Assessing academic writing: The construction and validation of an integrated task-based instrument to evaluate specific writing skills

Evaluación de la escritura académica: construcción y validación de un instrumento integrado basado en tareas para evaluar habilidades específicas de escritura

Andueza, Alejandra
 Alberto Hurtado University (Chile)

Abstract
Assessing writing skills is key to the development of instructional methods to help students new to university master academic writing. This article presents the construction and validation process of an instrument developed for that purpose. In so doing, we first discuss the theoretical construct of the evaluative test, then describe the process through which the instrument was developed and validated, and, lastly, present the results obtained and discuss some of its various implications for instruction

Keywords: Writing assessment; diagnostic assessment; academic writing; academic language; writing instruction.

Resumen
Evaluar las habilidades de escritura es clave para el desarrollo de métodos de instrucción que ayuden a los nuevos estudiantes con la escritura académica universitaria. Este artículo presenta el proceso de construcción y validación de un instrumento desarrollado para ese propósito. Al hacerlo, primero discutimos la construcción teórica de la prueba evaluativa, luego describimos el proceso a través del cual el instrumento fue desarrollado y validado, y, por último, presentamos los resultados obtenidos y discutimos algunas de sus diversas implicaciones para la instrucción.

Palabras clave: Evaluación de escritura; evaluación diagnóstica; escritura académica; lenguaje académico instrucción de escritura.

Academic writing is essential to tertiary education. Therefore, acquiring the skills needed for this specific type of writing is critical to academic success (Sundeen, 2014; Margolin, Ram & Mashiah, 2013). Writing can be a powerful tool to learn complex declarative knowledge (Andueza, 2016; Nückles, Hubner y Renkl, 2009; Klein y Rose, 2010) as well as an optimal means for students to master types of academic writing, or genres, and acquire the specific modes of thinking implied in those genres (Bazerman, 2012a, Bazerman, 2012b).

In Chile, even the more academically inclined students face major challenges when composing academic texts (Bitran, Zúñiga, Paulina, Padilla y Moreno, 2009). This can be explained, at least to some extent, by teachers' lack of training in writing instructional strategies (Sotomayor, Parodi, Coloma, Ibañez y Cavada, 2011). As a result, for most university students, gaining an adequate command of academic writing is an arduous process in which they must overcome diverse difficulties: a lack of awareness of cognitive and metacognitive strategies involved in the writing process, and unfamiliarity with academic genres and academic language features, etc.

Accordingly, over the last decade many Latin American universities have developed programs oriented toward enhancing academic writing (and reading) skills among students.

Corresponding author / Autor de contacto

Andueza, Alejandra. Alberto Hurtado University, Department of Language and Literature. Alameda 1869, 4to. Piso, Santiago de Chile   aandueza@uahurtado.cl

These programs have designed and implemented diverse initiatives: first-year composition courses, writing centers (with peer tutoring), faculty development, writing instruction interwoven into core courses, and team teaching (Carlino, 2012) in the hopes of providing students with effective scaffolding in the learning process of academic writing.

These universities expend a great deal of time and resources on such initiatives. However, there is not enough evidence to support their efficacy, since the effect on students' writing skills is seldom measured, analyzed, or interpreted through validated and systematic means. Therefore, academic writing programs and teachers must properly diagnose the writing skills of students new to university in order to inform instructional practice and, following implementation, assess its effectiveness in students' writing outcomes.

Nevertheless, diagnosing academic writing skills is not a simple task because no unifying theory exists on what academic writing should be as a theoretical construct. In fact, most instruments designed to assess academic writing do not specify how they were built or what the underlying theoretical framework entails. (Knoch, 2011). As a result, not much is known about how scales are composed, which is problematic considering that it is crucial in understanding evaluators' perception of writing ability (Dryer, 2013).

There are several published rubrics to assess writing (Boix, Dawes, Duraisingh, & Haynesbut, 2009; Muñoz & Valenzuela, 2015; Johnson & Riazi, 2017; Rakedzon & Baram-Tsabari, 2017; etc.) but, to our knowledge, none that describe "academic writing" as a theoretical construct. Giving the latter, the instrument presented in this paper may help ameliorate this problem in two ways. On one hand, we propose a description of academic writing as a theoretical construct and by categorizing it—for assessment purposes—into five dimensions and nine indicators, according to relevant theory. On the other hand, we account for its construction process through which we designed and validated an academic

writing test. Finally, we present the results of the instrument and offer several conclusions.

Writing assessment: developing an integrated task to assess academic writing

In the following section we discuss different types of writing tasks in terms of the validity and reliability of their content. Over the years, different approaches to writing assessment have been developed. Yancey (1999) identifies three overlapping waves: the first wave (1950-1970), in which the predominant type of writing evaluation consisted of objective tests with multiple-choice items; the second wave (1970-1986), which mainly included holistically scored essays; and the third wave (1986-to the present), which has been based on portfolio assessment.

Currently, none of these waves have produced a gold standard, with some gaining more predominance than others. Though institutions still use objective tests with multiple-choice items to assess writing, this practice is increasingly uncommon since it is fraught with major validity issues. More frequent are impromptu essay tests that include the use of prompts designed to guide student writing, scoring scales, and methods to calculate consistency among raters (Yancey, 1999). Likewise, portfolio assessment, although more valid than the other two methods, poses a reliability problem due to the lack of standardization—each student composes a unique portfolio with texts pertaining to different genres.

As a result many universities assess writing performance through impromptu essays scored on a scale. According to Jonsson & Svingby (2007), scoring rubrics have several benefits, such as: increased agreement among raters, valid operationalization of complex competencies, and promotion of learning. In their research, these two scholars concluded that rubrics enhance the reliability of scoring, especially if they are analytic, topic-specific, and complemented with rater training.

However, over the last several years there has been a tendency to replace impromptu essays with reading-to-write tasks, based on the

consensus that these kinds of instruments better reflect the demands of academic writing, improve test fairness, and address validity and authenticity issues (Gebril & Plakans, 2013; Knoch & Sitajalabhorn, 2013; Plakans & Gebril, 2012; Shin & Ewert, 2015). Knoch and Sitajalabhorn (2013), provide an insightful description of what integrated writing tasks should involve:

> We have argued that for a writing task to be integrated, (1) the source material should include a significant amount of language (which therefore excludes input based on visual stimuli only), (2) the writing product needs to draw on the ideas presented in the input text(s), and (3) the language presented in the source texts needs to undergo some language transformation before being used in the written product. Furthermore, the rating scale used to assess the written products needs to reflect these aspects to fully capture the construct being assessed by the task. The rubrics should provide raters with descriptors which are reflective of what test takers are required to do with the source material and should ideally be developed based on a review of sample performances by test takers. (Knoch & Sitajalabhorn, 2013: 306)

Academic writing as a theoretical construct

Despite ample research on academic writing, to our knowledge, there is no unifying theory to describe this type of writing for assessment purposes. In fact, scoring rubrics function as a *de facto* theoretical construct for writing tests (Knoch, 2011). In this context, we have carefully analyzed a large number of published scoring rubrics designed to assess academic writing through Knoch's taxonomy (2011), in which we detected five highly recurrent dimensions that we considered fundamental in academic writing across the disciplines. Those dimensions are (1) coherence, (2) academic language, (3) source domain, (4) genre domain, and (5) mechanics.

After determining which dimensions are usually considered characteristic of different types of academic writing, we embarked on the task of theoretically defining those dimensions

prior to the design of the scoring rubric. In doing so, we decided to categorize them into nine indicators that link to each dimensions: for coherence: (1) thematic progression and (2) consistency of ideas; for academic language: (3) informational density and (4) academic stance; for source domain: (5) text comprehension, (6) integration of ideas from source texts, and (7) citation; and for genre domain: (8) adjustment to genre and mechanics (9).

Academic language

We consider academic language a theoretical construct that defines a series of cross-discipline language skills. According to Uccelli, Dobbs & Scott (2013), academic language can be conceived of as a set of pragmatic solutions created to promote the study and generation of complex and abstract ideas. In other words, academic language refers to a set of language skills used to achieve certain communication goals, such as communicating precise meaning, condensing information, or explicitly pointing out ideational relations (Uccelli, Barr, Dobbs, Phillips Galloway, Meneses & Sánchez, 2014; Uccelli et al., 2013; Schleppegrell, 2006). For the purposes of operationalization, we focused on two aspects of academic language: informational density and academic stance. Certainly, to the authors mentioned above, academic language is a more complex theoretical construct that includes coherence and genre domain, among other elements; but for the purpose of rendering it operational, we decided to restrict it to these two key components. This naturally leads to the favoring of informational density, given that the remaining components are implied in the other dimensions of academic writing.

a. Informational density: This is a typical feature of academic language. Academic texts are expected to express information in a concise and direct mode, condensing a considerable amount of propositional content into few words, avoiding redundancy and ambiguity (Snow & Uccelli, 2009).

Academic texts use clause condensation strategies in order to aggregate knowledge and

condense information. The lexical aspects of academic language value explicitness, which implies the ability to construct precise meaning while avoiding ambiguity. Being explicit allows the writer to reflect their complete articulation of thought and clarity of meaning. Thus, explicitness has a cognitive dimension since it depends on prior experience and a combination of writer and reader background knowledge (Schleppegrell, 2001).

Nominalization is a grammatical resource that allows the writer to condense an extended explanation into a complex noun phrase. It "is the expression as a noun or noun phrase of what would more congruently be presented as a verb" (Schleppegrell, 2001, p. 443). Nominalization also helps synthetize information already presented in the text and in subsequent clauses.

This resource enhances lexical density and can function as embedded clauses, which allows for long, information-packed sentences (Snow & Uccelli, 2009. Ensuring high lexical diversity and precise lexical choices (which implied in discipline-specific texts) is key in order to achieve concision and density (Snow & Uccelli, 2009).

b. Academic stance: Analyzing stance implies examining how writers express meaning relative to attitude, degree of certainty, epistemic commitment, and engagement with the views of others (Aull & Lancaster, 2014; Uccelli et al., 2013). According to Hyland, stance refers "to the ways writers present themselves and convey their judgments, opinions, and commitments. It is the ways that writers intrude to stamp their personal authority onto their arguments or step back and disguise their involvement" (Hyland, 2005, p. 176).

This author defines three key components of stance. The first is 'evidentiality,' which deals with the level of commitment expressed by the writer regarding the veracity of the propositions he/she is putting forth, including accuracy, strength, and reliability. Evidentiality implies the expression of the writer's degree of certainty and confidence on what he/she is saying. That degree can range from categorical assurance to possibility or uncertainty. The second component is 'affect,' which relates to the writer's attitude toward what is being said, including emotions, perspectives, and beliefs. Affective factors reflect the writer's perspective and evaluation about that which he/she is writing. The third component is 'relation,' which involves the degree of intimacy or remoteness that the writer engages with the reader, as well as the way he/she discursively represents him/herself.

In academic writing, stance, in terms of evidentiality, is expected to be epistemically cautious, objective, and impassionate. In terms of affect and relation, stance is expected to be detached and authoritative (Schleppegrell, 2001; Schleppegrell, 2006; Snow & Uccelli, 2009).

Coherence

According to Charolles (1983, 1995, 2011), coherence is a general principle in the interpretation of human actions. Every time a person perceives someone else carrying out a series of actions, he/she assumes that there is a purpose, which justifies the carrying out of such actions. To Charolles (1983), the principle of coherence applies to discourse; when a sender conveys a message to a receiver, the receiver assumes that the message is intended to communicate something to someone, and that something is constituted as a whole.

Given that the sender usually hopes the receiver understands the message, he/she uses diverse mechanisms to establish links between ideas in order to facilitate comprehension (Charolles, 1983; Charolles, 1997; Sarda, Carter-Thomas, Faggard, Charolles, 2014). Those links—which allow the receiver to understand the message as a coherent entity—can be conceptualized into two key elements:

a. Thematic progression: According to Charolles (1979), all ideas in a coherent discourse must be connected to each other. This implies that every proposition must include given information (to maintain thematic continuity) and new information (to ensure semantic progression). This structure controls thematic progression within discourses and gives the reader a focal point

by which to interpret certain parts of the discourse in relation to others (Van Dijk, 1996).

b. Consistency of ideas: Charolles argues that, for a text to be coherent, all ideas expressed must be consistent with one another. This implies that no idea should contradict another previously stated idea; and that no idea should be inconsistent with the discursive frame to which it belongs.

Reading-to-write

One of the key aspects of academic writing is the incorporation of other texts within the text that is being written. Writing academic texts usually implies comprehending, selecting and editing information from a source text in order to integrate it within the new text (Shin & Ewert, 2015; Cumming, 2013). In so doing, the writer must make a considerable number of decisions, such as determining what information should be selected from the source text, how the selected information will meet the goals set out for the new text, what mechanisms will de used to integrate the information into the new text, and how that information will be organized, selected, and connected to the writer's own ideas. According to Knoch & Sitajalabhorn (2013), these are the main features of a reading-to-write task:

1. Selecting ideas from the input text.
2. Synthesizing ideas from one or more sources.
3. Transforming the language of the source text.
4. Choosing an organizational structure different from the input text.
5. Connecting the ideas from the input text with the writer's own ideas.

According to Kintsch (2004), to achieve this outcome, the writer must be able to (1) adequately read an academic text; (2) construct its propositional model (i.e., create a mental representation of its propositions, their coherent interrelations, and the organization of the text regardless of the way in which its propositions were expressed); and (3) create a situational model that demonstrates the ability to integrate those propositions with the reader's background knowledge . This is a necessary process in learning from texts, which is central to academic writing, given that it implies the ability to use information acquired from a text for use in new situations.

According to Hyland (2008), this ability to analyze and summarize information from diverse sources is crucial, given that academic writing requires the referencing of other authors' work. Those references are central to academic persuasion, not only because they create a theoretical framework that shows that the texts are undergirded by previous work in the discipline, but also because they validate the author as an insider, a member of the discursive community (Castelló, Bañales, Vega, 2011).

Genre domain

Another key aspect of academic writing is the ability to master academic genres, which requires a larger degree of analytical skills than others genres, such as explanatory or persuasive essays (Uccelli, Dobbs & Scott, 2013). According to Hyland (2008), one of the main characteristics of academic genres is the writer's attempt to predict adverse reactions to the claims he/she is establishing.

The writer's experience as a reader enables him/her to anticipate objections that might arise. In so doing, he/she must develop a strong line of reasoning, which entails an adequate command of argumentation. According to Uccelli et al. (2013), the argumentative practices of persuasive essays "consists of a writer's position or thesis about a topic, followed by organized stepwise argumentation that includes precise claims, data, warrants, counterargument, and rebuttals that lead to a well-justified conclusion" (Uccelli et al., 2013, pp. 38-39).

Method

In the following section we describe the construction and validation process of an instrument to assess specific academic writing skills through an integrated task. Accordingly, we first explain in some detail the design phase,

during which we developed the task and the scoring rubric. Secondly, we explain the rubric validation process using qualitative and quantitative methods.

Design phase

a. Task design

Once the theoretical construct was defined, we proceeded to design the instrument. In so doing, we developed an integrated reading-to-write task because, on the one hand, it better represents the demands students face in academic writing, and, on the other hand, reading provides students with relevant content about which to write (Shin & Ewert, 2015). To design the task, we reviewed many published evaluative instruments and writing manuals (Hacker, 2000) and laid out a set of instructions to specify what students would have to do. The task consisted of writing an academic argumentative essay on one of two possible topics.

To accomplish the task, students had to establish their point of view and sustain it with no less than three arguments. At least one of those arguments had to be supported by information contained in a source text. For that purpose, we selected two informative texts (one for each topic), of approximately 1,700 words each.

The prompt included a series of guidelines to clarify the writing process and specify the features that essay had to incorporate. Students were given a total of three hours to write the essay.

b. Rubric design

A first draft of a trait-based analytical scoring rubric was developed by two experts in order to render the theoretical construct operational in the form of nine traits: (1) thematic progression, (2) consistency of ideas, (3) informational density, (4) academic stance, (5) reading comprehension, (6) inclusion of ideas from the source text, (7) citation, (8) mechanics, and (9) adjustment to genre. Each trait specifies four performance levels.

Think-Aloud Study

Once the first version of the prompt and the scoring rubric were concluded, we conducted a think-aloud study with 11 participants, in which we asked them to perform the task and say aloud what were they thinking, following a structured protocol (Janssen; Braaksma & Rijlaarsdam, 2006). We also made an audio recording of the procedure. All test takers first read the task, then the source text, and finally wrote the essay. This study allowed us to verify (1) whether or not they adequately understood the prompt and source texts; (2) if they were able to write the essay as expected; and (3) if it would be possible to successfully apply the scoring rubric, among other components. This enabled us to identify problematic aspects of the prompt and rubric in order to solve such issues.

Expert panel evaluation

After making necessary modifications, we sent a second version of the instrument to a panel of experts for content validation. Four experts in the field of academic writing and/or academic language reviewed the prompt and scoring rubric, and responded with valuable feedback and recommendations to improve the instrument. We took those suggestions into consideration and incorporated them where possible.

Pilot study

Once the instrument was modified based on the expert panel's suggestions, we conducted a pilot study with a sample of 72 senior high school students. Four rating experts (academic writing teachers) scored all 72 essays. Prior to scoring, we required these experts to undergo three four-hour training sessions.

Test administration

Following the pilot study, we made amendments to the final prompt and scoring rubric (see appendix A and B). We then administered the instrument at a private Chilean university in 2015 at the beginning of the spring semester. A total of 319 first-year students from the Philosophy and Humanities Faculty (approximately 18-19 years of age) took the test voluntarily (162 women and 157 men). Participants were given a maximum of three

hours to complete the task, but most students required less time.

Results

After the administration of the test, we carried out the scoring process. The scorers were the same experts from the pilot study, therefore additional training sessions were not necessary. In the following section, we analyze inter-rater agreement levels and, then, conduct a descriptive and inferential analysis in order to determine the instrument's reliability.

Inter-rater agreement levels

We found high levels of inter-rater consistency (see table 1): 23% of essays were double-scored (n=73). For each trait, the highest percentage were matching scores (difference in score=0): thematic progression (54,8%), consistency of ideas (50,7%), informational density (68,5%), academic stance (60,3%), text comprehension (49,3%), integration of ideas from source texts (58,9%), citation (67,1%), mechanics (71,2%), adjustment to genre (53,4%). Differences in scores of ±1 were the second highest while differences in scores of ±2 was uncommon; and there were no differences of ±3.

Table 1. Inter-rater agreement levels

Dimension	Traits	Difference in scores	Frequency	Percentage
Coherence	Thematic progression	0	40	54,8
		±1	31	42,5
		±2	2	2,7
		±3	0	0,0
	Consistency of ideas	0	37	50,7
		±1	34	46,6
		±2	2	2,7
		±3	0	0,0
Academic language	Informational density	0	50	68,5
		±1	22	30,1
		±2	1	1,4
		±3	0	0,0
	Academic stance	0	44	60,3
		±1	26	35,6
		±2	3	4,1
		±3	0	0,0
Source domain	Text comprehension	0	36	49,3
		±1	30	41,1
		±2	7	9,6
		±3	0	0,0
	Integration of ideas from source texts	0	43	58,9
		±1	27	37,0
		±2	3	4,1
		±3	0	0,0
	Citation	0	49	67,1
		±1	23	31,5
		±2	1	1,4
		±3	0	0,0
Mechanics	Mechanics	0	52	71,2
		±1	18	24,7
		±2	3	4,1
		±3	0	0,0
Genre domain	Adjustment to genre	0	39	53,4
		±1	33	45,2
		±2	1	1,4
		±3	0	0,0

Descriptive analysis

The following table details the mean and standard deviation for each trait. Note that while *academic stance* and *text comprehension* were the highest rated traits, *mechanics* and *adjustment to genre* were the lowest.

Table 2. Mean and standard deviation

	Thematic progression	Consistency of ideas	Informational density	Academic stance	Text comprehension	Integration of ideas from source texts	Citation	Mechanics	Adjustment to genre
Mean	2.38133	2.26978	2.13898	2.70965	2.50105	2.36445	1.98365	1.49077	1.86551
St. Dv.	0.56709	0.46768	0.44291	0.58661	0.63676	0.62699	0.62177	0.6751	0.53165
N	316	316	316	316	316	316	316	316	316

Inferential analysis

The analysis conducted indicates an adequate level reliability (Cronbach's alpha for all elements=0.717). Table 3 shows a moderate correlation between most indicators. *Citation* and *mechanics* are the indicators that correlated the least. We believe that this is because those traits entailed more mechanic skills than the others.

Table 3. Pearson correlation between indicators

		Thematic progression	Consistency of ideas	Informational density	Academic stance	Text comprehension	Integration of ideas from source texts	Citation	Mechanics	Adjustment to genre
Thematic progression	Pearson Correlation Sig. (1-tailed)	1								
Consistency of ideas	Pearson Correlation	.457**	1							
	Sig. (1-tailed)	.000								
Informational density	Pearson Correlation	.353**	.417**	1						
	Sig. (1-tailed)	.000	.000							
Academic stance	Pearson Correlation	.244**	.173**	.217**	1					
	Sig. (1-tailed)	.000	.002	.000						
Text comprehension	Pearson Correlation	.442**	.330**	.323**	.352**	1				
	Sig. (1-tailed)	.000	.000	.000	.000					
Integration of ideas from source texts	Pearson Correlation	.393**	.360**	.243**	.158**	.460**	1			
	Sig. (1-tailed)	.000	.000	.000	.005	.000				
Citation	Pearson Correlation	.061	.085	-.046	.198**	.197**	.129*	1		
	Sig. (1-tailed)	.279	.133	.413	.000	.000	.022			
Mechanics	Pearson Correlation	.330**	.339**	.122*	.067	.146**	.212**	.047	1	
	Sig. (1-tailed)	.000	.000	.031	.233	.009	.000	.409		
Adjustment to genre	Pearson Correlation	.205**	.088	.172**	.218**	.262**	.208**	.111*	.094	1
	Sig. (1-tailed)	.000	.118	.002	.000	.000	.000	.049	.095	

**. Correlation is significant at the level of 0.01 (1-tailed).
*. Correlation is significant at the level of 0.05 (1-tailed).

Discussion and implications for instruction

Given that assessing academic writing skills can provide valuable data for the development and administration of writing programs, as well as contribute significantly to bolster instructional practices, we consider that the instrument presented in this paper can be useful for that purpose. In fact, we presented a theoretically grounded and empirically validated scoring scale, which constitutes an

effort to set in motion a complex set of skills for measurement purposes.

We believe this instrument can be used for diagnostic purposes, because, if taken prior to writing courses, it will allow teachers to evaluate the consistency between learning objectives and students' needs. Furthermore, the underlying theoretical construct may be useful in the development of writing courses, because it can help teachers determine which content should be taught (in terms of metacognitive declarative knowledge) in order to increase students' awareness of how academic writing should be.

Additionally, this instrument may be useful for measuring learning outcomes. For instance, if taken before and after instruction, it would be possible for teachers to measure the effectiveness of writing instruction strategies and for students to monitor their own progress in the development of writing skills. Furthermore, the scoring scale presented in this study can be adapted to several different writing tasks, and used as a formative assessment tool - if given to students prior to the writing process - which can foster self-regulatory skills thus increasing self regulated learning.

References

Andueza, A. (2016). La escritura como herramienta de aprendizaje significativo: un cuasiexperimento en la clase de ciencias. *Revista Complutense de Educación, 27*(2), 653-668. doi: https://doi.org/10.5209/rev_RCED.2016.v27.n2.46918

Aull, L. & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written Communication 2014*, *31*(2), 151–183. doi: https://doi.org/10.1177/0741088314527055

Bazerman, C. (2012a). Genre as social action. In J.P. Gee & M. Handford (Eds.), *The Routledge handbook of discourse analysis* (pp. 226-238). New York, U.S.A: Routledge.

Bazerman, C. (2012b). Writing, cognition and affect from the perspectives of sociocultural and historical studies of writing. In V. Wise (Ed.), *Past, present and future contributions of cognitive writing research to cognitive psychology* (pp.89-104). New York, U.S.A: Psychology Press.

Bitran, M., Zúñiga, D., Paulina, F., Padilla, O. y Moreno, R. (2009). Mejoría en las habilidades de comunicación escrita de estudiantes de medicina: Impacto de un taller de escritura. *Revista Médica de Chile , 137*, 617-624. doi: https://doi.org/10.4067/S0034-98872009000500004

Boix, V., Dawes, E., Wolfe, C. R., & Haynes, C. (2009). Targeted Assessment Rubric: An Empirically Grounded Rubric for Interdisciplinary Writing. *Journal of Higher Education*, 80(3), 334-353. doi: https://doi.org/10.1080/00221546.2009.11779016

Carlino, P. (2012). Section essay: Who takes care of writing in Latin American and Spanish universities? *Perspectives on writing*, 485.

Castelló, M; Bañales, G., & Vega, N (2011). Leer múltiples documentos para escribir textos académicos en la universidad: o cómo aprender a leer y escribir en el lenguaje de las disciplinas. *Pro-Posições*, 22, 1, p. 97-114. doi: https://doi.org/10.1590/S0103-73072011000100009

Charolles, M. (1978). Introduction aux problèmes de la cohérence des textes: Approche théorique et étude des pratiques pédagogiques. *Langue française*, (38), 7-41. doi: https://doi.org/10.3406/lfr.1978.6117

Charolles, M. (1983). Coherence as a principle in the interpretation of discourse. *Text-Interdisciplinary Journal for the Study of Discourse, 3*, 71–98. doi: https://doi.org/10.1515/text.1.1983.3.1.71

Charolles, M. (1995). Cohésion, cohérence et pertinence du discours. *Travaux de linguistique*, (29), 125-151.

Charolles, M. (1997). L'encadrement du discours. *Monographie Cahier de Recherche Linguistique* (6), 1-73.

Charolles, M. (2011). Cohérence et cohésion

du discours. *Dimensionen der Analyse Texten und Diskursivent-Dimensioni dell'analisi di Testi e Discoursi*, 153-173.

Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, *10*(1), 1-8. doi: https://doi.org/10.1080/15434303.2011.622016

Dryer, D. (2013). Scaling writing ability a corpus-driven inquiry. *Written Communication, 30*(1), 3-35. doi: https://doi.org/10.1177/0741088312466992

Gebril, A. & Plakans, L. (2013). Toward a transparent construct of reading to-write tasks: The interface between discourse features and proficiency. *Language Assessment Quarterly*, *10*(1), 9-27. doi: https://doi.org/10.1080/15434303.2011.642040

Hacker, D. (2000). *Rules for writers*. Boston: Bedford/St. Martin's.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2): 173–192. doi: https://doi.org/10.1177/1461445605050365

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching, 41*, 543-562. doi: https://doi.org/10.1017/S0261444808005235

Janssen, T., Braaksma, M. & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students: A think aloud study. *European Journal of Psychology of Education*, *21*(1), 35-52. doi: https://doi.org/10.1007/BF03173568

Johnson, R.C. & Riazi, A.M. (2017) Validation of a locally created and rated writing test used for placement in a higher education EFL program. *Assessing Writing* 32, 85–104. doi: https://doi.org/10.1016/j.asw.2016.09.002

Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130–144. doi: https://doi.org/10.1016/j.edurev.2007.05.002

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. *Theoretical models and processes of reading, 5*, 1270-1328. doi: https://doi.org/10.1598/0872075028.46

Klein, P., & Rose, M. (2010). Teaching argument and explanation to prepare junior students for writing to learn. *Reading Research Quarterly, 45*(4), 433-461. doi: https://doi.org/10.1598/RRQ.45.4.4

Knoch, U. (2011). Rating scales for diagnostic assessment of writing: what should they look like and where should the criteria come from. *Assessing Writing, 16*, 81-96. doi: https://doi.org/10.1016/j.asw.2011.02.003

Knoch, U., & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focussed definition for assessment purposes. *Assessing Writing, 18*, 300–308. doi: http://doi.org/10.1016/j.asw.2013.09.003

Margolin, B., Ram, D. & Mashiash, R. (2013). Building and validating a tool for assessing academic writing skills. *International Journal of Linguistics, 5*(4), 188. doi: https://doi.org/10.5296/ijl.v5i4.3498

Muñoz, C.& Valenzuela, J. (2015). Características psicométricas de una rúbrica para evaluar expresión escrita a nivel universitario. *Formación Universitaria*, 8(6), 75-84 doi: http://dx.doi.org/10.4067/S0718-50062015000600010

Nicola, D. J. & Macfarlane-Dickb, D. (2006) Formative assessment and selfregulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*. 31(2), 199–218 doi: https://doi.org/10.1080/03075070600572090

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and instruction, 19* (3), 259-271. doi: https://doi.org/10.1016/j.learninstruc.2008.05.002

Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated

second language writing tasks. *Assessing Writing, 17,* 18-34. doi: https://doi.org/10.1016/j.asw.2011.09.002

Rakedzon, T. & Baram-Tsabari, A. (2017). To make a long story short: A rubric for assessing graduate students' academic and popular science writing skills. *Assessing Writing* 32, 28–42. doi: http://dx.doi.org/10.1016/j.asw.2016.12.004 1075-2935

Sarda, L., Carter-Thomas, S., Fagard, B., & Charolles, M. (2014). Text structuring devices: An overview. *Pragmatic Approaches to Text Structuring,* 1-50.

Schleppegrell, M. (2001). Linguistic features of the language of schooling. *Linguistic and Education 12*(4), 431-459. doi: https://doi.org/10.1016/S0898-5898(01)00073-0

Schleppegrell, M. (2006). The challenges of academic language in school subjects. In Lindberg, I. & Sandwall, K. (Eds.), *Spraket och kunskapen: att lara pa sitt andrasprak i skola och hogskola* (pp.47-69). Gotenborg, Sweden: Gotenborgs universitet institutet for svenska som andrasprak.

Shin, S. Y., & Ewert, D. (2015). What accounts for integrated reading-to-write task scores? *Language Testing, 32*(2), 259-258. doi: https://doi.org/10.1177/0265532214560257

Snow, C. & Uccelli, P. (2009). The challenge of academic language. In Olson, D. Y Torrance, N. (eds.), *Cambridge handbook of literacy* (pp. 112-133). USA: Cambridge university press. doi: https://doi.org/10.1017/CBO9780511609664.008

Sotomayor, C., Parodi, G., Coloma, C., Ibáñez, R., y Cavada, P. (2011). La formación inicial de docentes de educación general básica en Chile: ¿Qué se espera que aprendan los futuros profesores en el área de Lenguaje y Comunicación?. *Revista de Investigación Educacional Latinoamericana*, 48(1), 28-41. doi: https://doi.org/10.7764/PEL.48.1.2011.3

Sundeen, T. (2014). Instructional rubrics: Effects of presentation options on writing quality. *Assessing Writing, 21*, 74-88. doi: https://doi.org/10.1016/j.asw.2014.03.003

Uccelli, P., Barr, C., Dobbs, C., Phillips Galloway, E., Meneses, A. & Sanchez, M. (2014). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics*, 1- 33. doi: https://doi.org/10.1017/S014271641400006X

Uccelli, P., Dobbs, C., & Scott, J. (2013). Mastering academic language organization and stance in the persuasive writing of high school students. *Written Communication, 30*(1), 36-62. doi: https://doi.org/10.1177/0741088312469013

Van Dijk, T. (1996). *La ciencia del texto*. Barcelona, España: Paidós.

Yancey, K. (1999). Looking back as we look forward: Historicizing writing assessment. *College Composition and Communication*, *50*(3), 483-503. doi: https://doi.org/10.2307/358862

APPENDIX A

ACADEMIC WRITING TEST

General information

I.   This test is designed to specifically assess the academic writing skills of undergraduate students.
II.   The test consists of writing an argumentative essay that supports your stance on ONE of the following topics:

TOPIC 1: Gender differences are a reality in the contemporary labor market. Some professions are primarily held by men while others by women. Write an essay on this topic by responding to the following question:

> Can it be said that some careers or professions are more adequate for men and others for women?

TOPIC 2: Numerous variables can affect the academic success of students in primary school as well as in university, such as the educational environment, motivation to learn, self-confidence, etc. Write an essay on this topic by responding to the following question:

> Do you believe a student's socioeconomic environment determines his/her academic performance?

NEXT STEPS:
- Choose one of the two topics and carefully read its corresponding academic text:
  o TOPIC 1: "What kinds of careers do boys and girls expect for themselves?"
  o TOPIC 2: "How do some students overcome their socio-economic background?"
- Underline and/or take notes of the text's most important ideas and information that can be incorporated into your argumentative essay. Remember that you must support your arguments with information from the text.
- Plan your essay. You may do so by using the provided scratch paper.
- Write your essay, then proofread it making modifications where you deem necessary.

INSTRUCTIONS:
1. Establish a clear hypothesis (or point of view).
2. Uphold your stance with at least three coherent and convincing arguments.
3. Support your arguments with information from the text you have read. In so doing, you must include the reference for each quote and add a bibliography at the end.
4. Consider at least one counterargument and refute it.
5. Offer a conclusion.
6. Use academic language.
7. Ensure that your writing is clear and coherent.
8. Revise your essay to make sure there are no spelling, punctuation, and/or grammatical errors.
9. Make sure that the essay is a minimum of ONE PAGE and a maximum of THREE PAGES.
10. Note that you have THREE HOURS to complete the test.

APPENDIX B

Table 4. Scoring Rubric

		4	3	2	1
Coherence	Thematic progression	The essay displays excellent thematic progression: there is a clear balance between semantic progression and thematic continuity, since the arguments coherently introduce new ideas based on prior information. In other words: • Every sentence in the text provides information (or is closely related to prior information); there are no digressions or thematic interruptions. • Every sentence includes new information (thematic progression). There is no unnecessary repetition of ideas. This allows the reader to connect all the different ideas within the text and follow a coherent line of argumentation.	The essay displays good thematic progression: there is a general balance between semantic progression and thematic continuity, although it occasionally gets lost due to: • Interruptions in thematic continuity (abrupt change of topic, loss of reference or digression). • Interruptions in semantic progression (new information is not included or old information is unnecessarily repeated). This allows the reader to connect most of the text's ideas because, overall, the argumentation can be followed, although there are several instances of problems of continuity.	The essay displays insufficient thematic progression: the balance between semantic progression and thematic continuity is precarious due to: • Interruptions in thematic continuity (abrupt change of topic, loss of reference or digression). • Interruptions in semantic progression (new information is not included or old information is unnecessarily repeated). This allows the reader to connect some of the text's ideas, but it is difficult to follow the argumentation due to problems of continuity.	The essay displays deficient thematic progression: there is no balance between semantic progression and thematic continuity due to an overabundance of: • Interruptions in thematic continuity (abrupt change of topic, loss of reference or digression). • Interruptions in semantic progression (new information is not included or old information is unnecessarily repeated). This prevents the reader from establishing connections within the text, since only several isolated ideas can be comprehended. It is almost impossible to follow the text's line of argumentation.
	Consistency of ideas	The reader can perfectly understand the text's meaning because all ideas are internally consistent. In other words: • No idea contradicts the meaning of another, previously introduced idea. • All ideas in the text are interrelated by the adequate and	The text's overall meaning is comprehensible and nearly all ideas are internally consistent. However, it does not allow for a fluid reading throughout. This is due to the following factors: • No idea contradicts the meaning of another, previously	The text's overall meaning is difficult to comprehend and various ideas are internally inconsistent, which prevent a fluid reading. This is due to the following factors: • At least one idea contradicts the meaning of another, previously introduced idea.	The overall meaning of the text is barely comprehensible and most ideas are internally inconsistent, which make it difficult to read. This is due to the following factors: • Various ideas contradict the meaning of other, previously introduced ideas. • The relationship between ideas is

		precise use of connectors. • All relationships between the ideas in the text can be unambiguously inferred (from punctuation or other language resources).	introduced idea. • Occasionally, the relationship between ideas is not adequately specified due to unspecific or inadequate markers. • Occasionally, relationships between the ideas in the text cannot be unambiguously inferred (due to punctuation mistakes or other language errors).	• In various occasions, the relationship between ideas is not adequately specified due to unspecific or inadequate markers. • In various occasions, relationships between the ideas in the text cannot be unambiguously inferred (due to punctuation mistakes or other language errors).	not adequately specified due to unspecific or inadequate markers. • Relationships between the ideas in the text cannot be unambiguously inferred (due to punctuation mistakes or other language errors).

Academic language		Information within the text is consistently condensed. The text conveys a good deal of information in a concise manner, using as few words as possible. This means that the text: - Expresses ideas directly and in a straightforward manner. - The text uses diverse and precise vocabulary. - The text uses nominalization to concentrate information or expand nominal groups through adjectives, clauses, or noun compliments, with the purpose of condensing ideas into one sentence.	Generally speaking, the information in the text is condensed. For the most part, the text conveys a good deal of information using few words, although: - Occasionally, ideas are not expressed directly (use of circumlocution; that is, unnecessary long-windedness). - There are occasional lexical imprecisions (vague lexicon or words used incorrectly). - Occasionally, words are repeated within the same paragraph when a synonym, pronoun, or elision could have been used to avoid repetition. - The text does not use nominalizations to concentrate information nor does it expand nominal groups through clauses.	The text's information is not very concise, since it conveys insufficient information using many words. This is because: - In various occasions, ideas are expressed through circumlocution. - In various occasions, there are lexical imprecisions (ambiguous or incorrect word use). - In various occasions, words are repeated within the same paragraph when a synonym, pronoun, or elision could have been used to avoid repetition. - The text does not use nominalizations to concentrate information nor does it expand nominal groups through clauses.	The text's information is dispersed, since it conveys little information using many words. This is because: - Most of the ideas are expressed through circumlocution. - The text mostly uses colloquial vocabulary in which lexical imprecisions predominate (ambiguous or incorrect word use). - Words are constantly repeated within the same paragraph when a synonym, pronoun, or elision could have been used to avoid repetition. - The text does not use nominalizations to concentrate information nor does it expand nominal groups through clauses.

| | | The text displays a clear academic stance. This means that it assertively and objectively presents information, which is expressed in the following manner:
• The text adopts a distanced tone. This means that it does not reflect an emotional, committed or passionate attachment with what is being said.
• The stance is epistemically cautious. In other words, it expresses a modest level of certainty regarding the truthfulness of what is being stated. | The text displays a mainly academic stance. This means that it assertively and objectively presents information, although:
• Occasionally, the text adopts an personal tone. This means that it reflects an emotional, committed or passionate attachment with what is being said.
• In several occasions, the text presents information in an epistemically careless way. In other words, it expresses ideas with an undue level of certainty (for example: presenting one's own ideas, emotions, judgment, or personal experiences as indisputable truths). | The text displays a mainly colloquial stance. In general, it lacks assertion and objectivity due to the following factors:
• Occasionally, the text adopts an personal tone. This means that it reflects an emotional, committed or passionate attachment with what is being said.
• In several occasions, the text presents information in an epistemically careless way. In other words, it expresses ideas with an undue level of certainty (for example: presenting one's own ideas, emotions, judgment, or personal experiences as indisputable truths). | The text displays a completely colloquial stance.
In general, it does not assertively or objectively present information (including errors) and constantly expresses the author's subjectivity. This is reflected in the following factors:
• The text adopts a tone that is very personal, expressive, prejudiced, and self-referential.
• The text often presents information in an epistemically careless way. In other words, it expresses ideas with an undue level of certainty (for example: presenting one's own ideas, emotions, judgment, or personal experiences as indisputable truths). |

Source domain	Comprehension of the writing prompt	The text adequately conveys ideas expressed in the writing prompt. In other words: • It always cites or paraphrases ideas from the referenced prompt, maintaining its original meaning.	The text somewhat conveys ideas expressed in the writing prompt. In other words: • It occasionally cites or paraphrases ideas from the referenced prompt, partially maintaining its original meaning.	The text does not adequately convey ideas expressed in the writing prompt. In other words: • It occasionally cites or paraphrases ideas from the referenced prompt, distorting its original meaning. AND/OR • It is impossible to evaluate the comprehension of the writing prompt, since the ideas cited are marginal or have little importance in the overall text.	The text does not refer to the content of the writing prompt. In other words: • It is impossible to identify ideas that come from the prompt. AND/OR • The author's text occasionally cites or paraphrases ideas from the referenced prompt, completely altering its original meaning.
	Incorporation of information from the writing prompt	The author adequately incorporates information extracted from the writing prompt. In other words: • Ideas taken from the prompt are added to the author's own text. • All ideas taken from the prompt contribute relevant information to the author's own text. • The author integrates all the ideas taken from the prompt with his/her own ideas.	In general, the author adequately incorporates information extracted from the writing prompt. In other words: • The author takes ideas from the prompt and adds them to his/her own text. • All the ideas taken from the prompt contribute relevant information to the author's own text. HOWEVER: • Occasionally, the author mentions an idea from the prompt, but does not integrate it into his/her own ideas.	The author inadequately incorporates information extracted from the writing prompt. In other words, • The author takes ideas from the prompt and adds them to his/her own text. HOWEVER: • Occasionally, the integrated ideas do not contribute relevant information to the author's own text (they seem to be chosen at random). • Occasionally, the author mentions an idea from the prompt, but does not integrate it into his/her own ideas.	The author does not integrate ideas extracted from the writing prompt with his/her own ideas. In other words: • The author does not take ideas from the prompt. OR • The ideas taken from the prompt have no relation whatsoever to what the author's text is attempting to convey.

	Reference to the writing prompt	The text properly cites the sources of all ideas extracted from the writing prompt. In other words: • Within the body of the text, all references adhere to only one citation style (APA, MLA, Chicago, etc.). • The bibliography appropriately references the prompt text adhering to the citation style used throughout the essay.	The text cites the sources of all ideas extracted from the writing prompt. However, the references are not completely adequate because: • Within the body of the text, references include information on locating the quote (author, year, page numbers, etc.). • The bibliography cites the prompt text by mentioning all information that helps to identify the text (author, year, article title, journal title, and source medium [web page, etc.]).	The text does not properly cite the sources of ideas extracted from the writing prompt. OR • In the body of the text, references include insufficient information for locating the quote in the prompt text. OR • No bibliography is included. OR • The prompt text is not mentioned in the body of the text but is included in a separate bibliography.	There is no mention of the prompt text. In other words: • Sources are not cited within the body of the text. AND • Sources are not cited in a separate bibliography.
Mechanics	Standard spelling, punctuation, and grammatical consistency	The text completely adheres to standard English punctuation and spelling. There are no typos or mistakes in abbreviations or in grammatical consistency.	The text adheres to standard English punctuation and spelling, but there are up to five mistakes in spelling, punctuation, abbreviations, or grammatical consistency.	The text somewhat adheres to standard English punctuation and spelling. In other words, there are up to ten mistakes in spelling, punctuation, abbreviations, or grammatical consistency.	The text does not adhere to standard English punctuation and spelling, since there are more than ten mistakes in spelling, punctuation, abbreviations, or grammatical consistency.

Genre domain	Compliance with characteristics of argumentative essays	The text complies with all the characteristics of an argumentative essay. In other words: • It clearly formulates a hypothesis or position regarding an issue (or, at least, this can be unambiguously inferred). And • The hypothesis is defended through organized argumentation (the order in which arguments are presented bolsters their comprehension) that includes at least three arguments supported by facts. And • The text poses at least one counterargument and adequately refutes it. And • The text concludes by reasserting and reformulating its point of view, as well as synthesizing its main arguments.	The text mostly complies with the characteristics of an argumentative essay, because: • The text formulates a hypothesis or position regarding an issue (or, at least, this can be unambiguously inferred). HOWEVER, • Although the hypothesis is defended by argumentation, it is not sufficiently organized (the arguments could be rearranged without altering the overall meaning of the text) into at least three separate arguments. AND/OR • The text does not pose any counterargument. AND • The text concludes by reasserting and reformulating its point of view, as well as synthesizing its main arguments.	The text complies with only certain characteristics of an argumentative essay, because: • The text formulates a hypothesis (or, at least, one can be inferred) regarding an issue. HOWEVER: • Although the hypothesis is defended by two or three arguments (that may or may not be organized), at least one of them is not directly related to the hypothesis or fails to defend it. AND/OR • The text does not pose any counterargument. AND • The text concludes by repeating its point of view or one of its main arguments OR the text simply DOES NOT have a conclusion.	The text does not comply with the characteristics of an argumentative essay, because: • It is impossible to identify a hypothesis or position regarding an issue. AND/OR • There is no argumentation, rather a list of unstructured ideas, some of which are tangentially related to the hypothesis but do not defend it. The essay organization appears more like an oral conversation rather than a structured argumentation. • The text may or may not pose a counterargument. • The text may or may not have a conclusion.

Authors / Autores	**To know more / Saber más**
Andueza, Alejandra aandueza@uahurtado.cl Alberto Hurtado University, Department of Language and Literature. Director of the Academic Writing Program at the Alberto Hurtado University. Her research interest includes, writing assessment, academic writing, writing instructional strategies, and metacognitive strategies during writing process Postal adresses: Alameda 1869, 4to. Piso, Santiago de Chile	0000-0002-7983-9823 Dialnet

Revista **EL**ectrónica de **I**nvestigación y **EV**aluación **E**ducativa
E-Journal of Educational Research, Assessment and Evaluation
[ISSN: 1134-4032]