuesta a una actividad que trabaje las habilidades correspondientes, se indica en las columnas las habilidades de pensamiento crítico trabajadas. En las filas se detallan las fases correspondientes. Tanto las fases de situación como la de argumentación e instancia pedagógica son comunes a todas las habilidades (por eso se extienden en el eje horizontal), mientras que las de metodología general (correspondiente a las habilidades de interpretación e inferencia) y puesta en relación (correspondiente a las habilidades de explicación, análisis y evaluación) se extienden a habilidades específicas. Otras fases, en tanto, son exclusivas de ciertas habilidades, pues corresponden a los elementos, actividades o sub-habilidades representativos de la habilidad de pensamiento crítico correspondiente (basado en el modelo presentado por Facione (1990)).

HABILIDADES DE PENSAMIENTO CRÍTICO					
	Interpretación (Comprender y expresar el significado de una gran variedad de experiencias, situaciones, datos, eventos, juicios, convenciones, creencias, reglas, procedimientos o criterios)	Inferencia (Identificar los elementos necesarios para sacar conclusiones razonables; para formar conjeturas e hipótesis; para considerar la información relevante y para extraer las consecuencias que fluyen de datos, declaraciones, principios, pruebas, juicios, creencias, opiniones, conceptos, descripciones, preguntas u otras formas de representación)	Explicación (Declarar los resultados del razonamiento en función de las consideraciones probatorias, conceptuales, metodológicas, de criterios y contextuales sobre los que se basaron los resultados; y presentar su razonamiento en forma de argumentos convincentes)	Análisis (Identificar la relación, tanto intencionada como real, entre enunciados, preguntas conceptos, descripciones o formas de representación que expresan creencias, juicios, experiencias, razones, información y opiniones)	Evaluación (Evaluar la credibilidad de planteamientos u otras representaciones que describen las percepciones, experiencias, situaciones, juicios, creencias de una persona y evaluar la coherencia lógica de las relaciones inferenciales entre planteamientos, descripciones, cuestiones u otras formas de representación)
MODELO DE FASES DISCURSIVAS	Situación Presentación o descripción del caso, consideraciones bibliográficas, antecedentes contextuales, conceptos o datos que permitan comprender los antecedentes o marco conceptual que sustentan la propuesta. Se espera que el estudiante extraiga estos elementos de manera explícita (es decir, que los reproduzca) de: la lectura bibliográfica exigida para desarrollar la actividad, la discusión de clases previa al desarrollo de la actividad, el caso o situación presentados en la actividad, entre otros. ***En las actividades que trabajen la habilidad de evaluación y que supongan la elección por parte del estudiante de una propuesta particular, en la fase de situación se debe mencionar cuál es la opción escogida.				
	Argumentación Justificación y explicación de las aseveraciones, hipótesis, metodologías, relaciones o conceptualizaciones propuestas en el desarrollo de la actividad. Se espera que la argumentación de la propuesta desarrollada por el estudiante esté sustentada en evidencia concreta, explícita y atinente al desarrollo de la actividad.				
	Instancia pedagógica Presentación de una posible resolución a una problemática desde una instancia pedagógica práctica. Se espera que el estudiante proponga una situación, diferente a las revisadas en el contexto de la actividad, en la que se detalle una estrategia a seguir por el docente para resolver un problema de aprendizaje o enseñanza en aula.				
	Metodología general Detalle de estrategias o lineamientos generales del método a emplear para acceder de manera inicial a la solución de la problemática presentada en la actividad. Estos elementos suponen una generalización de conocimientos por parte del estudiante, quien propone una metodología atinente al problema propuesto.		Puesta en relación Relación de un concepto disciplinar o curricular (presente en los contenidos del syllabus) atinente al desarrollo de la actividad con un contenido didáctico-pedagógico (relacionado con aspectos más bien didácticos de enseñanza o aprendizaje del concepto disciplinar correspondiente) que permita la comprensión del concepto en una instancia de aula.		
	Ejemplificación Presentación de un ejemplo disciplinar pedagógico que clarifique la metodología propuesta.	Hipótesis Presentación de una conjetura o suposición del estudiante a partir de la situación presentada en la actividad. Esta suposición es hecha a partir de datos extraídos del contenido de la actividad (pregunta, materia de clases, lecturas bibliográficas, entre otros) que sirven de base para iniciar una investigación o una argumentación.	Conceptualización Presentación y descripción de los conceptos disciplinares presentes en un caso pedagógico. Estos son propuestos por el estudiante a partir de la generalización de sus conocimientos previos.	Evaluación Explicitación y desarrollo de criterios evaluativos que permitan evaluar relaciones, elecciones y explicaciones a una problemática particular. Se espera que la evaluación desarrolle proposiciones atinentes al juicio o propuesta presentados. Los criterios de atinencia dependerán de la actividad correspondiente. ***Esta fase se encuentra presente en las actividades de análisis únicamente en actividades que supongan también evaluación.	

Figura 2-2 Modelo de fases discursivas en actividades que desarrollan las habilidades del pensamiento crítico de interpretación, análisis, evaluación, inferencia, y explicación

En línea con lo anterior, un docente en formación requiere no solo dominar la dimensión conceptual-disciplinar a nivel de contenido, sino también requiere desarrollar las habilidades de pensamiento crítico que le permitan por un lado ser pensador crítico y así potenciar sus prácticas pedagógicas, las cuales permitan enseñar a ser pensadores críticos a sus alumnos. (Liu, et al, 2018)

A su vez, para el desarrollo del pensamiento crítico en profesores en formación se requiere de diversas estrategias en la instrucción, por ejemplo, discusiones (Liu, et al, 2018), desarrollo de preguntas con uso de habilidades de orden superior y diversas tareas que incorporen el desarrollo del pensamiento crítico (Barnett, et al, 2012), y así también estrategias específicas de escritura que orienten al estudiante en esta dirección. Además, una instrucción con una estructura previa, es decir, el conocimiento de fases prototípicas previo a la escritura de la respuesta permitirá por un lado orientar al estudiante en la construcción, desarrollo y revisión de su respuesta lo que generará un aprendizaje significativo (Angeli y Valanides, 2009; Mayer, 2004 En Barnett, et al, 2012), y a su vez los estudiantes doten de mayor sentido el uso de las herramientas tecnológicas en el aula en un contexto de aprendizaje.

En Apéndice J, incluimos un ejemplo de cómo utilizar el Modelo de Figura 2-2, para actividades que desarrollan las habilidades del pensamiento crítico de interpretación, análisis, evaluación, inferencia, y explicación. Se detallan en este ejemplo, instancias de respuestas específicas y reales de los estudiantes, indicando las fases presentes en los textos.

2.4.4. Uso de tecnología

Al desarrollar la actividad de construcción de respuesta y evaluación de pares, a través de dispositivos en red (Tablet, or Laptop) se minimizó el esfuerzo logístico pues la transferencia de documentos era transparente tanto para los alumnos como para el profesor. Además, la plataforma permitía al profesor monitorear el trabajo en desarrollo

por los alumnos, que le permitía dar *feedback* al trabajo que estaban realizando, pudiendo proyectar en una pantalla el trabajo de un alumno en particular para analizarlo frente el curso. Por otro lado, al administrarse en forma automática todo el flujo de información por la red, era trivial mantener el anonimato de las respuestas de los alumnos evitando el sesgo social (Lin, 2018, Berndt et al., 2017). Asimismo, el uso de la tecnología permite mejores espacios de aprendizaje, que permite el desarrollo de tareas que antes eran inconcebibles. (Keane et al., 2016). Específicamente en nuestro caso, administrar el proceso de *peer feedback* en forma instantánea y anónima.

Aunque la plataforma utilizada fue concebida con una interfaz sencilla y fácil de usar, para una pronta aceptación y un fácil uso por parte de los estudiantes (Yilmaz, et al, 2018), los alumnos tuvieron algunas dificultades en su proceso de inducción. Esto se ilustra en el comentario de *Student 3*: “*me costaba al principio, conectarme a internet, abrir la página (...) no sabía si había enviado mi trabajo*”, o *Student 1*: “*cuando empezamos no entendía como tenía que evaluar a mi compañero, me costaba comprender si había escrito todo bien en la plataforma*”.

Desde el estudio cualitativo se observa la necesidad de los alumnos de contar con un bagaje tecnológico previo y mayores experiencias de aprendizaje con uso de tecnología, que permitan una apropiación de las herramientas tecnológicas. El hecho de que en la formación docente el uso de tecnología está insuficientemente presente (Cullen, et al, 2018), se observa con el comentario de *Student 2*: “*(...) me demoraba porque no estaba acostumbrada a trabajar con computador en las clases*”. Sin autoeficacia tecnológica, entendida como la habilidad para utilizar herramientas tecnológicas básicas, es difícil lograr los objetivos pedagógicos involucrados (Yilmaz, et al 2018).

2.5. Conclusiones

Esta investigación tuvo como objetivo responder a la pregunta sobre ¿cómo actividades con un modelo de andamiaje para la revisión de pares pueden desarrollar pensamiento crítico en formación de profesores? Los resultados muestran que en formación inicial

docente, la evaluación de pares basada en *feedback* anónimo a través del modelo de andamiaje facilitado por la tecnología, permite un avance significativo en el desarrollo de las habilidades de pensamiento crítico, y que tanto el *feedback* recibido como el *feedback* entregado son predictores de este avance. Esto confirma la relevancia de diseñar, en formación inicial docente, situaciones de aprendizaje que promuevan el desarrollo de pensamiento crítico (Arsal, 2015).

Considerando que la evaluación de pares *paper based* tiene un alta carga administrativa de trabajo con un gran número de estudiante (Hanrahan & Isaacs, 2001), la tecnología, a través de dispositivos interconectados de forma inalámbrica, permiten un manejo fluido y tranquilo para cursos de cualquier tamaño, siendo la única restricción el soporte de la red (*router* y ancho de banda). Este manejo transparente, asegura un proceso de revisión a ciegas, pero es necesario considerar que los estudiantes perciben que un estudiante altamente competente entrega un mejor *feedback* (Berndt & et al, 2018). Además, la tecnología permite al docente tener acceso al trabajo de cada uno de los estudiantes durante la clase, facilitando a éste guiar el proceso de retroalimentación, aclarar problemas de comprensión del contenido y estimular el compromiso de los estudiantes (Zhu & Carless, 2018).

Las limitaciones de este estudio refieren a un grupo pequeño de actividades, en este caso 3 actividades, la cantidad de alumnos que participan en cada actividad y las habilidades de un dominio específico. Para trabajo futuro se propone estudiar este modelo con mayor profundidad, realizar un mayor número de actividades, durante un período de tiempo más extenso, y en diferentes temáticas.

A. Detalle de actividades realizadas y habilidades de pensamiento crítico involucradas

ACTIVIDAD 1	
Analizar la secuencia de aprendizaje del texto “La democracia” a partir de las características clave para la enseñanza de la fluidez. Haz un cuadro resumen de los aspectos que están presentes y aquellos ausentes. No te olvides de fundamentar con evidencias. Da sugerencias para mejorar esta secuencia de aprendizaje. Ten presente la conexión con el modelo de transferencia progresiva de la responsabilidad.	
Tareas a realizar por el estudiante	1. Análisis de secuencia de aprendizaje a partir de un contenido disciplinar (características de la enseñanza de la fluidez). 2. Construcción de un cuadro resumen (elementos presentes y ausentes). 3. Fundamentación de los aspectos señalados con evidencias. 4. Propuesta de sugerencias. 5. Consideración de un contenido disciplinar previo (transferencia progresiva de la responsabilidad).
Contenido disciplinar	Características clave para la enseñanza de la fluidez
Habilidades de pensamiento crítico involucradas en el desarrollo de la actividad	Análisis y explicación
ACTIVIDAD 2	
Estás preparando tu discusión productiva para la comprensión y está evaluando cómo segmentar la información. Tienes estas tres propuestas para segmentar la información: Propuesta 1: Texto visual – título y párrafo inicial – apartado 1 – apartado 2 – apartado 3. Propuesta 2: Apartados 1, 2 y 3 --- texto visual – título y párrafo inicial Propuesta 3: Título y párrafo inicial – texto visual – apartado 3 – apartado 1 – apartado 2 ¿Qué propuesta escogerías? ¿Por qué? Luego elabora una pregunta para gatillar la discusión para cada segmento. Recuerda usar la tabla del plan de discusión. Pon por ahora el título o idea clave del segmento en la columna texto y completa la última columna con la pregunta específica para ese segmento. Recuerda elaborar una pregunta clave y no múltiples preguntas. No te olvides de que las preguntas deben apuntar a las ideas centrales del texto (foco) y promover la discusión (preguntas abiertas).	
Tareas a realizar por el estudiante	1. Evaluación de propuestas de segmentación de la información. 2. Elección de una opción y justificación de la decisión. 3. Elaboración y propuesta de una pregunta clave para discusión posterior. 4. Construcción e inferencia de un título e idea clave.
Contenido disciplinar	Producción textual y discusión productiva: segmentación y presentación de la información.
Habilidades de pensamiento crítico involucradas en el desarrollo de la actividad	Evaluación e inferencia
ACTIVIDAD 3	
Nagy (1988) planteó que un niño cada año debería aprender entre 1000-3000 palabras, Villalón (2008) plantea que deberían ser entre 2700 y 3000 palabras por año, ¿cómo enseñarías el vocabulario considerando que el año escolar tiene aproximadamente 36 semanas lo que quiere decir que los niños deberían aprender semanalmente aproximadamente 83 palabras?	
Tareas a realizar por el estudiante	1. Lectura y comparación de dos textos disciplinares. 2. Interpretación de variables e información de un problema pedagógico específico. 3. Identificación e inferencia de contenidos y herramientas conceptuales. 4. Elaboración de una propuesta y explicación.
Contenido disciplinar	Estrategias pedagógicas de enseñanza de vocabulario.
Habilidades de pensamiento crítico involucradas en el desarrollo de la actividad	Interpretación, inferencia y explicación.

Habilidad	Definición de la habilidad	Dimensión de la rúbrica
Interpretación	Comprender y expresar el significado de una gran variedad de experiencias, situaciones, datos, eventos, juicios, convenciones, creencias, reglas, procedimientos o criterios. Se consideran sub-habilidades de interpretación las siguientes actividades: categorizar, describir, distinguir, reconocer, identificar, ejemplificar, entre otras.	Los descriptores de esta dimensión evalúan si la respuesta la respuesta aborda de manera clara y justificada la categorización, decodificación o aclaración del significado o sentido consignado en la pregunta. Además, evalúan si la respuesta es consecuente con el formato de presentación de la información (texto, esquema, tabla, mapa conceptual, según corresponda).
Análisis	Identificar tanto la intencionada como la real relación entre enunciados, preguntas conceptos, descripciones o formas de representación que expresan creencias, juicios, experiencias, razones, información y opiniones. Se consideran sub-habilidades de análisis las siguientes actividades: relacionar, contrastar, examinar, proponer, entre otras.	Los descriptores de esta dimensión evalúan si la respuesta propone un análisis exhaustivo que reconozca las relaciones entre los enunciados, conceptos preguntas (u otros) supuestas en la pregunta. Además, evalúan la claridad, justificación y coherencia de estas relaciones, como también su sostén argumentativo (entrega de evidencia).
Evaluación	Evaluar la credibilidad de planteamientos u otras representaciones que describen las percepciones, experiencias, situaciones, juicios, creencias de una persona y evaluar la coherencia lógica de las relaciones inferenciales entre planteamientos, descripciones, cuestiones u otras formas de representación. Se consideran sub-habilidades de evaluación las siguientes actividades: justificar, argumentar, proponer, juzgar entre otras.	Los descriptores de esta dimensión evalúan si la respuesta presenta una evaluación de los planteamientos con criterios explícitos, claros, atingentes y creíbles y definidos a partir de la pregunta o respuesta evaluada. Además, evalúan la coherencia del planteamiento a través del el reconocimiento de temas (ideas centrales), intenciones (intereses o motivaciones) o representaciones (principales propuestas y visión particular).
Inferencia	Identificar los elementos necesarios para sacar conclusiones razonables; para formar conjeturas e hipótesis; para considerar la información relevante y para extraer las consecuencias que fluyen de datos, declaraciones, principios, pruebas, juicios, creencias, opiniones, conceptos, descripciones, preguntas u otras formas de representación. Se consideran sub-habilidades de inferencia las siguientes actividades: reconocer, reformular, extraer, parafrasear, explicar, entre otras.	Los descriptores de esta dimensión evalúan si la respuesta identifica y desarrolla los elementos conceptuales (inferidos del planteamiento de la pregunta) y las evidencias atingentes para sacar conclusiones o elaborar una propuesta para el planteamiento propuesto en la pregunta (conjetura, hipótesis). Además, evalúan la proposición de ideas principales y más relevantes supuestas, reconociendo causas y consecuencias de los datos analizados.
Explicación	Declarar los resultados del razonamiento en función de las consideraciones probatorias, conceptuales, metodológicas, de criterios y contextuales sobre los que se basaron los resultados; y presentar su razonamiento en forma de argumentos convincentes. Se consideran sub-habilidades de explicación las siguientes actividades: reportar, articular, anticipar, proveer, entre otras.	Los descriptores de esta dimensión evalúan si la respuesta presenta de manera detallada los resultados, procedimientos o argumentos requeridos en la pregunta. Cada ítem, según corresponda, se justifica con evidencia convincente y correctamente trabajada. Además, evalúan el detalle de los procedimientos, métodos, criterios, conceptos disciplinares, línea argumentativa o consideraciones relevantes para el razonamiento.

C-1. Rúbrica de R y NR

INDICADORES DE EVALUACIÓN	LOGRADO 4	EN DESARROLLO 3	INICIAL 2	NO LOGRADO 1
DESARROLLO DE LA HABILIDAD DE PENSAMIENTO CRÍTICO (DESCRIPTORES GENERALES)	La respuesta del estudiante evidencia un desarrollo adecuado de la habilidad implicada en la pregunta propuesta. Existen marcas textuales (lingüísticas y discursivas) y evidencia explícitas que sustentan el desarrollo de la habilidad correspondiente.	Se evidencia la habilidad implicada en la pregunta. Falta desarrollo en alguno de los elementos implicados en la pregunta. Existen marcas textuales y evidencia explícitas que sustentan el desarrollo de la habilidad correspondiente.	En la respuesta se evidencia parcialmente la habilidad implicada en la pregunta. Más de dos elementos implicados en la pregunta no se desarrollan a cabalidad. Existen marcas textuales y evidencias, pero no todas se presentan de manera explícita.	No se evidencia en la respuesta la habilidad implicada en la pregunta propuesta. Los elementos implicados en la pregunta no se desarrollan en la respuesta y/o la respuesta supone el desarrollo de otra habilidad.
a) Si se trata de una pregunta que implique INTERPRETACIÓN	La respuesta aborda de manera clara y justificada la categorización, decodificación o aclaración del significado o sentido consignado en la pregunta. Además, la respuesta es consecuente con el formato de presentación de la información (texto, esquema, tabla, mapa conceptual, según corresponda).	La respuesta aborda de manera clara la categorización, decodificación o aclaración del significado o sentido consignado en la pregunta, pero una de las categorías o parte del desarrollo se encuentra parcialmente justificada y comentada. La respuesta es consecuente con el formato de presentación de la información.	La respuesta aborda parcialmente la categorización, decodificación o aclaración del significado o sentido consignado en la pregunta, pues falta una de sus partes. La respuesta es consecuente con el formato de presentación de la información.	La respuesta no aborda la categorización, decodificación o aclaración del significado o sentido consignado en la pregunta. y/o el formato de presentación de la información no es consecuente con la habilidad supuesta en la pregunta.
b) Si se trata de una pregunta que implique ANÁLISIS	La respuesta propone un análisis exhaustivo que reconoce las relaciones entre los enunciados, conceptos, preguntas (u otros) supuestas en la pregunta. Las relaciones establecidas son claras (están correctamente delimitadas y planteadas) y coherentes (responden al tema y sentido global de la pregunta) y se encuentran correctamente identificadas (se menciona cuáles son), argumentadas (se proveen argumentos y razones que las defienden o refutan, y que respondan a la defensa de una postura) y justificadas con evidencia en el caso, situación o texto consignado en la pregunta.	La respuesta propone un análisis que reconoce la mayoría de las relaciones entre enunciados, conceptos, preguntas (u otros) supuestas en la pregunta. Las relaciones establecidas son claras y coherentes y se encuentran parcialmente identificadas, argumentadas y justificadas porque: no se menciona una de las relaciones o uno de los elementos supuestos en la pregunta o se identifican todos los elementos, pero no se analizan todos) o no se provee evidencia suficiente que justifique el examen de ideas efectuado.	La respuesta propone un análisis que reconoce algunas de las relaciones entre enunciados, conceptos, preguntas (u otros) supuestas en la pregunta. Las relaciones establecidas son claras y coherentes y se encuentran parcialmente identificadas, argumentadas y justificadas porque: no se menciona una de las relaciones o uno de los elementos supuestos en la pregunta y no se provee evidencia suficiente que justifique el examen de ideas efectuado.	La respuesta propone un análisis que no reconoce las relaciones entre enunciados, conceptos, preguntas (u otros) supuestas en la pregunta. Las relaciones establecidas no siempre son coherentes y no apuntan al foco(s) central(es) de la pregunta. o se identifican los elementos, pero no se vinculan ni analizan. o no se provee ninguna evidencia que justifique el examen de las ideas efectuado.
c) Si se trata de una pregunta que implique EVALUACIÓN	La respuesta presenta una evaluación de los planteamientos presentados o supuestos en la pregunta. Se reconoce los temas (ideas centrales), intenciones (intereses o motivaciones) o representaciones (principales propuestas y visión particular)	La respuesta presenta una evaluación de los planteamientos presentados o supuestos en la pregunta. Se reconoce los temas, intenciones o representaciones de dichos planteamientos y	La respuesta presenta una evaluación parcial de los planteamientos presentados o supuestos en la pregunta. Se reconoce la mayoría de los temas, intenciones o representaciones de dichos planteamientos y se evalúa su credibilidad a partir de	La respuesta no presenta una evaluación de los planteamientos presentados o supuestos en la pregunta. Sólo se identifican los temas, intenciones y representaciones, pero no se evalúan o no se definen criterios para evaluar la credibilidad de los

	de dichos planteamientos y se evalúa su credibilidad a partir de criterios específicos, atinentes (coherentes con los objetivos de la pregunta) y correctamente definidos.	se evalúa su credibilidad a partir de criterios específicos, sin embargo, uno de los criterios no es atinente para la evaluación propuesta o no está correctamente definido y explicado.	únicamente un criterio de evaluación y/o una justificación parcial de los criterios de evaluación establecidos.	planteamientos. 74
d) Si se trata de una pregunta que implique INFERENCIA	La respuesta identifica y desarrolla los elementos conceptuales (inferidos del planteamiento de la pregunta) y las evidencias atinentes para sacar conclusiones o elaborar una propuesta para el planteamiento propuesto en la pregunta (conjetura, hipótesis). Se abordan las ideas principales y más relevantes supuestas, reconociendo causas y consecuencias de los datos analizados.	La respuesta identifica y desarrolla los elementos conceptuales y las evidencias atinentes para sacar conclusiones o elaborar una propuesta para el planteamiento propuesto en la pregunta (conjetura, hipótesis). Se abordan la mayoría de las ideas supuestas, pero no siempre son principales o las más relevantes. Se reconocen algunas causas y consecuencias de los datos analizados.	La respuesta identifica y desarrolla los elementos conceptuales supuestos en el planteamiento de la pregunta, pero no entrega evidencia atinente en todos los casos que justifique las conjeturas o conclusiones propuestas. Se abordan las ideas principales y más relevantes supuestas, reconociendo causas y consecuencias de los datos analizados. o se abordan la mayoría de las ideas supuestas, pero no siempre son principales o las más relevantes. Se reconocen algunas causas y consecuencias de los datos analizados.	La respuesta no identifica ni desarrolla los elementos conceptuales supuestos en el planteamiento de la pregunta. y/o no entrega evidencia en ningún caso que justifique las conjeturas o conclusiones propuestas. y/o no se abordan las ideas principales y más relevantes supuestas.
e) Si se trata de una pregunta que implique EXPLICACIÓN	La respuesta presenta de manera detallada los resultados, procedimientos o argumentos requeridos en la pregunta. Cada ítem, según corresponda, se justifica con evidencia convincente (razonamientos lógicos que no incurran en generalidades injustificadas, valoraciones sin respaldo o apelación a afectividades) y correctamente trabajada (se sigue una línea expositiva y argumentativa, según corresponda), detallando por tanto los procedimientos, métodos, criterios, conceptos disciplinares, línea argumentativa o consideraciones relevantes para el razonamiento.	La respuesta presenta de manera detallada los resultados, procedimientos o argumentos requeridos en la pregunta. Cada ítem, según corresponda, se justifica con evidencia convincente y correctamente trabajada. Generalmente se detallan los procedimientos, métodos, criterios, conceptos disciplinares, línea argumentativa, pero no siempre se presentan las consideraciones relevantes para el razonamiento.	La respuesta presenta los resultados, procedimientos o argumentos requeridos en la pregunta, pero no se justifica con evidencia en todos los casos	La respuesta presenta los resultados, procedimientos o argumentos requeridos en la pregunta, pero no se justifica con evidencia en ningún caso ni se detallan los aspectos involucrados en el <i>nivel logrado</i> .

C-2. Rúbrica de F

Indicadores obligatorios Estos indicadores son aplicables a todos los escritos de la fase de argumentación	LOGRADO 4	EN DESARROLLO 3	INICIAL 2	NO LOGRADO 1
Criterios de evaluación	La respuesta presenta una evaluación de los planteamientos presentados o supuestos en el desarrollo de R. Se reconoce los temas (ideas centrales), intenciones	La respuesta presenta una evaluación de los planteamientos presentados o supuestos en R. Se reconoce los temas, intenciones o representaciones de dichos	La respuesta presenta una evaluación parcial de los planteamientos presentados o supuestos en R. Se reconoce la mayoría de los temas, intenciones o representaciones de dichos	La respuesta no presenta una evaluación de los planteamientos presentados o supuestos en R. Sólo se identifican los temas, intenciones y

	(intereses o motivaciones) o representaciones (principales propuestas y visión particular) de dichos planteamientos y se evalúa su credibilidad a partir de criterios específicos (el retroalimentador no pretende abordar todo el desarrollo de R, sino más bien orientar puntos e ideas específicas), atinentes (coherentes con los objetivos de la pregunta) y correctamente definidos.	planteamientos y se evalúa su credibilidad a partir de criterios específicos, sin embargo, uno de los criterios no es atinente para la evaluación propuesta o no está correctamente definido y explicado.	planteamientos y se evalúa su credibilidad a partir de únicamente un criterio de evaluación y/o una justificación parcial de los criterios de evaluación establecidos	representaciones, pero no se evalúan o no se definen criterios para evaluar la credibilidad de los planteamientos.
Claridad y pertinencia de la explicación	El retroalimentador presenta de manera detallada los resultados, procedimientos o argumentos resultantes de la lectura y análisis de R (integra el desarrollo del compañero, estableciendo un diálogo entre las ideas del retroalimentador y el evaluado). Cada ítem, según corresponda, se justifica con evidencia presente en el desarrollo de R (razonamientos lógicos que no incurran en generalidades injustificadas, valoraciones personales sin respaldo o apelación a afectividades) y correctamente trabajada (se sigue una línea expositiva y argumentativa, según corresponda), detallando por tanto los procedimientos, métodos, criterios, conceptos disciplinares, línea argumentativa o consideraciones relevantes para el razonamiento.	La respuesta presenta de manera detallada los resultados, procedimientos o argumentos resultantes de la lectura y análisis de R. Cada ítem, según corresponda, se justifica con evidencia presente en el desarrollo de R y correctamente trabajada. Generalmente se detallan los procedimientos, métodos, criterios, conceptos disciplinares, línea argumentativa, pero no siempre se presentan las consideraciones relevantes para el razonamiento.	La respuesta presenta los resultados, procedimientos o argumentos resultantes de la lectura y análisis de R, pero no se justifica con evidencia presente en el desarrollo de R en todos los casos (en este nivel se encuentran las retroalimentaciones que no dialogan con R bien porque el retroalimentador presenta su punto de vista, sin considerar el de R, bien porque se utiliza evidencia del desarrollo personal del retroalimentador, y no del escrito de R).	La respuesta presenta los resultados, procedimientos o argumentos resultantes de la lectura y análisis de R, pero no se justifica con evidencia en ningún caso ni se detallan los aspectos involucrados en el <i>nivel logrado</i> .
Intención de la argumentación	El comentario del retroalimentador tiene por intención motivar y orientar un <i>hacer</i> (o <i>rehacer</i> , según corresponda). Para esto, presenta preguntas (que orienten al escritor de R sobre los cambios requeridos o sobre aquellos aspectos que debe atender para el desarrollo de R2), llamados de atención (si el retroalimentador lo considera pertinente, podrá presentar a su evaluado nuevas formas de presentar la información contenida en R), explicaciones (sobre las partes del texto en las que el escritor de R deberá enfocarse más) o precauciones (si el retroalimentador considera que algunos de los aspectos señalados en R puede mal entenderse o interpretarse de manera ambigua). La carga de la argumentación es positiva.	El comentario del retroalimentador tiene por intención motivar y orientar un <i>hacer</i> o <i>rehacer</i> . Presenta preguntas, llamados de atención, explicaciones y/o precauciones que orientan al escritor de R sobre los puntos clave de la argumentación. La mayoría de los comentarios se encuentran justificados y/o proveen una estrategia para abordarlos en la escritura de R2. La carga de la argumentación es positiva.	El comentario del retroalimentador tiene por intención describir los aspectos logrados y no logrados de R. Para esto, señala preguntas, llamados de atención, explicaciones y/o precauciones que orientan al escritor de R sobre los puntos clave de la argumentación, mas no indica estrategias o justificaciones que orienten la escritura de R2. La carga de la argumentación es positiva.	El comentario del retroalimentador tiene por intención <i>corregir</i> el escrito de R. Los comentarios presentan los “errores de R” y no motivan ni inspiran nuevas reflexiones o nuevas interrogantes que orienten y ayuden al escritor de R. La carga de la argumentación es negativa.
Formulación de la retroalimentación	El léxico empleado es claro (no se presentan expresiones complejas y las palabras y las	El léxico empleado es claro, pero no es preciso en todos los casos porque:	El léxico empleado es parcialmente claro y preciso porque:	El léxico empleado no es claro ni preciso (no se cumple ninguno de

	construcciones sintácticas son entendibles) y preciso (no se utilizan unidades léxicas ambiguas, coloquiales o expresiones de la oralidad). Además, responde al dominio disciplinar correspondiente (se emplea un registro académico y/o tecnolecto cuando se trata de definiciones, descripciones u otras explicaciones relevantes para comprender los contenidos involucrados en la evaluación).	algunos comentarios presentan ambigüedades en el uso de ciertas palabras (uso de palabras generales como “cosa”, “algo”, “eso”, entre otras). y/o se evidencian algunas expresiones características de la oralidad que no responden al registro correspondiente y/o representan expresiones coloquiales e imprecisas de la oralidad (“bueno”, “o sea”, “es como cuando”).	todos los comentarios presentan variadas ambigüedades en el uso de ciertas palabras. y/o se evidencian variadas expresiones características de la oralidad que afectan el registro empleado.	los indicadores del nivel logrado).
Indicadores opcionales ¹	LOGRADO 4	EN DESARROLLO 3	INICIAL 2	NO LOGRADO 1
Solicitud de entrega de evidencia	El retroalimentador llama la atención sobre la necesidad de incluir evidencia o respaldar alguna de las ideas presentadas en R. Esta solicitud de evidencia no sólo reconoce qué se debe respaldar, sino que además indica el porqué es necesario justificar o sustentar la idea señalada (se señala el objetivo de la entrega de evidencia, es decir, el retroalimentador deberá indicar qué se debe sustentar (qué idea o segmento de R) y por qué es importante respaldarlo (además de identificar el qué se sugiere orientar al escritor de R con razones lógicas y coherentes con el foco de la pregunta). El retroalimentador, además, propone tipos de evidencia que resulten significativas para las ideas señaladas (las evidencias pueden ser: <i>bibliográficas</i> (si al retroalimentador le parece pertinente puede proponer a su compañero la consideración de una idea presentada en las fuentes bibliográficas implicadas en la pregunta), <i>situacionales</i> (se puede sugerir al compañero proveer evidencia presente en la situación, caso o contexto presentado en la pregunta) o <i>empíricas</i> (si se busca respaldar argumentos o ideas con datos, porcentajes, experimentos).	El retroalimentador llama la atención sobre la necesidad de incluir evidencia o respaldar alguna de las ideas presentadas en R. Esta solicitud de evidencia no sólo reconoce qué se debe respaldar, sino que además indica el porqué es necesario justificar o sustentar la idea señalada. No obstante, no provee mayor información sobre el tipo o contenido de dicha evidencia.	El retroalimentador llama la atención sobre la necesidad de incluir evidencia o respaldar alguna de las ideas presentadas en R. Sin embargo, sólo indica qué se debe respaldar, pero no señala la razón ni orienta al escritor de R sobre el tipo o contenido de dicha evidencia.	El retroalimentador únicamente llama la atención sobre la necesidad de incluir evidencia, pero no indica qué o por qué debe ser respaldada la información.

¹ La consideración de estos criterios dependerá del desarrollo y nivel de desempeño de R (el argumentador deberá evaluar la pertinencia de estos indicadores para el escrito de su compañero, es decir, deberá evaluar si el escrito de su compañero requiere que se comente alguno de estos aspectos).

Relectura de un argumento, presentación de un nuevo argumento y/o contraargumento	El retroalimentador propone nuevos argumentos o contraargumentos que responden a la secuencia argumentativa de R (son coherentes y lógicos con el contenido y desarrollo de R). Además, sugiere al escritor de R líneas de refutación (si se trata de la presentación de un contraargumento) y/o estrategias de integración del nuevo argumento, justificando además la relevancia de incluir esta nueva información).	El retroalimentador propone una nueva lectura de un argumento ya presente en R. El retroalimentador indica por qué es importante considerar esta nueva lectura. Tanto la sugerencia como su importancia se encuentran justificadas y presentan estrategias de integración de la nueva información.	El retroalimentador propone una nueva lectura de un argumento ya presente en R, no obstante, no indica por qué es importante considerar esta nueva lectura. y/o no se presentan estrategias de integración de la nueva información.	El retroalimentador realiza comentarios, pero estos no responden al desarrollo de R porque: presenta sus propias conclusiones o argumentos sin considerar el contenido del texto R. En este tipo de argumentaciones el retroalimentador presenta, generalmente, una postura contraria que no dialoga con R (se presentan argumentos del tipo “no es como tú lo mencionas, sino que es esto”).
Solicitud de reformulación o rehacer	El retroalimentador refiere a la necesidad de reorganizar el texto (refiere a aspectos formales y discursivos como la organización de párrafos, la inclusión de marcadores discursivos, revisión de aspectos tipográficos, entre otros²), pues esto ayudaría a mejorar la comprensión del mismo (el efecto de sentido al reformular sería más claro y entendible) y provee además pautas y estrategias específicas y que potencien o bien la transformación del formato o bien el flujo de la información (estas estrategias están justificadas más allá del “porque sí” o evaluaciones personales y presentan un objetivo claro y atingente al desarrollo de R). Todos los elementos señalados se encuentran correctamente descritos (se señalan los pasos, las sugerencias concretas, los procedimientos, las búsquedas requeridas, entre otros y éstos son atingentes al desarrollo de R).	El retroalimentador refiere a la necesidad de reorganizar el texto, pues esto ayudaría a potenciar la comprensión del mismo. Provee además pautas y estrategias justificadas, sin embargo, no todos los elementos se encuentran correctamente descritos (algunas estrategias sólo se mencionan de manera general y/o algunas estrategias se describen de manera confusa y/o las pautas y estrategias presentadas no son atingentes al desarrollo de R).	El retroalimentador solo refiere a la necesidad de reorganizar el texto, pero no explica cómo esto ayudaría a potenciar la comprensión del mismo y/o no provee pautas o estrategias para potenciar el desarrollo de la reformulación y el rehacer y/o provee pautas o estrategias, pero estas no están correctamente delimitadas y justificadas. y/o el retroalimentador sugiere mejoras o correcciones sobre el formato de presentación de la información.	El retroalimentador señala la reformulación en términos de “errores” y explica al escritor de R cómo corregirlos (o presenta una propuesta que asume como correcta).
Llamado de atención a la pregunta	El retroalimentador llama la atención sobre algún aspecto señalado en la pregunta que no fue considerado o abordado en R (indicar un contenido o tema faltante en R, presente en la pregunta, por ejemplo, la pregunta supone analizar un caso y además dar sugerencias, y en R sólo se	El retroalimentador llama la atención sobre algún aspecto señalado en la pregunta que no fue considerado o abordado en R. Si bien se explica cuál es el punto que falta por trabajar, la orientación al escritor de R es general y no presenta pautas o procedimientos	El retroalimentador sólo llama la atención a la necesidad de volver a revisar la pregunta, indicado qué aspecto o contenido de la pregunta debe ser reconsiderado (o integrado, en caso de ausencia). No se integra una orientación al escritor de R y tampoco se	El retroalimentador únicamente señala que se debe volver a revisar la pregunta, pero no indica qué ni por qué es necesario integrar esa información. y/o solo se señala la existencia de una

² En este nivel se encuentran comentarios que respondan a sugerencias sobre aspectos formales como los detallados en la parte 2 de la rúbrica.

	analiza). y/o ante una confusión de las instrucciones, el retroalimentador explica de otra forma la información de la pregunta (distinta a la presentada en la pregunta) el objetivo y contenidos implicados en la pregunta (indicar una habilidad u objetivo distinto de la pregunta (por ejemplo, la pregunta supone interpretar, y R sólo identifica o la pregunta supone crear y R sólo describe)). Además de señalarlo, orienta al escritor de R sobre su integración (la orientación presenta sugerencias, estrategias, pautas y/o procedimientos claros, precisos y bien definidos).	específicos. y/o ante una confusión de las instrucciones, el retroalimentador señala qué aspecto de la pregunta no fue comprendido a cabalidad, pero su explicación o reformulación no señala de manera específica los objetivos y contenidos (los señala parcialmente y/o de manera general y/o ambigua). y/o la orientación al escritor es parcial, pues: las sugerencias, estrategias, pautas y/o procedimientos no son lo suficientemente claros, precisos y bien definidos (ya sea porque se presentan de manera general o inespecífica o porque no detallan correctamente los pasos y procedimientos para aplicar las estrategias correspondientes).	presenta pautas o estrategias. y/o se señala la existencia de una confusión en la comprensión de las instrucciones de la pregunta, pero no se indica el objetivo de la misma ni los contenidos necesarios para comprender los requerimientos de la pregunta. y/o no se orienta al escritor de R con sugerencias, estrategias, pautas y/o procedimientos	confusión en las instrucciones, pero no se orienta ni reformula la explicación para la comprensión del escritor de R (se señala la existencia de un error y confusión de manera general, sin definir cuál es ni cómo sortearlo).
--	---	---	--	---

78

C-3. Rúbrica de integración de F en NR

Indicador	LOGRADO 4	EN DESARROLLO 3	INICIAL 2	NO LOGRADO 1
Integración de F en NR	El estudiante reescribe su R, considerando los comentarios presentes en la retroalimentación y, además, información nueva que surge tanto de su evaluación al escrito de su compañero como del autoexamen a su propio desempeño. Presenta en NR una reelaboración de R que generaliza los elementos relevantes analizados por el estudiante (es decir, no copia ni pega, sino reconstruye a partir de los resultados de su análisis y evaluación de los recursos correspondientes).	El estudiante evalúa la pertinencia de los comentarios del evaluador e integra a su escrito las partes de la retroalimentación que le parecen relevantes para su respuesta. Presenta en NR su respuesta original más la integración de los comentarios del retroalimentador (es decir, no copia y pega los comentarios, sino que los reelabora, añadiendo la nueva información de manera coherente a su respuesta original).	El estudiante reproduce textualmente los comentarios del argumentador (es decir, no integra ni elabora, solo copia y pega). Presenta en NR su respuesta original más los comentarios del argumentador (es decir, presenta la respuesta antigua con la nueva información).	El estudiante reproduce textualmente los comentarios del argumentador (es decir, no integra ni elabora, solo copia y pega). Presenta en NR únicamente la nueva información solicitada (es decir, no presenta la reelaboración de su R).

D. Cohen Kappa de instrumento de evaluación.

Tabla D1: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR).
Dimensión: Habilidad de Análisis.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	3	1	0	0	4
	Nivel Inicial	0	3	3	0	6
	Nivel En Desarrollo	0	4	3	1	8
	Nivel Logrado	0	0	1	9	10
Total Columnas		3	8	7	10	28

Tabla D2: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR).
Dimensión: Habilidad de Explicación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	4	0	0	0	4
	Nivel Inicial	0	37	1	0	38
	Nivel En Desarrollo	0	2	22	2	26
	Nivel Logrado	0	0	1	23	24
Total Columnas		4	39	24	25	92

Tabla D3: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR).
Dimensión: Habilidad de Interpretación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	3	0	0	0	3
	Nivel Inicial	0	26	0	0	26
	Nivel En Desarrollo	0	5	17	0	22
	Nivel Logrado	0	0	0	13	13
Total Columnas		3	31	17	13	64

Tabla D4: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR).
Dimensión: Habilidad de Evaluación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	7	0	0	0	7
	Nivel Inicial	0	12	0	0	12
	Nivel En Desarrollo	0	0	8	2	10
	Nivel Logrado	0	0	0	25	25
Total Columnas		7	12	8	27	54

Tabla D5: Matriz de nivel de acuerdo entre correctores en Rúbrica de Response (R) y New Response (NR).
Dimensión: Habilidad de Inferencia.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	4	2	0	0	6
	Nivel Inicial	0	54	2	0	56
	Nivel En Desarrollo	0	1	37	0	38
	Nivel Logrado	0	0	0	18	18
Total Columnas		4	57	39	18	118

Tabla D6: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Criterios de Evaluación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	4	0	0	0	4
	Nivel Inicial	0	2	0	0	2
	Nivel En Desarrollo	0	0	4	0	4
	Nivel Logrado	0	0	0	15	15
Total Columnas		4	2	4	15	25

Tabla D7: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Claridad y Pertinencia de la Explicación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	0	0	0	0	0
	Nivel Inicial	0	4	1	0	5
	Nivel En Desarrollo	0	0	9	0	9
	Nivel Logrado	0	0	0	11	11
Total Columnas		0	4	10	11	25

Tabla D8: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Intención de la Argumentación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	0	0	0	0	0
	Nivel Inicial	0	14	0	0	14
	Nivel En Desarrollo	0	0	4	0	4
	Nivel Logrado	0	0	0	7	7
Total Columnas		0	14	4	7	25

Tabla D9: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Formulación de la Retroalimentación.

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	0	0	0	0	0
	Nivel Inicial	0	1	0	0	1
	Nivel En Desarrollo	0	2	8	0	10
	Nivel Logrado	0	0	0	16	16
Total Columnas		0	3	8	16	27

Tabla D10: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Solicitud de Evidencia.

	Corrector 1					Total Filas
	Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	Habilidad No Presente	
Corrector 2	Nivel No Logrado	0	0	0	0	0
	Nivel Inicial	0	3	0	1	4
	Nivel En Desarrollo	0	0	4	0	4
	Nivel Logrado	0	1	0	6	7
	Habilidad No Presente	0	0	0	10	10
Total Columnas		0	4	4	7	10

Tabla D11: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).Dimensión: Relectura de un Argumento o Presentación de un Nuevo Argumento y/o Contraargumento.

	Corrector 1					Total Filas
	Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	Habilidad No Presente	
Corrector 2	Nivel No Logrado	0	0	0	0	0
	Nivel Inicial	0	0	0	0	0
	Nivel En Desarrollo	0	0	1	0	1
	Nivel Logrado	0	0	0	0	0
	Habilidad No Presente	0	0	0	1	23
Total Columnas		0	0	1	1	23

Tabla D12: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Solicitud de Reformulación.

	Corrector 1					Total Filas
	Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	Habilidad No Presente	
Corrector 2	Nivel No Logrado	1	0	0	0	1
	Nivel Inicial	0	1	0	0	1
	Nivel En Desarrollo	0	0	1	0	2
	Nivel Logrado	0	0	0	0	0
	Habilidad No Presente	0	1	1	0	19
Total Columnas		1	2	2	0	20

Tabla D13: Matriz de nivel de acuerdo entre correctores en Rúbrica de Feedback (F).
Dimensión: Llamado de Atención.

		Corrector 1					Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	Habilidad No Presente	
Corrector 2	Nivel No Logrado	0	0	0	0	0	0
	Nivel Inicial	0	1	0	0	0	1
	Nivel En Desarrollo	0	0	3	0	1	4
	Nivel Logrado	0	0	1	2	0	3
	Habilidad No Presente	0	1	1	0	15	17
Total Columnas		0	2	5	2	16	25

Tabla D14: Matriz de nivel de acuerdo entre correctores en Rúbrica de integración de Feedback (F) en New Response (NR).

		Corrector 1				Total Filas
		Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	
Corrector 2	Nivel No Logrado	4	0	0	0	4
	Nivel Inicial	0	2	0	0	2
	Nivel En Desarrollo	0	0	8	0	8
	Nivel Logrado	0	0	2	9	11
Total Columnas		4	2	10	9	25

Tabla D15: Frecuencias esperadas de Cohen Kappa de cada criterio de evaluación: Rúbrica de Response (R) y New Response (NR), Rúbrica de Feedback (F) y Rúbrica de integración de Feedback (F) en New Response (NR) según dimensiones correspondientes.

	Nivel No Logrado	Nivel Inicial	Nivel En Desarrollo	Nivel Logrado	Habilidad No Presente
Rúbrica de R y NR					
Análisis	0.43	1.71	2.00	3.57	—
Explicación	0.17	16.11	6.78	6.52	—
Interpretación	0.14	12.59	5.84	2.64	—
Evaluación	0.91	2.67	1.48	12.50	—
Inferencia	0.20	27.05	12.56	2.75	—
Rúbrica de F					
<i>Indicadores Obligatorios de F</i>					
Criterios de Evaluación	0.64	0.16	0.64	9.00	—
Claridad y pertinencia de la explicación	0.00	0.80	3.60	4.84	—
Intención de la argumentación	0.00	7.84	0.64	1.96	—
Formulación de la retroalimentación	0.00	0.04	3.20	8.96	—
<i>Indicadores Opcionales de F</i>					
Solicitud de evidencia	0.00	0.64	0.64	1.96	4.00
Relectura de un argumento o presentación de un nuevo argumento y/o contraargumento	0.00	0.00	0.04	0.00	22.08
Solicitud de reformulación	0.04	0.08	0.16	0.00	16.80
Llamado de Atención	0.00	0.08	0.80	0.24	10.88
Rúbrica de integración de F en NR	0.64	0.16	3.2	3.96	7.96

Nota: Frecuencia de cada nivel, se calcula como el producto del total de columnas por el total de filas, dividido por el gran total (Valores de la última fila y la última columna de las Tablas E1 a ETablas E7 a E14).

Tabla D16: Factor Kappa de cada dimensión evaluada por Rúbrica de R y NR, Rúbrica de F y Rúbrica de integración de F en NR.

	N	$\sum ef$	$\sum a$	Kappa
Rúbrica de R y NR				
Análisis	28	7.71	18	0.51
Explicación	92	29.59	86	0.90
Interpretación	64	21.22	59	0.88
Evaluación	54	17.56	52	0.95
Inferencia	118	42.56	113	0.93
Promedio	–	23.73	65	0.83
Rúbrica de F				
<i>Indicadores Obligatorios de F</i>				
Criterios de Evaluación	25	10.44	25	1.00
Claridad y pertinencia de la explicación	25	9.24	24	0.94
Intención de la argumentación	25	10.44	25	1.00
Formulación de la retroalimentación	25	12.20	23	0.84
<i>Indicadores Opcionales de F</i>				
Solicitud de evidencia	25	7.24	23	0.89
Relectura de un argumento o presentación de un nuevo argumento y/o contraargumento	25	22.12	24	0.65
Solicitud de reformulación	25	17.08	22	0.62
Llamado de Atención	25	12.00	21	0.69
Promedio	–	12.60	23.38	0.83
Rúbrica de integración de F en NR	25	7.96	23	0.88

Nota: N corresponde a la cantidad de respuestas donde se evaluó la dimensión, $\sum ef$ corresponde a la suma de frecuencias esperadas (Tabla E15) y $\sum a$ es la suma del número de acuerdos de los distintos criterios.

E. Tabulación de Reponse, New Response y Feedback para cada actividad.

Tabla E1: Tabulación de Response (R) y New Response (NR) de Actividad 1.

Estudiante	Estudiante Feedback	R _{análisis}	R _{explicación}	R ₁	NR _{análisis}	NR _{explicación}	NR ₁
Estudiante 1	Estudiante 26	3	3	3	4	4	4
Estudiante 2	—	—	—	—	—	—	—
Estudiante 3	Estudiante 19	3	3	3	4	4	4
Estudiante 4	—	—	—	—	—	—	—
Estudiante 5	—	—	—	—	—	—	—
Estudiante 6	—	—	—	—	—	—	—
Estudiante 7	—	—	—	—	—	—	—
Estudiante 8	Estudiante 17	1	2	1.5	1	2	1.5
Estudiante 9	—	—	—	—	—	—	—
Estudiante 10	Estudiante 6	1	2	1.5	4	4	4
Estudiante 11	Estudiante 1	2	3	2.5	3	4	3.5
Estudiante 12	—	—	—	—	—	—	—
Estudiante 13	Estudiante 9	4	3	3.5	4	4	4
Estudiante 14	—	—	—	—	—	—	—
Estudiante 15	—	—	—	—	—	—	—
Estudiante 16	Estudiante 24	3	3	3	4	4	4
Estudiante 17	—	—	—	—	—	—	—
Estudiante 18	Estudiante 25	3	3	3	4	4	4
Estudiante 19	Estudiante 8	2	2	2	4	4	4
Estudiante 20	—	—	—	—	—	—	—
Estudiante 21	—	—	—	—	—	—	—
Estudiante 22	—	—	—	—	—	—	—
Estudiante 23	—	—	—	—	—	—	—
Estudiante 24	Estudiante 13	3	4	3.5	4	4	4
Estudiante 25	Estudiante 10	2	2	2	2	3	2.5
Estudiante 26	Estudiante 21	3	2	2.5	4	4	4
Estudiante 27	—	—	—	—	—	—	—
Estudiante 28	—	—	—	—	—	—	—
Estudiante 29	Estudiante 18	3	4	3.5	4	4	4
Estudiante 30	—	—	—	—	—	—	—
Estudiante 31	—	—	—	—	—	—	—
Estudiante 32	—	—	—	—	—	—	—
Estudiante 33	—	—	—	—	—	—	—
Estudiante 34	—	—	—	—	—	—	—
Estudiante 35	—	—	—	—	—	—	—

Nota: Algunos estudiantes que entregan feedback no completaron la etapa de *New Response (NR)* por lo que el sistema no permitió obtener sus resultados, e.g., Estudiante 21 entrega feedback a Estudiante 26, pero Estudiante 21 no tiene resultados.

Tabla E2: Tabulación de Response (R) y New Response (NR) de Actividad 2.

Nombre	Estudiante Feedback	R _{interpretación}	R _{inferencia}	R _{explicación}	R ₂	NR _{interpretación}	NR _{inferencia}	NR _{explicación}	NR ₂
Estudiante 1	Estudiante 31	4	4	4	4.0	4	4	4	4.0
Estudiante 2	Estudiante 30	2	2	2	2.0	2	2	2	2.0
Estudiante 3	Estudiante 9	4	4	4	4.0	4	4	4	4.0
Estudiante 4	Estudiante 2	2	2	3	2.3	2	2	3	2.3
Estudiante 5	Estudiante 25	1	2	2	1.7	1	2	2	1.7
Estudiante 6	Estudiante 16	3	3	2	2.7	3	3	3	3.0
Estudiante 7	Estudiante 1	3	2	2	2.3	4	4	4	4.0
Estudiante 8	Estudiante 10	2	2	3	2.3	2	2	3	2.3
Estudiante 9	Estudiante 4	3	3	3	3.0	4	3	4	3.7
Estudiante 10	Estudiante 3	4	4	3	3.7	4	4	4	4.0
Estudiante 11	Estudiante 14	3	2	3	2.7	3	3	3	3.0
Estudiante 12	Estudiante 11	2	2	2	2.0	3	4	3	3.3
Estudiante 13	Estudiante 6	4	4	4	4.0	4	4	4	4.0
Estudiante 14	Estudiante 7	1	1	1	1.0	1	1	1	1.0
Estudiante 15	—	—	—	—	—	—	—	—	—
Estudiante 16	Estudiante 15	3	2	4	3.0	4	4	4	4.0
Estudiante 17	Estudiante 5	2	2	2	2.0	3	3	3	3.0
Estudiante 18	Estudiante 29	3	4	3	3.3	3	4	3	3.3
Estudiante 19	Estudiante 13	2	2	2	2.0	2	2	2	2.0
Estudiante 20	Estudiante 33	2	1	1	1.3	2	1	1	1.3
Estudiante 21	Estudiante 18	3	3	3	3.0	3	3	4	3.3
Estudiante 22	Estudiante 35	2	2	2	2.0	2	2	2	2.0
Estudiante 23	Estudiante 21	3	2	3	2.7	3	2	3	2.7
Estudiante 24	Estudiante 26	2	2	2	2.0	2	2	2	2.0
Estudiante 25	Estudiante 22	2	2	2	2.0	2	3	3	2.7
Estudiante 26	Estudiante 28	3	3	2	2.7	3	3	2	2.7
Estudiante 27	—	—	—	—	—	—	—	—	—
Estudiante 28	Estudiante 32	2	2	2	2.0	2	2	2	2.0
Estudiante 29	Estudiante 17	4	4	3	3.7	4	4	4	4.0
Estudiante 30	Estudiante 24	2	2	2	2.0	2	2	2	2.0
Estudiante 31	Estudiante 23	2	2	2	2.0	3	2	2	2.3
Estudiante 32	Estudiante 20	2	2	2	2.0	2	2	2	2.0
Estudiante 33	Estudiante 19	2	2	2	2.0	2	2	2	2.0
Estudiante 34	Estudiante 12	3	4	2	3.0	3	4	2	3.0
Estudiante 35	Estudiante 8	3	3	3	3.0	3	4	3	3.3

Nota: Algunos estudiantes que entregan feedback no completaron la etapa de *New Response (NR)* por lo que el sistema no permitió obtener sus resultados, e.g., Estudiante 15 entrega feedback a Estudiante 16, pero Estudiante 15 no tiene resultados.

Nombre	Nombre Feedback	R _{evaluación}	R _{inferencia}	R ₃	NR _{evaluación}	NR _{inferencia}	NR ₃
Estudiante 1	Estudiante 16	4	3	3.5	4	3	3.5
Estudiante 2	Estudiante 34	1	1	1	1	1	1
Estudiante 3	—	—	—	—	—	—	—
Estudiante 4	Estudiante 18	1	2	1.5	4	3	3.5
Estudiante 5	—	—	—	—	—	—	—
Estudiante 6	—	—	—	—	—	—	—
Estudiante 7	Estudiante 11	4	3	3.5	4	3	3.5
Estudiante 8	Estudiante 19	1	2	1.5	1	2	1.5
Estudiante 9	Estudiante 27	3	2	2.5	3	2	2.5
Estudiante 10	Estudiante 22	4	3	3.5	4	3	3.5
Estudiante 11	Estudiante 12	4	3	3.5	4	3	3.5
Estudiante 12	Estudiante 29	4	3	3.5	4	3	3.5
Estudiante 13	Estudiante 26	4	3	3.5	4	3	3.5
Estudiante 14	—	—	—	—	—	—	—
Estudiante 15	Estudiante 4	3	2	2.5	3	2	2.5
Estudiante 16	Estudiante 2	2	2	2	2	2	2
Estudiante 17	Estudiante 24	3	3	3	3	3	3
Estudiante 18	Estudiante 25	2	2	2	2	2	2
Estudiante 19	Estudiante 9	3	2	2.5	3	2	2.5
Estudiante 20	Estudiante 8	4	2	3	4	2	3
Estudiante 21	Estudiante 32	3	3	3	3	3	3
Estudiante 22	Estudiante 31	2	2	2	2	2	2
Estudiante 23	—	—	—	—	—	—	—
Estudiante 24	Estudiante 21	4	3	3.5	4	3	3.5
Estudiante 25	Estudiante 20	2	2	2	2	2	2
Estudiante 26	Estudiante 17	4	3	3.5	4	2	3
Estudiante 27	Estudiante 30	2	2	2	2	2	2
Estudiante 28	—	—	—	—	—	—	—
Estudiante 29	Estudiante 15	2	2	2	2	2	2
Estudiante 30	Estudiante 13	4	3	3.5	4	3	3.5
Estudiante 31	Estudiante 28	1	2	1.5	1	2	1.5
Estudiante 32	Estudiante 1	4	3	3.5	4	3	3.5
Estudiante 33	—	—	—	—	—	—	—
Estudiante 34	Estudiante 10	4	3	3.5	4	3	3.5
Estudiante 35	—	—	—	—	—	—	—

Nota: Algunos estudiantes que entregan feedback no completaron la etapa de *New Response (NR)* por lo que el sistema no permitió obtener sus resultados, e.g., Estudiante 28 entrega feedback a Estudiante 31, pero Estudiante 28 no tiene resultados.

Tabla E4: Tabulación del Avance entre R y NR

	Actividad 1			Actividad 2				Avance Actividad 3			89
	A _{análisis}	A _{explicación}	A ₁	A _{interpretación}	A _{inferencia}	A _{explicación}	A ₂	A _{evaluación}	A _{inferencia}	A ₃	
Estudiante 1	1	1	1.0	0	0	0	0.0	0	0	0.0	
Estudiante 2	—	—	—	0	0	0	0.0	0	0	0.0	
Estudiante 3	1	1	1.0	0	0	0	0.0	—	—	—	
Estudiante 4	—	—	—	0	0	0	0.0	3	1	2.0	
Estudiante 5	—	—	—	0	0	0	0.0	—	—	—	
Estudiante 6	—	—	—	0	0	1	0.3	—	—	—	
Estudiante 7	—	—	—	1	2	2	1.7	0	0	0.0	
Estudiante 8	0	0	0.0	0	0	0	0.0	0	0	0.0	
Estudiante 9	—	—	—	1	0	1	0.7	0	0	0.0	
Estudiante 10	3	2	2.5	0	0	1	0.3	0	0	0.0	
Estudiante 11	1	1	1.0	0	1	0	0.3	0	0	0.0	
Estudiante 12	—	—	—	1	2	1	1.3	0	0	0.0	
Estudiante 13	0	1	0.5	0	0	0	0.0	0	0	0.0	
Estudiante 14	—	—	—	0	0	0	0.0	—	—	—	
Estudiante 15	—	—	—	—	—	—	—	0	0	0.0	
Estudiante 16	1	1	1.0	1	2	0	1.0	0	0	0.0	
Estudiante 17	—	—	—	1	1	1	1.0	0	0	0.0	
Estudiante 18	1	1	1.0	0	0	0	0.0	0	0	0.0	
Estudiante 19	2	2	2.0	0	0	0	0.0	0	0	0.0	
Estudiante 20	—	—	—	0	0	0	0.0	0	0	0.0	
Estudiante 21	—	—	—	0	0	1	0.3	0	0	0.0	
Estudiante 22	—	—	—	0	0	0	0.0	0	0	0.0	
Estudiante 23	—	—	—	0	0	0	0.0	—	—	—	
Estudiante 24	1	0	0.5	0	0	0	0.0	0	0	0.0	
Estudiante 25	0	1	0.5	0	1	1	0.7	0	0	0.0	
Estudiante 26	1	2	1.5	0	0	0	0.0	0	-1	-0.5	
Estudiante 27	—	—	—	—	—	—	—	0	0	0.0	
Estudiante 28	—	—	—	0	0	0	0.0	—	—	—	
Estudiante 29	1	0	0.5	0	0	1	0.3	0	0	0.0	
Estudiante 30	—	—	—	0	0	0	0.0	0	0	0.0	
Estudiante 31	—	—	—	1	0	0	0.3	0	0	0.0	
Estudiante 32	—	—	—	0	0	0	0.0	0	0	0.0	
Estudiante 33	—	—	—	0	0	0	0.0	—	—	—	
Estudiante 34	—	—	—	0	0	0	0.0	0	0	0.0	
Estudiante 35	—	—	—	0	1	0	0.3	—	—	—	

Nota: El *Avance* representa el progreso que tienen los estudiantes entre *R* y *NR* y se calcula como $NR - R$.

Tabla E5: Tabulación de Feedback Received (FR) y Feedback Delivered (FD) de Actividad 1.

Estudiante	Estudiante Feedback	Indicadores Obligatorios	Indicadores Opcionales	FR ₁	FD ₁
Estudiante 1	Estudiante 26	11	6	17	23
Estudiante 2	—	—	—	—	—
Estudiante 3	Estudiante 19	14	7	21	—
Estudiante 4	—	—	—	—	—
Estudiante 5	—	—	—	—	—
Estudiante 6	—	—	—	—	—
Estudiante 7	—	—	—	—	—
Estudiante 8	Estudiante 17	0	0	0	15
Estudiante 9	—	—	—	—	—
Estudiante 10	Estudiante 6	12	2	14	20
Estudiante 11	Estudiante 1	13	10	23	—
Estudiante 12	—	—	—	—	—
Estudiante 13	Estudiante 9	15	7	22	24
Estudiante 14	—	—	—	—	—
Estudiante 15	—	—	—	—	—
Estudiante 16	Estudiante 24	16	5	21	—
Estudiante 17	—	—	—	—	—
Estudiante 18	Estudiante 25	14	8	22	20
Estudiante 19	Estudiante 8	11	4	15	21
Estudiante 20	—	—	—	—	—
Estudiante 21	—	—	—	—	—
Estudiante 22	—	—	—	—	—
Estudiante 23	—	—	—	—	—
Estudiante 24	Estudiante 13	16	8	24	21
Estudiante 25	Estudiante 10	16	4	20	22
Estudiante 26	Estudiante 21	14	2	16	17
Estudiante 27	—	—	—	—	—
Estudiante 28	—	—	—	—	—
Estudiante 29	Estudiante 18	16	4	20	—
Estudiante 30	—	—	—	—	—
Estudiante 31	—	—	—	—	—
Estudiante 32	—	—	—	—	—
Estudiante 33	—	—	—	—	—
Estudiante 34	—	—	—	—	—
Estudiante 35	—	—	—	—	—

Nota: algunos estudiantes no poseen Feedback Delivered, debido a que los estudiantes a los que les fue entregado dicho feedback no finalizaron la actividad y no fue posible acceder a éste.

Tabla E6: Tabulación de Feedback Received (FR) y Feedback Delivered (FD) de Actividad 2.

Estudiante	Estudiante Feedback	Indicadores Obligatorios	Indicadores Opcionales	FR ₂	FD ₂
Estudiante 1	Estudiante 31	15	3	18	15
Estudiante 2	Estudiante 30	16	0	16	10
Estudiante 3	Estudiante 9	14	0	14	13
Estudiante 4	Estudiante 2	10	0	10	16
Estudiante 5	Estudiante 25	0	0	0	9
Estudiante 6	Estudiante 16	16	4	20	13
Estudiante 7	Estudiante 1	11	4	15	17
Estudiante 8	Estudiante 10	0	0	0	11
Estudiante 9	Estudiante 4	13	3	16	14
Estudiante 10	Estudiante 3	10	3	13	–
Estudiante 11	Estudiante 14	11	3	14	14
Estudiante 12	Estudiante 11	11	3	14	16
Estudiante 13	Estudiante 6	13	0	13	20
Estudiante 14	Estudiante 7	14	3	17	14
Estudiante 15	–	–	–	–	–
Estudiante 16	Estudiante 15	10	4	14	20
Estudiante 17	Estudiante 5	9	0	9	–
Estudiante 18	Estudiante 29	15	0	15	13
Estudiante 19	Estudiante 13	16	4	20	–
Estudiante 20	Estudiante 33	13	0	13	–
Estudiante 21	Estudiante 18	13	0	13	–
Estudiante 22	Estudiante 35	0	0	0	14
Estudiante 23	Estudiante 21	0	0	0	14
Estudiante 24	Estudiante 26	0	0	0	–
Estudiante 25	Estudiante 22	14	0	14	–
Estudiante 26	Estudiante 28	0	0	0	–
Estudiante 27	–	–	–	–	–
Estudiante 28	Estudiante 32	0	0	0	
Estudiante 29	Estudiante 17	0	0	0	15
Estudiante 30	Estudiante 24	0	0	0	16
Estudiante 31	Estudiante 23	14	0	14	18
Estudiante 32	Estudiante 20	0	0	0	–
Estudiante 33	Estudiante 19	0	0	0	13
Estudiante 34	Estudiante 12	16	0	16	–
Estudiante 35	Estudiante 8	11	0	11	–

Nota: algunos estudiantes no poseen Feedback Delivered, debido a que los estudiantes a los que les fue entregado dicho feedback no finalizaron la actividad y no fue posible acceder a éste.

Tabla E7: Tabulación de Feedback Received (FR) y Feedback Delivered (FD) de Actividad 3.

Estudiante	Estudiante Feedback	Indicadores Obligatorios	Indicadores Opcionales	FR ₃	FD ₃
Estudiante 1	Estudiante 16	10	3	13	–
Estudiante 2	Estudiante 34	8	1	9	8
Estudiante 3	–	–	–	–	–
Estudiante 4	Estudiante 18	9	3	12	11
Estudiante 5	–	–	–	–	–
Estudiante 6	–	–	–	–	–
Estudiante 7	Estudiante 11	9	5	14	–
Estudiante 8	Estudiante 19	15	4	19	12
Estudiante 9	Estudiante 27	9	0	9	14
Estudiante 10	Estudiante 22	9	2	11	–
Estudiante 11	Estudiante 12	13	0	13	14
Estudiante 12	Estudiante 29	15	0	15	13
Estudiante 13	Estudiante 26	16	0	16	–
Estudiante 14	–	–	–	–	–
Estudiante 15	Estudiante 4	11	0	11	–
Estudiante 16	Estudiante 2	8	0	8	13
Estudiante 17	Estudiante 24	12	0	12	19
Estudiante 18	Estudiante 25	10	0	10	12
Estudiante 19	Estudiante 9	14	0	14	19
Estudiante 20	Estudiante 8	12	0	12	7
Estudiante 21	Estudiante 32	16	4	20	13
Estudiante 22	Estudiante 31	9	0	9	11
Estudiante 23	–	–	–	–	–
Estudiante 24	Estudiante 21	13	0	13	12
Estudiante 25	Estudiante 20	7	0	7	10
Estudiante 26	Estudiante 17	16	3	19	16
Estudiante 27	Estudiante 30	0	0	0	9
Estudiante 28	–	–	–	–	–
Estudiante 29	Estudiante 15	0	0	0	15
Estudiante 30	Estudiante 13	0	0	0	–
Estudiante 31	Estudiante 28	0	0	0	9
Estudiante 32	Estudiante 1	0	0	0	20
Estudiante 33	–	–	–	–	–
Estudiante 34	Estudiante 10	0	0	0	9
Estudiante 35	–	–	–	–	–

Nota: algunos estudiantes no poseen Feedback Delivered, debido a que los estudiantes a los que les fue entregado dicho feedback no finalizaron la actividad y no fue posible acceder a éste.

Tabla E8: Tabulación de Integración de Feedback Received en New Response NR(FR).

	$NR(FR)_1$	$NR(FR)_2$	$NR(FR)_3$
Estudiante 1	3	2	0
Estudiante 2	—	2	2
Estudiante 3	4	2	—
Estudiante 4	—	2	4
Estudiante 5	—	0	—
Estudiante 6	—	4	—
Estudiante 7	—	4	2
Estudiante 8	0	0	0
Estudiante 9	—	1	2
Estudiante 10	4	3	1
Estudiante 11	3	1	2
Estudiante 12	—	3	2
Estudiante 13	4	3	0
Estudiante 14	—	2	—
Estudiante 15	—	—	2
Estudiante 16	1	3	3
Estudiante 17	—	4	2
Estudiante 18	3	3	3
Estudiante 19	4	3	2
Estudiante 20	—	1	3
Estudiante 21	—	3	3
Estudiante 22	—	0	3
Estudiante 23	—	0	—
Estudiante 24	3	0	3
Estudiante 25	2	2	0
Estudiante 26	4	0	0
Estudiante 27	—	—	0
Estudiante 28	—	0	—
Estudiante 29	3	0	0
Estudiante 30	—	0	0
Estudiante 31	—	1	0
Estudiante 32	—	0	0
Estudiante 33	—	0	—
Estudiante 34	—	4	0
Estudiante 35	—	3	—

Tabla E9: Tabulación de promedios de actividades en Response (R), New Response (NR), Avance, Feedback

	$\bar{R}$	$\overline{NR}$	$\bar{A}$	$\overline{FR}$	$\overline{FD}$	$\overline{NR(FR)}$
Estudiante 1	3.5	3.8	0.3	16.0	19.0	1.7
Estudiante 2	1.5	1.5	0.0	12.5	9.0	2.0
Estudiante 3	3.5	4.0	0.5	17.5	13.0	3.0
Estudiante 4	1.9	2.9	1.0	11.0	13.5	3.0
Estudiante 5	1.7	1.7	0.0	0.0	9.0	0.0
Estudiante 6	2.7	3.0	0.3	20.0	13.0	4.0
Estudiante 7	2.9	3.8	0.8	14.5	17.0	3.0
Estudiante 8	1.8	1.8	0.0	6.3	12.7	0.0
Estudiante 9	2.8	3.1	0.3	12.5	14.0	1.5
Estudiante 10	2.9	3.8	0.9	12.7	20.0	2.7
Estudiante 11	2.9	3.3	0.4	16.7	14.0	2.0
Estudiante 12	2.8	3.4	0.7	14.5	14.5	2.5
Estudiante 13	3.7	3.8	0.2	17.0	22.0	2.3
Estudiante 14	1.0	1.0	0.0	17.0	14.0	2.0
Estudiante 15	2.5	2.5	0.0	11.0	12.1	2.0
Estudiante 16	2.7	3.3	0.7	14.3	16.5	2.3
Estudiante 17	2.5	3.0	0.5	10.5	19.0	3.0
Estudiante 18	2.8	3.1	0.3	15.7	15.0	3.0
Estudiante 19	2.2	2.8	0.7	16.3	20.0	3.0
Estudiante 20	2.2	2.2	0.0	12.5	7.0	2.0
Estudiante 21	3.0	3.2	0.2	16.5	13.0	3.0
Estudiante 22	2.0	2.0	0.0	4.5	12.5	1.5
Estudiante 23	2.7	2.7	0.0	0.0	14.0	0.0
Estudiante 24	3.0	3.2	0.2	12.3	16.5	2.0
Estudiante 25	2.0	2.4	0.4	13.7	16.0	1.3
Estudiante 26	2.9	3.2	0.3	11.7	16.5	1.3
Estudiante 27	2.0	2.0	0.0	0.0	9.0	0.0
Estudiante 28	2.0	2.0	0.0	0.0	12.1	0.0
Estudiante 29	3.1	3.3	0.3	6.7	15.0	1.0
Estudiante 30	2.8	2.8	0.0	0.0	16.0	0.0
Estudiante 31	1.8	1.9	0.2	7.0	13.5	0.5
Estudiante 32	2.8	2.8	0.0	0.0	20.0	0.0
Estudiante 33	2.0	2.0	0.0	0.0	13.0	0.0
Estudiante 34	3.3	3.3	0.0	8.0	9.0	2.0
Estudiante 35	3.0	3.3	0.3	11.0	15.4	3.0

Nota: el Apéndice F explica mediante un ejemplo cómo se calculan los promedios de R, NR, Avance, FR, FD e Integración de FR en NR.

Received (FR), Feedback Delivered (FD) e Integración de FR en NR (NR(FR)).

Tabla E10: Tabulación de Test 1 (T1) y Test 2 (T2).

Estudiante	T1	T2	Rango escala
Estudiante 1	5.7	5.6	1-7
Estudiante 2	4.8	4.6	1-7
Estudiante 3	4.7	6.9	1-7
Estudiante 4	5.0	6.0	1-7
Estudiante 5	5.5	5.0	1-7
Estudiante 6	5.1	4.7	1-7
Estudiante 7	4.8	6.1	1-7
Estudiante 8	5.3	5.7	1-7
Estudiante 9	4.7	4.4	1-7
Estudiante 10	4.8	5.2	1-7
Estudiante 11	5.4	6.0	1-7
Estudiante 12	5.5	5.9	1-7
Estudiante 13	6.2	6.2	1-7
Estudiante 14	5.0	0.0	1-7
Estudiante 15	5.4	5.9	1-7
Estudiante 16	4.8	5.3	1-7
Estudiante 17	5.7	5.3	1-7
Estudiante 18	5.5	5.3	1-7
Estudiante 19	5.5	6.4	1-7
Estudiante 20	5.2	5.2	1-7
Estudiante 21	6.2	5.1	1-7
Estudiante 22	4.3	5.2	1-7
Estudiante 23	5.7	5.1	1-7
Estudiante 24	5.0	6.0	1-7
Estudiante 25	6.0	6.4	1-7
Estudiante 26	4.7	5.3	1-7
Estudiante 27	4.9	6.1	1-7
Estudiante 28	4.3	4.6	1-7
Estudiante 29	4.5	5.9	1-7
Estudiante 30	4.1	6.6	1-7
Estudiante 31	5.1	4.4	1-7
Estudiante 32	4.3	5.5	1-7
Estudiante 33	5.6	5.5	1-7
Estudiante 34	5.1	5.5	1-7
Estudiante 35	5.1	6.6	1-7

Nota: T1 (*Mean* = 5.13, *Std. Dev.* = 0.53) y T2
(*Mean* = 5.41, *Std. Dev.* = 1.14)

F. Descripción mediante ejemplos y cálculo de variables consideradas en el modelo de regresión lineal

Para explicar cómo se obtuvieron las variables incluidas en el modelo, se considerará la siguiente situación hipotética: Estudiante 1 participó en actividades 1, 2 y 3; Estudiante 2 participó solamente en la actividad 2; y Estudiante 3 participó en las actividades 1 y 3.

E_x : Estudiante “X”

R_x : Response of Activity “X”

NR_x : New Response of Activity “X”

FR_x : Feedback Received of Activity “X”

FE_x : Feedback Entregado en la actividad “X”

1. **Promedio Response (R)**: es el promedio de las respuestas iniciales (R) de las actividades en las que participó cada estudiante.

Promedio Response Estudiante 1:

$$\overline{R} E_1 = \frac{R_1 + R_2 + R_3}{3}$$

Promedio Response Estudiante 2:

$$\overline{R} E_2 = R_2$$

Promedio Response Estudiante 3:

$$\overline{R} E_3 = \frac{R_1 + R_3}{2}$$

2. **Promedio New Response (NR)**: es el promedio de las respuestas finales (NR) de las actividades en las que participó cada estudiante.

Promedio New Response Estudiante 1:

$$\overline{NR} E_1 = \frac{NR_1 + NR_2 + NR_3}{3}$$

Promedio New Response Estudiante 2:

$$\overline{NR} E_2 = NR_2$$

Promedio New Response Estudiante 3:

$$\overline{NR} E_3 = \frac{NR_1 + NR_3}{2}$$

3. **Promedio Avance (NR – R):** Corresponde a la diferencia entre el promedio de Response (R) y el promedio de New Response(NR) de las actividades en las que participó el estudiante.

Avance Estudiante 1:

$$\overline{A}E_1 = \overline{NR}E_1 - \overline{R}E_1$$

Avance Estudiante 2:

$$\overline{A}E_2 = \overline{NR}E_2 - \overline{R}E_2$$

Avance Estudiante 3:

$$\overline{A}E_3 = \overline{NR}E_3 - \overline{R}E_3$$

4. **Promedio Feedback Received (FR):** es el promedio de los feedbacks que recibió el estudiante, de cada una de las actividades en las que participó.

Feedback Recibido Estudiante 1:

$$\overline{FR}E_1 = \frac{FR_1 + FR_2 + FR_3}{3}$$

Feedback Recibido Estudiante 2:

$$\overline{FR}E_2 = FR_2$$

Feedback Recibido Estudiante 3:

$$\overline{FR}E_3 = \frac{FR_1 + FR_3}{2}$$

5. **Promedio Feedback Delivered (FD):** es el promedio de los feedbacks entregados por el estudiante, de cada una de las actividades en las que participó.

Feedback Entregado Estudiante 1:

$$\overline{FD}E_1 = \frac{FD_1 + FD_2 + FD_3}{3}$$

Feedback Entregado Estudiante 2:

$$\overline{FD}E_2 = FD_2$$

Feedback Entregado Estudiante 3:

$$\overline{FD} E_3 = \frac{FD_1 + FD_3}{2}$$

6. **Promedio Integración de FR en NR (FRNR):** es el promedio de los puntajes asignados por la rúbrica de Integración de FR en NR a cada NR, de las actividades en las que participó cada estudiante.
Esta variable no fue considerada en el modelo.

Integración FR en NR Estudiante 1:

$$\overline{NR(FR)} E_1 = \frac{NR(FR)_1 + NR(FR)_2 + NR(FR)_3}{3}$$

Feedback Entregado Estudiante 2:

$$\overline{NR(FR)} E_2 = NR(FR)_2$$

Feedback Entregado Estudiante 3:

$$\overline{NR(FR)} E_3 = \frac{NR(FR)_1 + NR(FR)_3}{2}$$

G. Proceso de selección de variables modelo de regresión lineal de Avance

Tabla G1: Modelos engenerados con stepwise selection method en funcion de variable dependiente Avance

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
	β	Std. Error	β	Std. Error	β	Std. Error	β	Std. Error	β	Std. Error
$\overline{FR}$	0.5232***	0.1484	0.4240***	0.1427	0.4577***	0.1386	0.5380***	0.1424	0.6177***	0.1495
$\overline{FD}$	—	—	0.3644**	0.1427	0.3078**	0.1407	0.3948**	0.1456	0.4031***	0.1429
$T2$	—	—	—	—	0.2544*	0.1357	0.4075**	0.1591	0.4419***	0.1577
$\bar{R}$	—	—	—	—	—	—	-0.3047*	0.1778	-0.3316*	0.1752
$T1$	—	—	—	—	—	—	—	—	-0.2024	0.1359
Observaciones	35		35		35		35		35	
Adj R ²	0.2517		0.3589		0.4056		0.4405		0.4624	

Nota: modelo con variables estandarizadas ($Mean = 0$ y $Std. Deviation = 1$)

*p<0.1, **p<0.05, ***p<0.01

Tabla G2: Forward stepwise selection method para modelo de Avance (N=35)

Variable	β	p > t	Adj R ²	$[\Delta R]^2$	Prob > F
Step 1			.25	.25	.00
$\overline{FR}$	.52	.00			
Step 2			.36	.11	.00
$\overline{FR}$	.42	.00			
$\overline{FD}$	.36	.01			
Step 3			.41	.05	.00
$\overline{FR}$	.45	.00			
$\overline{FD}$	.30	.03			
$T2$	.25	.07			
Step 4			.40	.03	.00
$\overline{FR}$	.53	.00			
$\overline{FD}$	.39	.01			
$T2$	.40	.01			
$\bar{R}$	-.30	.09			
Step 5			.46	.02	.00
$\overline{FR}$	.62	.00			
$\overline{FD}$	.40	.00			
$T2$	.44	.00			
$\bar{R}$	-.33	.07			
$T1$	-.20	.14			

Nota: modelo con variables estandarizadas ($Mean = 0$ y $Std. Deviation = 1$). Para cada paso, se agregó al modelo la variable que hacía aumentar en un mayor valor el Adj R²

H. Datos utilizados en el modelo de regresión lineal.

Tabla H: Datos utilizados en STATA v. 12.0. para elaboración de la regresión lineal que predice el Avance de las habilidades de pensamiento crítico

Estudiante	$\bar{A}$	$\bar{R}$	$\overline{FR}$	$\overline{FD}$	T1	T2
Estudiante 1	0.3	3.5	16.0	19.0	5.7	5.6
Estudiante 2	0.0	1.5	12.5	9.0	4.8	4.6
Estudiante 3	0.5	3.5	17.5	13.0	4.7	6.9
Estudiante 4	1.0	1.9	11.0	13.5	5.0	6.0
Estudiante 5	0.0	1.7	0.0	9.0	5.5	5.0
Estudiante 6	0.3	2.7	20.0	13.0	5.1	4.7
Estudiante 7	0.8	2.9	14.5	17.0	4.8	6.1
Estudiante 8	0.0	1.8	6.3	12.7	5.3	5.7
Estudiante 9	0.3	2.8	12.5	14.0	4.7	4.4
Estudiante 10	0.9	2.9	12.7	20.0	4.8	5.2
Estudiante 11	0.4	2.9	16.7	14.0	5.4	6.0
Estudiante 12	0.7	2.8	14.5	14.5	5.5	5.9
Estudiante 13	0.2	3.7	17.0	22.0	6.2	6.2
Estudiante 14	0.0	1.0	17.0	14.0	5.0	0.0
Estudiante 15	0.0	2.5	11.0	12.1	5.4	5.9
Estudiante 16	0.7	2.7	14.3	16.5	4.8	5.3
Estudiante 17	0.5	2.5	10.5	19.0	5.7	5.3
Estudiante 18	0.3	2.8	15.7	15.0	5.5	5.3
Estudiante 19	0.7	2.2	16.3	20.0	5.5	6.4
Estudiante 20	0.0	2.2	12.5	7.0	5.2	5.2
Estudiante 21	0.2	3.0	16.5	13.0	6.2	5.1
Estudiante 22	0.0	2.0	4.5	12.5	4.3	5.2
Estudiante 23	0.0	2.7	0.0	14.0	5.7	5.1
Estudiante 24	0.2	3.0	12.3	16.5	5.0	6.0
Estudiante 25	0.4	2.0	13.7	16.0	6.0	6.4
Estudiante 26	0.3	2.9	11.7	16.5	4.7	5.3
Estudiante 27	0.0	2.0	0.0	9.0	4.9	6.1
Estudiante 28	0.0	2.0	0.0	12.1	4.3	4.6
Estudiante 29	0.3	3.1	6.7	15.0	4.5	5.9
Estudiante 30	0.0	2.8	0.0	16.0	4.1	6.6
Estudiante 31	0.2	1.8	7.0	13.5	5.1	4.4
Estudiante 32	0.0	2.8	0.0	20.0	4.3	5.5
Estudiante 33	0.0	2.0	0.0	13.0	5.6	5.5
Estudiante 34	0.0	3.3	8.0	9.0	5.1	5.5
Estudiante 35	0.3	3.0	11.0	15.4	5.1	6.6

Nota: Datos de Avance, R, FR, FD se obtienen de Tabla F9. T1 y T2 son las calificaciones obtenidas en el curso por los estudiantes.

I. Distribución del Avance

Tabla I1: Distribución del avance entre Response (R) y New Response (NR) en cada actividad, según habilidad de Pensamiento Crítico.

	Actividad 1	Actividad 2	Actividad 3
<i>Habilidad de Análisis</i>			
NR – R = 0	25.10%	-	-
NR – R = 1	58.3%	-	-
NR – R = 2	8.3%	-	-
NR – R = 3	8.3%	-	-
<i>Habilidad de Explicación</i>			
NR – R = 0	16.70%	85.00%	-
NR – R = 1	58.3%	13.3%	-
NR – R = 2	25.0%	1.7%	-
NR – R = 3	0.0%	0.0%	-
<i>Habilidad de Interpretación</i>			
NR – R = 0	-	80.00%	-
NR – R = 1	-	20.0%	-
NR – R = 2	-	0.0%	-
NR – R = 3	-	0.0%	-
<i>Habilidad de Inferencia</i>			
NR – R = 0	-	76.70%	96.30%
NR – R = 1	-	13.3%	3.7%
NR – R = 2	-	10.0%	0.0%
NR – R = 3	-	0.0%	0.0%
<i>Habilidad de Evaluación</i>			
NR – R = 0	-	-	96.30%
(NR) – (R) = 1	-	-	0.0%
(NR) – (R) = 2	-	-	0.0%
(NR) – (R) = 3	-	-	3.7%

J. Ejemplo de aplicación del modelo de andamiaje

Este ejemplo, cuyo objetivo es ilustrar el uso del Modelo de Figura 2, esta basado en respuestas específicas reales de los estudiantes, tomadas de distintas fases del desarrollo de la actividad.

1. Interpretación: Comprender y expresar el significado de una gran variedad de experiencias, situaciones, datos, eventos, juicios, convenciones, creencias, reglas, procedimientos o criterios.

Respuesta representativa de interpretación:

Situación: “Pensemos en un 3^{er} básico, creo que el primer paso para alcanzar la meta de 83 palabras semanales (según los autores) es evaluar el conocimiento de vocabulario académico de los alumnos mediante la lectura de un texto”.

Metodología general: “Para esto, es conveniente realizar un trabajo conjunto con los profesores de Historia y Ciencias Naturales (asignaturas donde típicamente se encuentran lecturas bastante especializadas) para que nos muestren los textos con los que trabajarán esa semana en sus clases y utilizarlos en la asignatura de Lenguaje para la enseñanza de vocabulario académico...”

Argumentación: “...esto sería beneficioso porque los niños estarían familiarizados con los textos, comprenderían mejor las otras asignaturas y realizarían una lectura contextualizada y con sentido”.

Instancia pedagógica: “Para la evaluación el docente realizará una lectura previa de los textos con el fin de decidir cuáles serán las palabras más adecuadas de enseñar, escogerá un mínimo de 5 y un máximo de 10 palabras por texto. Para esta elección se regirá por criterios tales como la utilidad de las palabras y su importancia para la comprensión global del texto. En clase, se leerá el texto en voz alta y se preguntará a los estudiantes qué palabras no conocen, el docente las escribirá en la pizarra y utilizará las que coincidan con su elección previa, las demás las dejará para una próxima clase. Entonces el profesor comunicará el objetivo de la clase, dará una definición amigable de las palabras en cuestión y pedirá a los alumnos que dibujen una situación para cada palabra con la cual puedan explicarle su significado a un compañero”.

Ejemplificación: “Por ejemplo, si la palabra es “deslumbrante” los niños dibujarán cosas que para ellos la identifiquen y al final de la clase compartirán sus dibujos y explicarán oralmente a sus compañeros la situación dibujada y cómo se relaciona con la palabra trabajada”.

2. Inferencia: Identificar los elementos necesarios para sacar conclusiones razonables; para formar conjeturas e hipótesis; para considerar la información relevante y para extraer las consecuencias que fluyen de datos, declaraciones, principios, pruebas, juicios, creencias, opiniones, conceptos, descripciones, preguntas u otras formas de representación.

Respuesta representativa de inferencia:

Situación: “Considerando la cantidad de palabras que deben aprender los estudiantes al año según esos autores...”

Hipótesis: “...resulta fundamental enseñar de forma explícita el vocabulario, es decir, no solo entregarle a los alumnos un texto que contenga las palabras que se desean trabajar, sino que a través de esa misma lectura se extraigan las palabras que se quiere enseñar...”

Argumentación: “...puesto que realizando este tipo de trabajo los estudiantes estarían aprendiendo dichas palabras desde un contexto (texto) entregándole un mayor significado al aprendizaje de estas”.

Situación: “Tomando en cuenta que la asignatura de Lenguaje y Comunicación tiene 6 horas semanales aproximadamente...”

Metodología general: “...se puede trabajar el vocabulario explícito de los estudiantes leyendo un texto al inicio de cada clase, que tenga coherencia con la materia que están pasando los alumnos, para luego extraer alrededor de 10 a 15 palabras (que varíen en su complejidad)...”

Instancia pedagógica: “...y estas sean enseñadas y practicadas a partir de múltiples estrategias, ya sea realizando preguntas a los estudiantes que lleven a las palabras a un contexto de los solicite a los estudiantes, solicitarles ejemplos donde ellos hayan visto dichas palabras, pedir que ellos elijan (entre 2 o 3 imágenes) donde se ven representadas las palabras, entre otras estrategias. La idea es que estas

estrategias varíen, para que así a los estudiantes no le resulte el aprendizaje del vocabulario una actividad tan aburrida y rutinaria.

3. Explicación: declarar los resultados del razonamiento en función de las consideraciones probatorias, conceptuales, metodológicas, de criterios y contextuales sobre los que se basaron los resultados; y presentar su razonamiento en forma de argumentos convincentes.

Respuesta representativa de explicación:

Situación: “En la secuencia de aprendizaje...”

Conceptualización: “se puede observar claramente dos de los cuatro principios para la enseñanza de la fluidez. En primer lugar, se observa el principio de lectura repetida del texto, el cual se lleva a cabo en la actividad cuando los estudiantes deben estar repitiendo la lectura del texto varias veces y de distinta forma, junto con la profesora, con un grupo de compañeros e independientemente. En segundo lugar, se evidencia en la actividad el principio de modelo lector...”

Argumentación: “...puesto que los estudiantes están recibiendo un modelamiento por parte del profesor, pues ella es quien al inicio de la actividad les lee en voz alta con el fin que los alumnos noten como se debe leer fluidamente el texto, para que luego de manera progresiva ellos leen de esta misma forma por sí solos”.

Conceptualización: “Los dos aspectos de enseñanza que quedan por mejorar corresponden a la retroalimentación y el contexto significativo”.

Argumentación: “El primer aspecto no se observa completamente, puesto que aunque la profesora les entrega consejos a los estudiantes, estos están preestablecidos y no se realizan en base a los aspectos logrados y por mejorar de cada uno de los alumnos...”

Instancia pedagógica: “...quizás sería más adecuado que la profesora realizara una retroalimentación constante cada uno de los estudiantes o por los grupos elaborados, y que esta sugerencia se realice en base a lo que los estudiantes están logrando y lo que les queda por alcanzar”.

Argumentación: “El segundo aspecto tampoco se observa claramente, ya que a pesar de que la profesora le da los nombres a los grupos en base a la lectura del texto (Atenienses y Espartanos), quizás para los estudiantes esto no se tan significativo si ellos aún no han logrado la comprensión de este”.

Instancia pedagógica: “Sería adecuado que la profesora antes de realizar la lectura modeladora a los estudiantes realizará una introducción que contextualizara el texto a los estudiantes y los haga relacionar sus conocimientos y experiencias con este”.

Puesta en relación: “En general, en la actividad se logra observar un modelo de transferencia progresiva, puesto que la profesora de a poco le va entregando la oportunidad a sus estudiantes de leer independientemente en base a los avances que ellos van teniendo en su fluidez lectora”.

4. Análisis: Identificar la relación, tanto intencionada como real, entre enunciados, preguntas conceptos, descripciones o formas de representación que expresan creencias, juicios, experiencias, razones, información y opiniones.

Respuesta representativa de análisis:

Situación: “De acuerdo a las características clave para la enseñanza de la fluidez...”

Puesta en relación: “se puede afirmar que en el guión de la clase se encuentra presente la existencia de un lector modelo, que en este caso sería la profesora...”

Argumentación: “...cuando en el primer punto se señala que “el profesor lee en voz alta el texto a los estudiantes””.

Puesta en relación: “También se observa la lectura repetida de textos...”

Argumentación: “...puesto que los estudiantes deben leer a coro, luego por grupos junto a la profesora, después por grupos sin la profesora; y, al final de la clase, nuevamente a coro. Por tanto, practican la lectura del mismo texto en varias ocasiones”.

Evaluación: “Por otra parte, uno de los aspectos que se encuentra débilmente presente, es la existencia de retroalimentación constante, pues si bien la profesora da consejos de lectura que contribuyen a que los alumnos piensen en cómo leen y perfeccionen su desempeño, en ninguna parte de la clase se enfatiza en los logros de los estudiantes. Además, la retroalimentación no se hace en función de los aspectos a mejorar que surgen en los alumnos mientras leen, sino que están predefinidos.

Adicionalmente, está ausente el contexto significativo que justifique la actividad, porque si bien esta tiene un foco claro en la mejora de la fluidez, no presenta un trasfondo que motive a leer una y otra vez sin que se torne tedioso”.

Instancia pedagógica: “Para mejorar, se podría enmarcar la actividad en una tutoría entre iguales, donde los alumnos tengan que asumir roles de evaluador y evaluado, lo que dota de sentido el escuchar constantemente leer a sus compañeros.

Puesta en relación: “Finalmente, esto se conecta con la transferencia progresiva de la responsabilidad...”

Argumentación: “en la medida en que cada actividad va variando en el grado de apoyo que la profesora da a los alumnos, ya que en un principio hay una instrucción explícita donde ella actúa como un modelo, luego apoya por grupos la lectura y finalmente, los alumnos leen entre ellos. Por ende, el control de la lectura pasa de tenerlo exclusivamente la docente, a ser compartido por esta y un grupo de alumnos, quienes leen de forma guiada, recibiendo apoyo de la profesora de vez en cuando”.

5. Evaluación: Evaluar la credibilidad de planteamientos u otras representaciones que describen las percepciones, experiencias, situaciones, juicios, creencias de una persona y evaluar la coherencia lógica de las relaciones inferenciales entre planteamientos, descripciones, cuestiones u otras formas de representación.

Respuesta representativa de evaluación:

Situación: “Yo utilizaría la propuesta 3”

Evaluación: “...porque creo que el título y párrafo inicial dan una idea global sobre qué tema se trabajará en la clase, luego el texto visual atrae mucho la atención de los estudiantes y les ayuda a centrarse en el ejemplo que se utilizará”.

Puesta en relación y argumentación: “Finalmente, sería bueno pasar en primero el apartado 3 del proceso porque este contiene información conocida para los alumnos (una planta adulta con tomate) y es necesaria para entender qué pasa en el apartado 1 (porque contiene la dificultad mencionada en clases, de información muy desconocida para los niños y algunas sub-etapas) y posteriormente analizar el apartado 2”.

Instancia pedagógica: “Preguntas:

Título y Párrafo inicial - En el párrafo 2 del texto dice que: “La planta de tomate, como toda planta que nace de una semilla, tiene tres etapas en su ciclo de vida”. ¿Habrán plantas que no nazcan de una semilla?

Texto Visual - Observando las imágenes, ¿qué crees que pasa con la flor de la planta de tomate una vez que crece el fruto?

Apartados - Si los tomates tienen semillas dentro de ellos y las flores nacen de semillas, ¿qué viene antes ¿la flor o el tomate? ¿Por qué?”

K. Tablas de correlaciones de variables eliminadas del modelo de regresión

Tabla K1: Correlación de $\bar{R}$ con variables descartadas del modelo por redundancia de información con respecto a $\bar{R}$

Actividad	Variable	<i>r</i>
Actividad 1 (N=13)	$R_{análisis}$	0.71 [*]
	$R_{explicación}$	0.57 ^{**}
	R_1	0.72 [*]
Actividad 2 (N=33)	$R_{interpretación}$	0.83 [*]
	$R_{inferencia}$	0.73 [*]
	$R_{explicación}$	0.60 ^{**}
Actividad 3 (N=27)	R_2	0.79 [*]
	$R_{evaluación}$	0.75 [*]
	$R_{inferencia}$	0.75 [*]
	R_3	0.78 [*]
Nota: (°) Factor de correlación fuerte, (**) Factor de correlación moderado		

Tabla K2: Correlación de $\overline{NR}$ con variables descartadas del modelo por redundancia de información con respecto a $\overline{NR}$

Actividad	Variable	<i>r</i>
Actividad 1 (N=13)	$NR_{análisis}$	0.81 [*]
	$NR_{explicación}$	0.83 [*]
	NR_1	0.83 [*]
Actividad 2 (N=33)	$NR_{interpretación}$	0.82 [*]
	$NR_{inferencia}$	0.82 [*]
	$NR_{explicación}$	0.73 [*]
Actividad 3 (N=27)	NR_2	0.84 [*]
	$NR_{evaluación}$	0.69 [*]
	$NR_{inferencia}$	0.69 [*]
	NR_3	0.73 [*]
Nota: (°) Factor de correlación fuerte, (**) Factor de correlación moderado		

Tabla K3: Correlación de $\bar{\mathbf{F}}$ con variables descartadas del modelo por redundancia de información con respecto a $\bar{\mathbf{F}}$

Variable	<i>r</i>
F_1	0.57 ^{**}
F_2	0.82 [*]
F_3	0.69 [*]
$NR(FR)_1$	0.50 ^{**}
$NR(FR)_2$	0.68 [*]
$NR(FR)_3$	0.42 ^{**}
$NR(NR)$	0.87 [*]
Nota: ([*]) Factor de correlación fuerte, (^{**}) Factor de correlación moderado	

L. Doble codificación de *Response* (R), *New Response* (NR) e Incorporación del *Feedback* en la *New Response* (NR(FR))

Tabla L1: Puntajes de la doble codificación de la rúbrica aplicada a *Response* (R) en la Actividad 1.

Student	Analysis		Explanation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	3	2	3	3
Student 2	—	—	—	—
Student 3	3	2	3	3
Student 4	—	—	—	—
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	—	—	—	—
Student 8	1	1	2	2
Student 9	3	3	2	2
Student 10	1	1	2	2
Student 11	2	3	3	3
Student 12	—	—	—	—
Student 13	4	3	3	3
Student 14	—	—	—	—
Student 15	—	—	—	—
Student 16	3	3	3	3
Student 17	2	2	2	2
Student 18	3	2	3	3
Student 19	2	2	2	2
Student 20	—	—	—	—
Student 21	—	—	—	—
Student 22	—	—	—	—
Student 23	—	—	—	—
Student 24	3	3	4	4
Student 25	2	2	2	2
Student 26	3	2	2	3
Student 27	—	—	—	—
Student 28	—	—	—	—
Student 29	3	3	4	4
Student 30	—	—	—	—
Student 31	—	—	—	—
Student 32	—	—	—	—
Student 33	—	—	—	—
Student 34	—	—	—	—
Student 35	—	—	—	—

Tabla L2: Puntajes de la doble codificación de la rúbrica aplicada a *Response* (R) en la Actividad 2.

Student	Explanation		Interpretation		Inference	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	4	4	4	4	4	4
Student 2	2	2	2	2	2	2
Student 3	4	4	4	4	4	4
Student 4	3	3	2	2	2	2
Student 5	2	2	1	1	2	2
Student 6	2	2	3	2	3	3
Student 7	2	2	3	2	2	2
Student 8	3	3	2	2	2	2
Student 9	3	3	3	3	3	2
Student 10	3	3	4	4	4	4
Student 11	3	2	3	3	2	2
Student 12	2	2	2	2	2	2
Student 13	4	4	4	4	4	4
Student 14	1	1	1	1	1	1
Student 15	—	—	—	—	—	—
Student 16	4	3	3	3	2	2
Student 17	2	2	2	2	2	2
Student 18	3	3	3	3	4	4
Student 19	2	2	2	2	2	2
Student 20	1	1	2	2	1	1
Student 21	3	2	3	2	3	3
Student 22	2	2	2	2	2	2
Student 23	3	3	3	3	2	2
Student 24	2	2	2	2	2	2
Student 25	2	2	2	2	2	2
Student 26	2	2	3	2	3	3
Student 27	—	—	—	—	—	—
Student 28	2	2	2	2	2	2
Student 29	3	4	4	4	4	4
Student 30	2	2	2	2	2	2
Student 31	2	2	2	2	2	2
Student 32	2	2	2	2	2	2
Student 33	2	2	2	2	2	2
Student 34	2	2	3	3	4	4
Student 35	3	3	3	3	3	3

Tabla L3: Puntajes de la doble codificación de la rúbrica aplicada a *Response* (R) en la Actividad 3.

Student	Inference		Evaluation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	3	3	4	4
Student 2	1	2	1	1
Student 3	—	—	—	—
Student 4	2	2	1	1
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	3	3	4	4
Student 8	2	2	1	1
Student 9	2	3	3	3
Student 10	3	3	4	4
Student 11	3	3	4	4
Student 12	3	3	4	4
Student 13	3	3	4	4
Student 14	—	—	—	—
Student 15	2	2	3	3
Student 16	2	2	2	2
Student 17	3	3	3	4
Student 18	2	2	2	2
Student 19	2	2	3	3
Student 20	2	2	4	4
Student 21	3	3	3	3
Student 22	2	2	2	2
Student 23	—	—	—	—
Student 24	3	3	4	4
Student 25	2	2	2	2
Student 26	3	3	4	4
Student 27	2	2	2	2
Student 28	—	—	—	—
Student 29	2	2	2	2
Student 30	3	3	4	4
Student 31	2	2	1	1
Student 32	3	3	4	4
Student 33	—	—	—	—
Student 34	3	3	4	4
Student 35	—	—	—	—

Tabla L4: Puntajes de la doble codificación de la rúbrica aplicada a *New Response* (R) en la Actividad 1.

Student	Analysis		Explanation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	4	4	4	4
Student 2	—	—	—	—
Student 3	4	4	4	4
Student 4	—	—	—	—
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	—	—	—	—
Student 8	1	2	2	2
Student 9	—	—	—	—
Student 10	4	4	4	4
Student 11	3	4	4	4
Student 12	—	—	—	—
Student 13	4	4	4	4
Student 14	—	—	—	—
Student 15	—	—	—	—
Student 16	4	4	4	4
Student 17	—	—	—	—
Student 18	4	4	4	4
Student 19	4	4	4	4
Student 20	—	—	—	—
Student 21	—	—	—	—
Student 22	—	—	—	—
Student 23	—	—	—	—
Student 24	4	4	4	4
Student 25	2	3	3	3
Student 26	4	4	4	4
Student 27	—	—	—	—
Student 28	—	—	—	—
Student 29	4	3	4	4
Student 30	—	—	—	—
Student 31	—	—	—	—
Student 32	—	—	—	—
Student 33	—	—	—	—
Student 34	—	—	—	—
Student 35	—	—	—	—

Tabla L5: Puntajes de la doble codificación de la rúbrica aplicada a *New Response* (R) en la Actividad 2.

Student	Explanation		Interpretation		Inference	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	4	4	4	4	4	4
Student 2	2	2	2	2	2	2
Student 3	4	4	4	4	4	4
Student 4	3	3	2	2	2	2
Student 5	2	2	1	1	2	2
Student 6	3	3	3	3	3	3
Student 7	4	4	4	4	4	4
Student 8	3	3	2	2	2	2
Student 9	4	4	4	4	3	3
Student 10	4	4	4	4	4	4
Student 11	3	3	3	3	3	3
Student 12	3	4	3	3	4	4
Student 13	4	4	4	4	4	4
Student 14	1	1	1	1	1	1
Student 15	—	—	—	—	—	—
Student 16	4	4	4	4	4	4
Student 17	3	3	3	3	3	3
Student 18	3	3	3	3	4	4
Student 19	2	2	2	2	2	2
Student 20	1	1	2	2	1	1
Student 21	4	4	3	3	3	3
Student 22	2	2	2	2	2	2
Student 23	3	3	3	3	2	2
Student 24	2	2	2	2	2	3
Student 25	3	3	2	2	3	3
Student 26	2	2	3	3	3	3
Student 27	—	—	—	—	—	—
Student 28	2	2	2	2	2	2
Student 29	4	4	4	4	4	4
Student 30	2	2	2	2	2	2
Student 31	2	2	3	2	2	2
Student 32	2	2	2	2	2	2
Student 33	2	2	2	2	2	2
Student 34	2	2	3	3	4	4
Student 35	3	3	3	3	4	4

Tabla L6: Puntajes de la doble codificación de la rúbrica aplicada a *New Response* (R) en la Actividad 3.

Student	Inference		Evaluation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	3	3	4	4
Student 2	1	2	1	1
Student 3	—	—	—	—
Student 4	3	3	4	4
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	3	3	4	4
Student 8	2	2	1	1
Student 9	2	2	3	3
Student 10	3	3	4	4
Student 11	3	3	4	4
Student 12	3	3	4	4
Student 13	3	3	4	4
Student 14	—	—	—	—
Student 15	2	2	3	3
Student 16	2	2	2	2
Student 17	3	3	3	4
Student 18	2	2	2	2
Student 19	2	2	3	3
Student 20	2	2	4	4
Student 21	3	3	3	3
Student 22	2	2	2	2
Student 23	—	—	—	—
Student 24	3	3	4	4
Student 25	2	2	2	2
Student 26	2	3	4	4
Student 27	2	2	2	2
Student 28	—	—	—	—
Student 29	2	2	2	2
Student 30	3	3	4	4
Student 31	2	2	1	1
Student 32	3	3	4	4
Student 33	—	—	—	—
Student 34	3	3	4	4
Student 35	—	—	—	—

Tabla L7: Puntajes de la doble codificación de la rúbrica aplicada a *Feedback* (F) en la Actividad 1.

Student	Assessment criteria		Clarity and relevance of the explanation		Argumentation		Syntax, lexis and register		Request for evidence		Reviewing an argument or presenting a new argument and/or counter-argument		Request to reword or rework		Highlighting the focus of the question		113
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	
Student 1	4	4	2	2	2	2	3	3	3	3	X	X	X	X	3	3	
Student 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 3	4	4	3	3	4	4	3	3	4	4	X	X	X	X	3	4	
Student 4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 10	4	4	2	2	2	2	4	4	X	X	X	X	2	X	X	X	
Student 11	4	4	3	3	3	3	3	3	4	4	X	X	2	2	4	4	
Student 12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 13	4	4	3	3	4	4	4	4	4	4	X	X	X	X	3	3	
Student 14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 16	4	4	4	4	4	4	4	4	3	3	X	X	2	2	X	X	
Student 17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 18	4	4	3	3	3	3	4	4	2	2	X	X	3	3	3	3	
Student 19	2	2	2	2	3	3	4	4	2	2	X	X	X	X	2	2	
Student 20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 24	4	4	4	4	4	4	4	4	4	4	X	X	X	X	4	4	
Student 25	4	4	4	4	4	4	4	4	4	4	X	X	X	X	X	X	
Student 26	4	4	3	4	3	3	4	4	X	X	X	X	X	X	2	3	
Student 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 29	4	4	4	4	4	4	4	4	X	X	X	X	X	X	4	4	
Student 30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Student 35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Note: E1 and E2 are Evaluator 1 and Evaluator respectively.

Tabla L8: Puntajes de la doble codificación de la rúbrica aplicada a *Feedback* (F) en la Actividad 2.

Student	Assessment criteria		Clarity and relevance of the explanation		Argumentation		Syntax, lexis and register		Request for evidence		Reviewing an argument or presenting a new argument and/or counter-argument		Request to reword or rework		Highlighting the focus of the question		114
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	
Estudiante 1	4	3	4	4	3	3	4	4	X	X	3	2	X	X	X	X	
Estudiante 2	4	4	4	4	4	4	4	4	X	X	X	X	X	3	X	X	
Estudiante 3	3	3	4	4	3	3	4	4	X	X	X	X	X	X	X	X	
Estudiante 4	2	4	3	4	1	1	4	4	X	X	X	X	X	X	X	X	
Estudiante 5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 6	4	4	4	4	4	4	4	4	4	4	X	X	X	X	X	X	
Estudiante 7	3	3	3	2	2	2	3	3	X	X	4	X	X	X	X	X	
Estudiante 8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 9	4	4	3	3	3	3	3	3	X	X	X	X	3	3	X	X	
Estudiante 10	3	3	3	3	2	2	2	2	3	3	X	X	X	X	X	X	
Estudiante 11	1	1	4	4	2	2	4	4	3	3	X	X	X	X	X	X	
Estudiante 12	1	1	4	4	2	2	4	4	3	3	X	X	X	X	X	X	
Estudiante 13	3	3	2	2	4	3	4	4	X	X	X	X	X	X	X	X	
Estudiante 14	4	4	4	3	2	2	4	3	X	X	3	2	X	X	X	X	
Estudiante 15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 16	2	2	3	3	2	2	3	3	2	2	X	X	X	X	2	2	
Estudiante 17	1	1	2	2	2	2	4	4	X	X	X	X	X	X	X	X	
Estudiante 18	4	3	4	3	3	2	4	4	X	X	X	X	X	X	X	X	
Estudiante 19	4	4	4	4	4	4	4	4	X	4	4	4	X	X	X	X	
Estudiante 20	3	4	4	4	2	3	4	4	X	X	X	3	X	X	X	3	
Estudiante 21	3	3	4	4	2	2	4	4	X	X	X	X	X	X	X	X	
Estudiante 22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 25	4	4	4	4	2	2	4	4	X	X	X	X	X	X	X	X	
Estudiante 26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 31	4	4	4	4	2	2	4	4	X	X	X	X	X	X	X	X	
Estudiante 32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 34	4	4	4	4	4	3	4	4	X	X	X	X	X	X	X	X	
Estudiante 35	3	3	3	3	2	2	3	3	X	X	X	X	X	X	X	X	

Note: E1 and E2 are Evaluator 1 and Evaluator respectively.

Tabla L9: Puntajes de la doble codificación de la rúbrica aplicada a *Feedback* (F) en la Actividad 3.

Student	Assessment criteria		Clarity and relevance of the explanation		Argumentation		Syntax, lexis and register		Request for evidence		Reviewing an argument or presenting a new argument and/or counter-argument		Request to reword or rework		Highlighting the focus of the question		115
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	
Estudiante 1	2	2	3	3	3	4	2	2	3	X	X	X	X	X	X	4	
Estudiante 2	2	2	3	3	1	2	2	1	X	X	1	2	X	X	X	X	
Estudiante 3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 4	1	1	3	3	2	2	3	3	X	X	X	X	1	1	2	X	
Estudiante 5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 7	2	2	3	2	2	2	2	2	X	X	X	X	2	2	3	3	
Estudiante 8	3	3	4	3	4	3	4	4	X	X	4	4	X	X	X	X	
Estudiante 9	1	1	3	3	3	2	2	2	X	X	X	X	X	X	X	X	
Estudiante 10	2	1	2	2	2	1	3	3	X	X	X	X	2	2	X	X	
Estudiante 11	3	4	3	3	4	4	3	3	X	X	X	X	X	X	X	X	
Estudiante 12	3	3	4	4	4	4	4	4	X	X	X	X	X	X	X	X	
Estudiante 13	4	4	4	3	4	2	4	3	X	X	X	X	X	X	X	X	
Estudiante 14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 15	2	2	3	3	3	3	3	4	X	2	X	X	X	X	X	X	
Estudiante 16	2	2	2	2	2	2	2	2	X	X	X	X	X	X	X	X	
Estudiante 17	3	4	3	4	3	3	3	3	X	X	X	X	X	X	X	X	
Estudiante 18	3	3	3	3	2	2	2	3	X	X	X	X	X	X	X	X	
Estudiante 19	3	2	4	3	3	2	4	3	X	X	X	X	X	X	X	X	
Estudiante 20	3	2	2	2	4	4	3	3	X	X	X	X	X	X	X	X	
Estudiante 21	4	4	4	4	4	4	4	4	X	X	4	4	X	X	X	X	
Estudiante 22	2	2	3	3	2	2	2	2	X	X	X	X	X	X	X	X	
Estudiante 23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 24	3	3	3	3	3	3	4	3	X	X	X	X	X	X	X	X	
Estudiante 25	1	1	3	3	1	1	2	2	X	X	X	X	X	X	X	X	
Estudiante 26	4	4	4	4	4	4	4	4	X	X	3	3	X	X	X	X	
Estudiante 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Estudiante 35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Note: E1 and E2 are Evaluator 1 and Evaluator respectively.

Tabla L10: Puntajes de la doble codificación de la rúbrica aplicada a la *Incorporación del Feedback en la New Response (NR(FR))* en la Actividad 1.

Student	Incorporation of FR into NR	
	Evaluator 1	Evaluator 2
Student 1	3	3
Student 2	-	-
Student 3	4	4
Student 4	-	-
Student 5	-	-
Student 6	-	-
Student 7	-	-
Student 8	-	-
Student 9	-	-
Student 10	4	4
Student 11	3	3
Student 12	-	-
Student 13	4	4
Student 14	-	-
Student 15	-	-
Student 16	1	1
Student 17	-	-
Student 18	3	3
Student 19	4	4
Student 20	-	-
Student 21	-	-
Student 22	-	-
Student 23	-	-
Student 24	3	3
Student 25	2	2
Student 26	4	4
Student 27	-	-
Student 28	-	-
Student 29	3	3
Student 30	-	-
Student 31	-	-
Student 32	-	-
Student 33	-	-
Student 34	-	-
Student 35	-	-

Tabla L11: Puntajes de la doble codificación de la rúbrica aplicada a la *Incorporación del Feedback en la New Response (NR(FR))* en la Actividad 2.

Student	Incorporation of FR into NR	
	Evaluador 1	Evaluador 2
Student 1	2	2
Student 2	2	3
Student 3	2	4
Student 4	2	2
Student 5	-	-
Student 6	4	4
Student 7	4	4
Student 8	-	-
Student 9	1	1
Student 10	3	3
Student 11	1	1
Student 12	3	3
Student 13	3	3
Student 14	2	2
Student 15	-	-
Student 16	3	3
Student 17	4	4
Student 18	3	3
Student 19	3	3
Student 20	1	1
Student 21	3	3
Student 22	-	-
Student 23	-	-
Student 24	-	-
Student 25	2	2
Student 26	-	-
Student 27	-	-
Student 28	-	-
Student 29	-	-
Student 30	-	-
Student 31	1	1
Student 32	-	-
Student 33	-	-
Student 34	4	4
Student 35	3	3

Tabla L12: Puntajes de la doble codificación de la rúbrica aplicada a la *Incorporación del Feedback en la New Response (NR(FR))* en la Actividad 3.

Student	Incorporation of FR into NR	
	Evaluador 1	Evaluador 2
Student 1	-	-
Student 2	2	2
Student 3	-	-
Student 4	4	4
Student 5	-	-
Student 6	-	-
Student 7	2	2
Student 8	-	-
Student 9	2	2
Student 10	1	1
Student 11	2	2
Student 12	2	2
Student 13	-	-
Student 14	-	-
Student 15	2	2
Student 16	3	3
Student 17	2	2
Student 18	3	3
Student 19	2	2
Student 20	3	2
Student 21	3	3
Student 22	3	3
Student 23	-	-
Student 24	3	3
Student 25	-	-
Student 26	-	-
Student 27	-	-
Student 28	-	-
Student 29	-	-
Student 30	-	-
Student 31	-	-
Student 32	-	-
Student 33	-	-
Student 34	-	-
Student 35	-	-

M. Guía para la Retroalimentación

HOW DO I COMPLETE THE DIFFERENT STAGES IN THE COLLABORATIVE PLATFORM?



To RESPOND...

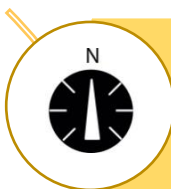


Identify the *skills* and *topics* addressed by the question:

ASK YOURSELF: What subject-matter knowledge are they asking me for? What information are they giving me? What information are they asking from me? Which critical thinking skill is being assessed? *The critical thinking skills are: interpretation, evaluation, analysis, inference and explanation.*

- **PREPARE YOURSELF:** The ideas, categories, arguments and relationships included in your response must be justified accordingly, using evidence from the situation, case, information or literature referenced in the question. To help you develop your response, prepare the evidence you'll need and think about how to integrate it so as to back up your ideas.

To GIVE MEANINGFUL FEEDBACK ...



CONGRATULATE, QUESTION AND GUIDE

Highlight the achievements in your classmate's response.

Guide them with strategies and tools for incorporating the suggested changes.

Provide new arguments or complement the information included in the response.



SPECIFY AND JUSTIFY

Comment on relevant and specific areas for improving the response.

Avoid giving unjustified personal views (or prescriptive corrections).

Recommend including evidence to support the ideas that are presented.



EXPLAIN

Mention which assessment criteria your comments fall under (formal, conceptual, structural).

Explain the question in other words if you see there has been a misunderstanding.

To REWRITE...

Interpret and make **inferences** based on the *concepts, suggestions* and *strategies* included in the feedback you received.

Evaluate and **analyze** how relevant these comments are and which strategies are best suited to your work.

Assess your work and include any aspects you found to be relevant in your new response.



To SELF-CHECK...

Does the response answer the question that is asked? Have you acknowledged and addressed all of the issues in the question?	
Does each paragraph contain a main idea that is both relevant and well-developed?	
Does the response use precise, varied, specialized and suitable vocabulary, considering both the register (academic) as well as the subject-matter knowledge?	
Have you avoided repetition and including unnecessary information?	
Have your views and opinions been justified and well-presented?	

BIBLIOGRAFÍA

Capítulo 1:

Anderson, L., & Krathwohl, D. (2001). A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy. New York: Longman., 137-175.

Asselin, M. (2016). Teaching information skills in the information age: An examination of trends in the middle grades. In *Librarians and Educators Collaborating for Success: The international Perspective* (p. 82).

Barkoczi, D., & Galesic, M. (2016). Social learning strategies modify the effect of network structure on group performance. *Nature communications*.

Bavelas, A. (1950). Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, 725-730.

Brabazon, T. (2016). *The University of Goolge: Education in the (post) information age*. Routledge.

Brame, C. (2013). Vanderbilt University Center for Teaching. Retrieved Febrero 13, 2018, from <http://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/>

Collier, K. (2012). *Agile analytics: A value-driven approach to business intelligence and data warehousing*. Addison-Wesley.

García, F., Portillo, J., Romo, J., & Benito, M. (2007). *Nativos digitales y modelos de aprendizaje*. SPDECE.

Gilboy, M., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of nutrition education and behavior*(47), 109-114.

Hajnal, P. I. (2018). *Civil society in the information age*. Routledge.

K. A., & M. C. (2009). *21st Century Skills and Competences for New Millennium Learners in OECD Countries*. OECD Publishing.

Kaufman, J. (2011). *The personal MBA: Master the art of business*. Penguin.

Leff, A., & Rayfield, J. (2001). Web-application development using the model/view/controller design pattern., (pp. 118-127).

MacMillan, J., Entin, E., & Serfaty, D. (2004). *Communication overhead: The hidden cost of team cognition. Team cognition: Process and performance at the interand intra-individual level*. American Psychological Association, Washington, DC. Available at http://www.aplima.com/publications/2004_MacMillan_EntinEE_Serfaty.pdf.

Manyica, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., et al. (2017, Enero). *UN FUTURO QUE FUNCIONA: AUTOMATIZACIÓN, EMPLEO Y PRODUCTIVIDAD*. Resumen Ejecutivo McKinsey Global Institute, 4.

Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., et al. (2017, Diciembre). *Jobs Lost, Jobs Gained: Workforce Transitions In A Time Of Automation*. McKinsey Global Institute.

Nielsen, J. (2003). *Usability 101: Introduction to usability*.

Partnership for 21st Century Learning. (2015, Mayo). *Framework for 21st Century Learning - P21*. Retrieved Febrero 12, 2018, from Partnership for 21st Century Learning: <http://www.p21.org/our-work/p21-framework>

Prensky, M. (2001). Digital Natives, Digital Immigrants. On the Horizon.

Statista. (2018, Septiembre 10). Development of the world population from 1950 to 2030 (in billions). Retrieved from The Statistics Portal: <https://www.statista.com/statistics/262875/development-of-the-world-population/>

Statista. (2018, Septiembre 10). Number of internet users worldwide from 2005 to 2017 (in millions). Retrieved from The Statistics Portal: <https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/>

Steen-Utheim, A., & Foldnes, N. (2017, Septiembre). A qualitative investigation of student engagement in a flipped classroom. *Teaching in Higher Education*, 1-18.

Vioreanu, D. (2018, Febrero). Distance Learning Course Options Get More Popular Every Year Where to study . Retrieved from Study Portals: <https://www.distancelearningportal.com/articles/393/distance-learning-course-options-get-more-popular-every-year.html>

Capítulo 2:

Airasian, P. W. (2001). *Classroom assessment: Concepts and applications*. McGraw-Hill, PO Box 548, Blacklick, OH 43003.

Alqassab, M., Strijbos, J. W., & Ufer, S. (2018). Training peer-feedback skills on geometric construction tasks: role of domain knowledge and peer-feedback levels. *European Journal of Psychology of Education*, 33(1), 11-30.

Akgun, A., & Duruk, U. (2016). The Investigation of Preservice Science Teachers' Critical Thinking Dispositions in the Context of Personal and Social Factors. *Science Education International*, 27(1), 3–15. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1100164.pdf>

Apedoe, X. S., Long, S. A., Morris, J. A., Wilson, A. A., Morris, R. J., Kroeger, S. D., ... & Strycker, J. (2017). Flipping Education. In *The Flipped College Classroom* (pp. 89-123). Springer, Cham.

Apedoe, X. S., Long, S. A., Morris, J. A., Wilson, A. A., Morris, R. J., Kroeger, S. D., ... (2017). Flipping Education. In *The Flipped College Classroom* (pp. 89-123). New York: Springer International Publishing.

Arsal, Z. (2015). The effects of microteaching on the critical thinking dispositions of pre-service teachers. *Australian Journal of Teacher Education* (Online), 40(3), 140.

Arsal, Z. (2017). The impact of inquiry-based learning on the critical thinking dispositions of pre-service science teachers. *International Journal of Science Education*, 39(10), 1326–1338. <https://doi.org/10.1080/09500693.2017.1329564>

Baker, K. M. (2016). Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education*, 17(3), 179–192. <https://doi.org/10.1177/1469787416654794>

Barak, M. (2017). Science Teacher Education in the Twenty-First Century: a Pedagogical Framework for Technology-Integrated Social Constructivism. *Research in Science Education*, 47(2), 283–303. <https://doi.org/10.1007/s11165-015-9501-y>

Berndt, M., Strijbos, J. W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, 33(1), 31-49.

Berndt, M., Strijbos, J. W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, 33(1), 31-49.

Betancourth Zambrano, S. (2015). (2015). Desarrollo del pensamiento crítico en docentes universita- rios. Una mirada cualitativa. *Revista. Revista Virtual Universidad Católica Del Norte*, 44, 238–252.

Bormann, J. (2014). Affordances of Flipped Learning and Its Effects on Student Engagement and Achievement, 1–38.

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712.

Brookhart, S. M. (2017). How to give effective feedback to your students. ASCD.

Brown, A. E. (2016). Critical Thinking to Justify an Answer in Mathematics Classrooms.

Cowan, J. (1999). Being an innovative university teacher (Buckingham, SRHE & Open University Press).

Cullen, T. A., & Karademir, T. (2018). Technology Experiences of Student Interns in a One to One Mobile Program. *International Journal of Mobile and Blended Learning (IJMBL)*, 10(2), 20-35.

De la Cruz, M. S. D., Kopec, M. T., & Wimsatt, L. A. (2015). Resident perceptions of giving and receiving peer-to-peer feedback. *Journal of graduate medical education*, 7(2), 208-213.

DeWaelsche, S. A. (2015). Critical thinking, questioning and student engagement in Korean university English courses. *Linguistics and Education*, 32, 131–147.
<https://doi.org/10.1016/J.LINGED.2015.10.003>

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, 69, 485-492.

Dunne, G. (2015). Beyond critical thinking to critical being: Criticality in higher education and life. *International Journal of Educational Research*, 71, 86–99.
<https://doi.org/10.1016/J.IJER.2015.03.003>

Ennis, R. H. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi*, 37(1), 165–184. <https://doi.org/10.1007/s11245-016-9401-4>

Escobar-Pérez, J., & Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición*, 6(1), 27-36.

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation. *Computers & education*, 106, 83-96.

Facione, P. A. (1990). Critical Thinking : A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction Executive Summary “ The Delphi Report. *The California Academic Press*, 423(c), 1–19. <https://doi.org/10.1016/j.tsc.2009.07.002>

Gibbs, G., & Simpson, C. (2004). Does your assessment support your students' learning. *Journal of Teaching and learning in Higher Education*, 1(1), 1-30.

Hammersley-Fletcher, L., & Hanley, C. (2016). The use of critical thinking in higher education in relation to the international student: Shifting policy and practice. *British Educational Research Journal*, 42(6), 978-992.

Hanrahan, S. J., & Isaacs, G. (2001). Assessing self-and peer-assessment: The students' views. *Higher Education Research & Development*, 20(1), 53-70.

Harland, T., Wald, N., & Randhawa, H. (2017). Assessment & Evaluation in Higher Education Student peer review: enhancing formative feedback with a rebuttal Student peer review: enhancing formative feedback with a rebuttal. *Assessment & EvAluAtion in HigHer EducAtion*, 42(5), 801–811.
<https://doi.org/10.1080/02602938.2016.1194368>

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.

Huisman, B., Saab, N., van Driel, J., & van den Broek, P. (2018). Peer feedback on academic writing: undergraduate students' peer feedback role, peer feedback perceptions and essay performance. *Assessment & Evaluation in Higher Education*, 1-14.

Hwang, G. J., Lai, C. L., Liang, J. C., Chu, H. C., & Tsai, C. C. (2018). A long-term experiment to investigate the relationships between high school students' perceptions of mobile learning and peer interaction and higher-order thinking tendencies. *Educational Technology Research and Development*, 66(1), 75-93.

Jong, M. S. Y. (2017). Empowering Students in the Process of Social Inquiry Learning through Flipping the Classroom. *Educational Technology & Society*, 20 (1), 306–322.

Keane, T., Keane, W. F., & Blicblau, A. S. (2016). Beyond Traditional Literacy: Learning and Transformative Practices Using ICT. *Education and Information Technologies*, 21(4), 769-781.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.

Kim, K., Sharma, P., Land, S. M., & Furlong, K. P. (2013). Effects of active learning on enhancing student critical thinking in an undergraduate general science course. *Innovative Higher Education*, 38(3), 223-235.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, 65(23), 2276-2284. (p. 2276)

Kostiainen, E., Ukskoski, T., Ruohotie-Lyhty, M., Kauppinen, M., Kainulainen, J., & Mäkinen, T. (2018). Meaningful learning in teacher education. *Teaching and Teacher Education*, 71, 66–77. <https://doi.org/10.1016/J.TATE.2017.12.009>

Lewicki, P., & Hill, T. (2006). Statistics: methods and applications. *Tulsa, OK. Statsoft*.

Li, L. (2017). The role of anonymity in peer assessment. *Assessment & Evaluation in Higher Education*, 42(4), 645-656.

Lin, G.-Y. (2018). Anonymous versus identified peer assessment via a Facebook-based learning application: Effects on quality of peer feedback, perceived learning, perceived fairness, and attitude toward the system. *Computers & Education*, 116, 81–92
<https://doi.org/10.1016/J.COMPEDU.2017.08.010>

Liu, J., McBride, R. E., Xiang, P., & Scarmardo-Rhodes, M. (2018). Physical education pre-service teachers' understanding, application, and development of critical thinking. *Quest*, 70(1), 12-27.

Lo, C. K., & Hew, K. F. (2017). Using" First Principles of Instruction" to Design Secondary School Mathematics Flipped Classroom: The Findings of Two Exploratory Studies. *Journal of Educational Technology & Society*, 20(1).

Luciano, R. D., Silva, B. L., Rosa, L. M., & Meier, H. F. (2018). Multi-objective optimization of cyclone separators in series based on computational fluid dynamics. *Powder Technology*, 325, 452-466.

Lund-Diaz, S., Montane, M., & Beery, P. (2016). “How”—The Key to Knowledge-Building Pedagogy Success in Supporting Paradigm Shifts for Student Growth and the 4Cs of Future Education. In *Google It* (pp. 353–362). New York, NY: Springer New York.

https://doi.org/10.1007/978-1-4939-6415-4_15

Mäkinen, T. (2018). Meaningful learning in teacher education. *Teaching and Teacher Education*, 71, 66–77. <https://doi.org/10.1016/J.TATE.2017.12.009>

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, 22(3), 276-282. (p. 276)

Mehring, J. G. (2015). An exploratory study of the lived experiences of Japanese undergraduate EFL students in the flipped classroom. Pepperdine University.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25), 1-10.

Mulder, R., Baik, C., Naylor, R., & Pearce, J. (2014). How does student peer review influence perceptions, engagement and academic outcomes? A case study. *Assessment & Evaluation in Higher Education*, 39(6), 657-677.

Murphy, C., Klotz, A. C., & Kreiner, G. E. (2017). Blue skies and black boxes: The promise (and practice) of grounded theory in human resource management research. *Human Resource Management Review*, 27(2), 291-305..

Nagy, W. (1988). Teaching vocabulary to improve reading comprehension. Urbana, IL: ERIC/RCS, NCTE & IRA

Newman, L. R., Roberts, D. H., & Schwartzstein, R. M. (2016). Peer Observation of Teaching Handbook. Shapiro Institute for Education and Research at Harvard Medical School and Beth Israel Deaconess Medical Center,
<https://hms.harvard.edu/sites/default/files/assets/Sites/Academy/files/MedEdPortalPeer%20observation%20handbook.pdf>

Nussbaum, M., & Diaz, A. (2013). Classroom logistics: Integrating digital and non-digital resources. *Computers & Education*, 69, 493-495.

Olasina, G. (2018). Human and social factors affecting the decision of students to accept e-learning. *Interactive Learning Environments*, 1-14.

Panadero, E., & Brown, G. T. L. (2017). Teachers' reasons for using peer assessment: positive experience predicts use. *European Journal of Psychology of Education*, 32(1), 133–156. <https://doi.org/10.1007/s10212-015-0282-5>

Planas-Lladó, A., Feliu, L., Castro, F., Fraguell, R. M., Arbat, G., Pujol, J., ... & Daunis-i-Estadella, P. (2018). Using peer assessment to evaluate teamwork from a multidisciplinary perspective. *Assessment & Evaluation in Higher Education*, 43(1), 14–30.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2), 119-144.

Saxton, E., Belanger, S., & Becker, W. (2012). The Critical Thinking Analytic Rubric (CTAR): Investigating intra-rater and inter-rater reliability of a scoring mechanism for critical thinking performance assessments. *Assessing Writing*, 17(4), 251-270.

Sinclair, S., Hack, T. F., Raffin-Bouchal, S., McClement, S., Stajduhar, K., Singh, P., ... & Chochinov, H. M. (2018). What are healthcare providers' understandings and experiences of compassion? The healthcare compassion model: a grounded theory study of healthcare providers in Canada. *BMJ open*, 8(3), e019701.

Spiller, D. (2012). Assessment matters: Self-assessment and peer assessment. The University of Waikato.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.

Stupples, E. J. N., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y. F., & Aubeeluck, A. V. (2017). Development of the Critical Thinking Toolkit (CriTT): A

measure of student attitudes and beliefs about critical thinking. *Thinking Skills and Creativity*, 23, 91–100. <https://doi.org/10.1016/j.tsc.2016.11.007>

Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35-39.

Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: teaching strategies. *Learning and instruction*, 14(4), 359-379.

Unruh, T., Peters, M. L., & Willis, J. (2016). Flip This Classroom: A Comparative Study. *Computers in the Schools*, 33(1), 38–58. <https://doi.org/10.1080/07380569.2016.1139988>

Vaughan, M. (2014). Flipping the Learning : An Investigation into the use of the Flipped Classroom Model in an Introductory Teaching Course. *Education Research and Perspectives*, 41, 25–41.

Van Gennip, N. A., Segers, M. S., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction*, 20(4), 280-290.

Vásquez-Colina, M. D., Russo, M. R., Lieberman, M., & Morris, J. D. (2017). A case study of using peer feedback in face-to-face and distance learning classes among pre-service teachers. *Journal of Further and Higher Education*, 41(4), 504-515.

Villalón, M., & Bedregal, P. (2017). Alfabetización inicial: claves de acceso a la lectura y escritura desde los primeros meses de vida. Ediciones UC.

Walshe, M., Keane, C., Gallagher, J., McDonald, K., O'Hanlon, R., & Ledwidge, M. (2016). 21 Communication difficulties in the care pathway of patients admitted to hospital with acute decompensated heart failure.

Wiggins, G. (2012). Seven keys to effective feedback. 2012, 70(1), 11-16.

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17-37.

Yousefi, S., & Mohammadi, M. (2016). Critical thinking and reading comprehension among postgraduate students. *Journal of Language Teaching and Research*, 7(4), 802-807.

Yang, Y. T. C., Chuang, Y. C., Li, L. Y., & Tseng, S. S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. *Computers & Education*, 63, 285-30

Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior*, 70, 251-260.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, 45(4), 477-501.

Zaidi, N. L. B., Grob, K. L., Monrad, S. M., Kurtz, J. B., Tai, A., Ahmed, A. Z., ... & Santen, S. A. (2018). Pushing Critical Thinking Skills With Multiple-Choice Questions: Does Bloom's Taxonomy Work?. *Academic Medicine*, 93(6), 856-859.

ANEXOS

A- ARTÍCULO TRADUCIDO

TECHNOLOGY-SCAFFOLDED PEER ASSESSMENT FOR DEVELOPING CRITICAL THINKING IN PRE-SERVICE TEACHER TRAINING

ABSTRACT

Critical thinking is rarely developed systematically in university classrooms and the case of pre-service teacher training is no different. Despite a trend towards constructivism, most institutions still continue to promote teacher-centered models. Developing critical thinking among trainee teachers is becoming increasingly important, as is giving and receiving feedback during the learning process. This paper therefore aims to answer the question “How can technology-scaffolded peer assessment activities develop critical thinking among pre-service teachers?” In order to do so, a series of practice and application activities were introduced using technology-scaffolded peer assessment. This was done in a flipped classroom model as part of a reading methods course taught at a school of education. Thirty students participated in three activities across a single semester. These activities took into account the students’ knowledge of both pedagogical theory and practice, as well as the development of one or more critical thinking skills. The technological scaffolding helped minimize the classroom logistics, while at the same time resolving any personal issues between peers as the tasks were assigned at random. Mixed-methods analysis revealed that technology-scaffolded peer assessment with anonymous feedback aided the significant development of critical thinking activities. It also showed that the feedback that was given and received was a predictor of the success of these activities. This study provides a model for developing critical thinking in the classroom, as well as guidance for pre-service teacher trainers on how to incorporate these practices.

Highlights

Pre-service teachers require subject-matter knowledge and critical thinking skills.

Randomized and anonymous peer assessment allows for the development of critical thinking.

The phases of responding, giving feedback, and re-drafting the response are key.

Receiving and giving feedback has the greatest impact on improvements in critical thinking.

Networked devices allow for seamless record management and anonymous peer review

1. Introduction

Critical thinking, one of the so-called 21st century skills, is understood as “purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference” (Facione, 1990, p.4). Activities that promote critical thinking place the student at the center of the learning process (DeWaelsche, 2015), with the teacher acting as a facilitator by providing the necessary scaffolding to develop critical thinking (Lajoie et al., 2001, in Kim, 2013).

Critical thinking is a fundamental requirement of professional life (Betancourth, 2014) and is considered to be one of the most desirable outcomes of higher education (Dunne, 2015). Although there have been attempts to make critical thinking the focus of higher education (Stupple, 2016), it is a skill that is rarely developed systematically in university classrooms (Hammersley, 2016). This is particularly relevant when considering how much students struggle with critical thinking (Ten Dam et al., 2004).

One of the reasons behind the failure to develop critical thinking skills is linked to the teaching methods. These tend to be traditional, teacher-centered approaches, where the teacher uses the class time to develop lower order thinking skills (such as remembering and understanding), while leaving higher order thinking skills (such as applying, analyzing and evaluating) for outside the classroom (Bormann, 2014). University classrooms rarely provide an environment in which students can become actively involved (Mehring, 2016) through student-centered learning (DeWaelsche, 2015). By having the students take on a passive role (Mehring, 2015), the teacher expects the students to practice independently (Unruh et al., 2016).

Initial teacher training should focus even more on fostering pre-service teachers’ pedagogical and critical thinking skills (Kostiainen et al., 2018). By combining critical thinking and subject-matter instruction, students can acquire both sets of skills more effectively (Ennis, 2018). Despite a trend towards constructivism in recent decades, teacher training institutions continue to promote teacher-centered models, where students are merely ‘vessels’ of knowledge (Vaughan, 2014). The limited development of future teachers’ critical thinking skills is due to several different factors. This includes a lack of exposure to critical thinking during their own schooling and at university (Arsal, 2017). Future teachers are rarely motivated to find different ways of solving a problem, preferring instead to memorize the solution step-by-step without understanding the *how* or *why* of the process (Brown, 2016). In order for future teachers to apply strategies that promote the development of critical thinking, they must first acquire and practice these skills themselves (Akgun et al., 2016).

One of the key mechanisms to help future teachers learn is feedback. The impact of feedback has been widely studied and is currently an area of focus within education (Boud & Molloy, 2013a; Evans, 2013; Nicol, 2010, in Winstone, 2017). This is because it is considered one of the factors that has the biggest impact on student learning (Hattie, 2007). Indeed, social constructivism considers giving and receiving feedback to be one of its pillars (Barak, 2017). In this sense, the student plays a leading role, as by giving and receiving feedback they become actively involved in the learning process (Harland et al., 2017). This also allows them to construct their own knowledge (Lin, 2018).

Giving feedback among peers is a valuable activity for learning (Harland et al., 2017) as the students can compare their own work with a classmate's and make suggestions based on their own experience (Baker, 2016). By helping their classmates develop their skills and understanding in this way (Winstone et al., 2017) the students are able to learn from the activity itself, as well as from their peers (van Gennip et al., 2010). However, students tend to think that they do not possess the necessary skills or do not feel prepared for carrying out peer assessment (Mulder et al., 2014). It is therefore essential that students should be guided in how to give feedback in order for this feedback to be high quality (Baker, 2016). In this case, teachers play a critical role in mediating the implementation of peer assessment by drawing attention to the most important areas and developing the necessary rubrics (Panadero & Brown, 2017).

Given the growing importance of critical thinking in pre-service teacher training, as well as the importance of feedback within the learning process, the aim of this paper is to answer the question "How can technology-scaffolded peer assessment activities develop critical thinking among pre-service teachers?"

2. Methodology

2.1. Chosen strategy: Scaffolded peer assessment

The peer assessment strategy analyzed in the present article was adopted as part of a flipped classroom model. The flipped classroom approach was used to transform learning in the classroom from a passive to an active process (Johnson et al., 2015). Within this model, the classroom time (i.e. teacher-student interaction) is maximized (Apedoe & Col, 2017). In this sense, the 'Connect', 'Comprehend' and 'Express' stages of the learning process are addressed through individual, out-of-class activities. During this time, video lectures and formative assessment are used to develop the students' own opinions and perspectives. Subsequently, the 'Wonder', 'Construct' and 'Reflect' stages

of the process are addressed through in-class work, with the students analyzing and evaluating a topic, with guidance from the teacher (Jong, 2017).

Out-of-class activities. We considered the use of videos and formative assessment as an activation activity (Lo & Hew, 2017). This included between 3 and 5 videos per week, each lasting between 5 and 8 minutes, with a total of 16 weeks of out-of-class activities. During these activities the students could watch, pause and rewind the videos as many times as they needed. Systematic feedback on student achievement has been shown to be more effective when given to students than when given to teachers (Faber et al., 2017). Therefore, formative assessment was included at the beginning of each in-class session as an activation activity. Each assessment included three questions, with sections on subject-matter knowledge and reasoning. Following the assessment, the students were then provided with personalized feedback. The reasoning section was designed to help the students develop their critical thinking skills, while at the same time reviewing the contents being evaluated (Lund-Diaz et al., 2016).

In-class work. The first objective of the in-class work was to review the out-of-class activities and provide feedback. Therefore, during the activation activities the teacher focused on the topics and higher-order skills that the students struggled with the most during the out-of-class activities, as informed by a report from the formative assessment for the teacher. Using this information, the teacher worked with the students on the required areas through an interactive Classroom Response System.

Technology-scaffolded peer assessment activities were then used for practice and application activities (Lo & Hew, 2017) during class. This involved the students working on networked devices (tablets or laptops). These activities were set within the context of teaching and learning reading skills and helped the future teachers develop their critical thinking skills. This includes their ability to develop a proposal and comprehend its significance [*interpretation*], to identify relationships among concepts [*analysis*], to assess proposals or possible solutions [*evaluation*], to clarify and articulate the most relevant information [*inference*] and to justify their reasoning [*explanation*] (Facione, 1990). This study analyzes the learning sequence that is followed during activities involving technology-scaffolded peer assessment.

2.1.1. Scaffolded peer assessment

The individual work completed by the students reveals the extent to which they acquire the pedagogical and practical knowledge, as well as the development of their critical thinking skills. This is done in three phases, Figure 1:

Phase 1: Developing a response. Based on Facione's framework for developing critical thinking (Facione, 1990; 1998), the students are encouraged to develop a series of skills when providing their response (R), R1, R2, and R3, respectively (Figure 1). This includes their ability to (1) acknowledge, interpret, integrate and develop the skills used to complete the task, (2) identify and understand the general or specialized pedagogical knowledge that is required (this is stated in the question and is also worked on during class through discussions of the reading material and its application to teaching/learning), (3) adapt their response to the appropriate format and ensure it meets the requirements of the task, and (4) prepare and present the necessary evidence in order to justify and support their response. In order to achieve this, the task sometimes required the students to read around the topic beforehand, as well as studying the content of the video lectures.

Phase 2: Giving feedback. The student's response (R) is sent to another student, who must evaluate the response and provide feedback in a second document. In Figure 1, Student 1 provides feedback to Student 3 (F3), while Student 2 provides feedback to Student 1 (F1), and so on. Following the teacher's recommendations, the student is expected to give feedback that is both effective and meaningful, building on the initial response (R) (Shute, 2008) by providing comments and suggestions that guide their classmate. The student giving the feedback should also show that they have considered and applied specific criteria when developing their response. Furthermore, the feedback that is given should include explanations and strategies for improving the initial response (R), providing clear, relevant and specific guidelines for doing so (Kluger & De Nisi, 1996). To help the future teachers with this process, a set of guidelines for giving feedback was developed, (Appendix M). These guidelines helped the students assess the information and its credibility using set criteria. They also helped them to work through their peers' logic and reasoning, while also detecting and commenting on any personal opinions and justifying this with evidence (Algassab et al., 2018). However, a person's opinion of their peers can have an effect on the feedback that is given and therefore on its reliability (Poulos & Mahony, 2008 in Berndt, Strijbos & Fischer, 2018). It was thus decided to make the feedback anonymous (and randomized) in order to reduce the impact of any socio-affective characteristics (Li, 2017). Furthermore, this also helped develop a positive attitude towards the task (Lin, 2018), as well as improving student motivation (Berndt et al., 2017). In this sense, the online platform designed for this study effectively allowed this process to be both anonymous and randomized.

Phase 3: Developing a new response. Based on their experience of giving and receiving feedback, the students then develop a new response (NR). In Figure 1, when developing their new response (NR1), Student 1 takes into account their initial response (R1), as well as the feedback received from Student 2 (F1) and their own experience of giving feedback (F3) to another student's response (R3). The process of developing a new response (NR) calls on the student's ability to think critically by evaluating and analyzing the relevance of the feedback that they received. They are also required to interpret and make inferences based on this feedback, as well as on their own experience of giving feedback (Boud, 2013). During this phase, the students are expected to acknowledge and incorporate any relevant suggestions made by their peers (F1). They are also expected to provide clear and logical evidence to justify any changes that they make to their response. In this sense, the student should build on their response (R) when developing their new response (NR), and not simply copy-paste the information. Instead, they should reword their response based on an analysis and evaluation of the relevant resources (Gibbs & Simpson, 2004).

During the activity, each of the students must develop a response, as well as providing feedback (Figure 1). Furthermore, each student is also involved with two other peers (i.e. the student to whom they give feedback and the student from whom they receive it). This process is managed by the digital platform with complete transparency. Peer assessment is key to our proposal as it actively involves the student with two of their peers (Phielix et al., 2010, in Alqassab et al., 2018). Furthermore, the proposal follows a constructivist approach as the new response builds on the knowledge provided by the student's peers. In this sense, peer assessment is an essential part of the process (Lin, 2018).

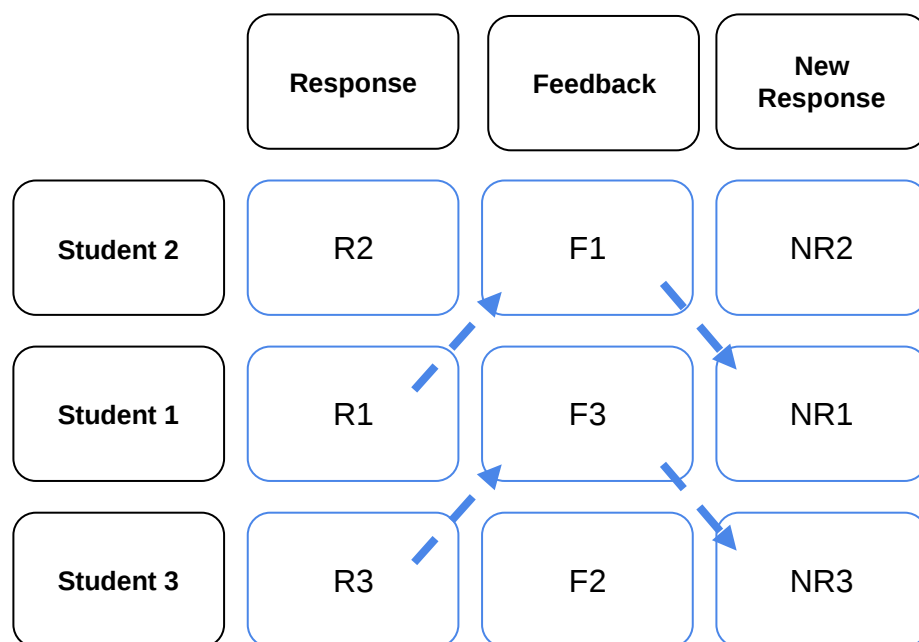


Figure 1: Activity for developing a response and engaging in peer assessment using a system of networked devices

The teacher plays a key role in this process. They not only define which tasks are to be completed; they also specify the criteria to be met during each phase and give feedback to the students during the different stages of the activity (Yorke, 2003).

The proposed scaffolding is facilitated by technology. Any interpersonal issues between peers are therefore eliminated as the tasks are assigned randomly using the technology. The technology also helps minimize the classroom logistics.

2.2. Context and participants

Communication is considered one of the most important 21st century skills (Keane et al., 2016). Previous studies have also revealed the relationship between critical thinking and reading comprehension skills (Yousefi & Mohammadi, 2016). Given this, we decided to focus our research on a reading methods class.

The study followed a mixed-methods approach and involved 30 students (2 males and 28 females) from a second-year (undergraduate) reading methods class taught at a

Chilean university. The class was part of a general education program, while the university's curriculum is notoriously hands-on. The age of the students ranged between 20 and 24 years, while the teacher who led the intervention was 42 years old. The teacher had been teaching the class for five years and is constantly looking for new methodologies to help optimize her students' learning. She also reported having previous experience of using technology in the classroom.

The main contents of the course included the following: theoretical and empirical knowledge of reader development; assessment and key practices for developing code-based skills; assessment and key practices for developing meaning-construction skills, with an emphasis on the role of academic language and vocabulary in the comprehension of complex texts; learning how to facilitate an academic, text-based discussion; and developing reading strategies for monitoring comprehension among 1st to 6th grade students. The classload included 10 hours of work per week, with three 80-minute sessions of class time and 6 hours of independent study.

The three 80-minute sessions of technology-scaffolded peer assessment were held in March, May and June. A description of the activities that were carried out can be found in Appendix A. Furthermore, a summary of the type of knowledge required by these activities, as well as the critical thinking skills that were developed and the number of sub-tasks involved, can be found in Table 1.

Table 1: *Summary of the activities*

	Activity 1	Activity 2	Activity 3
Knowledge required	- Pedagogical subject-matter knowledge	- Pedagogical subject-matter knowledge	- Pedagogical subject-matter knowledge - Specialized knowledge of generative learning: productive discussion
Critical thinking skills developed	Analysis and Explanation	Explanation, Interpretation and Inference	Inference and Evaluation
Number of sub-tasks	5	4	4

2.3. Assessment instruments

The effectiveness of the intervention was assessed using three rubrics (Appendix C) based on Facione's model of critical thinking and adapted by the researchers. The rubrics that were developed represent an analytical assessment model as they break a series of assessment criteria down into several indicators and levels of achievement (or levels of completion) (Airasian & Russell, 2001; Mertler, 2001). This type of assessment tool facilitates detailed analysis of the students' skills and abilities. It also allows the researchers to observe their work and progress, as well as producing objective responses that help reduce any subjectivity among the evaluators (Saxton, Belanger & Becker, 2012).

Developing these instruments involved a three-stage process, based on three fundamental aspects of pedagogy and assessment. The first of these, *Stage 1*, involved developing assessment criteria covering three dimensions: (1) the practical-textual dimension, which looked at how the students developed their responses; (2) critical thinking skills based on an analysis and adaptation of Facione's (1990) model of critical thinking, which looked at whether or not the students' work fit the model; and (3) feedback, which in turn comprised two sub-dimensions: a breakdown of the higher-order skill of 'evaluation' (along with other formal aspects) and a series of conceptual indicators based on the evaluator's intentions. The second stage, *Stage 2*, involved defining a series of levels of achievement to assess a learner's progressive development (Mertler, 2001). These levels included: fully achieved (4 points), developing (3 points), emerging (2 points) and not achieved (1 point). Finally, *Stage 3* involved developing a series of level descriptors. In this sense, the differences between levels could be described in two ways: a formal-quantitative description and an analytical-qualitative description. In the case of the formal-quantitative description, progressing from one level to another suggested a lack or only partial presence of the knowledge required by the activity (i.e. a lack of information). In the case of the analytical-qualitative description, on the other hand, the difference between levels addressed issues with the student's level of analysis or a failure to fully develop their response in conceptual terms.

Appendix B details the criteria used to define the different critical thinking skills (Interpretation, Analysis, Evaluation, Inference and Explanation), as well as the different dimensions included in the rubrics to assess the Initial Response (R) and the New Response (NR) (Table C-1, Appendix). In addition to this, two further rubrics were also developed to assess the Feedback (F) (Table C-2, Appendix) and the extent to which the Feedback was incorporated into the New Response (F in NR) (Table C-3, Appendix).

Appendix C-1 outlines the rubric used for assessing R and NR. The aim of this rubric is to assess whether the student's responses in stage 1 (R) and stage 3 (NR) demonstrate suitable development of the skills set out in the question. The rubrics themselves were based on the five higher order thinking skills that were covered in the activities (i.e. interpretation, analysis, evaluation, inference and explanation) (Facione, 1990). Consequently, if the question set by the teacher focused on interpretation and analysis, for example, these were the dimensions that were assessed in the student's response. Therefore, the final score for each activity depended on the number of skills involved in the question, e.g. if the question involved two skills, the maximum score was 8 (4 points for each skill).

Appendix C-2 shows the rubric used for assessing the Feedback (F), which is split into two main dimensions. The first of these dimensions consisted of the assessment criteria (i.e. the student's ability to propose specific criteria for their peers), the clarity and relevance of the explanation (i.e. the appropriateness of the comment and whether it effectively addressed the task at hand rather than being based on personal or unfounded opinions), the reasons given (i.e. the student's ability to provide constructive feedback, with strategies for improving the response rather than just a description), and the structure of the feedback (i.e. whether the feedback was logically sequenced).

Given its importance in evaluative terms, this dimension was considered compulsory and applied to all responses (**Wiggins, 2012**). As **Sadler (1989)** explains, certain essential elements are required in order for feedback to be effective and meaningful and to have an impact on student learning. This includes clearly developing a specific objective, as well as developing evaluative criteria and presenting explicit strategies. Doing so can help a student make progress and improve on their initial level.

The decision of whether or not to apply the indicators from the second dimension was optional. This is because it depended on the content of the initial response, R, or of the new response, NR, as well as the decisions made by the student giving the feedback (**Cowan, 1999**). These optional indicators can also be found in Appendix C-2. In this case, for example, if the student's response did not include the evidence required by the activity, the evaluator would assign the level 'not achieved' (i.e. no evidence was given). In short, the evaluator has to evaluate the applicability of each indicator when assessing the response (**Brookhart, 2013**).

Finally, Appendix C-3 shows the criteria used to assess the process of incorporating the comments from F into the new response, NR. According to **Facione (1990)**, this refers

to the concept of **metacognition**. The new response, NR, is a reformulation of the initial response, R. This reformulation is the product of a self-assessment based on external feedback (F), as well as on the student's own experience of giving feedback on another student's response. This process of incorporating the feedback into the new response therefore calls on the full range of critical thinking skills. In this sense, the student has to interpret and infer based on the comments received in F, evaluate and analyze how relevant the comments in F are to the initial response R, before explaining any potential changes between R and NR. In order to evaluate this process, the rubric for metacognition looks at how the students acknowledge and incorporate both the positive (i.e. relevant) and negative (i.e. irrelevant) feedback given by the evaluator in F. It also looks at whether the students give clear and coherent evidence to justify any potential changes (i.e. giving reasons for adding, removing or retaining certain information). Furthermore, it looks at whether the response addresses any potential weaknesses highlighted by the evaluator and explains the steps taken in order to improve the response (**Spiller, 2012**).

The rubrics that were developed for this study were validated by experts (Escobar-Perez et al., 2008). Three subject-matter experts validated the rubrics independently through a questionnaire using a Likert scale to indicate their level of agreement or disagreement with each descriptor.

Finally, the students were given two assessments. Test 1 (T1) corresponds to the students' grades from the test taken before activities 2 and 3 had been completed. Test 2 (T2) corresponds to the students' grades from the test taken after all three activities had been completed. The tests comprised 14 and 17 questions, respectively, including a mix of open-ended and multiple-choice questions.

2.4. Applying the rubrics

For double-coding, all of the students' responses (R, F and NR) were separately assessed by two evaluators (Appendix L). Firstly, the rubric was applied to each of the responses (R) for the three activities (Tables L-1, L-2 and L-3, Appendix L). Subsequently, each of the new responses (NR) to the three activities was then assessed (Tables L-4, L-5 and L-6, Appendix L). This was done so as to ensure that the assessment of the initial responses (R) would not hinder an objective assessment of the new responses (NR). Following this, the rubric was then applied to the feedback given for each of the three activities (Tables L-7, L-8 and L-9, Appendix L). Finally, the process of incorporating

this feedback (F) into the new responses (NR) was then assessed for each of the three activities (Tables L-10, L-11 and L-12, Appendix L).

When assessing the quality of an instrument, the validity and reliability of the results is key (Kimberlin & Winterstein, 2008). Our instrument was therefore validated using Cohen's kappa (McHugh, 2012). In the specific case of the rubrics used in this study (Appendix C), the level of agreement between the two evaluators regarding a particular response is compared against the levels of achievement (fully achieved, developing, emerging and not achieved). Appendix D outlines in detail the process for calculating Cohen's Kappa for each assessment made by the two evaluators. Tables D1 to D5 are grids showing the level of agreement between the evaluators, for each of the five critical thinking skills that were assessed for both R and NR. Tables D6 to D13 are grids showing the level of agreement when using the rubric to assess the feedback (F). Table D14, in turn, is a grid showing the level of agreement when using the rubric to assess how the students incorporated the feedback (F) into their new responses (NR). In order to find $\sum a$, we must determine the sum of the diagonals in such a grid. It is worth noting that there is only one table for each dimension of the rubric for R and NR as the same rubric is used for these two stages. Table D15 shows the expected frequencies for each criteria, as well as each of the dimensions included in the three rubrics. Finally, Table D16 shows that Cohen's for the dimensions assessed by the rubric for R and NR is 0.83. Furthermore, the dimensions assessed by the rubric for feedback (F) returned a Cohen's kappa of 0.76. While the rubric for assessing how the students incorporated F into NR returned a Cohen's kappa of 0.92.

Following this review and the process of double-coding, any differences were then analyzed and discussed, while potential issues with the criteria were resolved.

2.5. Quantitative analysis

Statistical analysis and linear regression modelling were performed using STATA (v. 12.0).

2.5.1. Stepwise Regression Model

To discover which factors have an impact on the development of critical thinking (Table 2), a linear regression model was developed using forward stepwise selection (Table 3). This method allows the relationship between different factors to be explored. Based on

this, a relative comparison is then made in order to identify the most relevant factors (Lewicki et al., 2006, in Luciando et al., 2018).

The model was built around the dependent variable *Improvement (I)*. This variable measures improvements in critical thinking skills during the activities, i.e. the differences in student achievement between the *Response (R)* and the *New Response (NR)*. The model also included the variables *Response (R)* and *Feedback (F)*, which was divided into the feedback each student received (*FR*) and the feedback they gave (*FG*). Furthermore, the model also included a variable to measure the extent to which students incorporated the feedback they received into their new response (*NR(FR)*), as well as their scores on Test 1 (*T1*) and Test 2 (*T2*). These tests measured the students' knowledge of the pedagogical subject-matter covered throughout the course.

Table 2 provides a breakdown of the different dimensions of the dependent and independent variables mentioned above. In this sense, the dependent variable *I* is divided into I_h , I_n and $\bar{I}$, while the independent variable *R* is divided into R_h , R_n and $\bar{R}$. Similarly, *NR* is divided into NR_h , NR_n and $\overline{NR}$, *F* is divided into F_n and $\bar{F}$, and *NR(FR)* is divided into $NR(FR)_n$ and $\overline{NR(FR)}$. *T1* and *T2* are not divided into any further dimensions. Table 2 outlines the objective of each variable, how each of the dimensions is calculated, and where a breakdown of the data can be found.

Table 2

Variables included in the linear regression model

Variable	Objective	Calculation method	Further details
I_h	Measure improvements in critical thinking skills for each activity	Difference between NR_h and R_h for each skill in each activity	Appendix E Table E4
I_n	Measure average improvements in critical thinking skills for each activity	Difference between the average for <i>new response</i> , NR_n , and the average for <i>response</i> , R_n	Appendix E Table E4
$\bar{I}$	Measure improvements in critical thinking for the average of the three activities	Difference between $\overline{NR}$ and $\bar{R}$	Appendix E Table E9
R_h	Measure the level of critical thinking in the initial response to each activity	The score obtained when applying the rubric from Appendix C-1 to the initial response to each activity	Appendix E Tables E1, E2 and E3
R_n	Measure the average level of critical thinking in the initial response to each activity	The average R_h for the critical thinking skills involved in each activity	Appendix E Tables E1, E2 and E3
$\bar{R}$	Measure the average of the initial responses for the three activities	Average of R_1 , R_2 and R_3	Appendix E Table E9

NR_h	Measure the level of critical thinking in the final response to each activity	The score obtained when applying the rubric from Appendix C-1 to the final response to each activity	Appendix E Tables E1, E2 and E3
NR_n	Measure the average level of critical thinking in the final response to each activity	The average NR_h for the critical thinking skills involved in each activity	Appendix E Tables E1, E2 and E3
$\overline{NR}$	Measure the average of the final responses for the three activities	Average of NR_1 , NR_2 and NR_3	Appendix E Table E9
F_n	Measure the compulsory and optional elements of the feedback for each activity, split into FR_n and FG_n	Sum of the compulsory and optional elements obtained when applying the rubric from Appendix C-2	Appendix E Table E5, E6 and E7
$\bar{F}$	Measure the average of the sum of the compulsory and optional elements for the three activities, Split into $\overline{FR}$ and $\overline{FG}$	Average of F_1 , F_2 and F_3 for FR and FG	Appendix E Table E9
$NR(FR)_n$	Measure the extent to which students incorporate the feedback (F) into their new response (NR) for each activity.	The score obtained when applying the rubric from Appendix C-3	Appendix E Table E8
$\overline{NR(FR)}$	Measure the extent to which students incorporate the feedback (F) into their new response (NR) on average for the three activities	Average of $NR(FR)_1$, $NR(FR)_2$ and $NR(FR)_3$	Appendix E Table E9
$T1$	Measure performance on the first course assessment	Scores on the first course assessment	Appendix E Table E10
$T2$	Measure performance on the second course assessment	Scores on the second course assessment	Appendix E Table E10

Note: For X_h "h" represents a critical thinking skill and for X_n "n" represents the number of the activity.

(*) Taking the model for Student 1 from Figure 1 as a reference

It is worth noting that not all of the students participated in all of the activities. Therefore, the variables that are calculated as an average of the three activities were actually calculated as an average of the activities in which each student participated. In this sense, if a student only participated in activities 1 and 3, the variable $\bar{R}$ was calculated as the average between R for activity 1 (R_1) and R for activity 3 (R_3). Appendix F shows how these variables were calculated for the three possible cases.

The variables $\bar{R}$, $\bar{F}$ (which is split into $\overline{FR}$ and $\overline{FG}$), $T1$ and $T2$ are independent variables, while $\bar{A}$ is the dependent variable for our model. As stated previously, the other dimensions of the variables were not included in the model and, therefore, were discarded. This includes the variables NR_h and NR_n , which are used to calculate $\overline{NR}$, as well as R_h and R_n , which are used to calculate $\bar{R}$. This is also the case for the variables A_h and A_n (used to calculate $\bar{A}$), as well as F_n and $NR(FR)_n$ (used to calculate $\bar{F}$ and $\overline{NR(FR)}$, respectively). $\overline{NR}$ is also discarded as it is used to calculate $\bar{A}$ ($\bar{A} = \overline{NR} - \bar{R}$). Therefore, the information provided jointly by $\bar{R}$ and $\overline{NR}$ is the same as the information provided by $\bar{R}$ and $\bar{A}$. Finally, the variable $\overline{NR(FR)}$ was also discarded as the

information it provides on $\bar{F}$ is redundant. Tables K1, K2 and K3 in Appendix K show the correlation between the discarded variables and the variables $\bar{R}$, $\overline{NR}$ and $\bar{F}$, respectively. In this case, the strong correlation between the variables is empirical evidence of the redundancy of the information that they provide (Taylor, 1990, in Olasina, 2018).

Based on the forward stepwise method, the variables were added to the model in the following order: $\overline{FR}$, $\overline{FD}$, $T2$, $\bar{R}$ and $T1$. The forward stepwise method establishes that the variables are added to the empty model one by one, based on a certain statistical criterion. At each stage, the most statistically significant variable is added to the model based on said criterion. This is repeated until none of the variables fulfill this criterion (Lewicki et al., 2006, in Luciendo et al., 2018). The selection criterion chosen for our model was the adjusted R^2 . In this sense, the variables were added to the model one by one, with the variable that provided the biggest increase in Adj R^2 added at each stage. As a result, the model that is obtained maximizes the explained variance. A description of each of the stages in which a new variable was added to the model can be found in Appendix G. The final model with all of the variables can be found in Table 3.

2.5.2. Descriptive statistics and t-test

In order to identify the variation between variables from one session to another, as well as during an activity, the descriptive statistics for R , NR , A , F (FR y FG) and $NR(FR)$ were analyzed. The statistics for R , NR and A can be found in Tables 4 and 5, while the statistics for F (split into FR and FG) and $NR(FR)$ can be found in Table 6.

Finally, the distribution of the variable *improvement* was analyzed (Table 7) in order to examine how the students developed their critical thinking skills throughout the different activities. To do so, the distribution of the variable *response* (R) was compared with the distribution of *new response* (NR) for each of the activities and each of the different critical thinking skills. This comparison then allowed us to identify the progress made by the students and how they moved from one level of achievement to another.

A t-test was used to determine whether the development of the students' critical thinking skills between R and NR was statistically significant. This test was applied for each of the dimensions, i.e. each of the five critical thinking skills included in this study ($I_h = NR_h - R_h$) (Table 4). The test was also applied for each of the activities ($I_n = NR_n - R_n$) and the average for the three activities ($\bar{I} = \overline{NR} - \bar{R}$) (Table 5).

2.6. Qualitative analysis

Qualitative analysis was used to understand the different student perspectives on the peer assessment process. This was done through an hour-long focus group involving 6 students who were randomly selected post-implementation. The session was recorded and transcribed word-for-word. The focus group covered both out-of-class and in-class activities and was guided by the question “What is the students’ opinion of learning about reading methods using the flipped classroom approach?” This study will focus on the in-class elements.

The analysis of the focus group was based on grounded theory (Murphy, 2017; Sinclair et al., 2017). This analysis started with a reading of the transcript. Following this, the data was then broken down line by line in order to identify units of meaning. Open coding was then performed, with the initial categories defined by asking the question “What is the text talking about?” and the properties defined by asking the question “What does it say about this?” This process was repeated until achieving theoretical saturation (i.e. the data no longer provided any additional information). Based on the results of the open coding, an initial hierarchical list and tree was developed for the phrases. Following this, each of the categories was analyzed based on the research question and the material was reassembled, leading to a recoding of the text and reordering of the list.

2.7. Ethical considerations

Before participating in the first activity, the students were provided with information on the purpose of the research, the specific activities it involved and the duration of these activities, as well as the risks and benefits of participating. The voluntary nature of participation in the study was also highlighted. Furthermore, the students were requested to sign an informed consent form.

For the qualitative analysis, the students were provided with information and an explanation of the purpose of the research, as well as the duration of the focus group. The students were informed that the information would only be used for research purposes. Furthermore, the students were asked for authorization to record the session and provided their written consent.

Finally, the study also received approval from the University’s ethics committee.

3 Results

3.1 Quantitative

3.1.1 Stepwise Regression Model

Table 3 shows the results for the linear regression model with the dependent variable *Improvement (I)* and independent variables *Feedback Received (FR)*, *Feedback Given (FG)*, *Test 2 (T2)*, *Response (R)* and *Test 1 (T1)*. Appendix G details the step-by-step process for selecting the independent variables included in the model. The first variable that was selected was *Feedback Received ($\overline{FR}$)*, which explained 25% of the variance ($Adj R^2 = .25$) and was statistically significant ($F(1, 33) = 12.44, p < .05$). The second variable was *Feedback Given ($\overline{FG}$)*, which explained an additional 11% of the variance ($Adj R^2 = .36$) and was also statistically significant ($F(2, 32) = 10.52, p < .05$). The third variable to be included in the model was the score on *Test 2 (T2)* (*Mean* = 5.41, *Std. Deviation* = 1.14). This variable explained a further 5% of the variance ($Adj R^2 = .41$) and was again statistically significant ($F(3, 31) = 8.73, p < .05$). The next variable to be added was *Response ($\overline{R}$)*, which explained an additional 3% of the variance ($Adj R^2 = .44$) and was statistically significant ($F(4, 30) = 7.69, p < .05$). The fifth and final variable added to the model was the score on *Test 1 (T1)* (*Mean* = 5.13, *Std. Deviation* = 0.53). This variable explained a further 2% of the variance ($Adj R^2 = .46$) and was statistically significant ($F(5, 29) = 6.85, p < .05$).

Table 3

Analysis of the variables included in the linear regression model to predict improvements in critical thinking skills (N = 35)

Variables	b	Std. Error	β	t	p-value
$\overline{FR}$	.18	0.1495	0.6177	4.13	< 0.001
$\overline{FG}$	.12	0.1429	0.4031	2.82	0.009
T2	.13	0.1577	0.4419	2.8	0.009
$\overline{R}$	-.10	0.1752	-0.3316	-1.89	0.068
T1	-.05	0.1359	-0.2024	-1.49	0.147

Note: model with standardized variables (*Mean* = 0 y *Std. Deviation* = 1).
The data used to build the model can be found in Appendix H.

The independent variables *Feedback Received* ($\overline{FR}$), *Feedback Given* ($\overline{FG}$), and *Test 2* (T_2) are all significant predictors of the dependent variable *Improvement* ($\bar{I}$), with *Feedback Received* having the biggest impact. *Average Response* ($\bar{R}$) and the scores on *Test 1* (T_1) are not predictors and have a $\beta < 0$.

3.1.2 Descriptive statistics and t-test

Table 4 shows that the *Improvement* between *R* and *NR* is statistically significant for all of the critical thinking skills required in Activity 1. This includes the skills of Analysis ($p\text{-value} = 0.0008$) and Explanation ($p\text{-value} = 0.0003$). This is also the case for the skills required in Activity 2 (i.e. Explanation ($p\text{-value} = 0.0025$), Interpretation ($p\text{-value} = 0.0119$) and Inference ($p\text{-value} = 0.0101$)). However, the *Improvement* in Activity 3 is not statistically significant (Inference ($p\text{-value} = 1.00$) and Evaluation ($p\text{-value} = 0.3265$)). In fact, only one student demonstrated an improvement between *R* and *NR* in Activity 3 (doing so for both of the required skills), while another student actually showed a decline in their inference skills. In this case, as the improvement made in inference by the former student matches the decline of the latter, the effect is nullified. The average progress made in this area for Activity 3 is therefore zero.

Table 4

Descriptive statistics for Response (R), New Response (NR) and Improvement (I) and t-tests to test the statistical significance of the Improvement (I) for each of the activities.

	Activity 1 (N = 13)			Activity 2 (N = 33)			Activity 3 (N = 27)		
	Mean	Std. Dev.	p-value (t-test)	Mean	Std. Dev.	p-value (t-test)	Mean	Std. Dev.	p-value (t-test)
$R_{analysis}$	2.54	0.88		—	—	—	—	—	—
$NR_{analysis}$	3.54	0.97		—	—	—	—	—	—
$I_{analysis}$	1.00	0.82	0.0008	—	—	—	—	—	—
$R_{explanation}$	2.77	0.73		2.48	0.80		—	—	—
$NR_{explanation}$	3.77	0.60		2.79	0.93		—	—	—
$I_{explanation}$	1.00	0.71	0.0003	0.30	0.53	0.0025	—	—	—
$R_{interpretation}$	—	—	—	2.58	0.83		—	—	—
$NR_{interpretation}$	—	—	—	2.76	0.90		—	—	—
$I_{interpretation}$	—	—	—	0.18	0.39	0.0119	—	—	—
$R_{inference}$	—	—	—	2.52	0.91		2.44	0.58	
$NR_{inference}$	—	—	—	2.82	0.98		2.44	0.58	
$I_{inference}$	—	—	—	0.30	0.64	0.0101	0.00	0.28	1.00
$R_{evaluation}$	—	—	—	—	—	—	2.93	1.14	
$NR_{evaluation}$	—	—	—	—	—	—	3.04	1.09	
$I_{evaluation}$	—	—	—	—	—	—	0.11	0.58	0.3265

Note: the statistics are taken from the table of results in Appendix E (Tables E1, E2, E3 and E4).

Analyzing the three sessions, there is an increase in the class' average score between *R* and *NR* for each of the three sessions (Table 5). There is also an improvement when averaging the three sessions. These increases are statistically significant in every case, with the exception of the third session (*I*₃). The first activity shows the greatest *Improvement* (*Mean* = 1.00 point), which is equivalent to moving from one level of achievement to the next (i.e. from not achieved to emerging, from emerging to developing and from developing to fully achieved).

Table 5

Descriptive statistics for Response (R), New Response (NR) and Improvement (I) and t-tests to test the statistical significance of the Improvement (I) for the class average in each activity as well as the overall average across the three activities.

	N	Mean	Std. Dev.	range of scale	p-value
<i>R</i> ₁	13	2.65	0.72	1 - 4	
<i>NR</i> ₁	13	3.65	0.77	1 - 4	
<i>I</i> ₁	13	1.00	0.68	0 - 3	0.0002
<i>R</i> ₂	33	2.53	0.77	1 - 4	
<i>NR</i> ₂	33	2.78	0.88	1 - 4	
<i>I</i> ₂	33	0.25	0.43	0 - 3	0.0017
<i>R</i> ₃	27	2.69	0.82	1 - 4	
<i>NR</i> ₃	27	2.74	0.79	1 - 4	
<i>I</i> ₃	27	0.05	0.40	0 - 3	0.4773
$\bar{R}$	35	2.54	0.62	1 - 4	
$\overline{NR}$	35	2.80	0.75	1 - 4	
$\bar{I}$	35	0.26	0.05	0 - 3	< 0.0001

Note: the statistics are taken from the table of results in Appendix E (Tables E1, E2, E3, E4 and E9).

With regards to *Feedback Received (FR)* and *Feedback Given (FG)*, these should be equal in terms of the mean and standard deviation as the feedback that one student receives is the same as the feedback given by another. Not all of the students completed the *New Response (NR)* and therefore could not be considered in the analysis. This is because the *NR* is required in order to calculate the *Improvement*. Table 6 therefore includes students who gave feedback (*FG*) to students who did not complete the *NR*. However, the feedback given by the students who did not complete the *NR* is not included in the table.

Table 6

Descriptive statistics for Feedback Received (FR), Feedback Given (FG) and Incorporating FR into NR (NR(FR)).

Variable	N	Mean	Std. Dev.	range of scale
FR_1	13	18.08	6.28	4 - 32
FG_1	9	20.33	2.83	4 - 32
$NR(FR)_1$	13	2.92	1.26	1 - 4
FR_2	33	9.67	7.29	4 - 32
FG_2	21	14.52	2.82	4 - 32
$NR(FR)_2$	33	1.67	1.47	1 - 4
FR_3	27	9.85	6.23	4 - 32
FG_3	21	12.67	3.62	4 - 32
$NR(FR)_3$	27	1.44	1.34	1 - 4
$\overline{FR}$	35	10.28	6.24	4 - 32
$\overline{FG}$	35	14.45	3.51	4 - 32
$\overline{NR(FR)}$	35	1.72	1.17	1 - 4

Note: Calculations based on information included in Appendix E, Tables E5, E6, E7, E8 and E9.

Table 7 shows the distribution of student performance, broken down by activity and skill. There is movement in each of the levels of achievement between R and NR for each of the skills required by Activity 1. Although the percentage of students listed as ‘Developing’ does not change between R and NR for Explanation and Interpretation in Activity 2, there is movement in the other levels of achievement (with the exception of Not Achieved). As highlighted previously, one student moves from ‘Emerging’ to ‘Developing’ for Inference in Activity 3, while another student moves in the opposite direction. The same student who improves their inference skills also moves from Not Achieved to Fully Achieved for Evaluation. A breakdown and distribution of the scale of the improvement between R and NR for each activity and skill can be found in Appendix I.

Table 7

Distribution of scores obtained by applying the rubric to R and NR, according to the criteria included in the rubric.

	Activity 1 (N = 13)		Activity 2 (N = 33)		Activity 3 (N = 27)	
	R	NR	R	NR	R	NR
<i>Analysis</i>						
Not achieved	15.38%	7.69%	-	-	-	-
Emerging	23.08%	7.69%	-	-	-	-
Developing	53.85%	7.69%	-	-	-	-
Fully achieved	7.69%	76.92%	-	-	-	-
<i>Explanation</i>						
Not achieved	0.00%	0.00%	6.06%	6.06%	-	-
Emerging	38.46%	7.69%	51.52%	36.36%	-	-
Developing	46.15%	7.69%	30.30%	30.30%	-	-
Fully achieved	15.38%	84.62%	12.12%	27.27%	-	-
<i>Interpretation</i>						
Not achieved	-	-	6.06%	6.06%	-	-
Emerging	-	-	45.45%	36.36%	-	-
Developing	-	-	33.33%	33.33%	-	-
Fully achieved	-	-	15.15%	24.24%	-	-
<i>Inference</i>						
Not achieved	-	-	6.06%	6.06%	3.70%	3.70%
Emerging	-	-	57.58%	39.39%	48.15%	48.15%
Developing	-	-	15.15%	21.21%	48.15%	48.15%
Fully achieved	-	-	21.21%	33.33%	0.00%	0.00%
<i>Evaluation</i>						
Not achieved	-	-	-	-	14.8%	11.1%
Emerging	-	-	-	-	22.2%	22.2%
Developing	-	-	-	-	18.5%	18.5%
Fully achieved	-	-	-	-	44.4%	48.2%

Note: the distribution of levels of achievement is taken from the results tables in Appendix E, Tables E1, E2 and E3.

3.2 Qualitative

For grounded theory, the category relating to the in-class work is “classroom experience”. This will be the central theme that ties together all of the data. This central category and the corresponding sub-categories can be found in Table 8.

Table 8*Central category and sub-categories*

Central Category	Sub-Category	Explanation
Classroom Experience	Importance of the teacher	This category focuses on the work done by the teacher in the classroom, i.e. the teacher's influence during the class. The students highlight the value of being able to do practical activities with the teacher, who guides and facilitates the students' learning.
	Importance of fellow students	Focuses on how the students as a group can influence the success of an in-class activity.
	Peer assessment	Focuses on the process of assessing peers during in-class activities.

As the main objective of this study is to explore the role of peer assessment, the qualitative analysis will focus on this sub-category.

3.2.1 Peer assessment and its properties

The properties of peer assessment that emerged as categories from the students' comments include the following: interest in and attitude towards the task, learning to peer assess, technology adoption, and building knowledge collectively through peer assessment and the development of higher order thinking skills (Table 9).

Table 9*Categories and comments*

Categories and explanations	Comments
<u>Interest in and attitude towards the task</u> The students suggest that their level of interest and attitude had a direct impact on the feedback they gave to their classmates.	<i>Student 2: "The quality of the feedback depends on the level of interest"</i> <i>Student 1: "It depends on the attitude; a positive attitude led to better feedback"</i> <i>Student 6: "It depends on the teacher, on the class, on the individual students; it all depends</i>

	<i>on their attitude”</i> <i>Student 4: “As the tasks were challenging, it made you want to give and receive feedback.”</i>
<u>Learning to peer assess</u> The students suggest that the level of feedback they gave to their classmates improved over time.	<i>Student 5: “When we started the class, I didn’t know how to give feedback and I had to learn. It was difficult to begin with, but after that I was able to do it”</i> <i>Student 4: “In the beginning I gave really bad feedback because I didn’t know how to do it”</i> <i>Student 1: “It improved over time”</i> <i>Student 6: “Giving and receiving feedback between peers was something that was new and easy to adopt”</i>
<u>Technology adoption</u> The students suggest that knowledge of the technology and the platform is essential.	<i>Student 3: “I found it difficult to begin with, to connect to the internet, open the website (...) I didn’t know if I’d sent my work”</i> <i>Student 2: “(...) It took me a while because I wasn’t used to working with a computer in class”</i> <i>Student 1: “When we started I didn’t understand how I was supposed to assess my classmate, I found it hard to tell whether I’d done everything right in the platform”</i>
<u>Building knowledge collectively through peer assessment</u> The students talk about how they developed their knowledge through peer assessment.	<i>Student 2: “It helps you to develop your knowledge as there’s someone else giving you advice on how to improve your answer”</i> <i>Student 6: “Everything was well put-together and made sense. The (peer review) activities were accessible and allowed you to apply what you’d learned in much greater depth”</i> <i>Student 1: “The activity involved everyone (...) so you develop a lot”</i> <i>Student 3: “Applying everything you know so that someone else can improve”</i> <i>Student 2: “It’s difficult to give feedback when you don’t know for sure what the answer is or if you haven’t studied the topic enough”</i> <i>Student 5: “The activities put you in the teacher’s position”</i> <i>Student 1: “Seeing another response allows</i>

	you to improve your own” and “When you had to review a terrible answer, it didn’t really help you because everything was missing... it was no help to you” <i>Student 6: “The system allows you to look at things from another perspective”</i> <i>Student 4: “It allows us to improve our responses, as well as to give and receive advice”</i> <i>Student 5: “There was a lot to the feedback: seeing things from another perspective (...) and, well, the advice”</i>
--	--

4 Discussion

4.1 Developing higher-order thinking skills

The rubrics used to assess interpretation, analysis, evaluation, inference and explanation in the initial and final responses reveal a significant improvement in higher order thinking skills. There was a significant improvement between the initial response (R) and the new response (NR) in two of the three activities (Table 4). The lack of improvement in the third activity may be due to the teacher asking the students to focus more on their reasoning than on the feedback. This can be seen in Table 7, which shows that practically all of the students maintained their initial level for both skills in Activity 3 (Inference and Evaluation). Other possible factors at play include the time spent on each phase, as well as the discussion and scaffolding provided by the teacher between phases. It is worth highlighting the importance of the class structure. In this sense, it is important to consider the orchestration of the resources that are used (Nussbaum et al., 2013; Dillenbourg, 2013) when planning the class. This is reflected in a comment made by one of the students “*Student 6: Everything was well set-up and made sense. The peer review activities were accessible and allowed you to apply what you’d learned in much greater depth*”.

It is also worth noting that improvements were made in all skills except Evaluation (i.e. Analysis, Explanation, Interpretation and Inference) during the activities (Table 4). This was largely due to the feedback process. Future studies are required to explore whether Evaluation can also be developed using the proposed mechanism, or whether this skill is simply more complex than the others (Zaidi et al., 2018).

4.2 Influence of peer assessment on the development of higher order thinking skills

The linear regression model shows that feedback received has the greatest impact on improvements in critical thinking skills ($\overline{FR} \beta = 0.62$, $p\text{-value} = <0.001$) (Table 3). This is consistent with the students' comments, e.g. Student 3, who highlights the benefit of "*applying everything you know so that someone else can improve*". Good feedback shows a student how to improve their response, while peer assessment can provide the support that is needed for a student to improve (Vásquez-Colina et al., 2017).

Furthermore, as previously highlighted by the literature, students play an active role in peer assessment, both in the assessment of themselves as well as others (Planas-Lladó et al., 2018). One student, Student 2, claimed that peer assessment "*helps you to develop your knowledge as there's someone else giving you advice on how to improve your answer*". Finally, feedback not only boosts motivation; it also depends on it (Lai & Hwang, 2015; Tseng & Tsai, 2007; Hwang, Hung & Chen, 2013). This issue is raised by Student 1, who suggests that "*it depends on the attitude; a positive attitude led to better feedback*".

The linear regression model also reveals that feedback given is a significant predictor of improvements in critical thinking skills ($\overline{FD} \beta = 0.40$, $p\text{-value} = <0.01$) (Table 3). This can be explained by the fact that students compare their own work with the work they review and subsequently incorporate this knowledge (Huisman et al., 2018). This is demonstrated by Student 1, who provides two different perspectives on the matter: "*seeing another response allows you to improve your own*" and "*when you had to review a terrible response, it didn't really help you because everything was missing... it was no help to you*".

We agree with Newman et al. (2018) that in order to give quality feedback, students must be given guidelines. This includes suggestions such as looking beyond the case at hand and not just highlighting issues within the response. Giving feedback to a peer is a complex skill that needs to be developed by providing opportunities for learning (Sluijsmans et al., 2004, in Alqassab et al., 2017). It is therefore important to teach the students how to give feedback beforehand, thus enhancing their skills and confidence in this area (De la Cruz et al., 2015). Comments made during the focus group reveal that peer assessment is not a common skill among pre-service teachers and must therefore be developed, not least because it helps foster collaboration among students (Alqassab et al., 2018). For example, Student 4 suggests that "*In the beginning I gave really bad feedback because I didn't know how to do it*", while Student 5 claims that "*When we*

started the class, I didn't know how to give feedback and I had to learn. It was difficult to begin with, but after that I was able to do it". Student 1, on the other hand, suggests that it was something that *"improved over time"*. Furthermore, working on peer assessment activities helps future teachers improve their evaluation skills (Sluijsmans et al., 2004, in Algassab et al., 2018).

The results for $\overline{FR}$ and $\overline{FG}$ in the regression model (Table 3) reaffirm the importance of building knowledge collectively by giving and receiving feedback (Lin, G, 2018). This can also be seen in the comments made by Student 1, who suggests that *"the activity involved everyone"*, and Student 2, who felt that *"it helps you to develop your knowledge as there's someone else giving you advice on how to improve your answer"*.

Finally, the average for the variable Response ($\bar{R}$ in Table 3) has a $\beta = -0.33$. This may be due to the fact that students who have a higher starting level tend to have less room for improvement as they are closer to the highest possible level of achievement (Yang et al., 2013).

4.3 Scaffolding model for pre-service teachers

In order for students to make effective use of peer assessment they need training in the specific model (Hanrahan & Isaacs, 2001). A scaffolding model for pre-service teachers was developed based on an analysis of the responses by students who achieved the highest level (i.e., Fully achieved). This model included guidelines for developing critical thinking skills that allow the students to structure their thoughts and enrich their pedagogical knowledge.

This discursive phases model, Figure 2, has several purposes. The first is to guide teacher trainers in how to work with specific and progressive indicators that provide students with the tools to develop their responses. It also allows them to evaluate the complexity of a task. In this sense, the phases and topics to be covered guide the teacher trainer on the learning objectives that the students are expected to achieve in the allotted time. A second purpose is to realize which stages or processes are involved in all of the skills. In other words, there are certain discursive phases that are shared by all of the skills and which must therefore be mastered by the students in order for them to develop said skills. In this sense, the proposed model allows the student to develop a previous structure for the response, which in turn allows some of the phases to be generalized. For example, all of the skills require a *contextual* phase, as well as a *teaching moment*. These elements must therefore be developed with previous guidance from the teacher during in-class discussions. The third purpose of the model is to differentiate between the specific tasks required by each critical thinking skill. This strengthens the guidelines for writing the response, as well as for giving feedback. In this case, for example, an activity

that focuses on developing an explanation will necessarily have to include a “conceptualization” phase (inherent to this skill) in order to be successful.

Figure 2 summarizes the phases or stages of writing that a response to an activity ought to include when working on each skill. The critical thinking skills are included as columns, while the corresponding phases are described in the rows below. The contextual phase, as well as the argumentation and the pedagogical instance, are shared by all of the skills and therefore extend across the table. Phases involving general methodology (which corresponds only to interpretation and inference) and connecting topics (which corresponds to explanation, analysis and evaluation) only appear under specific skills. Furthermore, other phases are unique to certain skills as they correspond to elements, activities or sub-skills that are representative of said skills (based on Facione’s (1990) model for critical thinking).

CRITICAL THINKING SKILLS					
	Interpretation (Comprehend and express the meaning or significance of a wide variety of experiences, situations, data, events, judgments, conventions, beliefs, rules, procedures, or criteria)	Inference (Identify and secure elements needed to draw reasonable conclusions; to form conjectures and hypotheses; to consider relevant information and to reduce the consequences flowing from data, statements, principles, evidence, judgments, beliefs, opinions, concepts, descriptions, questions, or other forms of representation)	Explanation (State the results of one's reasoning; to justify that reasoning in terms of the evidential, conceptual, methodological, criteriological and contextual considerations upon which one's results were based; and to present one's reasoning in the form of cogent arguments)	Analysis (Identify the intended and actual inferential relationships among statements, questions, concepts, descriptions, or other forms of representation intended to express belief, judgment, experiences, reasons, information, or opinions)	Evaluation (Assess the credibility of statements or other representations which are accounts or descriptions of a person's perception, experience, situation, judgment, belief, or opinion; and to assess the logical strength of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation)
DISCURSIVE PHASES MODEL	Contextual Presenting or describing the case, bibliographical references, background information, concepts or data which aid understanding of the context or conceptual framework underpinning the proposal. The student is expected to extract these elements (i.e. reproduce them). In order to do so, the information should be taken from the required reading for the activity, the in-class discussions before the class, or the case or situation presented in the activity, among others. *** In <i>evaluation</i> activities that require the student to choose a specific proposal, this choice must be stated during the contextual phase.				
	Argumentation Justifying and explaining the claims, hypotheses, methodologies, relationships or conceptualizations proposed in response to the activity. The student's reasoning is expected to be based on explicit, relevant and concrete evidence.				
	Pedagogical instance Presenting a possible solution to a problem through a practical teaching moment. The student is expected to propose a situation in which they describe a strategy to be followed by the teacher in order to solve a teaching/learning issue in the classroom. This situation should be different to those reviewed during the activity.				
	General methodology Describes the strategies or general outline of the method to be used as an initial approach to solving the problem presented by the activity. These elements assume a general level of knowledge from the student, who must propose a suitable method for solving the problem that is presented.		Connecting topics Relating relevant subject-matter knowledge (included in the syllabus) to pedagogical subject-matter knowledge (related more specifically to the teaching or learning of the corresponding subject matter) in order to aid the comprehension of a particular concept in the classroom.		
	Exemplification	Hypothesis	Conceptualization	Evaluation	

	Presenting an example of pedagogical content knowledge to clarify the proposed methodology.	Presenting a belief or assumption based on the situation presented by the activity. This assumption is made based on information taken from the content of the activity (i.e. exercises, class material and recommended reading, among others). This information can then be used as the basis for the student's research or reasoning.	Presenting and describing the subject-matter knowledge present in a case study. These are proposed by the student based on their previous general knowledge.	Specifying and defining evaluation criteria that allow the relationships, choices and explanations for a given problem to be analyzed. The evaluation is expected to lead to a relevant response to the proposal. The criteria for evaluating the relevance of the proposals will depend on the activity. *** This phase is only found in analysis activities that also involve evaluation.
--	---	---	--	--

Figure 2 Discursive phases model for activities that develop the critical thinking skills of interpretation, analysis, evaluation, inference and explanation

Considering the above, it is clear that pre-service teachers require more than just subject-matter knowledge. They must also develop the critical thinking skills that will allow them to be better critical thinkers, as well as improving their own teaching practices. Doing so will then allow them to teach their own students to become critical thinkers themselves (Liu et al., 2018).

Furthermore, different teaching strategies are required in order to develop critical thinking among pre-service teachers. This includes holding discussions (Liu et al., 2018) and setting questions and tasks that require the use of higher order thinking skills (Barnett et al., 2012), as well as using specific writing strategies that guide the student in how to develop a written response. Furthermore, providing students with a structure before they submit their response, as shown in Figure 2, will help them to build, develop and review their work. This in turn will lead to significant learning outcomes (Angeli & Valanides, 2009; Mayer, 2004, in Barnett et al., 2012), while giving greater purpose to the use of technology in the classroom.

Appendix J includes an example of how to use the model presented in Figure 2. This model is applicable for activities that develop critical thinking skills such as interpretation, analysis, evaluation, inference and explanation. This example was built with selected student responses, highlighting the different phases that are present in the text.

4.4 Use of technology

The logistics of the activity were minimized by using networked devices (tablets or laptops) to have students submit their responses and complete the peer assessment.

Doing so ensured that the transfer of documents was transparent for both students and teacher. Furthermore, the platform allowed the teacher to monitor the students' work and give feedback, with the possibility of projecting a particular student's work on the screen to be analyzed by the whole class. Furthermore, as the information flow was automated it was easy to make the students' responses anonymous and avoid any social bias (Lin, 2018; Berndt et al., 2017). This experience is an example of how the use of technology allows for better learning spaces and activities that were previously inconceivable (Keane et al., 2016). In our specific case, this means managing the peer assessment process both instantly and anonymously.

The platform that was used was designed with a simple and user-friendly interface in order to optimize the student experience (Yilmaz et al., 2018). Nevertheless, the students still faced some difficulties during the induction process. This can be seen in the comments made by Student 3, who said that *"I found it difficult to begin with, to connect to the internet, open the website (...) I didn't know if I'd sent my work"* and Student 1, who claimed that *"When we started I didn't understand how I was supposed to assess my classmate, I found it hard to tell whether I'd done everything right in the platform"*.

The qualitative study reveals the need for students to have previous experience with technology (and learning through technology) in order to facilitate appropriation of technology for their professional life. The fact that the use of technology is insufficient in initial teacher training (Cullen et al., 2018) can be seen in the comment made by Student 2, who suggests that *"(...) it took me a while because I wasn't used to working with a computer in class"*. Without mastering the technology (i.e. knowing how to use basic technological tools) it is hard to achieve the corresponding teaching objectives (Yilmaz et al., 2018).

5 Conclusions

The aim of this study was to answer the research question "How can technology-scaffolded peer assessment activities develop critical thinking among pre-service teachers?" The results, with pre-service teachers reveal that networked devices (tablets or laptops) that are scaffolded so that students develop a response, give anonymous feedback to a peer and develop a new response based on their initial response and the feedback process itself, leads to significant improvements in critical thinking skills. They also show that the level of feedback received and feedback given are predictors of said improvements. This confirms the importance of designing learning situations in pre-service teacher training to promote the development of critical thinking (Arsal, 2015).

Considering that paper-based peer evaluation has a high administrative load with a large number of students (Hanrahan & Isaacs, 2001), technology, through wirelessly interconnected devices, allows for smooth record management for any classroom size, with the only restriction being the capacity of the network (router and bandwidth). This system also allows for an anonymous peer review process. This is essential as students tend to assume that higher performing students give better feedback (Berndt et al., 2018), which is not always the case. Furthermore, the technology allows a teacher to access their students' work during class, thus facilitating their role in guiding the peer review process by clarifying meaning and encouraging active engagement (Zhu & Carless, 2018).

The limitations of this study are all related to its scale. The research was conducted with a limited number of activities, a small number of students and a specific set of skills. Future work should look to test this model further, by carrying out more activities over a longer period of time and in different subjects.

6 References

- Airasian, P. W. (2001). Classroom assessment: Concepts and applications. McGraw-Hill, PO Box 548, Blacklick, OH 43003.
- Alqassab, M., Strijbos, J. W., & Ufer, S. (2018). Training peer-feedback skills on geometric construction tasks: role of domain knowledge and peer-feedback levels. *European Journal of Psychology of Education*, 33(1), 11-30.
- Akgun, A., & Duruk, U. (2016). The Investigation of Preservice Science Teachers' Critical Thinking Dispositions in the Context of Personal and Social Factors. *Science Education International*, 27(1), 3–15. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1100164.pdf>
- Apedoe, X. S., Long, S. A., Morris, J. A., Wilson, A. A., Morris, R. J., Kroeger, S. D., ... & Strycker, J. (2017). Flipping Education. In *The Flipped College Classroom* (pp. 89-123). Springer, Cham.
- Apedoe, X. S., Long, S. A., Morris, J. A., Wilson, A. A., Morris, R. J., Kroeger, S. D., ... (2017). Flipping Education. In *The Flipped College Classroom* (pp. 89-123). New York: Springer International Publishing.
- Arsal, Z. (2015). The effects of microteaching on the critical thinking dispositions of pre-service teachers. *Australian Journal of Teacher Education* (Online), 40(3), 140.
- Arsal, Z. (2017). The impact of inquiry-based learning on the critical thinking dispositions of pre-service science teachers. *International Journal of Science Education*, 39(10), 1326–1338. <https://doi.org/10.1080/09500693.2017.1329564>
- Baker, K. M. (2016). Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education*, 17(3), 179–192. <https://doi.org/10.1177/1469787416654794>
- Barak, M. (2017). Science Teacher Education in the Twenty-First Century: a Pedagogical Framework for Technology-Integrated Social Constructivism. *Research in Science Education*, 47(2), 283–303. <https://doi.org/10.1007/s11165-015-9501-y>
- Berndt, M., Strijbos, J. W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, 33(1), 31-49.

- Berndt, M., Strijbos, J. W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, 33(1), 31-49.
- Betancourth Zambrano, S. (2015). (2015). Desarrollo del pensamiento crítico en docentes universita- rios. Una mirada cualitativa. *Revista. Revista Virtual Universidad Católica Del Norte*, 44, 238-252.
- Bormann, J. (2014). Affordances of Flipped Learning and Its Effects on Student Engagement and Achievement, 1-38.
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712.
- Brookhart, S. M. (2017). How to give effective feedback to your students. ASCD.
- Brown, A. E. (2016). Critical Thinking to Justify an Answer in Mathematics Classrooms.
- Cowan, J. (1999). Being an innovative university teacher (Buckingham, SRHE & Open University Press).
- Cullen, T. A., & Karademir, T. (2018). Technology Experiences of Student Interns in a One to One Mobile Program. *International Journal of Mobile and Blended Learning (IJMBL)*, 10(2), 20-35.
- De la Cruz, M. S. D., Kopec, M. T., & Wimsatt, L. A. (2015). Resident perceptions of giving and receiving peer-to-peer feedback. *Journal of graduate medical education*, 7(2), 208-213.
- DeWaelsche, S. A. (2015). Critical thinking, questioning and student engagement in Korean university English courses. *Linguistics and Education*, 32, 131-147.
<https://doi.org/10.1016/J.LINGED.2015.10.003>
- Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, 69, 485-492.
- Dunne, G. (2015). Beyond critical thinking to critical being: Criticality in higher education and life. *International Journal of Educational Research*, 71, 86-99.
<https://doi.org/10.1016/J.IJER.2015.03.003>
- Ennis, R. H. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi*, 37(1), 165-184. <https://doi.org/10.1007/s11245-016-9401-4>

- Escobar-Pérez, J., & Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición*, 6(1), 27-36.
- Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation. *Computers & education*, 106, 83-96.
- Facione, P. A. (1990). Critical Thinking : A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction Executive Summary “ The Delphi Report. *The California Academic Press*, 423(c), 1–19.
<https://doi.org/10.1016/j.tsc.2009.07.002>
- Gibbs, G., & Simpson, C. (2004). Does your assessment support your students’ learning. *Journal of Teaching and learning in Higher Education*, 1(1), 1-30.
- Hammersley- Fletcher, L., & Hanley, C. (2016). The use of critical thinking in higher education in relation to the international student: Shifting policy and practice. *British Educational Research Journal*, 42(6), 978-992.
- Hanrahan, S. J., & Isaacs, G. (2001). Assessing self-and peer-assessment: The students’ views. *Higher Education Research & Development*, 20(1), 53-70.
- Harland, T., Wald, N., & Randhawa, H. (2017). Assessment & Evaluation in Higher Education Student peer review: enhancing formative feedback with a rebuttal Student peer review: enhancing formative feedback with a rebuttal. *Assessment & Evaluation in Higher Education*, 42(5), 801–811.
<https://doi.org/10.1080/02602938.2016.1194368>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Huisman, B., Saab, N., van Driel, J., & van den Broek, P. (2018). Peer feedback on academic writing: undergraduate students’ peer feedback role, peer feedback perceptions and essay performance. *Assessment & Evaluation in Higher Education*, 1-14.
- Hwang, G. J., Lai, C. L., Liang, J. C., Chu, H. C., & Tsai, C. C. (2018). A long-term experiment to investigate the relationships between high school students’ perceptions of mobile learning and peer interaction and higher-order thinking tendencies. *Educational Technology Research and Development*, 66(1), 75-93.
- Jong, M. S. Y. (2017). Empowering Students in the Process of Social Inquiry Learning through Flipping the Classroom. *Educational Technology & Society*, 20 (1), 306–322.

- Keane, T., Keane, W. F., & Blicblau, A. S. (2016). Beyond Traditional Literacy: Learning and Transformative Practices Using ICT. *Education and Information Technologies*, 21(4), 769-781.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Kim, K., Sharma, P., Land, S. M., & Furlong, K. P. (2013). Effects of active learning on enhancing student critical thinking in an undergraduate general science course. *Innovative Higher Education*, 38(3), 223-235.
- Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, 65(23), 2276-2284. (p. 2276)
- Kostiainen, E., Ukskoski, T., Ruohotie-Lyhty, M., Kauppinen, M., Kainulainen, J., & Mäkinen, T. (2018). Meaningful learning in teacher education. *Teaching and Teacher Education*, 71, 66–77. <https://doi.org/10.1016/J.TATE.2017.12.009>
- Lewicki, P., & Hill, T. (2006). Statistics: methods and applications. *Tulsa, OK. Statsoft*.
- Li, L. (2017). The role of anonymity in peer assessment. *Assessment & Evaluation in Higher Education*, 42(4), 645-656.
- Lin, G.-Y. (2018). Anonymous versus identified peer assessment via a Facebook-based learning application: Effects on quality of peer feedback, perceived learning, perceived fairness, and attitude toward the system. *Computers & Education*, 116, 81–92 <https://doi.org/10.1016/J.COMPEDU.2017.08.010>
- Liu, J., McBride, R. E., Xiang, P., & Scarmardo-Rhodes, M. (2018). Physical education pre-service teachers' understanding, application, and development of critical thinking. *Quest*, 70(1), 12-27.
- Lo, C. K., & Hew, K. F. (2017). Using "First Principles of Instruction" to Design Secondary School Mathematics Flipped Classroom: The Findings of Two Exploratory Studies. *Journal of Educational Technology & Society*, 20(1).
- Luciano, R. D., Silva, B. L., Rosa, L. M., & Meier, H. F. (2018). Multi-objective optimization of cyclone separators in series based on computational fluid dynamics. *Powder Technology*, 325, 452-466.
- Lund-Diaz, S., Montane, M., & Beery, P. (2016). "How"—The Key to Knowledge-Building Pedagogy Success in Supporting Paradigm Shifts for Student Growth and

- the 4Cs of Future Education. In *Google It* (pp. 353–362). New York, NY: Springer New York.
https://doi.org/10.1007/978-1-4939-6415-4_15
- Mäkinen, T. (2018). Meaningful learning in teacher education. *Teaching and Teacher Education*, 71, 66–77. <https://doi.org/10.1016/J.TATE.2017.12.009>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, 22(3), 276-282. (p. 276)
- Mehring, J. G. (2015). An exploratory study of the lived experiences of Japanese undergraduate EFL students in the flipped classroom. Pepperdine University.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25), 1-10.
- Mulder, R., Baik, C., Naylor, R., & Pearce, J. (2014). How does student peer review influence perceptions, engagement and academic outcomes? A case study. *Assessment & Evaluation in Higher Education*, 39(6), 657-677.
- Murphy, C., Klotz, A. C., & Kreiner, G. E. (2017). Blue skies and black boxes: The promise (and practice) of grounded theory in human resource management research. *Human Resource Management Review*, 27(2), 291-305..
- Nagy, W. (1988). Teaching vocabulary to improve reading comprehension. Urbana, IL: ERIC/RCS, NCTE & IRA
- Newman, L. R., Roberts, D. H., & Schwartzstein, R. M. (2016). Peer Observation of Teaching Handbook. Shapiro Institute for Education and Research at Harvard Medical School and Beth Israel Deaconess Medical Center,
<https://hms.harvard.edu/sites/default/files/assets/Sites/Academy/files/MedEdPortal/Peer%20observation%20handbook.pdf>
- Nussbaum, M., & Diaz, A. (2013). Classroom logistics: Integrating digital and non-digital resources. *Computers & Education*, 69, 493-495.
- Olasina, G. (2018). Human and social factors affecting the decision of students to accept e-learning. *Interactive Learning Environments*, 1-14.
- Panadero, E., & Brown, G. T. L. (2017). Teachers' reasons for using peer assessment: positive experience predicts use. *European Journal of Psychology of Education*, 32(1), 133–156. <https://doi.org/10.1007/s10212-015-0282-5>

- Planas-Lladó, A., Feliu, L., Castro, F., Fraguell, R. M., Arbat, G., Pujol, J., ... & Daunis-i-Estadella, P. (2018). Using peer assessment to evaluate teamwork from a multidisciplinary perspective. *Assessment & Evaluation in Higher Education*, 43(1), 14-30.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2), 119-144.
- Saxton, E., Belanger, S., & Becker, W. (2012). The Critical Thinking Analytic Rubric (CTAR): Investigating intra-rater and inter-rater reliability of a scoring mechanism for critical thinking performance assessments. *Assessing Writing*, 17(4), 251-270.
- Sinclair, S., Hack, T. F., Raffin-Bouchal, S., McClement, S., Stajduhar, K., Singh, P., ... & Chochinov, H. M. (2018). What are healthcare providers' understandings and experiences of compassion? The healthcare compassion model: a grounded theory study of healthcare providers in Canada. *BMJ open*, 8(3), e019701.
- Spiller, D. (2012). *Assessment matters: Self-assessment and peer assessment*. The University of Waikato.
- Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.
- Stuppel, E. J. N., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y. F., & Aubeeluck, A. V. (2017). Development of the Critical Thinking Toolkit (CriTT): A measure of student attitudes and beliefs about critical thinking. *Thinking Skills and Creativity*, 23, 91–100. <https://doi.org/10.1016/j.tsc.2016.11.007>
- Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35-39.
- Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: teaching strategies. *Learning and instruction*, 14(4), 359-379.
- Unruh, T., Peters, M. L., & Willis, J. (2016). Flip This Classroom: A Comparative Study. *Computers in the Schools*, 33(1), 38–58. <https://doi.org/10.1080/07380569.2016.1139988>
- Vaughan, M. (2014). Flipping the Learning : An Investigation into the use of the Flipped Classroom Model in an Introductory Teaching Course. *Education Research and Perspectives*, 41, 25–41.

- Van Gennip, N. A., Segers, M. S., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction*, 20(4), 280-290.
- Vásquez-Colina, M. D., Russo, M. R., Lieberman, M., & Morris, J. D. (2017). A case study of using peer feedback in face-to-face and distance learning classes among pre-service teachers. *Journal of Further and Higher Education*, 41(4), 504-515.
- Villalón, M., & Bedregal, P. (2017). Alfabetización inicial: claves de acceso a la lectura y escritura desde los primeros meses de vida. Ediciones UC.
- Walshe, M., Keane, C., Gallagher, J., McDonald, K., O'Hanlon, R., & Ledwidge, M. (2016). 21 Communication difficulties in the care pathway of patients admitted to hospital with acute decompensated heart failure.
- Wiggins, G. (2012). Seven keys to effective feedback. 2012, 70(1), 11-16.
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17-37.
- Yousefi, S., & Mohammadi, M. (2016). Critical thinking and reading comprehension among postgraduate students. *Journal of Language Teaching and Research*, 7(4), 802-807.
- Yang, Y. T. C., Chuang, Y. C., Li, L. Y., & Tseng, S. S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. *Computers & Education*, 63, 285-300.
- Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior*, 70, 251-260.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, 45(4), 477-501.
- Zaidi, N. L. B., Grob, K. L., Monrad, S. M., Kurtz, J. B., Tai, A., Ahmed, A. Z., ... & Santen, S. A. (2018). Pushing Critical Thinking Skills With Multiple-Choice Questions: Does Bloom's Taxonomy Work?. *Academic Medicine*, 93(6), 856-859.

7. Appendix

A. Description of the activities and the critical thinking skills they involve

ACTIVITY 1	
Analyze the learning sequence in the text “The Democracy” based on the key elements of teaching fluency. Make a summary table with the elements that are present and those that are missing. Don’t forget to back this up with evidence. Give suggestions on how to improve this learning sequence. Think about how it relates to the model for the gradual release of responsibility.	
Tasks to be completed by the student	1. Analyze the learning sequence based on the subject-matter knowledge (i.e. the elements of teaching fluency). 2. Create a summary table (i.e. elements that are present and missing). 3. Back the table up with evidence. 4. Make suggestions. 5. Consider previous subject-matter knowledge (i.e. the gradual release of responsibility).
Subject-matter knowledge	Key elements of teaching fluency
Critical thinking skills involved in the activity	Analysis and explanation
ACTIVITY 2	
Nagy (1988) suggested that a child should learn between 1,000 and 3,000 words per year. Villalón (2008) suggests that number should be between 2,700 and 3,000. Considering that the school year lasts approximately 36 weeks (i.e. the children should learn approximately 83 words per week), how would you teach this vocabulary?	
Tasks to be completed by the student	5. Read and compare two recommended texts. 6. Interpret the information and variables involved in a specific pedagogical problem. 7. Identify and infer the main information and conceptual tools. 8. Develop a proposal and provide an explanation.
Subject-matter content	Pedagogical strategies for teaching vocabulary.
Critical thinking skills involved in the activity	Interpretation, inference and explanation.
ACTIVITY 3	
You are preparing a productive discussion to aid comprehension and are evaluating how to segment the information. You have these three proposals for segmenting the information: Proposal 1: Visual text – title and opening paragraph – section 1 – section 2 – section 3 Proposal 2: Sections 1, 2 and 3 – visual text – title and opening paragraph Proposal 3: Title and opening paragraph – visual text – section 3 – section 1 – section 2 Which proposal would you choose? Why? Now think of a question to trigger a discussion for each segment. Remember to use the table from the discussion plan. For now, put the title or main idea of the segment in the main column and fill out the final column with the specific question for that segment. Remember to think of just one question and not several. Don’t forget that the questions should focus on the main ideas of the text (i.e. the focus) and encourage discussion (i.e. open-ended questions).	
Tasks to be completed by the student	5. Evaluate the proposals for segmenting the information. 6. Select a proposal and justify the decision. 7. Come up with a key question for subsequent discussion. 8. Come up with the title and the main idea.
Subject-matter content	Textual production and productive discussion: segmenting and presenting information
Critical thinking skills involved in the activity	Evaluation and inference

B. Summary of the critical thinking skills, indicators and descriptors for the rubrics applied to R and NR

Skill	Definition of the skill	Dimension of the rubric
Interpretation	Comprehend and express the meaning or significance of a wide variety of experiences, situations, data, events, judgments, conventions, beliefs, rules, procedures, or criteria. The sub-skills of interpretation include: categorizing, describing, distinguishing, recognizing, identifying and exemplifying, among others.	The descriptors for this dimension evaluate whether the response clearly and justifiably categorizes, decodes or clarifies the significance or meaning of the question. They also evaluate whether the format used to present the information (e.g. text, diagram, table, concept map etc.) is consistent with the response.
Analysis	Identify the intended and actual inferential relationships among statements, questions, concepts, descriptions, or other forms of representation intended to express belief, judgment, experiences, reasons, information, or opinions. The sub-skills of analysis include: relating, contrasting, examining and proposing, among others.	The descriptors for this dimension evaluate whether the response provides an exhaustive analysis of the relationships between the statements, concepts and issues raised by the question. They also evaluate the clarity, justification and coherence of these relationships, as well as the supporting evidence.
Evaluation	Assess the credibility of statements or other representations which are accounts or descriptions of a person's perception, experience, situation, judgment, belief, or opinion; and to assess the logical strength of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation. The sub-skills of evaluation include: justifying, reasoning, proposing and judging, among others.	The descriptors for this dimension evaluate whether the response provides a critical assessment of the proposal set out in the question, using clear, explicit, relevant and credible criteria. They also evaluate whether the response analyzes the coherence of said proposal by examining its main ideas (focus), intentions (interests or motivations) and beliefs (principles and perspectives).
Inference	Identify and secure elements needed to draw reasonable conclusions; to form conjectures and hypotheses; to consider relevant information and to deduce the consequences flowing from data, statements, principles, evidence, judgments, beliefs, opinions, concepts, descriptions, questions, or other forms of representation. The sub-skills of inference include: recognizing, reformulating, extracting, paraphrasing and explaining, among others.	The descriptors for this dimension evaluate whether the response identifies and builds on the conceptual aspects of the question and provides relevant evidence with which to draw conclusions or develop a proposal in response to said question (conjecture, hypothesis etc.). They also evaluate the main ideas and assumptions that are proposed, analyzing whether or not they acknowledge the causes and effects of the data.
Explanation	State the results of one's reasoning; to justify that reasoning in terms of the evidential, conceptual, methodological, criteriological and contextual considerations upon which one's results were based; and to present one's reasoning in the form of cogent arguments. The sub-skills of explanation include: reporting, articulating, anticipating and presenting, among others.	The descriptors for this dimension evaluate whether the response clearly sets out the results, procedures or arguments required by the question. Each item is justified with compelling evidence and developed appropriately. They also evaluate the procedures, methods, criteria, subject-matter knowledge, lines of argument or main factors considered in their reasoning.

C. Assessment instruments: rubric for assessing R, NR, F and incorporation of F into NR

C-1. Rubric applied to R and NR

ASSESSMENT INDICATORS	FULLY ACHIEVED 4	DEVELOPING 3	EMERGING 2	NOT ACHIEVED 1
DEVELOPMENT OF CRITICAL THINKING (GENERAL DESCRIPTORS)	The student's response shows evidence of suitable development of the relevant critical thinking skill. There are textual clues (both linguistic and discursive), as well as explicit evidence demonstrating the development of said skill.	There is evidence of the relevant skill. Certain aspects of the question have not been developed. There are textual clues and explicit evidence demonstrating the development of said skill.	There is partial evidence of the relevant skill. More than two aspects of the question have not been fully developed. There are textual clues and some evidence demonstrating the development of said skill, though this is not always explicit.	There is no evidence of the relevant skill. The aspects of the question have not been developed in the response and/or the response develops a different skill than the one required by the question.
a) If the question involves INTERPRETATION	The response clearly and justifiably categorizes, decodes or clarifies the significance or meaning of the question. Furthermore, the format used to present the information (e.g. text, diagram, table, concept map etc.) is also consistent with the response.	The response clearly and justifiably categorizes, decodes or clarifies the significance or meaning of the question, but one of the categories or some of the work is only partially justified and discussed. The format used to present the information is consistent with the response.	The response only partially categorizes, decodes or clarifies the significance or meaning of the question, as a part of it is missing. The format used to present the information is consistent with the response.	The response does not categorize, decode or clarify the significance or meaning of the question and/or the format in which the information is presented is not consistent with the skill required by the question.
b) If the question involves ANALYSIS	The response provides an exhaustive analysis of the relationships between the statements, concepts and issues raised by the question. These relationships are clear (correctly defined and well set out) and coherent (in line with the topic and general idea of the question). They are also correctly identified (mentioned in the response), well-reasoned (arguments are given for or against in order to support the position) and justified using evidence from the case study, situation or text that is referenced in the question.	The response provides an analysis of the majority of the relationships between the statements, concepts and issues raised by the question. The relationships are clear and coherent but have only been partially identified, argued or justified because: one of the relationships or aspects of the question is not mentioned or all of the aspects are identified but not all have been analyzed or not enough evidence is provided to justify the analysis.	The response provides an analysis of some of the relationships between the statements, concepts and issues raised by the question. The relationships are clear and coherent but have only been partially identified, argued or justified because: one of the relationships or aspects of the question is not mentioned or not enough evidence is provided to justify the analysis.	The response fails to acknowledge the relationships between the statements, concepts and issues raised by the question. These relationships are not always coherent and do not relate to the main focus of the question. or the aspects of the question are identified but they are not analyzed or no connections are made or no evidence is provided to justify the analysis.
c) If the question involves EVALUATION	The response provides a critical assessment of the proposal described or implied in the question.	The response provides a critical assessment of the proposal described or implied in the	The response provides a partial assessment of the proposal described or implied in the question.	The response does not provide an assessment of the proposal described or implied in the question.

	The main ideas (focus), intentions (interests or motivations) and beliefs (principles and perspectives) in said proposal are acknowledged, and its credibility is assessed based on specific, coherent and well-defined criteria (i.e. in line with the objectives of the question).	question. The main ideas, intentions and beliefs in said proposal are acknowledged and its credibility is assessed based on specific criteria. However, one of these criteria is not relevant to the assessment or has not been well-defined or explained.	The majority of the main ideas, intentions and beliefs in said proposal are acknowledged and its credibility is assessed using a single criterion and/or the criteria used are otherwise only partially justified.	The main ideas, intentions and beliefs are identified but not assessed, or no criteria are defined for assessing the credibility of the proposal.
d) If the question involves INFERENCE	The response identifies and builds on the conceptual aspects of the question and provides relevant evidence in order to draw conclusions or develop a proposal in response to the question (conjecture, hypothesis etc.). The main ideas and assumptions are addressed, acknowledging the causes and effects of the data.	The response identifies and builds on the conceptual aspects of the question and provides relevant evidence in order to draw conclusions or develop a proposal in response to the question (conjecture, hypothesis etc.). The majority of the ideas and assumptions are addressed, though these are not always central or relevant. Some of the causes and effects of the data are acknowledged.	The response identifies and builds on the conceptual aspects of the question but relevant evidence is not always provided in support of the conclusions or conjectures that are made. The main ideas and assumptions are addressed, acknowledging the causes and effects of the data or the majority of the ideas and assumptions are addressed, though these are not always central or relevant. Some of the causes and effects of the data are acknowledged.	The response does not identify or build on the conceptual aspects of the question and/or no evidence is provided to support any of the conclusions or conjectures that are made and/or the main ideas and assumptions are not addressed.
e) If the question involves EXPLANATION	The response clearly sets out the results, procedures or arguments required by the question. Each item is justified with compelling evidence (logical reasoning that does not resort to unjustified generalizations, unsubstantiated claims or emotional appeals) and developed appropriately (following a line of inquiry or argument). The response also details the procedures, methods, criteria, subject-matter knowledge, lines of argument or main factors considered in their reasoning.	The response clearly sets out the results, procedures or arguments required by the question. Each item is justified with convincing evidence and developed appropriately. In general, the response also details the procedures, methods, criteria, subject-matter knowledge and lines of argument included in their reasoning, though it does not always explain the main factors behind these.	The response sets out the results, procedures or arguments required by the question, but these are not always justified with evidence.	The response sets out the results, procedures or arguments required by the question, but no evidence is provided as justification and the elements included in the descriptor for <i>fully achieved</i> are not present.

C-2. Rubric applied to F

Compulsory indicators These indicators are applicable to all written work in the argumentation phase	FULLY ACHIEVED 4	DEVELOPING 3	EMERGING 2	NOT ACHIEVED 1
Assessment criteria	The feedback provides a critical assessment of the proposal described or implied in R. The main ideas (focus), intentions (interests or motivations) and beliefs (principles and perspectives) in said proposal are acknowledged, and their credibility is assessed based on specific, coherent and well-defined criteria (i.e. the student giving the feedback does not try to address everything in R, instead providing guidance on specific points and ideas, while the feedback itself is coherent with the objectives of the activity).	The feedback provides an assessment of the proposal described or implied in R. The main ideas, intentions and beliefs in said proposal are acknowledged, and their credibility is assessed based on specific criteria. However, one of the criteria is not relevant or is not well-defined and explained.	The feedback provides a partial assessment of the proposal described or implied in R. Most of the main ideas, intentions and beliefs in said proposal are acknowledged, and their credibility is assessed based on a single criterion and/or criteria that are only partially justified.	The feedback does not provide an assessment of the proposal described or implied in R. The main ideas, intentions and beliefs are identified but are not assessed, or there are no criteria defined for assessing the credibility of the proposal.
Clarity and relevance of the explanation	The student giving the feedback clearly sets out the results, procedures or arguments that come from reading and analyzing R (they draw on their peer's work, thus establishing a dialogue between their own ideas and the ideas of the student they are assessing). Each item is justified accordingly with evidence from R (logical reasoning that does not resort to unjustified generalizations, unsubstantiated claims or emotional appeals), therefore outlining the procedures, methods, criteria, subject-matter knowledge, lines of argument or main factors considered in their reasoning.	The feedback clearly sets out the results, procedures or arguments that come from reading and analyzing R. Each item is justified accordingly with evidence from R and is developed appropriately. The procedures, methods, criteria, subject-matter knowledge and lines of argument are generally outlined, although the main factors considered in their reasoning are not always explained.	The feedback sets out the results, procedures or arguments that come from reading and analyzing R, though this is not always justified with evidence from R (this level includes cases where the feedback does not interact well with R as the student giving the feedback only considers their personal opinion, and not the opinion set out in R, often drawing on their own work rather than their peer's).	The feedback sets out the results, procedures or arguments that come from reading and analyzing R, but no evidence is provided as justification and the elements included in the descriptor for <i>fully achieved</i> are not present.
Argumentation	The comments made by the student giving the feedback aim to motivate and guide their classmate in wording (or rewording) their response. To do so, they ask questions (which guide the author of R in the changes they must make or areas they must improve when developing R2), highlight certain areas (e.g. suggest different ways of presenting the information),	The comments made by the student giving the feedback aim to motivate and guide their classmate in wording or rewording their response. They ask questions, highlight certain areas, provide explanations and/or warn the student in order to focus their attention on key areas. Most of the comments are justified and/or provide	The comments made by the student giving the feedback aim to describe the strengths and weaknesses of R. To do so, they ask questions, highlight certain areas, provide explanations and/or warn their classmate in order to focus their attention on key areas, although they do not provide suggestions for improving R2. The outcome of the process is	The comments made by the student giving the feedback only aim to <i>correct</i> R. The comments highlight the "errors" found in R and do not motivate or inspire further reflection or questions that might guide the author of R. The outcome of the process is negative.

	provide explanations (regarding the sections of the text that require more attention) or warn the student (if they feel certain aspects of R may be ambiguous or misinterpreted). The outcome of the process is positive.	suggestions for improving R2. The outcome of the process is positive.	positive.	
Syntax, lexis and register	The lexis used is clear (there are no complex expressions, while the wording and syntax is easy to understand) and precise (the text avoids ambiguity, colloquialisms and features of spoken language). The text also uses the appropriate register (i.e. an academic register and/or technical jargon when dealing with descriptions or other relevant explanations that aid comprehension of the topics addressed).	The lexis used is clear but not always precise because: certain comments are ambiguous (e.g. they include words such as 'it', 'something' and 'that', among others). and/or it includes some features of spoken language that are not in keeping with the register and/or are colloquialisms (e.g. "well", "I mean" "like").	The lexis used is not always clear or precise because: all of the comments involve some level of ambiguity in the use of certain words and/or it uses a wide range of features of spoken language, which affect the register.	The lexis used is neither clear nor precise (none of the criteria for <i>fully achieved</i> are met).
Optional indicators³	FULLY ACHIEVED 4	DEVELOPING 3	EMERGING 2	NOT ACHIEVED 1
Request for evidence	The student giving the feedback highlights the need to include evidence or back up some of the ideas presented in R. This request for evidence not only acknowledges which ideas need further support, but also explains why (the aim of providing the evidence is well explained, i.e. the student giving the feedback must highlight which ideas or areas of R require further support and explain why this is important, guiding the author of R with logical, coherent arguments that focus on the question). Furthermore, the student giving the feedback suggests types of evidence that would be relevant to the ideas that have been highlighted (this evidence may be: <i>bibliographical</i> (they may suggest their classmate consider an idea from the literature related to the question), <i>contextual</i> (they may suggest that their classmate provide evidence	The student giving the feedback highlights the need to include evidence or back up some of the ideas presented in R. This request for evidence not only acknowledges which ideas need further support, but also explains why. However, no further information is provided regarding the type or content of such evidence.	The student giving the feedback highlights the need to include evidence or back up some of the ideas presented in R. However, they only suggest which areas need further support, without explaining why or providing any further information on the type or content of such evidence.	The student giving the feedback only highlights the need to include evidence, but does not explain where or why.

³ The decision regarding whether or not to consider these criteria will depend on the level of R (the evaluator should assess the relevance of these indicators based on whether or not the text requires any comment on these aspects).

	from the context or case study contained within the question) or <i>empirical</i> (if ideas or arguments are to be backed up with data, percentages or experimental evidence)).			
Reviewing an argument or presenting a new argument and/or counter-argument	The student giving the feedback proposes new arguments or counter-arguments that build on the line of argument used in R (they are consistent with the work developed in R). Furthermore, they provide the author of R with certain rebuttals (if providing a counter-argument) and/or strategies for incorporating new arguments (explaining the importance of including this new information).	The student giving the feedback suggests reviewing one of the arguments included in R and explains why this is important. Both the suggestion as well as its relevance are justified, while strategies are provided for incorporating this new information.	The student giving the feedback suggests reviewing one of the arguments included in R. However, they do not explain why this is important and/or do not provide any strategies for incorporating this new information.	The student giving the feedback makes certain comments but these are not relevant to R because: they simply provide their own conclusions or arguments, instead of building on R. In this case, the student giving the feedback tends to be in opposition to R and does not engage with the response (the arguments are along the lines of “it’s not like that, it’s like this”).
Request to reword or rework	The student giving the feedback refers to the need to reorganize the text (they reference both formal and discursive aspects, such as the structure of the paragraphs, the inclusion of discourse markers, and a review of spelling and punctuation, among others ⁴) in order to improve its readability (the reworded response would be clearer and easier to understand). They also provide guidelines and specific strategies to enhance or change the format or presentation of the information (these strategies are justified with more than just a “because I say so” or personal opinions, presenting instead a clear objective that is relevant to R). All of the above elements are described in detail (with the steps to follow, specific recommendations, procedures and required research, among others. These are also relevant to R).	The student giving the feedback refers to the need to reorganize the text in order to improve its readability. They also provide guidelines and well-founded strategies, but not all of these are described in detail (some of the strategies are only mentioned in general terms and/or their description is somewhat confusing and/or the guidelines and strategies that are presented are not relevant to R).	The student giving the feedback only refers to the need to reorganize the text, without explaining how this will help improve its readability and/or no guidelines or strategies are provided for enhancing the rewording or reworking of the response and/or guidelines or strategies are provided but these are not well defined or justified and/or the student giving the feedback only suggests superficial improvements to the format or presentation of the information.	The student giving the feedback only highlights the “errors” in R and explains how to correct these (they assume that their own suggestions are correct).
Highlighting the focus of the question	The student giving the feedback highlights certain aspects of the questions that were not addressed in R	The student giving the feedback highlights certain aspects of the question that were not addressed in R.	The student giving the feedback only highlights the need to review the question, indicating which area needs	The student giving the feedback only suggests that the question should be reviewed, without

⁴ This level includes comments in reference to the formal aspects of the response detailed in section 2 of the rubric.

	(indicating which topic or area is missing in R, e.g. the question asks the student to analyze a case and give suggestions but R only analyzes) and/or should the instructions have been misunderstood, the student giving the feedback rephrases the question (explaining it in different terms) or its aim/focus (highlighting a skill or objective that has not been addressed, e.g. the question asks the student to interpret but the author of R only identifies, or the question asks the student to create but the response only describes). As well as highlighting the issue, the feedback also guides the author in how to incorporate the new information (the feedback provides clear and well-defined suggestions, strategies, guidelines and/or procedures).	Although they explain which area needs working on, the guidance that is provided is quite general, with no specific guidelines or procedures included in the feedback and/or should the instructions have been misunderstood, the student giving the feedback highlights which aspect has not been fully understood, but the explanation does not specifically identify the aim/focus of the question (it is only partially identified and/or identified in general/ambiguous terms) and/or the guidance for the author is incomplete as the suggestions, strategies, guidelines and/or procedures are not well defined or clear enough (either because they are presented too generally or because they do not correctly explain the procedures needed in order to implement the relevant strategies).	working on (or including). No guidance is provided to the author of R, with no guidelines or strategies included in the feedback and/or they identify a misunderstanding of the instructions but do not explain the objective of the question nor its focus so as to help the author understand its requirements and/or the author of R is not provided with any guidance in the form of suggestions, strategies, guidelines and/or procedures.	explaining which aspects or why it is important and/or they identify a misunderstanding of the instructions but do not rephrase the question or provide any guidance to help the author of R understand its requirements (they highlight the error or misunderstanding in general terms without saying what it is nor how to fix it).
--	---	--	--	--

C-3. Rubric applied to the incorporation of F into NR

Indicator	FULLY ACHIEVED 4	DEVELOPING 3	EMERGING 2	NOT ACHIEVED 1
Incorporating F into NR	The student rewrites their response, taking into account the comments made in the feedback, as well as any new information to emerge from the assessment of their peer's work or a self-assessment of their own. NR represents a reworked version of R, in which the relevant aspects analyzed by their peer have been addressed (i.e. they do not simply copy-paste the feedback, instead they rework their response based on an analysis and evaluation of this new information).	The student evaluates the relevance of the comments made in the feedback and includes the parts that they feel are most relevant. NR includes the original response, plus their response to the feedback (i.e. they do not simply copy-paste the comments, instead they rework them and add any new information to their original response in a coherent way).	The student literally copies the comments made in the feedback (i.e. they do not incorporate or build on them, they simply copy-paste them). NR includes the original response, plus the comments made in the feedback (i.e. the original response plus the new information).	The student literally copies the comments made in the feedback (i.e. they do not incorporate or build on them, they simply copy-paste them). NR only provides the new information prompted by the feedback (i.e. it does not include the original response).

D. Cohen's kappa for the assessment instrument

Table D1

Level of agreement between evaluators for the Rubric applied to the Response (R) and New Response (NR). Dimension: Analysis

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	2	1	0	0	3
	Emerging	0	3	2	0	5
	Developing	0	4	4	1	9
	Fully achieved	0	0	2	9	11
	Total Columns	2	8	8	10	28

Table D2

Level of agreement between evaluators for the Rubric applied to the Response (R) and New Response (NR). Dimension: Explanation

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	4	0	0	0	4
	Emerging	0	36	1	0	37
	Developing	0	2	23	2	27
	Fully achieved	0	0	1	25	26
	Total Columns	4	38	25	27	94

Table D3

Level of agreement between evaluators for the Rubric applied to the Response (R) and New Response (NR). Dimension: Interpretation

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	4	0	0	0	4
	Emerging	0	27	0	0	27
	Developing	0	5	17	0	22
	Fully achieved	0	0	0	13	13
	Total Columns	4	32	17	13	66

Table D4

Level of agreement between evaluators for the Rubric applied to the Response (R) and New Response (NR). Dimension: Evaluation

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	7	0	0	0	7
	Emerging	0	12	0	0	12
	Developing	0	0	8	2	10
	Fully achieved	0	0	0	25	25
Total Columns		7	12	8	27	54

Table D5

Level of agreement between evaluators for the Rubric applied to the Response (R) and New Response (NR). Dimension: Inference

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	4	2	0	0	6
	Emerging	0	55	3	0	58
	Developing	0	1	37	0	38
	Fully achieved	0	0	0	18	18
Total Columns		4	58	40	18	120

Table D6

Level of agreement between evaluators for the Rubric applied to the Feedback (F). Dimension: Assessment Criteria

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	6	0	0	0	6
	Emerging	1	8	0	1	10
	Developing	0	2	10	3	15
	Fully achieved	0	0	2	22	24
Total Columns		7	10	12	26	55

Table D7

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Clarity and Relevance of the Explanation*

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	0	0	0	0	0
	Emerging	0	8	0	0	8
	Developing	0	2	18	3	23
	Fully achieved	0	0	5	19	24
	Total Columns	0	10	23	22	55

Table D8

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Argumentation*

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	2	1	0	0	3
	Emerging	1	18	1	0	20
	Developing	0	3	10	1	14
	Fully achieved	0	1	3	14	18
	Total Columns	3	23	14	15	55

Table D9

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Syntax, lexis and register*

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	0	0	0	0	0
	Emerging	1	7	1	0	9
	Developing	0	0	12	1	13
	Fully achieved	0	0	4	29	33
	Total Columns	1	7	17	30	55

Table D10

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Request for Evidence*

		Evaluator 1					Total Rows
		Not achieved	Emerging	Developing	Fully achieved	Skill not present	
Evaluator 2	Not achieved	0	0	0	0	0	0
	Emerging	0	3	0	0	0	3
	Developing	0	0	5	0	1	6
	Fully achieved	0	0	0	6	0	6
	Skill not present	0	1	0	1	38	40
	Total Columns	0	4	5	7	39	55

Table D11

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Reviewing an Argument or Presenting a New argument and/or Counter-Argument*

		Evaluator 1					Total Rows
		Not achieved	Emerging	Developing	Fully achieved	Skill not present	
Evaluator 2	Not achieved	0	1	0	0	0	1
	Emerging	0	0	0	0	0	0
	Developing	0	2	1	0	0	3
	Fully achieved	0	0	0	3	1	4
	Skill not present	0	0	1	0	46	47
	Total Columns	0	3	2	3	47	55

Table D12

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Request to Reword or Rework*

		Evaluator 1					Total Rows
		Not achieved	Emerging	Developing	Fully achieved	Skill not present	
Evaluator 2	Not achieved	1	0	0	0	0	1
	Emerging	0	4	0	0	1	5
	Developing	0	0	2	0	0	2
	Fully achieved	0	0	0	0	0	0
	Skill not present	0	0	1	0	46	47
	Total Columns	1	4	3	0	47	55

Table D13

*Level of agreement between evaluators for the Rubric applied to the Feedback (F).
Dimension: Highlighting the Focus of the Question*

		Evaluator 1					Total Rows
		Not achieved	Emerging	Developing	Fully achieved	Skill not present	
Evaluator 2	Not achieved	0	0	0	0	0	0
	Emerging	0	2	1	0	1	4
	Developing	0	0	4	1	0	5
	Fully achieved	0	0	0	3	0	3
	Skill not present	0	0	1	1	41	43
	Total Columns	0	2	6	5	42	55

Table D14

Level of agreement between evaluators for the Rubric applied to the incorporation of the Feedback into the New Response FR(NR).

		Evaluator 1				Total Rows
		Not achieved	Emerging	Developing	Fully achieved	
Evaluator 2	Not achieved	6	0	0	0	6
	Emerging	0	13	1	1	15
	Developing	0	1	18	0	19
	Fully achieved	0	0	0	10	10
	Total Columns	6	14	19	11	50

Table D15

Expected frequencies for Cohen's kappa for each assessment criteria: Rubric applied to the Response (R) and New Response (NR), Rubric applied to the Feedback (F) and Rubric applied to the incorporation of the Feedback (F) into the New Response (NR), for each of the relevant dimensions.

	Not achieved	Emerging	Developing	Fully achieved	Skill not present
<i>Rubric applied to R and NR</i>					
Analysis	0.21	1.43	2.57	3.93	–
Explanation	0.17	14.96	7.18	7.47	–
Interpretation	0.24	13.09	5.67	2.56	–
Evaluation	0.91	2.67	1.48	12.50	–
Inference	0.20	28.03	12.67	2.70	–
<i>Rubric applied to F</i>					
<i>Compulsory Indicators</i>					
Assessment criteria	0.76	1.82	3.27	11.35	–
Clarity and relevance of the explanation	0.00	1.45	9.62	9.60	–
Argumentation	0.16	8.36	3.56	4.91	–
Syntax, lexis and register	0.00	1.15	4.02	18.00	–
<i>Optional Indicators</i>					
Request for evidence	0.00	0.22	0.55	0.76	28.36
Reviewing an argument or presenting a new argument and/or counter-argument	0.00	0.00	0.11	0.22	40.16
Request to reword or rework	0.02	0.36	0.11	0.00	40.16
Highlighting the focus of the question	0.00	0.15	0.55	0.27	32.84
<i>Rubric applied to the incorporation of F into NR</i>					
	0.72	4.20	7.22	2.20	–

Note: The frequency for each level is calculated by multiplying the total number of columns by the total number of rows and dividing by the overall total (The values included in the final row and final column in Tables D1 to D14).

Table D16

Cohen's kappa for each dimension assessed by the Rubric for R and NR, the Rubric for F and the Rubric for F into NR.

	N	$\sum ef$	$\sum a$	Kappa
<i>Rubric applied to R and NR</i>				
Analysis	28	8.14	18	0.50
Explanation	94	29.78	88	0.91
Interpretation	66	21.56	61	0.89
Evaluation	54	17.56	52	0.95
Inference	120	43.60	114	0.92
<i>Mean</i>	–	24.13	66.60	0.83
<i>Rubric applied to F</i>				
<i>Compulsory Indicators</i>				
Assessment criteria	55	17.20	46	0.77
Clarity and relevance of the explanation	55	20.67	45	0.71
Argumentation	55	17.00	44	0.71
Syntax, lexis and register	55	23.16	48	0.78
<i>Optional Indicators</i>				
Request for evidence	55	29.89	52	0.84
Reviewing an argument or presenting a new argument and/or counter-argument	55	40.49	50	0.66
Request to reword or rework	55	40.65	53	0.86
Highlighting the focus of the question	55	33.80	50	0.73
<i>Mean</i>	–	27.86	48.50	0.76
Rubric applied to the incorporation of F into NR	50	14.34	47	0.92

Note: *N* corresponds to the number of responses in which the dimension was assessed, $\sum ef$ corresponds to the sum of the expected frequencies (Table E15) and $\sum a$ is the sum of the number of coincidences for the different criteria.

E. Scores for Response, New Response and Feedback for each activity

Table E1

Scores for Response (R) and New Response (NR) for Activity 1.

Student	Student Feedback	R _{analysis}	R _{explanation}	R _l	NR _{analysis}	NR _{explanation}	NR _l
Student 1	Student 26	3	3	3	4	4	4
Student 2	—	—	—	—	—	—	—
Student 3	Student 19	3	3	3	4	4	4
Student 4	—	—	—	—	—	—	—
Student 5	—	—	—	—	—	—	—
Student 6	—	—	—	—	—	—	—
Student 7	—	—	—	—	—	—	—
Student 8	Student 17	1	2	1.5	1	2	1.5
Student 9	—	—	—	—	—	—	—
Student 10	Student 6	1	2	1.5	4	4	4
Student 11	Student 1	2	3	2.5	3	4	3.5
Student 12	—	—	—	—	—	—	—
Student 13	Student 9	4	3	3.5	4	4	4
Student 14	—	—	—	—	—	—	—
Student 15	—	—	—	—	—	—	—
Student 16	Student 24	3	3	3	4	4	4
Student 17	—	—	—	—	—	—	—
Student 18	Student 25	3	3	3	4	4	4
Student 19	Student 8	2	2	2	4	4	4
Student 20	—	—	—	—	—	—	—
Student 21	—	—	—	—	—	—	—
Student 22	—	—	—	—	—	—	—
Student 23	—	—	—	—	—	—	—
Student 24	Student 13	3	4	3.5	4	4	4
Student 25	Student 10	2	2	2	2	3	2.5
Student 26	Student 21	3	2	2.5	4	4	4
Student 27	—	—	—	—	—	—	—
Student 28	—	—	—	—	—	—	—
Student 29	Student 18	3	4	3.5	4	4	4
Student 30	—	—	—	—	—	—	—
Student 31	—	—	—	—	—	—	—
Student 32	—	—	—	—	—	—	—
Student 33	—	—	—	—	—	—	—
Student 34	—	—	—	—	—	—	—
Student 35	—	—	—	—	—	—	—

Nota: Some Students that gave feedback did not finish the *New Response (NR)* therefore the system was not able to obtain their feedback results, e.g., Student 21 gave feedback to Student 26, but Student 21 does not have feedback result.

Table E2

Scores for Response (R) and New Response (NR) for Activity 2.

Student	Student Feedback	R _{interpretation}	R _{inference}	R _{explanation}	R ₂	NR _{interpretation}	NR _{inference}	NR _{explanation}	NR ₂
Student 1	Student 31	4	4	4	4.0	4	4	4	4.0
Student 2	Student 30	2	2	2	2.0	2	2	2	2.0
Student 3	Student 9	4	4	4	4.0	4	4	4	4.0
Student 4	Student 2	2	2	3	2.3	2	2	3	2.3
Student 5	Student 25	1	2	2	1.7	1	2	2	1.7
Student 6	Student 16	3	3	2	2.7	3	3	3	3.0
Student 7	Student 1	3	2	2	2.3	4	4	4	4.0
Student 8	Student 10	2	2	3	2.3	2	2	3	2.3
Student 9	Student 4	3	3	3	3.0	4	3	4	3.7
Student 10	Student 3	4	4	3	3.7	4	4	4	4.0
Student 11	Student 14	3	2	3	2.7	3	3	3	3.0
Student 12	Student 11	2	2	2	2.0	3	4	3	3.3
Student 13	Student 6	4	4	4	4.0	4	4	4	4.0
Student 14	Student 7	1	1	1	1.0	1	1	1	1.0
Student 15	—	—	—	—	—	—	—	—	—
Student 16	Student 15	3	2	4	3.0	4	4	4	4.0
Student 17	Student 5	2	2	2	2.0	3	3	3	3.0
Student 18	Student 29	3	4	3	3.3	3	4	3	3.3
Student 19	Student 13	2	2	2	2.0	2	2	2	2.0
Student 20	Student 33	2	1	1	1.3	2	1	1	1.3
Student 21	Student 18	3	3	3	3.0	3	3	4	3.3
Student 22	Student 35	2	2	2	2.0	2	2	2	2.0
Student 23	Student 21	3	2	3	2.7	3	2	3	2.7
Student 24	Student 26	2	2	2	2.0	2	2	2	2.0
Student 25	Student 22	2	2	2	2.0	2	3	3	2.7
Student 26	Student 28	3	3	2	2.7	3	3	2	2.7
Student 27	—	—	—	—	—	—	—	—	—
Student 28	Student 32	2	2	2	2.0	2	2	2	2.0
Student 29	Student 17	4	4	3	3.7	4	4	4	4.0
Student 30	Student 24	2	2	2	2.0	2	2	2	2.0
Student 31	Student 23	2	2	2	2.0	3	2	2	2.3
Student 32	Student 20	2	2	2	2.0	2	2	2	2.0
Student 33	Student 19	2	2	2	2.0	2	2	2	2.0
Student 34	Student 12	3	4	2	3.0	3	4	2	3.0
Student 35	Student 8	3	3	3	3.0	3	4	3	3.3

Nota: Some Students that gave feedback did not finish the *New Response (NR)* therefore the system was not able to obtain their feedback results, e.g., Student 15 gave feedback to Student 16, but Student 15 does not have feedback result.

Table E3

Scores for Response (R) and New Response (NR) for Activity 3.

Student	Student Feedback	R _{evaluation}	R _{inference}	R ₃	NR _{evaluation}	NR _{inference}	NR ₃
Student 1	Student 16	4	3	3.5	4	3	3.5
Student 2	Student 34	1	1	1	1	1	1
Student 3	—	—	—	—	—	—	—
Student 4	Student 18	1	2	1.5	4	3	3.5
Student 5	—	—	—	—	—	—	—
Student 6	—	—	—	—	—	—	—
Student 7	Student 11	4	3	3.5	4	3	3.5
Student 8	Student 19	1	2	1.5	1	2	1.5
Student 9	Student 27	3	2	2.5	3	2	2.5
Student 10	Student 22	4	3	3.5	4	3	3.5
Student 11	Student 12	4	3	3.5	4	3	3.5
Student 12	Student 29	4	3	3.5	4	3	3.5
Student 13	Student 26	4	3	3.5	4	3	3.5
Student 14	—	—	—	—	—	—	—
Student 15	Student 4	3	2	2.5	3	2	2.5
Student 16	Student 2	2	2	2	2	2	2
Student 17	Student 24	3	3	3	3	3	3
Student 18	Student 25	2	2	2	2	2	2
Student 19	Student 9	3	2	2.5	3	2	2.5
Student 20	Student 8	4	2	3	4	2	3
Student 21	Student 32	3	3	3	3	3	3
Student 22	Student 31	2	2	2	2	2	2
Student 23	—	—	—	—	—	—	—
Student 24	Student 21	4	3	3.5	4	3	3.5
Student 25	Student 20	2	2	2	2	2	2
Student 26	Student 17	4	3	3.5	4	2	3
Student 27	Student 30	2	2	2	2	2	2
Student 28	—	—	—	—	—	—	—
Student 29	Student 15	2	2	2	2	2	2
Student 30	Student 13	4	3	3.5	4	3	3.5
Student 31	Student 28	1	2	1.5	1	2	1.5
Student 32	Student 1	4	3	3.5	4	3	3.5
Student 33	—	—	—	—	—	—	—
Student 34	Student 10	4	3	3.5	4	3	3.5
Student 35	—	—	—	—	—	—	—

Nota: Some Students that gave feedback did not finish the *New Response (NR)* therefore the system was not able to obtain their feedback results, e.g., Student 28 gave feedback to Student 31, but Student 28 does not have feedback result.

Table E4*Scores for Improvement between R and NR*

	Activity 1			Activity 2				Avance Activity 3		
	I _{analysis}	I _{explanation}	I ₁	I _{interpretation}	I _{inference}	I _{explanation}	I ₂	I _{evaluation}	I _{inference}	I ₃
Student 1	1	1	1.0	0	0	0	0.0	0	0	0.0
Student 2	–	–	–	0	0	0	0.0	0	0	0.0
Student 3	1	1	1.0	0	0	0	0.0	–	–	–
Student 4	–	–	–	0	0	0	0.0	3	1	2.0
Student 5	–	–	–	0	0	0	0.0	–	–	–
Student 6	–	–	–	0	0	1	0.3	–	–	–
Student 7	–	–	–	1	2	2	1.7	0	0	0.0
Student 8	0	0	0.0	0	0	0	0.0	0	0	0.0
Student 9	–	–	–	1	0	1	0.7	0	0	0.0
Student 10	3	2	2.5	0	0	1	0.3	0	0	0.0
Student 11	1	1	1.0	0	1	0	0.3	0	0	0.0
Student 12	–	–	–	1	2	1	1.3	0	0	0.0
Student 13	0	1	0.5	0	0	0	0.0	0	0	0.0
Student 14	–	–	–	0	0	0	0.0	–	–	–
Student 15	–	–	–	–	–	–	–	0	0	0.0
Student 16	1	1	1.0	1	2	0	1.0	0	0	0.0
Student 17	–	–	–	1	1	1	1.0	0	0	0.0
Student 18	1	1	1.0	0	0	0	0.0	0	0	0.0
Student 19	2	2	2.0	0	0	0	0.0	0	0	0.0
Student 20	–	–	–	0	0	0	0.0	0	0	0.0
Student 21	–	–	–	0	0	1	0.3	0	0	0.0
Student 22	–	–	–	0	0	0	0.0	0	0	0.0
Student 23	–	–	–	0	0	0	0.0	–	–	–
Student 24	1	0	0.5	0	0	0	0.0	0	0	0.0
Student 25	0	1	0.5	0	1	1	0.7	0	0	0.0
Student 26	1	2	1.5	0	0	0	0.0	0	-1	-0.5
Student 27	–	–	–	–	–	–	–	0	0	0.0
Student 28	–	–	–	0	0	0	0.0	–	–	–
Student 29	1	0	0.5	0	0	1	0.3	0	0	0.0
Student 30	–	–	–	0	0	0	0.0	0	0	0.0
Student 31	–	–	–	1	0	0	0.3	0	0	0.0
Student 32	–	–	–	0	0	0	0.0	0	0	0.0
Student 33	–	–	–	0	0	0	0.0	–	–	–
Student 34	–	–	–	0	0	0	0.0	0	0	0.0
Student 35	–	–	–	0	1	0	0.3	–	–	–

Note: The *Improvement* shows how much progress the students make between R and NR and is calculated as $NR - R$.

Table E5*Scores for Feedback Received (FR) and Feedback Given (FG) for Activity 1.*

Student	Student Feedback	Compulsory Indicators	Optional Indicators	FR ₁	FG ₁
Student 1	Student 26	11	6	17	23
Student 2	–	–	–	–	–
Student 3	Student 19	14	7	21	–
Student 4	–	–	–	–	–
Student 5	–	–	–	–	–
Student 6	–	–	–	–	–
Student 7	–	–	–	–	–
Student 8	Student 17	0	0	0	15
Student 9	–	–	–	–	–
Student 10	Student 6	12	2	14	20
Student 11	Student 1	13	10	23	–
Student 12	–	–	–	–	–
Student 13	Student 9	15	7	22	24
Student 14	–	–	–	–	–
Student 15	–	–	–	–	–
Student 16	Student 24	16	5	21	–
Student 17	–	–	–	–	–
Student 18	Student 25	14	8	22	20
Student 19	Student 8	11	4	15	21
Student 20	–	–	–	–	–
Student 21	–	–	–	–	–
Student 22	–	–	–	–	–
Student 23	–	–	–	–	–
Student 24	Student 13	16	8	24	21
Student 25	Student 10	16	4	20	22
Student 26	Student 21	14	2	16	17
Student 27	–	–	–	–	–
Student 28	–	–	–	–	–
Student 29	Student 18	16	4	20	–
Student 30	–	–	–	–	–
Student 31	–	–	–	–	–
Student 32	–	–	–	–	–
Student 33	–	–	–	–	–
Student 34	–	–	–	–	–
Student 35	–	–	–	–	–

Note: some of the students do not have a score for Feedback Given as the students who were supposed to receive said feedback did not finish the activity and therefore the feedback could not be retrieved.

Table E6*Scores for Feedback Received (FR) and Feedback Given (FG) for Activity 2.*

Student	Student Feedback	Compulsory Indicators	Optional Indicators	FR ₂	FG ₂
Student 1	Student 31	15	3	18	15
Student 2	Student 30	16	0	16	10
Student 3	Student 9	14	0	14	13
Student 4	Student 2	10	0	10	16
Student 5	Student 25	0	0	0	9
Student 6	Student 16	16	4	20	13
Student 7	Student 1	11	4	15	17
Student 8	Student 10	0	0	0	11
Student 9	Student 4	13	3	16	14
Student 10	Student 3	10	3	13	–
Student 11	Student 14	11	3	14	14
Student 12	Student 11	11	3	14	16
Student 13	Student 6	13	0	13	20
Student 14	Student 7	14	3	17	14
Student 15	–	–	–	–	–
Student 16	Student 15	10	4	14	20
Student 17	Student 5	9	0	9	–
Student 18	Student 29	15	0	15	13
Student 19	Student 13	16	4	20	–
Student 20	Student 33	13	0	13	–
Student 21	Student 18	13	0	13	–
Student 22	Student 35	0	0	0	14
Student 23	Student 21	0	0	0	14
Student 24	Student 26	0	0	0	–
Student 25	Student 22	14	0	14	–
Student 26	Student 28	0	0	0	–
Student 27	–	–	–	–	–
Student 28	Student 32	0	0	0	
Student 29	Student 17	0	0	0	15
Student 30	Student 24	0	0	0	16
Student 31	Student 23	14	0	14	18
Student 32	Student 20	0	0	0	–
Student 33	Student 19	0	0	0	13
Student 34	Student 12	16	0	16	–
Student 35	Student 8	11	0	11	–

Note: some of the students do not have a score for Feedback Given as the students who were supposed to receive said feedback did not finish the activity and therefore the feedback could not be retrieved.

Table E7*Scores for Feedback Received (FR) and Feedback Given (FG) for Activity 3.*

Student	Student Feedback	Compulsory Indicators	Optional Indicators	FR ₃	FG ₃
Student 1	Student 16	10	3	13	–
Student 2	Student 34	8	1	9	8
Student 3	–	–	–	–	–
Student 4	Student 18	9	3	12	11
Student 5	–	–	–	–	–
Student 6	–	–	–	–	–
Student 7	Student 11	9	5	14	–
Student 8	Student 19	15	4	19	12
Student 9	Student 27	9	0	9	14
Student 10	Student 22	9	2	11	–
Student 11	Student 12	13	0	13	14
Student 12	Student 29	15	0	15	13
Student 13	Student 26	16	0	16	–
Student 14	–	–	–	–	–
Student 15	Student 4	11	0	11	–
Student 16	Student 2	8	0	8	13
Student 17	Student 24	12	0	12	19
Student 18	Student 25	10	0	10	12
Student 19	Student 9	14	0	14	19
Student 20	Student 8	12	0	12	7
Student 21	Student 32	16	4	20	13
Student 22	Student 31	9	0	9	11
Student 23	–	–	–	–	–
Student 24	Student 21	13	0	13	12
Student 25	Student 20	7	0	7	10
Student 26	Student 17	16	3	19	16
Student 27	Student 30	0	0	0	9
Student 28	–	–	–	–	–
Student 29	Student 15	0	0	0	15
Student 30	Student 13	0	0	0	–
Student 31	Student 28	0	0	0	9
Student 32	Student 1	0	0	0	20
Student 33	–	–	–	–	–
Student 34	Student 10	0	0	0	9
Student 35	–	–	–	–	–

Note: some of the students do not have a score for Feedback Given as the students who were supposed to receive said feedback did not finish the activity and therefore the feedback could not be retrieved.

Table E8

Scores for the Rubric applied to the incorporation of the Feedback into the New Response $NR(FR)$.

	$NR(FR)_1$	$NR(FR)_2$	$NR(FR)_3$
Student 1	3	2	—
Student 2	—	2	2
Student 3	4	2	—
Student 4	—	2	4
Student 5	—	—	—
Student 6	—	4	—
Student 7	—	4	2
Student 8	—	—	—
Student 9	—	1	2
Student 10	4	3	1
Student 11	3	1	2
Student 12	—	3	2
Student 13	4	3	—
Student 14	—	2	—
Student 15	—	—	2
Student 16	1	3	3
Student 17	—	4	2
Student 18	3	3	3
Student 19	4	3	2
Student 20	—	1	3
Student 21	—	3	3
Student 22	—	—	3
Student 23	—	—	—
Student 24	3	—	3
Student 25	2	2	—
Student 26	4	—	—
Student 27	—	—	—
Student 28	—	—	—
Student 29	3	—	—
Student 30	—	—	—
Student 31	—	1	—
Student 32	—	—	—
Student 33	—	—	—
Student 34	—	4	—
Student 35	—	3	—

Table E9

Mean scores for the activities for Response (R), New Response (NR), Improvement (I), Feedback Received (FR), Feedback Given (FG) and Incorporation of the Feedback into the New Response NR (NR(FR)).

	$\bar{R}$	$\overline{NR}$	$\bar{I}$	$\overline{FR}$	$\overline{FG}$	$\overline{NR(FR)}$
Student 1	3.5	3.8	0.3	16.0	19.0	1.7
Student 2	1.5	1.5	0.0	12.5	9.0	2.0
Student 3	3.5	4.0	0.5	17.5	13.0	3.0
Student 4	1.9	2.9	1.0	11.0	13.5	3.0
Student 5	1.7	1.7	0.0	0.0	9.0	0.0
Student 6	2.7	3.0	0.3	20.0	13.0	4.0
Student 7	2.9	3.8	0.8	14.5	17.0	3.0
Student 8	1.8	1.8	0.0	6.3	12.7	0.0
Student 9	2.8	3.1	0.3	12.5	14.0	1.5
Student 10	2.9	3.8	0.9	12.7	20.0	2.7
Student 11	2.9	3.3	0.4	16.7	14.0	2.0
Student 12	2.8	3.4	0.7	14.5	14.5	2.5
Student 13	3.7	3.8	0.2	17.0	22.0	2.3
Student 14	1.0	1.0	0.0	17.0	14.0	2.0
Student 15	2.5	2.5	0.0	11.0	12.1	2.0
Student 16	2.7	3.3	0.7	14.3	16.5	2.3
Student 17	2.5	3.0	0.5	10.5	19.0	3.0
Student 18	2.8	3.1	0.3	15.7	15.0	3.0
Student 19	2.2	2.8	0.7	16.3	20.0	3.0
Student 20	2.2	2.2	0.0	12.5	7.0	2.0
Student 21	3.0	3.2	0.2	16.5	13.0	3.0
Student 22	2.0	2.0	0.0	4.5	12.5	1.5
Student 23	2.7	2.7	0.0	0.0	14.0	0.0
Student 24	3.0	3.2	0.2	12.3	16.5	2.0
Student 25	2.0	2.4	0.4	13.7	16.0	1.3
Student 26	2.9	3.2	0.3	11.7	16.5	1.3
Student 27	2.0	2.0	0.0	0.0	9.0	0.0
Student 28	2.0	2.0	0.0	0.0	12.1	0.0
Student 29	3.1	3.3	0.3	6.7	15.0	1.0
Student 30	2.8	2.8	0.0	0.0	16.0	0.0
Student 31	1.8	1.9	0.2	7.0	13.5	0.5
Student 32	2.8	2.8	0.0	0.0	20.0	0.0
Student 33	2.0	2.0	0.0	0.0	13.0	0.0
Student 34	3.3	3.3	0.0	8.0	9.0	2.0
Student 35	3.0	3.3	0.3	11.0	15.4	3.0

Note: Appendix F provides an example to explain how the averages of R, NR, Improvement, FR, FG and Incorporation of FR into NR are calculated.

Table E10

Scores for Test 1 (T1) and Test 2 (T2).

Student	T1	T2	Range of scale
Student 1	5.7	5.6	1-7
Student 2	4.8	4.6	1-7
Student 3	4.7	6.9	1-7
Student 4	5.0	6.0	1-7
Student 5	5.5	5.0	1-7
Student 6	5.1	4.7	1-7
Student 7	4.8	6.1	1-7
Student 8	5.3	5.7	1-7
Student 9	4.7	4.4	1-7
Student 10	4.8	5.2	1-7
Student 11	5.4	6.0	1-7
Student 12	5.5	5.9	1-7
Student 13	6.2	6.2	1-7
Student 14	5.0	0.0	1-7
Student 15	5.4	5.9	1-7
Student 16	4.8	5.3	1-7
Student 17	5.7	5.3	1-7
Student 18	5.5	5.3	1-7
Student 19	5.5	6.4	1-7
Student 20	5.2	5.2	1-7
Student 21	6.2	5.1	1-7
Student 22	4.3	5.2	1-7
Student 23	5.7	5.1	1-7
Student 24	5.0	6.0	1-7
Student 25	6.0	6.4	1-7
Student 26	4.7	5.3	1-7
Student 27	4.9	6.1	1-7
Student 28	4.3	4.6	1-7
Student 29	4.5	5.9	1-7
Student 30	4.1	6.6	1-7
Student 31	5.1	4.4	1-7
Student 32	4.3	5.5	1-7
Student 33	5.6	5.5	1-7
Student 34	5.1	5.5	1-7
Student 35	5.1	6.6	1-7

Note: T1 (*Mean* = 5.13, *Std. Dev.* = 0.53) and
T2 (*Mean* = 5.41, *Std. Dev.* = 1.14)

F. Description, examples and calculation of the variables included in the linear regression model

To explain how the variables that were included in the model were obtained, we will consider the following hypothetical situation: Student 1 participated in activities 1, 2 and 3; Student 2 only participated in activity 2; and Student 3 participated in activities 1 and 3.

E_x : Student “X”

R_x : Response to Activity “X”

NR_x : New Response to Activity “X”

FR_x : Feedback Received in Activity “X”

FG_x : Feedback Given in Activity “X”

7. **Average for Response (R)**: the average score for the initial responses (R) to the activities in which each student participated.

Average Response for Student 1:

$$\overline{R} E_1 = \frac{R_1 + R_2 + R_3}{3}$$

Average Response for Student 2:

$$\overline{R} E_2 = R_2$$

Average Response for Student 3:

$$\overline{R} E_3 = \frac{R_1 + R_3}{2}$$

8. **Average for New Response (NR)**: the average score for the final responses (NR) to the activities in which each student participated.

Average New Response for Student 1:

$$\overline{NR} E_1 = \frac{NR_1 + NR_2 + NR_3}{3}$$

Average New Response for Student 2:

$$\overline{NR} E_2 = NR_2$$

Average New Response for Student 3:

$$\overline{NR} E_3 = \frac{NR_1 + NR_3}{2}$$

9. **Average for Improvement (NR – R):** the difference between the average Response (R) and average New Response (NR) for the activities in which the student participated.

Improvement by Student 1:

$$\overline{A}E_1 = \overline{NR}E_1 - \overline{R}E_1$$

Improvement by Student 2:

$$\overline{A}E_2 = \overline{NR}E_2 - \overline{R}E_2$$

Improvement by Student 3:

$$\overline{A}E_3 = \overline{NR}E_3 - \overline{R}E_3$$

10. **Average for Feedback Received (FR):** the average score for the feedback received by the student for the activities in which they participated.

Feedback Received by Student 1:

$$\overline{FR}E_1 = \frac{FR_1 + FR_2 + FR_3}{3}$$

Feedback Received by Student 2:

$$\overline{FR}E_2 = FR_2$$

Feedback Received by Student 3:

$$\overline{FR}E_3 = \frac{FR_1 + FR_3}{2}$$

11. **Average for Feedback Given (FG):** the average score for the feedback given by the student for the activities in which they participated.

Feedback Given by Student 1:

$$\overline{FD}E_1 = \frac{FD_1 + FD_2 + FD_3}{3}$$

Feedback Given by Student 2:

$$\overline{FD}E_2 = FD_2$$

Feedback Given by Student 3:

$$\overline{FD} E_3 = \frac{FD_1 + FD_3}{2}$$

12. **Average for Incorporating FR into NR (FRNR):** the average score assigned using the Rubric for the incorporation of F into NR for each New Response for the activities in which each student participated.
This variable was not included in the model.

Incorporation of FR into NR by Student 1:

$$\overline{NR(FR)} E_1 = \frac{NR(FR)_1 + NR(FR)_2 + NR(FR)_3}{3}$$

Incorporation of FR into NR by Student 2:

$$\overline{NR(FR)} E_2 = NR(FR)_2$$

Incorporation of FR into NR by Student 3:

$$\overline{NR(FR)} E_3 = \frac{NR(FR)_1 + NR(FR)_3}{2}$$

G. Process for selecting variables for the linear regression model for *Improvement*

Table G1

Models generated using the stepwise selection method with the dependent variable Improvement

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
	β	Std. Error	β	Std. Error	β	Std. Error	β	Std. Error	β	Std. Error
$\overline{FR}$	0.5232***	0.1484	0.4240***	0.1427	0.4577***	0.1386	0.5380***	0.1424	0.6177***	0.1495
$\overline{FD}$	—	—	0.3644**	0.1427	0.3078**	0.1407	0.3948**	0.1456	0.4031***	0.1429
$T2$	—	—	—	—	0.2544*	0.1357	0.4075**	0.1591	0.4419***	0.1577
$\bar{R}$	—	—	—	—	—	—	−0.3047*	0.1778	−0.3316*	0.1752
$T1$	—	—	—	—	—	—	—	—	−0.2024	0.1359
Observaciones	35		35		35		35		35	
Adj R ²	0.2517		0.3589		0.4056		0.4405		0.4624	

Note: model using standardized variables (*Mean* = 0 and *Std. Deviation* = 1)

*p<0.1, **p<0.05, ***p<0.01

Table G2

Forward stepwise selection method for the model for Improvement (N=35)

Variable	β	p > t	Adj R ²	$[\Delta R]^2$	Prob > F
Step 1			.25	.25	.00
$\overline{FR}$	.52	.00			
Step 2			.36	.11	.00
$\overline{FR}$	.42	.00			
$\overline{FD}$	.36	.01			
Step 3			.41	.05	.00
$\overline{FR}$	.45	.00			
$\overline{FD}$	.30	.03			
$T2$	.25	.07			
Step 4			.40	.03	.00
$\overline{FR}$	.53	.00			
$\overline{FD}$	.39	.01			
$T2$	.40	.01			
$\bar{R}$	−.30	.09			
Step 5			.46	.02	.00
$\overline{FR}$	.62	.00			
$\overline{FD}$	.40	.00			
$T2$	.44	.00			
$\bar{R}$	−.33	.07			
$T1$	−.20	.14			

Note: model using standardized variables (*Mean* = 0 and *Std. Deviation* = 1). At each stage, the variable that led to the biggest increase in the Adj R was added to the model.

H. Data used in the linear regression model

Table H1

Data used in STATA v. 12.0 to develop a linear regression model to predict Improvements in students' critical thinking skills.

Student	$\bar{I}$	$\bar{R}$	$\bar{FR}$	$\bar{FG}$	T1	T2
Student 1	0.3	3.5	16.0	19.0	5.7	5.6
Student 2	0.0	1.5	12.5	9.0	4.8	4.6
Student 3	0.5	3.5	17.5	13.0	4.7	6.9
Student 4	1.0	1.9	11.0	13.5	5.0	6.0
Student 5	0.0	1.7	0.0	9.0	5.5	5.0
Student 6	0.3	2.7	20.0	13.0	5.1	4.7
Student 7	0.8	2.9	14.5	17.0	4.8	6.1
Student 8	0.0	1.8	6.3	12.7	5.3	5.7
Student 9	0.3	2.8	12.5	14.0	4.7	4.4
Student 10	0.9	2.9	12.7	20.0	4.8	5.2
Student 11	0.4	2.9	16.7	14.0	5.4	6.0
Student 12	0.7	2.8	14.5	14.5	5.5	5.9
Student 13	0.2	3.7	17.0	22.0	6.2	6.2
Student 14	0.0	1.0	17.0	14.0	5.0	0.0
Student 15	0.0	2.5	11.0	12.1	5.4	5.9
Student 16	0.7	2.7	14.3	16.5	4.8	5.3
Student 17	0.5	2.5	10.5	19.0	5.7	5.3
Student 18	0.3	2.8	15.7	15.0	5.5	5.3
Student 19	0.7	2.2	16.3	20.0	5.5	6.4
Student 20	0.0	2.2	12.5	7.0	5.2	5.2
Student 21	0.2	3.0	16.5	13.0	6.2	5.1
Student 22	0.0	2.0	4.5	12.5	4.3	5.2
Student 23	0.0	2.7	0.0	14.0	5.7	5.1
Student 24	0.2	3.0	12.3	16.5	5.0	6.0
Student 25	0.4	2.0	13.7	16.0	6.0	6.4
Student 26	0.3	2.9	11.7	16.5	4.7	5.3
Student 27	0.0	2.0	0.0	9.0	4.9	6.1
Student 28	0.0	2.0	0.0	12.1	4.3	4.6
Student 29	0.3	3.1	6.7	15.0	4.5	5.9
Student 30	0.0	2.8	0.0	16.0	4.1	6.6
Student 31	0.2	1.8	7.0	13.5	5.1	4.4
Student 32	0.0	2.8	0.0	20.0	4.3	5.5
Student 33	0.0	2.0	0.0	13.0	5.6	5.5
Student 34	0.0	3.3	8.0	9.0	5.1	5.5
Student 35	0.3	3.0	11.0	15.4	5.1	6.6

Note: Data of Improvement (I), Response (R), Feedback Received (FR), Feedback Given (FG) are retrieved from Table E9, while T1 and T2 are scores of the course assessment.

I. Distribution of Improvement

Table I1

Distribution of improvement between Response (R) and New Response (NR) for each activity, broken down by critical thinking skill.

	Activity 1	Activity 2	Activity 3
<i>Analysis</i>			
NR – R = 0	25.10%	-	-
NR – R = 1	58.3%	-	-
NR – R = 2	8.3%	-	-
NR – R = 3	8.3%	-	-
<i>Explanation</i>			
NR – R = 0	16.70%	85.00%	-
NR – R = 1	58.3%	13.3%	-
NR – R = 2	25.0%	1.7%	-
NR – R = 3	0.0%	0.0%	-
<i>Interpretation</i>			
NR – R = 0	-	80.00%	-
NR – R = 1	-	20.0%	-
NR – R = 2	-	0.0%	-
NR – R = 3	-	0.0%	-
<i>Inference</i>			
NR – R = 0	-	76.70%	96.30%
NR – R = 1	-	13.3%	3.7%
NR – R = 2	-	10.0%	0.0%
NR – R = 3	-	0.0%	0.0%
<i>Evaluation</i>			
NR – R = 0	-	-	96.30%
(NR) – (R) = 1	-	-	0.0%
(NR) – (R) = 2	-	-	0.0%
(NR) – (R) = 3	-	-	3.7%

J. Sample use of the scaffolding model

The aim of this example is to show how the Model in Figure 2 can be applied. The example is based on authentic student responses, taken from different stages of the activity.

1. Interpretation: Comprehend and express the meaning or significance of a wide variety of experiences, situations, data, events, judgments, conventions, beliefs, rules, procedures, or criteria.

Representative response for interpretation:

Context: “Let’s take the case of a 3rd grade group. I think that the first step towards meeting the target of 83 words per week (according to the authors) is to assess the students’ knowledge of academic vocabulary by having them read a text”.

General methodology: “In order to do, it would be good to work with the History and Science teachers (subjects that typically involve specialized reading) so that they show us the texts they will be working with this week in class and then we can use them in our Language Arts class to teach academic vocabulary...”

Argumentation: “...this would be beneficial because the children would familiarize themselves with the texts and understand the other subjects more, while their reading activity would also be more contextualized and meaningful”

Pedagogical instance: “For the assessment, the teacher will read the texts beforehand in order to decide which are the most appropriate words for teaching, selecting a minimum of 5 and a maximum of 10 words per text. This selection will be made based on criteria such as how useful the words are and how important they are to the overall comprehension of the text. The text will be read out loud in class and the students will be asked which words they don’t know. The teacher will then write these words on the board and use the ones that coincide with the words they selected previously. The remaining words will be left for a future class. The teacher will then communicate the objective of the class and give an easy-to-understand definition of the words. The students will be asked to draw a picture for each word, which they will then use to explain the meaning of the word to a classmate.”

Exemplification: “For example, if the word is “dazzling”, the children will draw things that they think represent this word and at the end of the class they will show their drawings to their classmates and explain how they relate to the word in question.”

2. Inference: Identify and secure elements needed to draw reasonable conclusions; to form conjectures and hypotheses; to consider relevant information and to deduce the consequences flowing from data, statements, principles, evidence, judgments, beliefs, opinions, concepts, descriptions, questions, or other forms of representation.

Representative response for inference:

Context: “Given the number of words that according to the authors the students must learn each year...”

Hypothesis: “...this vocabulary must be taught explicitly, i.e. not just giving the students a text containing the target words, but instead using that text to extract the words that we want to teach...”

Argumentation: “...because by doing this kind of exercise the students will be learning said words within a context (text) that will make the learning more meaningful for them.”

Context: “Given that there are approximately 6 lessons of Language Arts each week...”

General methodology: “...the students’ vocabulary can be developed explicitly at the beginning of every class, relating these to the topics that they are studying and choosing approximately 10 to 15 words (of varying levels of complexity)...”

Pedagogical instance: “...and then teaching and practicing these using different strategies. This could include asking the students questions that will provide the words with a context, asking them for examples of where they have seen the words used, or asking them to choose between 2 or 3 images that represent the words, among others. The idea is to mix up these strategies so that acquiring vocabulary does not become a monotonous and boring activity for the students.

3. Explanation: State the results of one's reasoning; to justify that reasoning in terms of the evidential, conceptual, methodological, criteriological and contextual considerations upon which one's results were based; and to present one's reasoning in the form of cogent arguments.

Representative response for explanation:

Context: “In the learning sequence...”

Conceptualization: “you can clearly see two of the four principles of teaching fluency. Firstly, there is the principle of repeated reading of the text, which is present in the activity when the students have to read the text several times in different ways: with the teacher, with a group of classmates and on their own. Secondly, there is the principle of modelled reading...”

Argumentation: “...given that the students receive a model from the teacher, as she is the one who reads out loud at the beginning of the activity so that the students can see how a text should be read fluently before they progressively start to read like this themselves.”

Conceptualization: “The two aspects of teaching fluency that need to be improved are feedback and meaningful context.”

Argumentation: “The first of these aspects is not entirely evident because, although the teacher gives the students some tips, these are pre-defined and are not based on the students’ achievements or areas of improvement...”

Pedagogical instance: “...perhaps it would be better if the teacher were to give ongoing feedback to each of the students or groups, based on the students’ achievements and what they still need to achieve.”

Argumentation: “The second of these two aspects is also not particularly evident because, although the teacher gives the groups names based on the text (Athenians and Spartans), these are perhaps not so meaningful for the students if they haven’t yet managed to understand the text.”

Pedagogical instance: “It would be good if the teacher could give the students some context before doing the modelled reading so that they can relate it to their own knowledge and experiences.”

Connection: “In general, the activity demonstrates a gradual release of responsibility as the teacher increasingly gives the students the opportunity to read on their own based on the progress they make with their reading fluency.”

4. Analysis: Identify the intended and actual inferential relationships among statements, questions, concepts, descriptions, or other forms of representation intended to express belief, judgment, experiences, reasons, information, or opinions.

Representative response for analysis:

Context: “Based on the key principles of teaching fluency...”

Connection: “we can see that the class script includes the presence of a model reader, which in this case would be the teacher...”

Argumentation: “...as the first point says that “the teacher reads the text out loud to the students””.

Connection: “Repeated reading of texts is also present...”

Argumentation: “...as the students participate in choral reading, then read in groups with the teacher, then in groups without the teacher, before another choral reading at the end of the class. They therefore read the same text several times.”

Evaluation: “Furthermore, another of the principles that is only partially present is ongoing feedback. This is because although the teacher gives the students tips that help them think about their reading and improve their performance, at no point does she focus on their achievements. Furthermore, the feedback is not based on issues that arise as the students are reading; instead the tips are pre-defined. Additionally, the activity is missing a meaningful context to justify its purpose. Although there is a focus on improving fluency, there is no background information to help motivate the students to read the text over and over without it becoming tedious.”

Pedagogical instance: “The activity could be improved by framing it with peer-led tutorials, where the students would alternate between evaluating and being evaluated, thus adding greater purpose to the idea of having to constantly listen to their classmates read a text.”

Connection: “Finally, this is connected to the idea of the gradual release of responsibility...”

Argumentation: “as the level of support that the teacher provides changes with each activity, starting with explicit instruction and modelled reading, before supporting the groups and, finally, allowing them to read by themselves. Therefore, the control of the reading activity goes from belonging exclusively to the teacher, to being shared by the teacher and a group of students, who engage in guided reading and receive only occasional support from the teacher.”

5. Evaluation: Assess the credibility of statements or other representations which are accounts or descriptions of a person’s perception, experience, situation, judgment, belief, or opinion; and to assess the logical strength

of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation.

Representative response for evaluation:

Context: “I would use proposal 3”

Evaluation: “...because I think that the title and the opening paragraph give a good overview of the topic that will be covered in class, then the visual text attracts the students’ attention and helps them to focus on the example that will be used.”

Connection and argumentation: “Finally, it would be good to start with section 3 of the process as it contains information that is familiar to the students (an adult plant with a tomato) and is essential for them to understand what happens in section 1. I would then analyze section 2.”

Pedagogical instance: “Questions:

Title and opening paragraph: In section 2 of the text it says that “The tomato plant, like any plant that comes from a seed, has three stages in its life cycle.” Do you think that there are any plants that don’t come from a seed?

Visual text: Looking at the images, what do you think happens to the tomato plant flower once the fruit has grown?

Sections 1-3: If tomatoes have seeds inside them and flowers come from seeds, what comes first: the flower or the tomato? Why?

K. Tables with the correlations of the variables discarded from the regression model

Table K1:

Correlation between $\bar{R}$ and the variables discarded from the model due to the redundancy of the information in $\bar{R}$

Activity	Variable	<i>r</i>
Activity 1 (N=13)	$R_{analysis}$	0.71 [*]
	$R_{explanation}$	0.57 ^{**}
	R_1	0.72 [*]
Activity 2 (N=33)	$R_{interpretation}$	0.83 [*]
	$R_{inference}$	0.73 [*]
	$R_{explanation}$	0.60 ^{**}
	R_2	0.79 [*]
Activity 3 (N=27)	$R_{evaluation}$	0.75 [*]
	$R_{inference}$	0.75 [*]
	R_3	0.78 [*]
Note: (°) Denotes a strong correlation, (**) Denotes a moderate correlation		

Table K2:

Correlation between $\overline{NR}$ and the variables discarded from the model due to the redundancy of the information in $\overline{NR}$

Activity	Variable	<i>r</i>
Activity 1 (N=13)	$NR_{analysis}$	0.81 [*]
	$NR_{explanation}$	0.83 [*]
	NR_1	0.83 [*]
Activity 2 (N=33)	$NR_{interpretation}$	0.82 [*]
	$NR_{inference}$	0.82 [*]
	$NR_{explanation}$	0.73 [*]
	NR_2	0.84 [*]
Activity 3 (N=27)	$NR_{evaluation}$	0.69 [*]
	$NR_{inference}$	0.69 [*]
	NR_3	0.73 [*]
Note: (°) Denotes a strong correlation, (**) Denotes a moderate correlation		

Table K3:

Correlation between $\bar{F}$ and the variables discarded from the model due to the redundancy of the information in $\bar{F}$

Variable	<i>r</i>
F_1	0.57 ^{**}
F_2	0.82 [*]
F_3	0.69 [*]
$NR(FR)_1$	0.50 ^{**}
$NR(FR)_2$	0.68 [*]
$NR(FR)_3$	0.42 ^{**}
$\overline{NR(NR)}$	0.87 [*]

Note: (^{*}) Denotes a strong correlation,
(^{**}) Denotes a moderate correlation

L. Double coding of Response (R), New Response (NR) and Incorporation of Feedback into New Response (NR(FR))

Table L1:

Double-Coding scores for the Rubric applied to Response (R) for Activity 1.

Student	Analysis		Explanation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	3	2	3	3
Student 2	—	—	—	—
Student 3	3	2	3	3
Student 4	—	—	—	—
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	—	—	—	—
Student 8	1	1	2	2
Student 9	3	3	2	2
Student 10	1	1	2	2
Student 11	2	3	3	3
Student 12	—	—	—	—
Student 13	4	3	3	3
Student 14	—	—	—	—
Student 15	—	—	—	—
Student 16	3	3	3	3
Student 17	2	2	2	2
Student 18	3	2	3	3
Student 19	2	2	2	2
Student 20	—	—	—	—
Student 21	—	—	—	—
Student 22	—	—	—	—
Student 23	—	—	—	—
Student 24	3	3	4	4
Student 25	2	2	2	2
Student 26	3	2	2	3
Student 27	—	—	—	—
Student 28	—	—	—	—
Student 29	3	3	4	4
Student 30	—	—	—	—
Student 31	—	—	—	—
Student 32	—	—	—	—
Student 33	—	—	—	—
Student 34	—	—	—	—
Student 35	—	—	—	—

Table L2:*Double-Coding scores for the Rubric applied to Response (R) for Activity 2.*

Student	Explanation		Interpretation		Inference	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	4	4	4	4	4	4
Student 2	2	2	2	2	2	2
Student 3	4	4	4	4	4	4
Student 4	3	3	2	2	2	2
Student 5	2	2	1	1	2	2
Student 6	2	2	3	2	3	3
Student 7	2	2	3	2	2	2
Student 8	3	3	2	2	2	2
Student 9	3	3	3	3	3	2
Student 10	3	3	4	4	4	4
Student 11	3	2	3	3	2	2
Student 12	2	2	2	2	2	2
Student 13	4	4	4	4	4	4
Student 14	1	1	1	1	1	1
Student 15	—	—	—	—	—	—
Student 16	4	3	3	3	2	2
Student 17	2	2	2	2	2	2
Student 18	3	3	3	3	4	4
Student 19	2	2	2	2	2	2
Student 20	1	1	2	2	1	1
Student 21	3	2	3	2	3	3
Student 22	2	2	2	2	2	2
Student 23	3	3	3	3	2	2
Student 24	2	2	2	2	2	2
Student 25	2	2	2	2	2	2
Student 26	2	2	3	2	3	3
Student 27	—	—	—	—	—	—
Student 28	2	2	2	2	2	2
Student 29	3	4	4	4	4	4
Student 30	2	2	2	2	2	2
Student 31	2	2	2	2	2	2
Student 32	2	2	2	2	2	2
Student 33	2	2	2	2	2	2
Student 34	2	2	3	3	4	4
Student 35	3	3	3	3	3	3

Table L3:*Double-Coding scores for the Rubric applied to Response (R) for Activity 3.*

Student	Inference		Evaluation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	3	3	4	4
Student 2	1	2	1	1
Student 3	—	—	—	—
Student 4	2	2	1	1
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	3	3	4	4
Student 8	2	2	1	1
Student 9	2	3	3	3
Student 10	3	3	4	4
Student 11	3	3	4	4
Student 12	3	3	4	4
Student 13	3	3	4	4
Student 14	—	—	—	—
Student 15	2	2	3	3
Student 16	2	2	2	2
Student 17	3	3	3	4
Student 18	2	2	2	2
Student 19	2	2	3	3
Student 20	2	2	4	4
Student 21	3	3	3	3
Student 22	2	2	2	2
Student 23	—	—	—	—
Student 24	3	3	4	4
Student 25	2	2	2	2
Student 26	3	3	4	4
Student 27	2	2	2	2
Student 28	—	—	—	—
Student 29	2	2	2	2
Student 30	3	3	4	4
Student 31	2	2	1	1
Student 32	3	3	4	4
Student 33	—	—	—	—
Student 34	3	3	4	4
Student 35	—	—	—	—

Table L4:*Double-Coding scores for the Rubric applied to New Response (NR) for Activity 1.*

Student	Analysis		Explanation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	4	4	4	4
Student 2	—	—	—	—
Student 3	4	4	4	4
Student 4	—	—	—	—
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	—	—	—	—
Student 8	1	2	2	2
Student 9	—	—	—	—
Student 10	4	4	4	4
Student 11	3	4	4	4
Student 12	—	—	—	—
Student 13	4	4	4	4
Student 14	—	—	—	—
Student 15	—	—	—	—
Student 16	4	4	4	4
Student 17	—	—	—	—
Student 18	4	4	4	4
Student 19	4	4	4	4
Student 20	—	—	—	—
Student 21	—	—	—	—
Student 22	—	—	—	—
Student 23	—	—	—	—
Student 24	4	4	4	4
Student 25	2	3	3	3
Student 26	4	4	4	4
Student 27	—	—	—	—
Student 28	—	—	—	—
Student 29	4	3	4	4
Student 30	—	—	—	—
Student 31	—	—	—	—
Student 32	—	—	—	—
Student 33	—	—	—	—
Student 34	—	—	—	—
Student 35	—	—	—	—

Table L5:*Double-Coding scores for the Rubric applied to New Response (NR) for Activity 2.*

Student	Explanation		Interpretation		Inference	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	4	4	4	4	4	4
Student 2	2	2	2	2	2	2
Student 3	4	4	4	4	4	4
Student 4	3	3	2	2	2	2
Student 5	2	2	1	1	2	2
Student 6	3	3	3	3	3	3
Student 7	4	4	4	4	4	4
Student 8	3	3	2	2	2	2
Student 9	4	4	4	4	3	3
Student 10	4	4	4	4	4	4
Student 11	3	3	3	3	3	3
Student 12	3	4	3	3	4	4
Student 13	4	4	4	4	4	4
Student 14	1	1	1	1	1	1
Student 15	—	—	—	—	—	—
Student 16	4	4	4	4	4	4
Student 17	3	3	3	3	3	3
Student 18	3	3	3	3	4	4
Student 19	2	2	2	2	2	2
Student 20	1	1	2	2	1	1
Student 21	4	4	3	3	3	3
Student 22	2	2	2	2	2	2
Student 23	3	3	3	3	2	2
Student 24	2	2	2	2	2	3
Student 25	3	3	2	2	3	3
Student 26	2	2	3	3	3	3
Student 27	—	—	—	—	—	—
Student 28	2	2	2	2	2	2
Student 29	4	4	4	4	4	4
Student 30	2	2	2	2	2	2
Student 31	2	2	3	2	2	2
Student 32	2	2	2	2	2	2
Student 33	2	2	2	2	2	2
Student 34	2	2	3	3	4	4
Student 35	3	3	3	3	4	4

Table L6:*Double-Coding scores for the Rubric applied to New Response (NR) for Activity 3.*

Student	Inference		Evaluation	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
Student 1	3	3	4	4
Student 2	1	2	1	1
Student 3	—	—	—	—
Student 4	3	3	4	4
Student 5	—	—	—	—
Student 6	—	—	—	—
Student 7	3	3	4	4
Student 8	2	2	1	1
Student 9	2	2	3	3
Student 10	3	3	4	4
Student 11	3	3	4	4
Student 12	3	3	4	4
Student 13	3	3	4	4
Student 14	—	—	—	—
Student 15	2	2	3	3
Student 16	2	2	2	2
Student 17	3	3	3	4
Student 18	2	2	2	2
Student 19	2	2	3	3
Student 20	2	2	4	4
Student 21	3	3	3	3
Student 22	2	2	2	2
Student 23	—	—	—	—
Student 24	3	3	4	4
Student 25	2	2	2	2
Student 26	2	3	4	4
Student 27	2	2	2	2
Student 28	—	—	—	—
Student 29	2	2	2	2
Student 30	3	3	4	4
Student 31	2	2	1	1
Student 32	3	3	4	4
Student 33	—	—	—	—
Student 34	3	3	4	4
Student 35	—	—	—	—

Table L7:

Double-Coding scores for the Rubric applied to Feedback (F) for Activity 1.

Student	Assessment criteria		Clarity and relevance of the explanation		Argumentation		Syntax, lexis and register		Request for evidence		Reviewing an argument or presenting a new argument and/or counter-argument		Request to reword or rework		Highlighting the focus of the question	
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2
Student 1	4	4	2	2	2	2	3	3	3	3	X	X	X	X	3	3
Student 2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 3	4	4	3	3	4	4	3	3	4	4	X	X	X	X	3	4
Student 4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 10	4	4	2	2	2	2	4	4	X	X	X	X	2	X	X	X
Student 11	4	4	3	3	3	3	3	3	4	4	X	X	2	2	4	4
Student 12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 13	4	4	3	3	4	4	4	4	4	4	X	X	X	X	3	3
Student 14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 16	4	4	4	4	4	4	4	4	3	3	X	X	2	2	X	X
Student 17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 18	4	4	3	3	3	3	4	4	2	2	X	X	3	3	3	3
Student 19	2	2	2	2	3	3	4	4	2	2	X	X	X	X	2	2
Student 20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 24	4	4	4	4	4	4	4	4	4	4	X	X	X	X	4	4
Student 25	4	4	4	4	4	4	4	4	4	4	X	X	X	X	X	X
Student 26	4	4	3	4	3	3	4	4	X	X	X	X	X	X	2	3
Student 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 29	4	4	4	4	4	4	4	4	X	X	X	X	X	X	4	4
Student 30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Student 35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: E1 and E2 are Evaluator 1 and Evaluator respectively.

Table L8:

Double-Coding scores for the Rubric applied to Feedback (F) for Activity 2.

Student	Assessment criteria		Clarity and relevance of the explanation		Argumentation		Syntax, lexis and register		Request for evidence		Reviewing an argument or presenting a new argument and/or counter-argument		Request to reword or rework		Highlighting the focus of the question	
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2
Estudiante 1	4	3	4	4	3	3	4	4	X	X	3	2	X	X	X	X
Estudiante 2	4	4	4	4	4	4	4	4	X	X	X	X	X	3	X	X
Estudiante 3	3	3	4	4	3	3	4	4	X	X	X	X	X	X	X	X
Estudiante 4	2	4	3	4	1	1	4	4	X	X	X	X	X	X	X	X
Estudiante 5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 6	4	4	4	4	4	4	4	4	4	4	X	X	X	X	X	X
Estudiante 7	3	3	3	2	2	2	3	3	X	X	4	X	X	X	X	X
Estudiante 8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 9	4	4	3	3	3	3	3	3	X	X	X	X	3	3	X	X
Estudiante 10	3	3	3	3	2	2	2	2	3	3	X	X	X	X	X	X
Estudiante 11	1	1	4	4	2	2	4	4	3	3	X	X	X	X	X	X
Estudiante 12	1	1	4	4	2	2	4	4	3	3	X	X	X	X	X	X
Estudiante 13	3	3	2	2	4	3	4	4	X	X	X	X	X	X	X	X
Estudiante 14	4	4	4	3	2	2	4	3	X	X	3	2	X	X	X	X
Estudiante 15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 16	2	2	3	3	2	2	3	3	2	2	X	X	X	X	2	2
Estudiante 17	1	1	2	2	2	2	4	4	X	X	X	X	X	X	X	X
Estudiante 18	4	3	4	3	3	2	4	4	X	X	X	X	X	X	X	X
Estudiante 19	4	4	4	4	4	4	4	4	X	4	4	4	X	X	X	X
Estudiante 20	3	4	4	4	2	3	4	4	X	X	X	3	X	X	X	3
Estudiante 21	3	3	4	4	2	2	4	4	X	X	X	X	X	X	X	X
Estudiante 22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 25	4	4	4	4	2	2	4	4	X	X	X	X	X	X	X	X
Estudiante 26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 31	4	4	4	4	2	2	4	4	X	X	X	X	X	X	X	X
Estudiante 32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 34	4	4	4	4	4	3	4	4	X	X	X	X	X	X	X	X
Estudiante 35	3	3	3	3	2	2	3	3	X	X	X	X	X	X	X	X

Note: E1 and E2 are Evaluator 1 and Evaluator respectively.

Table L9:

Double-Coding scores for the Rubric applied to Feedback (F) for Activity 3.

Student	Assessment criteria		Clarity and relevance of the explanation		Argumentation		Syntax, lexis and register		Request for evidence		Reviewing an argument or presenting a new argument and/or counter-argument		Request to reword or rework		Highlighting the focus of the question	
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2
Estudiante 1	2	2	3	3	3	4	2	2	3	X	X	X	X	X	X	4
Estudiante 2	2	2	3	3	1	2	2	1	X	X	1	2	X	X	X	X
Estudiante 3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 4	1	1	3	3	2	2	3	3	X	X	X	X	1	1	2	X
Estudiante 5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 7	2	2	3	2	2	2	2	2	X	X	X	X	2	2	3	3
Estudiante 8	3	3	4	3	4	3	4	4	X	X	4	4	X	X	X	X
Estudiante 9	1	1	3	3	3	2	2	2	X	X	X	X	X	X	X	X
Estudiante 10	2	1	2	2	2	1	3	3	X	X	X	X	2	2	X	X
Estudiante 11	3	4	3	3	4	4	3	3	X	X	X	X	X	X	X	X
Estudiante 12	3	3	4	4	4	4	4	4	X	X	X	X	X	X	X	X
Estudiante 13	4	4	4	3	4	2	4	3	X	X	X	X	X	X	X	X
Estudiante 14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 15	2	2	3	3	3	3	3	4	X	2	X	X	X	X	X	X
Estudiante 16	2	2	2	2	2	2	2	2	X	X	X	X	X	X	X	X
Estudiante 17	3	4	3	4	3	3	3	3	X	X	X	X	X	X	X	X
Estudiante 18	3	3	3	3	2	2	2	3	X	X	X	X	X	X	X	X
Estudiante 19	3	2	4	3	3	2	4	3	X	X	X	X	X	X	X	X
Estudiante 20	3	2	2	2	4	4	3	3	X	X	X	X	X	X	X	X
Estudiante 21	4	4	4	4	4	4	4	4	X	X	4	4	X	X	X	X
Estudiante 22	2	2	3	3	2	2	2	2	X	X	X	X	X	X	X	X
Estudiante 23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 24	3	3	3	3	3	3	4	3	X	X	X	X	X	X	X	X
Estudiante 25	1	1	3	3	1	1	2	2	X	X	X	X	X	X	X	X
Estudiante 26	4	4	4	4	4	4	4	4	X	X	3	3	X	X	X	X
Estudiante 27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Estudiante 35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: E1 and E2 are Evaluator 1 and Evaluator respectively.

Table L10:

Double-Coding scores for the Rubric applied to the incorporation of Feedback into New Response (NR(FR)) for Activity 1.

Student	Incorporation of FR into NR	
	Evaluator 1	Evaluator 2
Student 1	3	3
Student 2	-	-
Student 3	4	4
Student 4	-	-
Student 5	-	-
Student 6	-	-
Student 7	-	-
Student 8	-	-
Student 9	-	-
Student 10	4	4
Student 11	3	3
Student 12	-	-
Student 13	4	4
Student 14	-	-
Student 15	-	-
Student 16	1	1
Student 17	-	-
Student 18	3	3
Student 19	4	4
Student 20	-	-
Student 21	-	-
Student 22	-	-
Student 23	-	-
Student 24	3	3
Student 25	2	2
Student 26	4	4
Student 27	-	-
Student 28	-	-
Student 29	3	3
Student 30	-	-
Student 31	-	-
Student 32	-	-
Student 33	-	-
Student 34	-	-
Student 35	-	-

Table L11:

Double-Coding scores for the Rubric applied to the incorporation of Feedback into New Response (NR(FR)) for Activity 2.

Student	Incorporation of FR into NR	
	Evaluador 1	Evaluador 2
Student 1	2	2
Student 2	2	3
Student 3	2	4
Student 4	2	2
Student 5	-	-
Student 6	4	4
Student 7	4	4
Student 8	-	-
Student 9	1	1
Student 10	3	3
Student 11	1	1
Student 12	3	3
Student 13	3	3
Student 14	2	2
Student 15	-	-
Student 16	3	3
Student 17	4	4
Student 18	3	3
Student 19	3	3
Student 20	1	1
Student 21	3	3
Student 22	-	-
Student 23	-	-
Student 24	-	-
Student 25	2	2
Student 26	-	-
Student 27	-	-
Student 28	-	-
Student 29	-	-
Student 30	-	-
Student 31	1	1
Student 32	-	-
Student 33	-	-
Student 34	4	4
Student 35	3	3

Table L12:

Double-Coding scores for the Rubric applied to the incorporation of Feedback into New Response (NR(FR)) for Activity 3.

Student	Incorporation of FR into NR	
	Evaluador 1	Evaluador 2
Student 1	-	-
Student 2	2	2
Student 3	-	-
Student 4	4	4
Student 5	-	-
Student 6	-	-
Student 7	2	2
Student 8	-	-
Student 9	2	2
Student 10	1	1
Student 11	2	2
Student 12	2	2
Student 13	-	-
Student 14	-	-
Student 15	2	2
Student 16	3	3
Student 17	2	2
Student 18	3	3
Student 19	2	2
Student 20	3	2
Student 21	3	3
Student 22	3	3
Student 23	-	-
Student 24	3	3
Student 25	-	-
Student 26	-	-
Student 27	-	-
Student 28	-	-
Student 29	-	-
Student 30	-	-
Student 31	-	-
Student 32	-	-
Student 33	-	-
Student 34	-	-
Student 35	-	-

M. Set of the guidelines for giving feedback

HOW DO I COMPLETE THE DIFFERENT STAGES IN THE COLLABORATIVE PLATFORM?



To RESPOND...



Identify the *skills* and *topics* addressed by the question:

ASK YOURSELF: What subject-matter knowledge are they asking me for? What information are they giving me? What information are they asking from me? Which critical thinking skill is being assessed? *The critical thinking skills are: interpretation, evaluation, analysis, inference and explanation.*

• **PREPARE YOURSELF:** The ideas, categories, arguments and relationships included in your response must be justified accordingly, using evidence from the situation, case, information or literature referenced in the question. To help you develop your response, prepare the evidence you'll need and think about how to integrate it so as to back up your ideas.

To GIVE MEANINGFUL FEEDBACK ...



CONGRATULATE, QUESTION AND GUIDE

Highlight the achievements in your classmate's response.

Guide them with strategies and tools for incorporating the suggested changes.

Provide new arguments or complement the information included in the response.



SPECIFY AND JUSTIFY

Comment on relevant and specific areas for improving the response.

Avoid giving unjustified personal views (or prescriptive corrections).

Recommend including evidence to support the ideas that are presented.



EXPLAIN

Mention which assessment criteria your comments fall under (formal, conceptual, structural).

Explain the question in other words if you see there has been a misunderstanding.

To REWRITE...

Interpret and make **inferences** based on the *concepts, suggestions* and *strategies* included in the feedback you received.

Evaluate and **analyze** how relevant these comments are and which strategies are best suited to your work.

Assess your work and include any aspects you found to be relevant in your new response.



To SELF-CHECK...

Does the response answer the question that is asked? Have you acknowledged and addressed all of the issues in the question?	
Does each paragraph contain a main idea that is both relevant and well-developed?	
Does the response use precise, varied, specialized and suitable vocabulary, considering both the register (academic) as well as the subject-matter knowledge?	
Have you avoided repetition and including unnecessary information?	
Have your views and opinions been justified and well-presented?	

B- CARTA DE RECEPCIÓN DE ARTÍCULO

24/1/2019

Correo de Pontificia Universidad Católica de Chile - Track your recent Co-Authored submission to CAE



VICENTE MARTIN <vrmartin@uc.cl>

Track your recent Co-Authored submission to CAE

Computers & Education <eesserver@eesmail.elsevier.com>

8 de enero de 2019, 21:48

Responder a: Computers & Education <cae@elsevier.com>

Para: vrmartin@uc.cl

*** Automated email sent by the system ***

Dear Dr. Vicente Martin,

You have been listed as a Co-Author of the following submission:

Journal: Computers & Education

Title: Technology-scaffolded peer assessment for developing critical thinking in pre-service teacher training

Corresponding Author: Camila Barahona

Co-Authors: Miguel Nussbaum; Vicente Martin; Alejandra Meneses; Silvana Arriagada

To be kept informed of the status of your submission, register or log in (if you already have an Elsevier profile).

Register here: <https://ees.elsevier.com/cae/default.asp?acw=&pg=preRegistration.asp&user=coauthor&fname=Vicente&lname=Martin&email=vrmartin@uc.cl>

Or log in: <https://ees.elsevier.com/cae/default.asp?acw=&pg=login.asp&email=vrmartin@uc.cl>

If you did not co-author this submission, please do not follow the above link but instead contact the Corresponding Author of this submission at cebaraho@uc.cl.

Thank you,

Computers & Education

