Research Article

Liliana Tolchinsky*, Elisa Rosado and Melina Aparici

Internal and external appraisals of analytical writing. A proposal for assessing development and potential improvement

https://doi.org/10.1515/iral-2023-0012
Received January 15, 2023; accepted June 17, 2023; published online August 11, 2023

Abstract: This paper introduces Developing analytical writing (DAW), a model of text analysis aimed at capturing how social expectancies of analytical writing become part of students' linguistic literacy. DAW proposes a multilayer analysis of text features on repeated text production prompted by the same and different topics, control for pedagogical input, and researcher- and reader-based evaluations of text quality. We revise DAW previous implementations that delved into lexical, syntactic-discursive, and structural aspects of analytical essays as indexes for developing writing proficiency. We focus thereafter on a current application of DAW to assess content-related dimensions: explicitness of a central standpoint, flexibility of writers' positioning, and expression of reflective thinking. Analyses revealed that the probability of explicitness and reflexiveness increases from elementary to higher levels of schooling, while the probability of flexible positioning appeared conditioned by text topic and pedagogical input. The three dimensions showed stronger effect of instruction on high schoolers' than on elementary and university students' texts. Regression analyses support a foundational role of development and a significant contribution of the assessed dimensions to an appraisal of text quality. DAW facilitates distinguishing analytical text features that improve in the course of development from those sensitive to pedagogical scaffolding.

***Corresponding author: Liliana Tolchinsky**, Department of Catalan Philology and General Linguistics, University of Barcelona, 525 Gran Via de les Corts Catalanes, 08007, Barcelona, Spain, E-mail: ltolchinsky@ub.edu. https://orcid.org/0000-0001-7893-6587
Elisa Rosado, Department of Linguistic and Literary Education, University of Barcelona, Barcelona, Spain, E-mail: erosado@ub.edu
Melina Aparici, Department of Cognitive, Developmental and Educational Psychology, Autonomous University of Barcelona, Barcelona, Spain, E-mail: melina.aparici@uab.es

Keywords: analytical writing; content-related features; discourse features; elementary to university level; teaching implications; text analysis

1 Developing analytical writing

This paper presents a model of text analysis to account for developing facets of analytical texts construction. Developing analytical writing (DAW) focuses on individual essays "because they are arguably the most powerful outcome measures – of transfer of skill from the social to the individual plane – and also because they are the measures of most familiarity and direct interest to educators" (Kuhn and Crowell 2011, p. 1). The "transfer of skill" is an acculturation process implying appropriation of the characteristics of a discourse type, its specific rhetorical patterns, and motives. The characteristics of individual essays "can be neither described, predicted, nor analyzed without resorting to a characterization of the discourse type to which that essay belongs" (Longacre 1996, p. 7).

Analytical types of discourse are topic-oriented, aimed at exposing information on a certain matter. Unlike narrative discourse, analytical discourse privileges thematic and logical relations over chronological ones and, when dealing with debatable topics, it embraces both an expository and an argumentative component. Expositive elements supply reflections on the topic or description of data, while argumentative elements assert and support the writer's standpoint to persuade the audience of its validity (Tolchinsky et al. 2017).

The extent of novice writers' awareness about what is expected from a given discourse-type, and what linguistic resources they mobilize to fit expectations, is open to a wide variety. DAW's goal is to capture how social expectancies of analytical writing are realized in essays spanning the school-age years from elementary through high school and on to university level, in the domain of *later language development* (Berman 2004; Tolchinsky 2004). DAW would help to trace the development and potentiality of improvement of the features of written expression that play a role in *linguistic literacy* (Ravid and Tolchinsky 2002).

In what follows we elaborate on why focus on analytical written discourse and deploy the theoretical basis of our approach. We describe thereafter the main features and previous applications of DAW aimed at tracking the developmental route to analytical writing as foundational to teaching. By developmental route we refer to the changes that occur across age/school levels[1] in those skills we assumed to be responsible for analytical writing proficiency. Lastly, we report a new application of

1 We refer to age/school level groups because in mainstream schools age and school level are entangled.

DAW that delves into content-related dimensions committed to analytical essays' goals and well-formedness.

1.1 Analytical writing

The overall purpose of a discourse, its "notional structures", may encode in the form of differing surface structures or genres (Longacre 1996). Analytic notional structures are realized in diverse genres such as essays, descriptions, or reports in which communicative functions – explaining, describing, arguing – are differently weighted (Schleppegrell 2004). While some analytical realizations are mainly expository and focus on analyzing the topic, others are mainly argumentative presenting two- or more-sided perspectives on the topic. The diversity of analytical genres allows students to build and express command on a certain topic, and to shape a (well) supported standpoint when dealing with controversial topics.

Skills involved in critical thinking play a crucial role (Driver et al. 2000). Writers must be willing to be well-informed about the topic addressed to accumulate evidence for grounding a solid standpoint while putting personal biases aside (Willingham 2007). Analytical skills facilitate the writer to work out what is relevant to the problem in hand. Those are useful skills especially in educational contexts, where argumentation is formally introduced, practiced, and assessed. Attaining proficiency in analytical essays contributes to academic success. We assume that the very process of analytical text production has an epistemic effect: it activates writers' topic knowledge and affects their perception of how much they know about it.

1.2 Theoretical framework of DAW

Pivotal to our approach to text analysis is Longacre's (1996) crucial remark on the futility "[…] to look in a discourse for a feature which is not characteristic of the type to which that discourse belongs" (p. 7). In line with Berman and associates (Berman and Verhoeven 2002) we applied functionally motivated quantifiable criteria for characterizing different facets and quality of text construction in the light of specific type of discourse.

Genre distinction proved to play a key role in vocabulary, grammar, discourse structure and content of expository and narrative texts (Berman and Nir-Sagiv 2007).

Our approach is also nurtured by Mann and Thompson's (1988) Rhetorical Structure Theory which conceptualizes text structure as a hierarchical organization with compulsory foregrounded (nuclei) and other backgrounded (satellites) elements. For example, the sequence of events is nuclear in narrative texts – there

is no narrative without events – while evaluative components, although adding to text richness, are peripheral. In our analysis, claims and grounds are nuclear components whereas counterclaims following a claim, although useful for the communicative purpose of argumentation, are dispensable. This distinction is useful for examining texts produced by novice writers and allows us to appreciate writers' awareness of genre constraints.

We also consider several aspects from Toulmin's pioneer model of argumentation. Toulmin (1958) points at the role played by different elements of an argument. *Claims*, the writer's assertions on the topic, and *grounds*, appealed to as support of the claims, are the main elements of an argument. Other no obligatory elements such as *qualifiers* and *scopes* indicate the strength of the claim and limit its span.

Finally, for text parsing, we resort to a proposal of rhetorical moves (RMs) as strands of discourse that fulfill a concrete communicative goal which are identified applying pragmatic, semantic and syntactic criteria (Vilar Weber and Tolchinsky 2021). This approach is particularly useful for analyzing texts in terms of the functional/communicative structures that typically make up particular genres.

1.3 Main features of DAW

The model provides a methodological framework to account for *developmental changes* – resulting from ordinary developmental course – and *micro-developmental changes* – occurring in the short run in the context of circumscribed pedagogical scaffolding – in text construction. It was designed to assess analytical text writing but is applicable to any type of discourse. The main features of the model are the following: 1) repeated text writing controlling for topic content and pedagogical input; 2) focus on linguistic expression of written products; 3) multilayered text-analysis; and 4) researcher and reader-based appraisals of the texts. Applications of the model open crucial questions (and provide some answers) on learning and teaching of writing.

1.3.1 Control for topic and pedagogical input

When attempting to trace the developmental route in discourse-related language proficiencies such as text construction, it is imperative to control for pedagogical input and content topic to distinguish changes that result from age/schooling developments from those tied to a certain topic and/or pedagogical scaffolding (Aparici et al. 2021; Vilar Weber and Tolchinsky 2021). Writing development is an ambiguous term since "it can refer to the ordinary developmental course of learning to write, or to the systematic (or less so) instruction for developing those skills"

(Applebee 2000, p. 2). This "ordinary developmental course" is part and parcel of the general linguistic and cognitive development of the literate mind.

To control for topic and pedagogical input, we asked our participants for repeated text writing on same and different topics in the context of an identical set of activities. The choice of topics resulted from piloting a range of topics to warrant varied and rich responses across schooling levels. The selection of classroom activities aimed to raise participants' awareness of the main features of the target genre.

Table 1 displays an outline of the classroom activities.

Table 1: Outline of classroom activities.

Purpose and activities
Lesson 1 – Sharing a purpose Students are informed of project purpose. The teacher offers examples of topics for students to discuss based on own experiences. Text 1 on *Freedom of dress* *Lesson 2* – Self-evaluation of topic knowledge After the teacher gives the prompt, students respond to pre- and post-writing questionnaires to evaluate how much they know about the topic. Production of text 2 on *Freedom of movement*. *Lesson 3* – Contrasting standpoints on the target topic Students read individually and with the teachers two texts on *Freedom of movement* expressing divergent standpoints and supporting evidence. In dyads, students discuss and underline texts standpoint and supporting evidence and complete a workout sheet identifying these elements. Production of text 3: second text on *Freedom of movement*. *Lesson 4* – Highlighting relevant text features In dyads, students evaluate classmates' texts using a rubric and score the author's topic-knowledge, clarity of standpoint, quality of evidence, organization, vocabulary, and grammar. *Lesson 5* – Highlighting relevant text features Students visualize the scores obtained in their texts to raise awareness on their weak and strong points. Each component of the rubric is illustrated by authentic (anonymous) selected samples from text 3 for students to reason about the received scores. *Lesson 6* – Motivating repeated writing on same topic The project coordinators send a video message to motivate students for writing another text on the same topic. Production of text 4: third text on *Freedom of movement*. *Lesson 7* – Evaluation of gains persistence Production of text 5 one month after lesson 6: on *Rewards and punishments*

As seen, students participated in oral discussions, topic-related reading, analysis, and evaluation of texts targeting features of analytical essays. The goal was not to test the quality of the implemented activities, but to ensure homogeneity of the pedagogical input. This two-faceted assessment of macro- and micro-developmental changes – those that occur after the sequence of classroom activities – puts us in a privileged position to distinguish linguistic features prone to improvement in the short run from those requiring long developmental windows or a different treatment.

1.3.2 Focus on the written product

DAW applies on the written product rather than on the characteristics of the writing process (e.g., Hayes 1996), or writing as an online activity (e.g., Alves and Limpo 2015), we assume that the properties of the written product always result from someone *making decisions* as to what words are to be put together.

Written language is analyzed as a *special style of discourse* (Berman 2016). Although features of *transcription* or *notation* like spelling, or punctuation reflect knowledge of both language structure and discursive function (Graham and Harris 2000; Lukeman 2007), they are not the focus of our analyses.

1.3.3 Multilayered text-analysis

When producing a text, the writer needs to concurrently attend to content, discourse organization, syntactic packaging, and lexical selection. Each of the writer's concerns was turned into a layer of analysis. That is, our analysis of the written text features ranges across different layers of language use. We analyzed from lexical category and morphological complexity of words, type of phrases, intra- and inter-clausal connectivity, up to overall text rhetorical structure. Only expert writers can separate out these various layers of text-embedded features correcting or improving local issues that may affect the whole text.

Different linguistic units are involved at different levels of analysis of written language, from word via phrase and clause combinations, and up to rhetorical moves. These are by no means obvious or agreed: they need to be specified in advance for different languages. For instance, in *written texts*, a word is defined operationally as any string of characters separated from the next string by a space. Yet, even when dealing with typologically close languages, text analyzers must agree on how to count words or deal with segmentation errors, compounds, multiword expressions, or collocations (Berman 2002; Berman and Ravid 2009). A critical unit of analysis is the *clause* defined both syntactically and as a "single predication expressing a unified situation (event, or state)" (Berman and Slobin 1994, pp. 660–662). The clause has proved as a reliable unit of discourse in different languages.

Proficient writers' decisions go top–down since a writing schema embracing genre and stylistic constraints stored in Long Term Memory controls the writing

process guiding syntactic and lexical choices (Swales 1990; Vande Kopple 1998). Novice writers' decisions, in contrast, depend more on the writer's previous local decision than on a central writing schema (McCutchen et al. 1997). Consequently, writers' skill to build a text structure that fully reflects the communicative purpose of a genre is a major indicator of genre appropriation (Allen et al. 2019).

Various units have been proposed for understanding the structure of a text (e.g., paragraphs, T-Units, clause packages). These units are defined and labeled based on syntactic and scriptal criteria (e.g., indentation) with functional and thematic coherence further considered. Rhetorical moves (RM) are another type of text unit. RMs refer to spans of text that do a particular job; they are defined for the function they fulfill in the text to enact specific genre expectancies as interpreted by the writer (Swales 1990). Swales' model has been applied to analyze the discourse structure of academic and professional genres in corpus-based analyses (Upton and Cohen 2009).

Aparici and Perera (2001) applied Britton's (1994) proposal of RM analysis to chart developmental changes in expository texts' structure of Spanish elementary, junior, and high-school students, and adults. They identified three types of RMs: advance, to introduce information, expansion, to elaborate on it, and unification, to summarize information. The number of RMs increased with schooling, but only expansion and unification moves increased throughout the four groups. This suggests that younger students interpreted expository genre expectancies as putting more information in the text while older and more experienced writers responded to the need to elaborate on and summarize the introduced information.

Analysis of RMs fits well to accomplish DAW's goal of capturing how social expectancies of analytical writing become part of participants' linguistic literacy. We proceeded top-down (Upton and Cohen 2009). We first established the socially agreed communicative functions of an essay on controversial topics (e.g., express a standpoint, argue to support it, elaborate on the topic), and then examined their enactment in the texts' RM. For delimitating the RM, we resort to syntactic, scriptal, and thematic criteria (Vilar Weber and Tolchinsky 2021). In so doing we characterize the text structure in terms of two indicators of writing performance in a particular genre (1) internal differentiation (i.e., the presence of text spans performing different jobs), and (2) completeness of structure (i.e., the extent to which the extant text spans embrace the expected communicative functions).

1.4 Researcher and reader-based text appraisals

Research has been devoted to identifying the text features that support subjective impressions of text quality (Aparici et al. 2021). The assessment of these features by researchers that have designed and participated in the research process provide an

internal appraisal of quality. In DAW the internal appraisal is performed by means of qualitative and quantitative assessments of text-embedded features (e.g., text structure) that, according to previous studies, indicate text quality (McNamara et al. 2010). This researcher-based assessment has been complemented with an external appraisal performed by judges who neither design nor participate in the research process (McMaster and Espin 2007). In DAW, teachers and graduates in language-related studies provide a reader-based appraisal by means of analytical rubrics embracing different aspects of text (e.g., lexical richness, text organization) followed by global scoring of text quality. This two-sided approach is an additional strength of our model.

1.5 Applications of DAW

We followed the above-described methodological framework in previous studies aimed at tracing how students appropriate the expected features of analytical writing from elementary to higher education, and which (if any) of these features are sensitive to pedagogical scaffolding. Across studies, we assessed different sets of text-embedded features that were both assumed to fulfill rhetorical aims (Langacker 1999), and to serve as indicators of developing writing proficiency. The researcher-based assessment was complemented with a reader-based one to test which features best explained it.

In Aparici et al. (2021) we focused on lexical, syntactic, and discursive indicators. We also tracked students' productivity as a predictor of writing quality and a proxy of number of ideas (Weston et al. 2011), and found that up to university they use more words and clauses for responding to the prompt. The implemented classroom activities also affected productivity, but not at university level.

At a lexical level, we traced word length, lexical diversity, lexical density, and use of adjectives and nominalizations. Across schooling, students used more sophisticated and diverse words, as well as more adjectives. In contrast, nominalizations decreased with school level, while lexical density did not show any development. Classroom activities affected positively lexical diversity and word length, but only among the youngest groups.

We found similar divergences in the syntax and discourse domains: the use of different syntactic linkages (e.g., coordination, subordination) increased across schooling, but syntactic density (e.g., number of words per discourse-level units) remained stable after elementary school (Crossley and McNamara 2010). Pedagogical scaffolding affected only the use of relative clauses and syntactic density, and only in elementary school.

At the discourse level we assessed the use of Discourse Markers creating texts' texture (Choi 2007) to find that experienced writers produced better-connected discourse units, although differences were found only between elementary and higher school, and were not affected by the implemented activities.

In Rosado et al. (2021) we examined high school and university students use of intra- and interclausal connectivity devices, and productivity as indicators of text quality. Students contained significantly longer texts with age. However, younger students' texts produced a higher proportion of intraclausal connectors than older ones, suggesting that text cohesion does not necessarily reside in a profusion of connectors (Crossley et al. 2011). The analysis revealed a "trade-off effect" whereby conjunctions were most frequent in high school, while optional discourse markers were more frequent in university.

Parallel studies inquire into the structural features of analytical essays of Spanish monolinguals (Vilar Weber and Tolchinsky 2021) and Catalan/Spanish bilinguals (Tolchinsky et al. 2021; Tolchinsky et al. submitted). We assessed (1) completeness of text structure (i.e., the extent to which the argumentative component – assertive and argumentative moves – and the non-argumentative component – expository moves – are realized), and (2) the relative weight of the non-argumentative component. We found an unequal pattern of development in bilinguals and monolinguals. While the texts of bilingual elementary school students included assertive, argumentative, and expository moves, monolinguals completeness of TS was not attained until high school. Moreover, no further development in completeness of structure was observed in the two samples. Only the expository component, containing evidence and justifying reasons, showed a steady increase, especially among bilinguals. We attributed this difference to bilinguals' enhanced metalinguistic awareness (Hsin and Snow 2017) which may raise their concern for potential addressees to justify their standpoint.

To sum, the reported applications of DAW enabled us to distinguish linguistic features that improved in the course of ordinary development (e.g., text and word length, syntactic/discursive connectivity), from fewer features that were also sensitive to pedagogical scaffolding (e.g., lexical diversity; use of relative clauses), and showed a higher potential development. Changes in writing quality were facilitated by a developmental basis rather than induced only by pedagogy.

An exception to this rather clear developmental pattern of indicators of writing proficiency was text structure that showed mixed findings. The presence of the expository component in analytical essays continued to grow up to university and was sensitive to pedagogy. In contrast, completeness of structure grew with schooling among monolingual students but was already attained among elementary school bilinguals, while appearing resistant to pedagogical work across the board.

Moreover, by applying a distributional approach to text structure we failed to capture the features of overall text organization that make a difference in text quality. To clarify, once the main communicative function of rhetorical moves (e.g., to assert an opinion, to provide evidence) was identified, we assessed the distribution of moves in the texts to chart the text structure. We disregarded the differential qualities of moves performing a similar function, and how they relate to each other. For instance, whether the asserted opinion favors or opposes the writer's standpoint, or the provided data support the writers' opinion. We speculated that dwelling on the differential qualities of moves would enable us to capture developmental differences in text well-formation that were not detected in previous analyses. To explore this hunch, we revisited our corpus of essays and applied a qualitative, content-related analysis of text structure in light of relevant features of analytical writing.

2 Inquiry into content-related dimensions of text structure

Analytical essays benefit from writers' unequivocal expression of standpoint on the topic. In Longacre's (1996) words, a discourse "[…] involves *progress*, i.e., a well-formed discourse is going somewhere. The progress of a discourse typically issues in some sort of climactic development […] a *rhetorical underlining*. It's as if you took a pencil and underlined certain lines of what you are writing. The importance of rhetorical underlining must not be underestimated. It is one of the simplest and most universal devices for marking the important point …" (p. 39).

In analytical essays an explicit formulation of the writer's standpoint provides the "climactic development" that in well-formed discourse indicates where the discourse is going to. In previous studies we showed that students' views appeared unsupported in assertive moves and supported in argumentative moves. Focusing on their differential qualities would enable us to detect whether any of these moves function as the explicit standpoint that signals the discourse progress, and so establishes a hierarchical differentiation that implies central control.

Nonetheless, writing about a debatable topic requires capability to express alternative views "weighing positive and negative attributes of contrasting positions on the topic" (Kuhn and Crowell 2011). Children and even adolescents are not generally capable of maintaining a two-sided argument (by claims and counterclaims), nor are they as flexible in alternating perspectives. They typically concentrate on expounding their own claims, disregarding opposite ones. For example, 9-year-olds asked to reach agreement about a historical claim spent the

bulk (81 %) of their utterances on expressing and justifying their own claims but were unable to produce counterclaims (Ferretti and Graham 2019). Similarly, a group of young adolescents and community-college students arguing about capital punishment concentrated on claims supporting their own position, whereas young adults tended to also address their opponent's arguments, often through counterclaims (Felton and Kuhn 2001).

Two factors may explain this propensity to such one-sided positioning in writing. On the one hand, counterclaims, and other devices that attenuate writers' commitment to their own perspective (Caffi 2013), are peripheral to it. In Mann and Thompson's (1988) conceptualization they are satellite elements. In Toulmin's argumentation model they are identified as qualifiers and cover a wide set of elements that establish the claim's scope. Stavans et al. (2019) found that Hebrew-speaking students aged 7–12 years writing on different topics incorporated 'peripheral' (introduction and conclusion) later than 'core' elements (claims and grounds).

On the other hand, deviating from own contentions to consider other possibilities requires decentering, as "the capacity to shift experiential perspective from within one's subjective experience onto that experience" (Bernstein et al. 2015, p. 1). This capacity not only enables to detach from and reflect on own experience, but also to consider the addressee's perspective. Although Piaget and Inhelder (1969) regarded this ability to be attained by around age 12, we all realize how difficult it is to move from one's own *perspective* and consider various aspects of a given state of affairs. Previous exploration of a subsample of the present corpus (Tolchinsky et al. 2023) showed that dwelling on the differential qualities of moves facilitates a grasp of novice writers' skills to address the pros and cons of a given topic. That is, whether all the moves are aligned with the writers' positioning, or rather, mitigate or contradict the writer's view.

From cognitive and sociocultural perspectives, writing, as other forms of external representation, is not only a communicative but also an epistemological tool (Chen 2019). The mental activities involved in text writing—planning, revising, rereading, editing, etc. — impact not only writers' knowledge of the topic at issue, but also their reasoning, how they use and think about language. In previous studies (Tolchinsky et al. 2021; Vilar Weber and Tolchinsky 2021) we computed the distribution of expository moves disregarding their differential quality, but while some contained examples or empirical data that serve as evidence, other expressed writers' reflective thinking (i.e., rumination on the topic or own knowledge), metatextual commentaries, or rhetorical questions to elicit attention on particular aspects of the topic. Here we will consider the different subtypes of expository moves indicating reflective thinking.

2.1 Goals and expectations

In this study, we aimed at depicting development and potentiality of improvement of the explicitness of writers' standpoint, the quality of writers' positioning, and the expression of reflective thinking by one between subject factor – school level (Elementary, High school, University) – and one within subject factor – Text (T1, T2, T4) – as well as by the interactions of school level by Text. Specifically, we asked what changes these dimensions of analytical writing experience across school levels, and which of these changes indicate increased skillfulness in the expression of standpoint, positioning, and reflective thinking.

For T1, students were prompted to write about freedom of dressing code, a hot topic for adolescents and young adults. For T2 and T4, students had to write about freedom of move from own country, a debated topic at a time of strong migration throughout Europe. T1 and T2, on different topics, had to be produced before the classroom activities while T4, on the same topic than T2, was produced after these. This enabled us to discuss the possible influence of topic versus the influence of pedagogical input.

We expected a positive effect of school level on the explicit expression of a central standpoint, a greater incidence of explicit expression with schooling. We expected a more flexible positioning as well: students' texts would include not only claims or aspects of the topic that support their own view, but also elements attenuating and even opposing their own views. We also anticipated increasing production of reflective thinking as students advance in schooling. Given that the targeted dimensions are genre-constrained, we expect these would be more sensitive to genre-oriented instruction than to topic. Finally, informed by previous studies (e.g., Tolchinsky et al. 2021), we expected that external evaluation of text quality will be positively affected by school level and pedagogical work.

2.2 Methods

2.2.1 Participants

The study included 212 participants from elementary school through university level (65 (35 girls) from Grade 6 (M = 11.6 years, SD = 0.29), 78 (45 girls) from Grade 10 (M = 15.8 years, SD = 0.58) and 69 (42 girls) from second year of university (M = 21 years, SD = 2.40). Participants were recruited and tested in Castilla-León and Castilla-La Mancha, Spain, in two elementary public schools, three public high schools, and one university faculty. Schools and participants were chosen following these criteria: 1) they had more than one class per grade, 2) the teachers were open

to working on the project, and 3) the students came from a middle-class SES background and showed no known behavioral or linguistic issues. All parents provided their written consent for their children's participation in the study. The corpus comprised 636 texts (3 per participant). External evaluation was performed on a subsample of the corpus: 243 texts produced by 135 participants equally distributed by Text and school level.

Because of the importance of sociodemographic factors and parental education for children's literacy outcomes (Bialystok 2018), we gathered data on these factors using a detailed questionnaire (Aparici and García 2018) including occupational training requirements and SES characteristics. The three participant groups were homogeneous with regard to parents' education and fathers' employment (for sociodemographic details, see Tolchinsky et al. 2021).

2.2.2 Tasks and procedures

All participants engaged in text writing assignments as part of a series of three-stage activities in which each participant produced five texts: Time 1 (T1) before the onset of the sequence of classroom activities; T2 at the onset of the sequence; T3 and T4 during the sequence; and T5 a month later (see Table 1). In the present study we will focus in T1, T2 – both before classroom activities but on different topics – and T4 – after classroom activities, on the same topic than T2.

For the three school levels, an identical set of classroom activities was used as described in the introduction. Only the younger participants' instructions were slightly modified by the teachers. After attending three training sessions and following comprehensive instructions, regular teachers implemented the activities in language classes. To ensure the reliability of the application, at least two classroom observations per group were conducted.

2.2.3 Transcription and coding

We used identical transcription and coding procedures as in earlier research (Aparici et al. 2021; Vilar Weber and Tolchinsky 2021). Text Structure (TS) was manually coded. Each text was divided into rhetorical moves (RMs), i.e., strands of discourse accomplishing specific communicative goals. Several indicators helped define each move frontier: change in the communicative goal, change of topic, punctuation (periods), shift from affirmative to negative modality, and/or presence of discourse markers.

Three main types of moves were identified: *assertive, argumentative*, and *expository*. The first two contain a claim (assertion) expressing the writer's opinions on different aspects of the topic. While *assertive moves* did not include grounds

within the same move, *argumentative moves* contain both a claim (assertion) and grounds (e.g., facts, examples, experiences that function as evidence of the claim) within the same move. *Expository moves* contain no claims (no assertions), only examples, descriptions of facts, reflections on the topic, or definition of terms related or not to the topic that do not function as evidence of a claim. For each main type we further identified subtypes in terms of their differential qualities.

Coding was subject to inter-rater agreement between one of the authors and independent raters. Agreement on 100 texts (5 % of the corpus) was 0.89 for segmentation into RMs, and 0.94 with a Cohen's κ 0.92 for RM-type identification.

2.3 Dimensions and strategies of analysis

We focused on three textual dimensions assumed to indicate well-formedness of analytical essays: (1) explicitness of a central standpoint; (2) quality of students' positioning, and (3) expression of reflective thinking. Each dimension was built on the subtypes regarding the differential qualities of rhetorical moves.

2.3.1 Explicitness of a central standpoint

If verbal expression(s) of standpoint subsuming local claims already presented in assertive and/or argumentative moves appeared in the text, this was considered as explicitly formulated (coded as 1); if not, the text was coded as 0. For texts lacking an explicit formulation, the analysts *inferred* the student's standpoint to guide further coding (see dimension 2). For statistical analysis we fitted a negative binomial distribution and used Generalized Estimating Equations (GEE) model for non-normal distribution and repeatedly measured texts (Hardin and Hilbe 2012).

2.3.2 Quality of the writer's positioning

We refer to whether local claims – realized in subtypes of assertive and argumentative moves- and expository moves – providing data or examples – are aligned with the writer's view, or rather, supply attenuating or diverging elements. Quality of positioning was assessed by the relation between two composite outcomes: *supporting devices* and *mitigating devices*. While the first was built of multiple categories of elements aligned with the writer's own opinion on the topic, the second was built of different rhetorical devices that attenuate the writer's view: spans of discourse (moves) that function as *neutral* or *counter-claims*, and *qualifiers*, extra-clausal or intra-clausal elements (hedges, conditionals) that attenuate the strength of a statement.

For the analyses of these two categorical outcomes, we applied a *success over trial* modeling approach. This approach assessed school level (SL) and Text effects on the rates of supporting devices over the total number of devices. For statistics, we fitted a negative binomial distribution and a Generalized Estimating Equations (GEE) model for non-normal distribution and repeatedly measured texts. Post-hoc marginal means were the predicted numbers of supporting devices over all devices, namely, the proportions or corrected probabilities to favor the students' views.

2.3.3 Expression of reflective thinking

Reflectiveness was established by the occurrence of three subtypes of expository moves: metatextual commentaries, rhetorical questions, and commentaries about the writer's own knowledge or relevance of the topic (see Table 2). These subtypes of expository moves were distinguished from expository moves containing facts or examples as supporting evidence, not included in this dimension. Reflectiveness was measured by a composite outcome built on the addition of the subtypes of moves containing reflective commentaries. Statistics were computed on this composite outcome by applying again a (GEE) Negative binomial procedure.

Table 2: Coding of content-related text dimensions.

	La libertad de movimiento **The freedom to move**	
M1	*Después de los textos leídos y las reflexiones que hemos hecho* *puedo decir que sé más sobre este tema* After the texts read and the reflections we have made I can say that I know more about this topic	Expository move containing a commentary about the writer's own knowledge on the topic. Coded as expression of reflective thinking.
M2	*Desde mi punto de vista la libertad de movimiento tendría que ser una de las más importantes* From my point of view freedom of movement would have to be one of the most important	Assertive move expressing the student ungrounded view on the importance of the topic.
M3	*Actualmente en casi todos los países hay extranjeros* *que normalmente no son un incordio* *sino ayudan a su crecimiento* Currently in almost all countries there are foreigners who usually are not a nuisance but help its growth	Expository move providing data on positive aspects of the topic and supporting the writer's view. Coded among the supporting devices.

Table 2: (continued)

	La libertad de movimiento **The freedom to move**	
M4	*aunque hay otros* *que no ayudan* *y solo incordian* *como los que se dedican a mafias o al crimen organizado* […] although there are others that do not help and they only bother such as those who are dedicated to mafias or organized crime	Expository move offering data on a negative aspect of the situation. Coded among the mitigating devices.
M5	*También hay otro tipo de inmigrantes* *pero que tienen que irse de sus casas por necesidad* *como son los refugiados* *estos lo hacen* *para huir de las guerras en sus países* […] *Como es el caso de miles de personas* *que huyen de la guerra en África* […] There are also other types of immigrants but (that) they have to leave their homes out of necessity such (as) are the refugees These do it to flee from the wars in their countries As is the case of thousands of people who flee from the war in Africa	Expository move providing reasons that lead to migration with illustrating examples. Coded among the supporting devices.
M6	*También están por otra parte* *las personas que viven en los países de acogida.* *que normalmente son respetuosos* There are also on the other hand people who live in host countries. Who are usually respectful	Expository move offering data on the positive attitudes of the hosting population. It favors the writer's standpoint. Coded among the supporting devices.
M7	*aunque una parte de la población no quiera acogerles por miedo a las diferencias sociales y étnicas* […] *y los discriminan* although part of the population does not want (to) welcome them out of fear of social and ethnic differences and they discriminate them	Expository move offering data on negative attitudes of the hosting population. It opposes to the writer's standpoint. Coded among the mitigating devices.

Table 2: (continued)

	La libertad de movimiento **The freedom to move**	
M8	*A mi parecer las personas tienen el derecho de* *poder ir a otros países* *por la razón que quieran* *sin justificarle nada a nadie* *porque la tierra no es de nadie* *y nadie debería controlar el paso de la gente* I think people have the right to go to other countries for whatever reason they want without justifying anything to anyone because the earth belongs to no one and no one should control the transit of people	Argumentative move subsuming previously presented reasons. It functions as a rhetorical underlining of the writer standpoint. Coded as explicit central standpoint.

Table 2 presents a text produced by a high-school boy and illustrates how the three dimensions were targeted. The text appears segmented into 8 rhetorical moves and into clauses within each move.

The text opens with a reflective commentary on the student's own knowledge of the topic after participating in the classroom activities (M1). Both in the second (M2) and last move (M8) the student expresses his view on the topic. However, M2 embraces an unsupported claim on the conditioned importance of freedom to move, while M8 conveys an unequivocal expression of standpoint subsuming previously presented reasons. The next two moves (M3 and M4) share the function of presenting relevant data but differ in their alignment with the explicit standpoint. While M3 sustains the writer's view and was coded as *supporting*, M4 defies this view and was therefore coded as a *mitigating device*. A similar interplay between supporting and mitigating devices is conveyed in the next three moves.

For every dimension, analyses were run (across texts) for simple effects and interactions: first on the three texts (T1, T2, T4) and thereafter on T1 and T2, independently of T4, for testing the effect of topic; then on T2 and T4, independently of T1, for testing the effect of pedagogical work.

2.3.4 External evaluation

Four teachers evaluated the 243 texts (38 % of this study's corpus) by means of an analytic rubric that included a "global score" expressed in a Likert scale (1 minimal to 5 maximal score) for each text.

2.4 Results

Analyses were aimed at determining the impact of school level and pedagogical input on three textual dimensions: explicitness of the writer's standpoint, quality of the writer's positioning, and expression of reflective thinking.

2.4.1 Explicitness of a central standpoint

Over a total number of 636 texts, 572 texts had assertive and argumentative moves, and 64 texts contained only expository moves excluded from the present analysis. Only 29 % (163 texts) contained one or more explicit central claims unequally distributed by text: 23 % (45) of T1 contained such central claims, 27 % (48) of T2, and 36 % (70) of T4. Most of these texts contained 1 move that worked to summarize the viewpoint of the writer, very few texts contained 2, and still fewer contained 3 such moves. For statistics, we computed presence (1) or absence (0) of explicit formulation of a central standpoint, irrespective of their number. Table 3 shows results on the probability of explicit expression of standpoint by school level and Text, as well as interactions SL by Text.

Table 3: Effect of school level (SL), Text (T1, T2, T4) and Interaction SL by Text on explicitness of the writer's standpoint. GEE Binomial model results. Marginal means and (SE).

		Texts ***n* = 572**		
School level	**All**	**T1** ***n* = 195**	**T2** ***n* = 180**	**T4** ***n* = 197**
Elementary *n* = 65	0.10 (0.025)[a]	0.09 (0.037)	0.07 (0.034)	0.16 (0.046)
High School *n* = 78	0.39 (0.037)[b]	0.28 (0.053)	0.30 (0.056)	0.59 (0.058)
University *n* = 69	0.32 (0.41)[b]	0.31 (0.059)	0.42 (0.066)	0.28 (0.057)
Total *n* = 212		0.20^{A} (0.030)	0.23^{AB} (0.032)	0.33^{B} (0.036)
$Wald_{SL\text{-}all}$ 50.200, p = 0.000; $Wald_{text\text{-}all}$ 9.969, p = 0.007; $Wald_{SL\ x\ text}$ 14.868, p = 0.005				
T1 versus T2 $Wald_{SL}$ 19.673, p = 000				
T2 versus T4 $Wald_{SL}$ 30.5945; $Wald_{text}$ 3.992, p = 0.046; $Wald_{SL\ x\ text}$ 12.689, p = 0.002				

Note. Latin letters for marginal mean ranking by pairwise comparisons with Bonferroni correction. Uppercase indicates ranking across time points and lowercase ranking across SLs where "a" means the lowest mean. Marginal means are predicted probabilities.

There was a main effect of SL and Text on the probability to express an explicit point of view across the three school levels and texts. Pairwise comparisons showed that elementary school students had a significantly lower probability to express an explicit standpoint than highschoolers and university students; differences between the older groups were not significant. As for Text effects, pairwise comparisons across the three texts showed significant differences between T1 and the two subsequent texts, but not between T2 and T4.

The comparison between T1 and T2 (before classroom activities but on different topics), independently of T4, showed a unique significant effect of SL (Wald 19.673, p = 0.000) but not of Text. Again, pairwise comparisons pointed at a significant difference only between elementary students' texts and the older groups, but no significant differences between the two older groups or significant interaction SL by Text. In contrast, the comparison between T2 and T4 (on the same topic, before and after classroom activities), independently from T1, showed both a significant effect of SL and of Text. The probability to express an explicit standpoint was significantly lower at elementary than at high-school or university level, and at T2 than at T4.

The interaction Text by SL was also significant (Wald 12.689, p = 0.002). Pairwise comparisons of interaction effects across SL showed that university students had higher probability to show explicitness than high and elementary school students in T2. However, in T4 probability of explicitness was higher in high school than in elementary school and university. Across texts, T4 displayed a higher level of explicitness that T2 in elementary school and high school, but not in university.

2.4.2 Quality of the writer's positioning

We computed 3,661 assertive, argumentative, and expository moves across all texts, 81.6 % (2,988) contained devices *supporting* the writer's standpoint and only 18.4 % (673) contained *mitigating* devices. Table 4 shows the distribution of supporting and mitigating devices by SL and Text.

Table 5 displays results of the *success over trial* modeling approach to assess SL and Text effects on the rates of supporting devices over the total number of supporting plus mitigating devices. Low rates mean higher probability of flexible positioning.

There was no effect of SL or Text on the rate of supporting devices over the total number of devices across the three texts, but only a significant interaction SL by Text. Pairwise comparisons of estimated marginal means of the interaction effects (Bonferroni correction) of SLs across texts showed that only university students had a significantly a lower probability to produce a higher rate of supporting devices in T1 than in T4. And, across SLs, highschoolers showed a lower probability to produce a higher rate of supporting devices than university students only in T4.

Table 4: Mean number (SD) and percentages of moves that contain supporting and mitigating devices by SL and Text.

	Texts *n* = 633					
	Text 1 *n* = 210		Text 2 *n* = 211		Text 4 *n* = 212	
School level	Supporting devices	Mitigating devices	Supporting devices	Mitigating devices	Supporting devices	Mitigating devices
Elementary *n* = 65	3.87 (SD 2.36); 85.2	0.67 (SD 1.00); 14.7	3.00 (SD 1.53); 79.9	0.75 (SD 1.02); 20.00	3.95 (SD 1.86); 80.8	0.94 (SD 1.39); 19.2
High School *n* = 78	4.01 (SD 2.06); 81.9	0.88 (SD 1.15); 18.03	4.40 (SD 1.95); 81.00	1.02 (SD 1.19); 18.9	5.28 (SD 2.10); 77.5	1.53 (SD 152); 22.4
University *n* = 69	5.44 (SD 2.98); 77.3	1.60 (SD 1.75); 22.6	6.12 (SD 3.22); 82.5	1.29 (SD 1.36); 17.5	6.22 (SD 2.63); 88.6	0.80 (1.06); 11.4
Total *n* = 212	4.72 (SD 2.55); 80.8	1.05 (SD 1.39); 19.1	4.53 (SD 2.64); 97.8	1.02 (SD 1.21); 2.2	5.18 (SD 2.39); 82.3	1.10 (SD 1.35); 17.6

Table 5: Effect of school level, Text (T1, T2, T4) and Interaction SL by Text on quality of the writer's positioning. GEE Binomial model results. Marginal means and (SE).

	Texts			
School level	All *n* = 633	T1 *n* = 210[a]	T2 *n* = 211[a]	T4 *n* = 212
Elementary *n* = 65	0.82 (0.019)	0.85 (0.028)	0.80 (0.032)	0.81 (0.027)
High School *n* = 78	0.80 (0.017)	0.82 (0.027)	0.81 (0.023)	0.78 (0.023)
University *n* = 69	0.83 (0.015)	0.77 (0.028)	0.82 (0.022)	0.89 (0.017)
Total *n* = 212		0.81 (0.017)	0.81 (0.014)	0.82 (0.013)
$Wald_{SL\ x\ text}$ 16.492, $p = 0.002$				
T1 versus T2				
T2 versus T4 $Wald_{SL}$ 8.262, $p = 0.016$; $Wald_{SL\ x\ text}$ 6.903, $p = 0.032$				

Note: [a]Two T1 texts and one T2 could not be coded because neither explicit nor inferred standpoint could be established.

No effect of SL, Text, or interaction SL by Text was found across T1 and T2. In contrast, a significant effect of SL emerged across T2 and T4. A one-sided aligned positioning was lower at elementary than at high school and university level, but

differed significantly only from university level. The interaction was significant (Wald 18.890, p = 0.002). Pairwise comparisons of estimated marginal means of interaction effects across SLs indicated than highschoolers had less probability to show a one-sided positioning than elementary and university students, especially in T4.

2.4.3 Expression of reflective thinking

From a total of 592 texts that contained expository moves, 61 % (359) of texts contained at least one reflective move (M = 1.106, SD 1.345 range 0.00–11.00) while 39 % (233) did not contain reflective moves but only other types of expository moves. Table 6 shows the mean number of moves (SD) and counts that express reflective thinking by SL and Text. Table 7 shows effects of SL, Text, and interaction SL by Text on the presence of reflective thinking in the texts.

There was a simple effect of SL on the probability to express reflective thinking across the three texts and school levels. Pairwise comparisons showed that highschoolers and university students had a significant higher probability of expressing reflective thinking than elementary school students, but no significant differences among them were found. As for Text effects, although the probability was higher in T4, the difference did not attain significance. There was a significant interaction SL by Text. Pairwise comparison of marginal means of interaction effects showed that, across SLs only for highschoolers the probability to express reflective thinking was significantly lower in T2 than in T4.

The comparison between T1 and T2, independently of T4, also yielded a significant effect of SL. Pairwise comparisons pointed at a higher probability of reflective thinking in high school and university than in elementary school, but no significant

Table 6: Mean number (SD), and counts of moves that express reflective thinking by Text and SL.

	Texts			
School level	**Text 1**	**Text 2**	**Text 4**	**Total**
Elementary n = 65	0.94 (SD 1.10); 52	0.44 (SD 0.63); 52	0.53 (SD 0.63); 57	0.63 (0.86); 161
High School n = 78	1.03 (SD 1.08); 70	0.92 (SD 1.15); 76	1.62 (SD 1.94); 69	1.17 (SD 1.24); 224
University n = 69	1.23 (SD 1.44); 69	1.62 (SD 1.94); 69	1.36 (SD 1.46); 69	1.40 (SD 1.63); 207
Total	1.08 (SD 1.23); 191	1.04 (SD 1.47); 197	1.20 (SD 1.32); 204	1.10 (SD 1.35); 592

Table 7: Effect of school level and Text (T1, T2, T4), and Interaction school level by Text on reflective thinking. GEE Negative binomial. Marginal means and (SE).

	Texts			
School level	**All** ***n*** **= 592**	**T1** ***n*** **= 191**	**T2** ***n*** **= 197**	**T4** ***n*** **= 204**
Elementary *n* = 65	0.64 (0.07)[a]	0.93 (0.152)	0.42 (0.099)	0.52 (0.08)
High School *n* = 78	1.18 (0.09)[b]	1.02 (0.13)	0.92 (0.13)	1.54 (0.16)
University *n* = 69	1.36 (0.14)[b]	1.23 (0.17)	1.62 (0.23)	1.36 (0.17)
Total *n* = 212		1.00 (0.08)	0.94 (0.09)	1.10 (0.08)
$Wald_{SL\text{-}all}$ = 31.278, *p* = 0.000; $Wald_{SL\ x\ text}$ 24.825, *p* = 0.000				
T1 versus T2 $Wald_{SL}$ = 17.009, *p* = 000; $Wald_{SL\ x\ text}$ = 11.403, *p* = 0.003				
T2 versus T4 $Wald_{SL}$ = 46.154, *p* = 0.000; $Wald_{SL\ x\ text}$ =11.403, *p* = 0.00				

Note. Latin letters for marginal mean ranking by pairwise comparisons with Bonferroni correction. Lowercase ranking across SLs where "a" means the lowest mean Marginal means are predicted probabilities.

differences between the two older groups. There was no effect of Text, but a significant interaction Text by SL was found. Pairwise comparison of marginal means of interaction effects showed that significant differences due to school level appeared only in T2. Elementary school students' probability to produce reflective commentaries was lower than highschoolers' and university students', but only in T2.

The same comparison between T2 and T4, independently of T1, yielded an effect of SL. Pairwise comparisons indicated that highschoolers' and university students' probabilities to express reflective thinking were significantly higher than elementary school students'; no significant differences among them were found. Although the probability to express reflective thinking was higher in T4 compared to T2 in every SL, there was not a simple effect of Text. The interaction Text by SL was significant. Across SLs, only highschoolers showed a significant higher probability of including reflective commentaries in T4 than in T2.

2.4.4 External evaluation

Text global evaluation was expressed in a Likert scale (1 minimal to 5 maximal) for each text (*M* = 3.43, SD 0.98; range 0.00–5.00). Table 8 presents the effect of SL, Text, and interaction SL by Text on Global evaluation (GE).

There was a simple effect of SL and Text on the external evaluation global score across the three school levels and texts. Pairwise comparisons showed that

Table 8: Effect of School level and Text (T1, T2, T4), and Interaction School level by Text on global evaluation. GEE Normal model results. Marginal means and Standard error (SE).

	Texts *n* = 243			
School level	**All**	**T1** *n* = 81	**T2** *n* = 81	**T4** *n* = 81
Elementary *n* = 65	3.00 (0.10)[a]	2.92 (0.119)	2.79 (0.136)	3.27 (0.123)
High School *n* = 78	3.77 (0.09)[b]	3.70 (0.117)	3.56 (0.166)	4.02 (0.091)
University *n* = 69	3.72 (0.07)[b]	3.63 (0.106)	3.78 (0.099)	3.67 (0.271)
Total *n* = 212		3.42 (0.065)[A]	3.38 (0.08)[A]	3.70 (0.127)[B]
$Wald_{SL}$ = 38.157, p = 0.000; $Wald_{text}$ = 18.354, p = 0.000; $Wald_{SL\ x\ text}$ = 14.130, p = 0.007.				
T1 versus T2 $Wald_{SL}$ = 39.197, p = 0.000 p. 22				
T2 versus T4 $Wald_{SL}$ = 32.456, p = 0.000; $Wald_{text}$ = 15.732, p = 0.000; Wald = 14.433 p = 0.001				

Note. Latin letters for marginal mean ranking by pairwise comparisons with Bonferroni correction. Uppercase indicates ranking across time points and lowercase ranking across SLs where "a" means the lowest mean. Marginal means are predicted probabilities.

elementary school students had significantly higher probabilities to get lower scores than highschoolers and university students, without significant differences among them. As for Text effects, pairwise comparisons across the three texts showed significant differences between T4 and the two previous texts, but no significant differences between them. Texts produced at T4 got the highest scores.

Across T1 and T2 there was only significant effect of SL but not of Text. The different topic had no effect on the global score. The same comparison between T2 and T4 (after classroom activities but on the same topic than T2) showed significant main effect of SL and significant interaction Text by SL. Pairwise comparison of interaction effects indicated that only elementary school students had significant lower probabilities to get higher scores than highschoolers and university students in T2. However, in T4 elementary school texts got only lower scores than highschoolers but did not differ significantly from university students. University students did not improve their probabilities to get higher scores from T2 to T4. Although highschoolers' T4 got higher scores than T2, this improvement did not attain significance.

To test the extent to which the analyzed text dimensions explained global score on text quality we ran a linear regression analysis using SLs and the analyzed dimensions as independent variables and GE as dependent variable.

Table 9: Linear regression of school level and Text dimension on external text evaluation.

Variables	*B*	SE	*β*	*t*	R^2	ΔR^2
					0.175***	0.166
High versus elementary school	0.596	0.107	0.276***	5.578		
University versus elementary	0.504	0.114	0.232***	4.431		
Explicit central standpoint	0.348	0.105	0.147**	3.322		
Reflective thinking	0.076	0.034	0.099**	2.226		
Quality of positioning	0.035	0.017	0.093*	1.999		

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Results showed that together SL and the three text dimensions significantly explained 17.5 % of the global evaluation variance (F (5, 458) = 19.479; $p < 0.000$), see Table 9. The standardized Beta indicated that SLs showed a significant role in explaining global evaluation ($\beta = 0.28$, $p < 0.001$; $\beta = 0.23$, $p < 0.001$; respectively), approximately half the explained variance of global evaluation. But explicitness, reflectiveness, and positioning complemented the explained variation of global evaluation ($\beta = 0.15$, $p = 0.001$; $\beta = 0.10$, $p = 0.026$; $\beta = 0.09$, $p = 0.046$; respectively), although only to a lower extent.

2.5 Summary and discussion

One of DAW's goals is to detect the developmental route of text features that indicate students' appropriation of genre relevant features. To this end, we have traced development in three text dimensions committed to fulfill analytical writing communicative goals and text well-formedness: Explicitness of a central standpoint, Flexible positioning, and Expression of Reflective thinking. Explicit formulation of central standpoint is a core element in essays on controversial topics (Mann and Thompson 1988), and establishes a hierarchical differentiation of text components (Longacre 1996). While the Expression of (R)reflective thinking plays a crucial role in writing analytically, Flexible positioning is an additional asset in such essays (Kuhn and Crowell 2011). Overall, the results of the study showed how challenging is for novice writers to fulfill analytical genre expectancies regarding the realization of the argumentative component. The formulation of a central standpoint was realized in less than 30 % of the texts, and 80 % of the texts contained only devices *supporting* own position rather than deviating from own contention to build a flexible positioning. The realization of the expository component was less difficult, and 60 % of the texts contained at least one reflective move.

Analyses of the differential qualities of the texts' rhetorical moves disclosed a similar developmental pattern in the probability to express a central standpoint and to include reflective commentaries. Both skills developed with schooling although significant differences were found only between elementary and higher school levels across texts. In contrast, the development of a flexible positioning was relatively conditioned by Text. University students displayed a higher probability to deviate from own contentions than younger schoolers but only in T1, when writing on freedom of a dressing code. Moreover, highschoolers showed a more flexible positioning than elementary and university students, but only after their participation in the classroom activities. These findings fit partially our expectation of a steady development across schooling for the three dimensions. The greatest developmental gains occurred toward high school and slowed down afterward. This stronger improvement during the years of learning to compose texts replicates previous findings for other text-features (Aparici et al. 2021).

In line with expectations, differences in topic did not affect the pattern of responses. Nevertheless, the development of a Flexible positioning – clearer when writing about a dressing code than when writing about freedom to move – hints at a facilitating effect of certain topics on decentering. This finding calls for research controlling for the relative influence of the tested factors (pedagogical input and topic-content) on the quality of analytical essays.

An additional goal of DAW's was to detect potential improvement of genre relevant features. To this end we traced changes in the assessed dimensions after the implementation of pedagogical work. Unlike what we found regarding the developmental pattern, the three dimensions were similarly impinged by pedagogical scaffolding: it benefited mainly high school students and less university students. High schoolers were more prone to highlight their standpoint, attenuate their one-sided perspective and reflect on the target topic after their participation in the classroom activities. Elementary school students showed a relative improvement in the three dimensions although not significantly so.

Against expectations, pedagogical work or, perhaps, repeated text writing on the same topic seemed to negatively impact university students: their probabilities to underlie a standpoint, to show a flexible positioning and provide reflections on the topic decreased after classroom activities. The three dimensions assessed showed potentiality for improvement up and to the university level. However, university students would need pedagogical activities different than those used in earlier school levels to impinge their improvement.

Finally, results of regression analyses supported the foundational role of development which explained the highest proportion of variance, and the significant contribution of the three assessed dimensions. Flexible positioning, though, contributed less than Explicit central standpoint and Reflective thinking to reader's

appreciation of text quality. This finding hints at a certain dissonance between researchers' and evaluators' criteria regarding the features that better contribute to the quality of analytical essays.

3 Revisiting DAW

The ultimate aim of our studies was to capture how social expectancies of analytical writing become part of participants linguistic literacy. To this aim we developed a methodological framework to objectivize subjective, unavoidable appraisal(s) of analytical essays quality in the light of analytical genres constraints. The ultimate aim of our studies was to capture how social expectancies of analytical writing become part of participants' linguistic literacy. To this aim, we developed a methodological framework to objectivize subjective, unavoidable appraisal(s) of analytical essay quality in light of analytical genre constraints. This methodological framework enabled us to address the specific questions that guided our inquiry across studies: what text features make a difference in the appraisal of quality, what is their development across school levels, and what are their differential sensitivities to pedagogical input so as to detect potential improvement. Applying this framework we can contribute to teaching and learning analytical writing.

3.1 DAW's basic tenets

A basic tenet underlies the developmental focus and the control for pedagogical input as distinguishing features of our framework: tracing the trajectory on a particular domain of knowledge is valuable for teaching and learning. This value would increase if we could identify not only what students have to learn in order to become proficient in a certain domain, but also what they *bring* to the learning process.

Children's growing sensitivity to different types of discourse is part of their pragmatic development (e.g., Donovan and Smolkin 2002). Over the years from childhood to adolescence and beyond, the heterogeneity of discourse practices increases. During this transition, young people are required to gain command in different types of discourse. This involves learning discourse conventions that result both from experience with written language and individuals' cognitive development (Uccelli et al. 2012). We assumed that if we promote classroom activities that raise participants' awareness of the particular motives and rhetorical patterns of analytical writing, we would be able to identify the insights learners gain through their participation in such practices. That is, we would be able to

assess how pedagogical input impinges on how skills are transferred from social expectancies of genre to individuals, as part of their ongoing development (Kuhn and Crowell 2011).

A second set of tenets grounded two additional features of DAW: focus on the written product and multilayered analysis. Following Longacre's (1996) proposal we assumed that the quality of a text can only be appreciated in light of the characteristic type of discourse that the text it belongs to. Discourse realizations in specific genres compel individual writer' decisions concerning the amount and structuring of information, as well as the selection of devices at every level of language (Alamargot and Chanquoy 2001). The features of the written product account for the writer's decisions. And, as semantic units, written texts proper functioning result from a delicate interplay between different linguistic levels – from words to overall text organization.

Consequently, we took a multilayered approach both in the different studies dwelling on text features at different language levels, and in the very process of text analysis. We focused on different levels of the written product – from the lexical, through syntactic-discursive to textual overall organization. However, when centering on one level, we were attentive to how the others were affected. For example, the relevance of particular syntactic constructions (e.g., use of relative clauses) was interpreted according to the role they fulfilled in the text. Likewise, when centered on featuring text structure, we paid attention to the incidence of modal expressions (i.e., qualifiers).

Finally, what novice and expert writers pursue with writing in academic settings prompted us to include a reader-based text appraisal as an additional feature of DAW. After all, writers' ultimate goal is to persuade their readers—teachers, colleagues, reviewers, editors — that they qualify as good enough to be members of the target discourse community. Besides our researcher-based assessment, we resorted to a reader-based assessment and tested the relevance of the selected features in the eyes of the actual judges/readers in natural academic settings, that is, teachers. This approach yielded crucial implications for research and teaching.

3.2 DAW's applications and (possible) contribution to teaching

Across studies, we assessed diverse text-embedded features as indicators of developing writing proficiency. In Aparici et al. (2021) we examined lexical and syntactic-discursive indicators of quality. We could tell apart linguistic features that improved in the course of development from those that show malleability to pedagogical scaffolding i.e., higher potentiality for improvement by means of

pedagogical work. However, none of the identified features showed a simple effect of pedagogical input. In the domain of analytical writing, pedagogical action impinges on development rather than the other way around. The sensitivity to pedagogical scaffolding shown by the improvement of most indicators of quality in a short time span of a genre-oriented instruction highlights the significance of pedagogical input for developing analytical writing proficiency but provided age/school level be considered when choosing which features to focus pedagogical work on.

The findings described for the impact pedagogical work has on proficient writing (expression of a central standpoint, mitigation of own contentions, and supply of reflective elaboration) confirmed the possibilities of pedagogical input, but also exposed some limitations. The applied classroom activities enhanced developmental gains in proficiency, but neither at every school level nor for every dimension.

Beyond the identification of indicators of essay writing proficiency, additional findings of this and previous studies might be helpful for educators. The first concerns the impact of pedagogical input on specific lexical and syntactic-discursive features rather than on particular language levels (e.g., the effect on lexical diversity but not on lexical density); this calls for diversified pedagogical intervention within each language domain. The second relates to the inverse relation between the experience of the writers and the efficacy of classroom activities. Less experienced writers obtained greater benefits than more experienced university students. This finding points at the need to tailor pedagogical work to suit the specific needs of writers to warrant further improvement. Again, this conclusion is reinforced by the above-reported results on the greatest developmental gains during elementary school for Explicit central standpoint, Flexible positioning and Reflective thinking.

An additional finding that calls for reflection concerns the strong and generalized impact of productivity on external assessment of text quality. Across schooling, producing more text made a better impression on external judges than appropriate text cohesion or high lexical richness. This puts into questions current advice on the benefits of concise writing (Williams and Bizup 2017). Being concise may serve to grab and hold reader's attention but does not appear to be highly appreciated by regular teachers.

3.3 Future research

Our incursion in the qualitative, content-related differences of rhetorical moves enabled us to grasp the specific challenges that students must face to gain proficiency in analytical writing, in particular, the importance of expressing a central standpoint and a flexible positioning by providing pros and cons to their own views. Students'

difficulties to face these challenges, as reflected in the results we obtained, underlie the need of raising students' awareness on how explicit expression of standpoint affects the "climatic development" (Longacre 1996) of their essays, and providing adequate pedagogical scaffolding to this end.

In light of the fruitful implications that we drew from applying a qualitative, content-related approach to the specific nature of rhetorical moves, future research should extend this approach to explore the type and quality of the grounds that students provide for the claims included in argumentative moves.

The applications of the model served to show developmental changes in the way novice (and more experienced) writers mobilize their linguistic repertoire to fit readers' expectations on analytical essays. Two extensions are needed to probe DAW's level of generality. We should extend this methodological framework to different languages and/or linguistically diverse populations to determine how language-specific or dependent on participants' linguistic conditions the results are. Finally, only by extending this model to other types of discourse will we be able to demonstrate its useful gauge for the study of rhetorical flexibility.

References

Alamargot, Denis & Lucile Chanquoy. 2001. *Through the models of writing*. Dordrecht-Boston-London: Kluwer Academic Publishers.

Allen, Laura K., D. Aaron Likens & S. Danielle McNamara. 2019. Writing flexibility in argumentative essays: A multidimensional analysis. *Reading and Writing* 32(6). 1607–1634.

Alves, Rui & Teresa Limpo. 2015. Progress in written language bursts, pauses, transcription, and written composition across schooling. *Scientific Studies of Reading* 19(5). 374–392.

Aparici, Melina, Rocío Cuberos, Naymé Salas & Elisa Rosado. 2021. Linguistic indicators of text quality in analytical texts: Developmental changes and sensitivity to pedagogical work. *Journal for the Study of Education and Development* 44(1). 9–46.

Aparici, Melina & Marta García. 2018. *A usage–based assessment of multilingual/multiliterate knowledge*. Symposium paper presented at the 1st Literacy Summit, University of Porto, 1–3 November.

Aparici, Melina & Joan Perera. 2001. Variedad y distribución de los movimientos retóricos en textos expositivos escritos [Variety and distribution of rhetorical moves in expository texts]. In *Proceedings Congreso Internacional de la Asociación Española de Lingüística Aplicada, Alcalá de Henares, Madrid* 125–130.

Applebee, Arthur N. 2000. Alternative models of writing development. In Roselmina Indrisano & James R. Squire (eds.), *Perspectives on writing: Research, theory, and practice*, 90–110. Newark, DE: International Reading Association.

Berman, Ruth A. & Dorit Ravid. 2009. Becoming a literate language user: Oral and written text construction across adolescence. In R. David Olson & Nancy Torrance (eds.), *Cambridge handbook of literac*, 92–111. Cambridge: Cambridge University Press.

Berman, Ruth A. & Dan I. Slobin (eds.). 1994. *Relating events in narrative: A crosslinguistic. developmental study*. Hillsdale, NJ: Lawrence Erlbaum.

Berman, Ruth A. 2002. Crosslinguistic comparisons in later language development. In Sven Strömqvist (ed.), *The diversity of languages and language learning*, 25–44. Lund: Center for Languages and Literature.

Berman, Ruth A. 2004. Between emergence and mastery: The long developmental route of language acquisition. In A. Ruth Berman (ed.), *Language development across childhood and adolescence: Psycholinguistic and crosslinguistic perspectives, Trends in Language Acquisition Research (TILAR)* 3, 9–34. Amsterdam: John Benjamins..

Berman, Ruth A. 2016. Linguistic literacy and later language development. In Joan Perera, Melina Aparici, Elisa Rosado & Nayme Salas (eds.), *Written and spoken language development across the lifespan: essays in honor of Liliana Tolchinsky*, 181–200. New York: Springer Verlag.

Berman, Ruth A. & Bracha Nir-Sagiv. 2007. Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes* 43. 79–120.

Berman, Ruth A. & Ludo Verhoeven. 2002. Cross-linguistic perspectives on the development of text production abilities in speech and writing. *Written Languages and Literacy* 5(Special issue). 1–289.

Bernstein, Amit, Yuval Hadash, Yael Lichtash, Galia Tanay, Kathrine Shepherd & David M. Fresco. 2015. Decentering and related constructs: A critical review and metacognitive processes model. *Perspectives on Psychological Science* 10(5). 599–617.

Bialystok, Ellen. 2018. Bilingual education for young children: Review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism* 21(6). 666–679.

Britton, Bruce K. 1994. Understanding expository text: Building mental structures to induce insights. In Morton Gernsbacher (ed.), *Handbook of psycholinguistics*, 641–674. New York: Academic Press.

Caffi, Claudia. 2013. Mitigation. In Marina Sbisà & Ken Turner (eds.), *Pragmatics of speech actions*. (Handbooks of Pragmatics [HOPS] 2), 257–286. Berlin & New York: Mouton de Gruyter.

Chen, Ying-Chih. 2019. Writing as an epistemological tool: Perspectives from personal, disciplinary, and sociocultural landscapes. In Vaughan Prain & Brian Hand (eds.), *Theorizing the future of science education research*. Contemporary Trends and Issues in Science Education, 1878-0482, 49. New York: Springer Verlag.

Choi, Inji. 2007. How and when do children acquire the use of Discourse Markers? In *CamLing2007: Proceedings of the 5th University of Cambridge Postgraduate Conference in Language Research* 40–47.

Crossley, Scott A. & S. Danielle McNamara. 2010. Cohesion, coherence, and expert evaluations of writing proficiency. In Stella Ohlsson & Richard Catrambone (eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 984–989. Austin, TX: Cognitive Science Society.

Crossley, Scott A., L. Jennifer Weston, Susan McLain Sullivan & Danielle McNamara. 2011. The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication* 28(3). 282–311.

Donovan, Carol & Laura Smolkin. 2002. Children's genre knowledge: An examination of K–5 students' performance on multiple tasks providing differing levels of scaffolding. *Reading Research Quarterly* 37(4). 428–465.

Driver, Rosalind, Paul Edward Newton & Jonathan Osborne. 2000. Establishing the norms of scientific argumentation in classrooms. *Science Education* 84(3). 287–312.

Felton, Mark K. & Deanna Kuhn. 2001. The development of argumentative discourse skill. *Discourse Processes* 32(2–3). 135–153.

Ferretti, Ralph, P. & Steve Graham. 2019. Argumentative writing: Theory, assessment, and instruction. *Reading and Writing* 32(6). 1345–1357.

Graham, Steve S. & Karen Harris, R. 2000. The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist* 35(1). 3–12.

Hardin, James W. & Joseph M. Hilbe. 2012. *Generalized estimating equations*. Boca Raton, FL: CRC Press. Available at: https://www.perlego.com/es/book/1605883/generalized-estimating-equations-pdf.

Hayes, John R. 1996. A New Framework for understanding cognition and affect in writing. In C. Michael Levy & Sarah Ransdell (eds.), *The science of writing: Theories, methods, individual differences and applications*, 1–27. Mahwah, NJ: Lawrence Erlbaum.

Hsin, Lisa & Catherine Snow. 2017. Social perspective taking: A benefit of bilingualism in academic writing. *Reading and Writing* 30(6). 1193–1214.

Kuhn, Deanna & Amanda Crowell. 2011. Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science* 22(4). 545–552.

Langacker, Ronald W. 1999. *Grammar and conceptualization*. Berlin & New York: Mouton de Gruyter.

Longacre, Robert E. 1996. *The grammar of discourse. Topics in language and linguistics*, 2nd edn. New York: Springer Verlag.

Lukeman, Noah. 2007. A dash of style: The art and mystery of punctuation. New York: W.W Norton.

Mann, William C. & Sandra A. Thompson. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3). 243–281.

McCutchen, Deborah, Mardean Francis & Shannon Kerr. 1997. Revising for meaning: Effects of knowledge and strategy. *Journal of Educational Psychology* 89(4). 667–676.

McMaster, Kristen & Christine Espin. 2007. Technical features of curriculum–based measurement in writing. *The Journal of Special Education* 41(2). 68–84.

McNamara, Danielle S., A. Scott Crossley & M. Philip McCarthy. 2010. Linguistic features of writing quality. *Written Communication* 27(1). 57–86.

Piaget, Jean. & Barbel Inhelder. 1969. *The psychology of the child*. NY, New York: Basic Books.

Ravid, Dorit & Liliana Tolchinsky. 2002. Developing linguistic literacy: A comprehensive model. *Journal of Child Language* 29(2). 417–447.

Rosado, Elisa, Iban Mañas, Irene Yúfera & Melina Aparici. 2021. El desarrollo de la escritura analítica: Aprender a enlazar la información, aprender a posicionarse [The development of analytical writing: Learning to link information, learning to position yourself]. *Pensamiento Educativo* 58(2). 1–18.

Schleppegrell, Mary J. 2004. *The language of schooling. A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum.

Stavans, Anat, Batia Seroussi & Sara Zadunaisky Ehrlich. 2019. Literacy-related abilities' effects on argumentative text quality structure. *Journal of Literacy Research* 51(3). 315–335.

Swales, John. 1990. *Genre analysis: English for academic and research settings*. Cambridge: Cambridge University Press.

Tolchinsky, Liliana. 2004. The nature and scope of later language development. In A. Ruth Berman (ed.), *Language development across childhood and adolescence: Psycholinguistic and crosslinguistic perspectives, trends in language acquisition research (TILAR)*, 3, 233–247. Amsterdam: John Benjamins.

Tolchinsky, Liliana, Melina Aparici & Elisa Rosado. 2017. Escribir para pensar y persuadir [Write for thinking and persuading]. *Textos de Didáctica de la Lengua y la Literatura* 76. 14–21.

Tolchinsky, Liliana, Melina Aparici & Hugo Vilar Weber. 2021. Macro– and micro–developmental changes in analytical writing of bilinguals from elementary to higher education. *International Journal of Bilingual Education and Bilingualism* 25(1). 1–16.

Tolchinsky, Liliana, María Dolores Alonso-Cortes Fradéjas, Teresa Llamazares Prieto & Mercedes López Aguado. 2023. The development of rhetorical preferences in the analytical writing of Spanish students from elementary to higher education. In Alina Galvao Spinillo & Carmen Sotomayor (eds.), *Development of writing skills in children in diverse cultural contexts. Contributions to teaching and learning*. New York: Springer Verlag.

Tolchinsky, Liliana, Elisa Rosado, Hugo Vilar Weber, Melina Aparici & Rocío Cuberos. submitted. Portraying analytical texts of monolingual and bilingual students from elementary to higher education.

Toulmin, Stephen E. 1958. *The uses of argument*. Cambridge: Cambridge University Press.

Uccelli, Paola, L. Christina Dobbs & Jessica Scott. 2012. Mastering academic language: Organization and stance in the persuasive writing of high school students. *Written Communication* 30(1). 36–62.

Upton, Thomas A. & Mary Ann Cohen. 2009. An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies* 11(5). 585–605.

Vande Kopple, William J. 1998. Relative clauses in spectroscopic articles in the physical review, beginnings and 1980: Some changes in patterns of modification and a connection to a possible shift in style. *Written Communication* 15(2). 170–202.

Vilar Weber, Hugo & Liliana Tolchinsky. 2021. The rhetorical structure of analytical writing: A developmental approach. *Text & Talk* 42(1). 131–152.

Weston, Jennifer L., Scott A. Crossley, Philip M. McCarthy & Danielle S. McNamara. 2011. Number of words versus number of ideas: Finding a better predictor of writing quality. In *Proceedings of the 24th International Florida Artificial Intelligence Research Society*, 335–340. Available at: http://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS11/paper/download/2618/318.

Williams, Joseph M. & Joseph Bizup. 2017. *Style: Lessons in clarity and grace*, 12th edn. London: Pearson.

Willingham, Daniel T. 2007. Critical thinking: Why is it so hard to teach? *American Educator* 31. 8–19.