
6-3-2021

Questioning Standards of Evaluation in Educational Research: Do Educational Researchers Ventriloquize Learners' Voices in L2 Education?

Anastasia A. Boldireff

Concordia University, Montreal, Quebec, Canada, anastasia.boldireff@mail.concordia.ca

Follow this and additional works at: <https://nsuworks.nova.edu/tqr>

 Part of the [Educational Methods Commons](#), [Language and Literacy Education Commons](#), [Quantitative, Qualitative, Comparative, and Historical Methodologies Commons](#), and the [Social Statistics Commons](#)

Recommended APA Citation

Boldireff, A. A. (2021). Questioning Standards of Evaluation in Educational Research: Do Educational Researchers Ventriloquize Learners' Voices in L2 Education?. *The Qualitative Report*, 26(6), 1724-1735. <https://doi.org/10.46743/2160-3715/2021.4691>

This Article is brought to you for free and open access by the The Qualitative Report at NSUWorks. It has been accepted for inclusion in The Qualitative Report by an authorized administrator of NSUWorks. For more information, please contact nsuworks@nova.edu.



Qualitative Research Graduate Certificate
Indulge in Culture
Exclusively Online • 18 Credits
LEARN MORE

NSU
NOVA SOUTHEASTERN
UNIVERSITY

NOVA SOUTHEASTERN

Questioning Standards of Evaluation in Educational Research: Do Educational Researchers Ventriloquize Learners' Voices in L2 Education?

Abstract

Learners are not stakeholders in their own education. Adhering to the quantitative gold standard in English as a Second Language (ESL) deprives the learner from having a voice in their learning process. This paper addresses voicelessness and ventriloquism in ESL, ventriloquism referring to the act of voicing the thoughts of another person, in this case the system overriding the learners' experiences. This article addresses this problem, aligning itself with the Platinum standard while challenging the quantitative gold standard in ESL research. This paper offers resonance and semantic reliability as evaluative measures in educational research taken from literary criticism. The notion of resonance has been addressed in the literature on qualitative research since the dawn of the narrative turn; I address how resonance can be used in educational research.

Keywords

qualitative research, validity, narrative inquiry, applied linguistics, ESL/EFL

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Questioning Standards of Evaluation in Educational Research: Do Educational Researchers Ventriloquize Learners' Voices in L2 Education?

Anastasia A. Boldireff
Concordia University, Montreal, Quebec, Canada

Learners are not stakeholders in their own education. Adhering to the quantitative gold standard in English as a Second Language (ESL) deprives the learner from having a voice in their learning process. This paper addresses voicelessness and ventriloquism in ESL, ventriloquism referring to the act of voicing the thoughts of another person, in this case the system overriding the learners' experiences. This article addresses this problem, aligning itself with the Platinum standard while challenging the quantitative gold standard in ESL research. This paper offers resonance and semantic reliability as evaluative measures in educational research taken from literary criticism. The notion of resonance has been addressed in the literature on qualitative research since the dawn of the narrative turn; I address how resonance can be used in educational research.

Keywords: qualitative research, validity, narrative inquiry, applied linguistics, ESL/EFL

Introduction: The Gold Standard

Qualitative research is uncommon in English as a Second Language (ESL) education as statistical validity remains the gold standard in research (Hammersley, 2015). This paper's position aligns with what Khagram and Thomas (2020) call the "*Platinum standard*" which esteems both quantitative and qualitative research equally; the former standard based on experimental methods; the latter, based on case studies, comparative methods and triangulation. Inclusion of both methods strengthens the application from theory to practice. Each standard currently competes for dominance in research; however, the qualitative gold standard should be equally valued in ESL. This paper addresses the implications of student voicelessness in light of this quantitative gold standard. I question the lack of attention to a qualitative approach in second language (L2) education to claim that stories are relevant in ESL. Learners' voices and experiences are decontextualized and dehumanized when data or a researchers' objective interpretation are presented in the stead of learners' or teachers' personal stories (Hanauer, 2012; Iida, 2017). I discuss the ramifications of validity remaining divorced from a pluralistic interpretation (Altrichter, 1986). I suggest a more inclusive method of evaluation to include *resonance* in ESL education from literary criticism. I concurrently wish to argue that literary criticism is valuable to educational researchers for its emphasis on hermeneutic analyses and objective criticism through what I wish to call *semantic reliability*. I posit that ESL research would benefit from consistency in the use of their language to mean what they say and to question everything that is said. This essay argues for why emotions and learners' voices should be communicated along with statistical findings that quantify them through measurements of resonance, which should be valued in ESL, and educational research at large.

Defining Terms and Hermeneutically Analyzing Shifts in Their Meaning

Statistical significance, as a gold standard, has extended to the social sciences and educational research (Hammersley, 2015). This standard is problematic in educational research because the factors that researchers are attempting to evaluate, such as learning processes, teaching practices, learner motivation and language acquisition, are difficult to quantify (S. Shaw, personal communication, April 4, 2019). Before addressing my argument, I wish to define operational terms including *validity*, *reliability*, *evaluation*, *ventriloquism*, and *resonance*.

Quantitative research looks at measuring variables which impact learning with numbers to confirm predictable generalizations in closed-ended questions which yield simplistic confirmatory results, whereas qualitative research looks at interrogating the classroom in an open-ended investigation through in-depth inquisition through surveys, interviews, observation, analyses of artefacts or discourse. Both methods value *scientific validity*, meaning: “the quality of being well-founded on fact or established on sound principles and thoroughly applicability to the case or circumstances,” (OED Online) and require statistical *reliability*: “the degree to which repeated measurements of the same subject under identical conditions yields consistent results,” (OED Online). However, attempts of applying statistical reliability inappropriately to qualitative research to call it mixed-methods often results in a poor mixed-method study failing both the objectives of a quantitative study and the meaning in a qualitative study. Alternative methodologies, epistemologies, and measurements of evaluation and even interpretations of “validity” are currently being questioned (Dirkx, 2006). The very meaning of what evaluation is, and what education is, is what is on trial and at stake; the current definition and use of education is displaced from the etymology of evaluation which results in a shift of semantic use and interpretation. *Evaluation* is defined as “the action of appraising or valuing; a calculation or statement of value” (OED). I question what is valued in the classroom, in curriculum, or policy, and ultimately, in educational research.

The quantitative push in ESL Educational research currently values *validity* and *reliability* in academic studies on educational research to be the status quo; whereas educators and the very education system values efficiency and results over process and learning. This state of affairs may have to do with the fact that literature was removed from the ESL curriculum when Topping (1968) argued that literary language lacked structural complexity, did not conform to standard grammatical rules, and held a remote cultural perspective. Indeed, Edmondson’s (1997) criticized educationalist researchers’ ability to implement scientific studies using poetry. This stance has not shifted: Topping’s declaration corresponds to both an era when linguists became the authority on ESL education, silencing the literary scholar (Widdowson, 1983) thus emphasizing a quantitative approach to language learning. This period corresponds to when ESL curriculum was first created to reflect the Grammar-Translation approach, highlighting grammar and vocabulary as the main priorities in L2 instruction (Richards, 2001) de-prioritizing other things such as culture, pronunciation, or literacy instruction among others (Carter, 2007; Kuru Gönen, 2018). Literature’s disappearance from language curriculum ushered in a focus on “utilitarian output” (Kramsch & Kramsch, 2000, p. 6). It wasn’t until the linguistic turn – 1990s – when Vytgotskian and Bahktian theories in pragmatics threatened traditional linguists (Carter, 2007). This turn both accelerated a concern for the ownership of English and a professional need to become bilingual (Carter, 2007). As a result, the very definition of *education* is undergoing a shift in semantic meaning from: “The process of bringing up a child, with reference to forming character, shaping manners and behaviour” (OED, 1) to meaning “the systematic instruction, teaching, or training in various academic and non-academic subjects given to or received by a child, typically at a school; the course of scholastic instruction a person receives in his or her lifetime” (OED, 4a).

Education has then shifted from a learning process and to a system (Hanauer, 2012; Kramsch, 2006). Thus, in ESL, the final product, that is, the learner's passing grade related to their knowledge of conjugations is more important than if they can speak the language fluently or accurately.

By contrast, literary studies teach and esteem *resonance* as the measurement and standard of a text or piece of art being *valid*. Resonance is understood as "a corresponding or sympathetic response; the power or quality of evoking suggesting images, memories, and emotions, an allusion, connotation or overtone" (OED, 2). I believe that resonance can be applied to ESL. Emotional resonance could become a measurement of evaluation to inform current research and educational practices so that we better understand learners' frustrations, successes, motivations and ultimately, their progress.

In remaining consistent with hermeneutic analysis, *ventriloquism* means both: "to produce sounds in the manner of a ventriloquist" (Oxford Dictionary), as in a form of another person, or to imitate, or pantomime them; it's newer use however is defined as "voicing (the thoughts) of another person" (Oxford Dictionary). It is the latter definition that I want the reader to bear in mind throughout this paper as this use aids the allegorical story, I wish to portray of the L2 learners being represented through numerical statistical analyses. The exemplary sentence the dictionary uses to contextualize understanding is: "Instead of having their preferences ventriloquized by fathers, husbands, or brothers, suffragists wanted to express their own political opinions" (*Oxford Dictionary Online*).

Experiences are not adequately expressed through language; everyday feelings are transient and non-coherent and difficult to evaluate as a standardized human experience; thus, I recommend that the way *resonance* is measured in art be reflected in the reporting of research in education:

The criterion of good art is its power to command one's contemplation and reveal a feeling that one recognizes as real, with the same "click of recognition" with which an artistic knows that a form is true. (Langer, 1953, p. 405)

I argue that Langer's theory of connotational semantics applies to both educational advocacy and successful implementation of educational interventions in ESL. Researchers' and policy makers' collective inability to connect and engage with learner's experience and teachers' experiences in addition to their proliferation of statistical numbers falls flat to implement change because these numbers are disconnected from human voices. I am not arguing to discredit the use of statistics in ESL or English as a Foreign Language (EFL) but propose the inclusion of resonance as an approach for evaluation in education. I explore gaps in empirical scholarship which ventriloquize subject's voices and divorces their emotions for the benefit of measurements. In taking second language education as metonym for educational research as a whole; I aim for the inclusion of learners' and teachers' voices in L2 education and resonance as a new approach or standard of evaluation. Education has changed from process to a system. This system, I contend, dehumanizes the learner's voice and experience in the classroom in the very research which aims to improve the L2 learning experiences by tending to rely on the gold standard: the voices of learners and teachers are both needed in educational research in order to make informed decisions.

Voicelessness in ESL

Voicelessness is not only a problem in educational research but in the ESL/EFL classroom. Iida (2017) reports his experience learning to write as a Japanese ESL student in graduate school. He claims that "voicelessness" was his problem in L2 writing. Iida has

transitioned from an ESL learner to a professor who uses poetry to teach ESL, he cites the leading problem that ESL students are expected to learn to write academic papers. In an earlier paper, he writes that students do not have an opportunity to develop their voice, which he defines as: “the articulation of their personal needs, interests and ideas – in a social context that presumes an audience – the teachers, classmates, and even community at large” (Iida, 2010, p. 28 cf. Iida, 2008). Current scholarship in Second Language Acquisition (SLA) recognizes what Block (2003) labels as the “Social Turn.” Bylund et al. (2013) writes that “scientific disciplines are commonly characterized by recurring changes or transformations in the beliefs, values, standards, and methods that constitute exemplary ways of conducting research. The field of SLA is not exempt from this trend” (p. 69). Current scholarship in SLA has observed two growing trends: the first the “Social Turn” (Block, 2003; Bylund et al., 2013) and the “Bilingual Turn” (Bylund et al., 2013; Ortega, 2010). The Social Turn looks to emphasize social and cultural contexts in L2 learning; whereas the Bilingual Turn looks to counter the assumption that bilingual speakers and L2 language learners should acquire and be held to the same standards as monolingual native speakers.

While voicelessness is an issue in reporting research, so is the appropriation of the learners’ and subjects’ voices in speaking for them. Garcia (2009) argues for a new modern approach of translanguaging which moves from having bilingual children fit into either-or dialectics to be able to experience and share their culture as well as their language. The learner’s voices need to be heard and their experiences should not be solely enumerated. Kim and Duff (2012) cite Norton (2010) who claims that “when individuals invest in a target language, it is also an investment in their social identity, which continually changes over time” (p. 84). Kim and Duff (2012) studied children of Korean immigrant families who were schooled in Canada and labeled generation 1.5. Scholars have adopted notions of “hybridity” or “third spaces” for these students who have transnational identities. They are marginalized because they are “stuck”: “marginalized to both the new and old worlds, for while they straddle both worlds, they are in some profound sense fully part of neither of them” (p. 84). Speaking English for these Korean immigrants is interpreted as an act of betrayal or a lack of allegiance to the L1 selves. As one of the students, Yellina says: “the words *ESL student* were reminiscent of being below English literacy standards,” (Kim & Duff, 2012, p. 89, *author’s emphasis*) and a label of being substandard to her Canadian peers. I present studies such as these to argue that student’s voices are needed in educational research to help researchers and practitioners create interventions to aid them better; their voices are necessary because it helps us connect with their plight.

This reliance on validity being purely statistical problematizes research which involves learners by dismissing extraneous variables such as gender, or even age, or learning environments. As Dirkx (2006) writes:

It brushes aside qualitative methodologies and alternative epistemologies that have emerged in educational research over the past 30 years as “unscientific.” Proponents of EBR [Evidence Based Research] largely ignore the ways in which knowledge and processes of coming to know are deeply embedded within the social, cultural, political, and economic relationships that constitute the various contexts of educational research. (p. 274)

Presenting educational research is messy because it involves human elements such as learning, motivation, participation, and emotional involvement, among others, which make it difficult to control for (S. Shaw, personal communication, April 4, 2019). As such, there is a lot that can be taken from literary criticism to improve the manner in which ESL/EFL research is conducted. Artistic and literary hypotheses are as necessary and valid as scientific ones

probing questions of human existence (Bruner, 1986). Literature is a manner of observing and reporting events; however, literary evaluation is not concerned with veridicality of hard facts, or even details of a story, like science is with statistics, rather it is concerned with verisimilitude measured through resonance (Seilman & Larson, 1989). These hypotheses vary, where artistic hypotheses do not presume, constrain or separate the reader/subject/participant/learner from the control/poem/learning style/language in order to measure the relationship, instead, it focuses specifically on the subject; it desires to “fit different human perspectives and that they be recognizable as true to conceivable experience” (Bruner, 1986; Larson, 1989). The way that a text is measured to be literary or not has largely to do with evaluating the “charged experience” that the reader experiences through relations to *personal resonance* (Spiro, 1982). Art and science seem to be antithetical for the purpose of evaluation, but I argue that there is value in both methods of evaluation and that for a study to be truly valid that the numbers need to correspond to a phenomenon connected to very real human experiences; and that for the data to be resonant it needs to reflect these experiences. Indeed, Wassiliwizky and Menninghaus (2021) discuss the value of combining different fields and established disciplines in his paper on empirical aesthetics which combines “the fields of psychology, neurosciences, art-related disciplines (such as art history, literary studies, musicology, design), and philosophy at its core, and with anthropology, ethnology, sociology, and linguistics” (p. 10) for the purpose of understanding the mechanism and processes of aesthetic experience. Wassiliwizky and Menninghaus (2021) write that there is complementary effort to map individual variability of aesthetic evaluation and that this is only meaningful against the backdrop of inter-individually shared perceptions and evaluations. I don’t understand why academics of both science and arts have joined forces to understand aesthetic experiences which is recognized as an individual experience, however, the experience of learning a language cannot be given the same respect – as a collection of inter-individual experiences as opposed to a system of learning?

While it is no secret that literary scholarship has been rejected from language teaching methodology (Carter, 2007, Kuru Gönen, 2018), the benefits of applying the stylistic approach to language instruction allows for instructors to teach language awareness and provide learners an opportunity to improve their language capacity, cultural awareness and personal growth. This approach values resonance in the learning process (Carter & Long, 1991; Kuru Gönen, 2018). Resonant research may have more to do with clashing epistemological methodologies than the struggle of research between arts and sciences. Opler (1945) wrote of themes in qualitative analysis being expressions in data. Connection or resistance to language learning and its culture should be an opportunity for educational researchers to collect learners’ experiences of acceptance or ostracization. However, learners in L2 education are not valued in educational research to speak about their own opportunities. Ryan and Bernard (2003) discuss what themes are and how they are a foundational element for qualitative researchers. They write that the collection of themes is “qualitative analysis” (Ryan & Bernard, 2003, p. 88). I believe that critical qualitative analysis can address voicelessness in ESL and between contextualize second language acquisition.

Ventroliquizing the Learner

“One death is a tragedy; one million is a statistic” (Stalin, 1879-1953 Soviet Dictator, 2016). This statement holds true for both the learners and the researchers who evaluate, calculate and report these learner’s experiences as their findings. In evidence-based research, scholars are currently questioning the gold standard. Statistical evidence is still lauded as true validity and reliability. Some researchers have even gone so far to claim that the science is now retrograde, as it lacks the human experience (Dirkx, 2006). In literary studies and criticism, one is taught to use resonance as a method of evaluation; question whether or not the scenario

is believable, whether one is able to resonate and connect with the protagonist (Widdowson, 1983). Reporting numbers lacks learners' stories, and this is consistent with how education and the classroom has changed. The ESL/EFL classroom, at least, has been converted into an assembly line for learning: "the language learner at the center of this system becomes nothing more than an intellectual entity involved in an assessable cognitive process" (Hanauer, 2012, p. 105). The classroom has lost touch with "the flesh and blood individuals who are doing the learning" (Kramersch, 2006, p. 98). The classroom has been dehumanized and replaced with a Western can-do mentality which looks to assess and evaluate and numerate learning and learners:

The experiences, emotions and symbolic transformations inherent in the process of learning a language are erased and superseded with Systemic Functional Linguistics (SFL) teaching by the overriding emphasis placed on the communicative and cognitive aims of language usage, to such an extent that it seems natural to avoid any discussion of the human in the classroom and to emphasize the learning and testing of a decontextualized code. (Hanauer, 2012, p. 106)

Not only is the ESL/EFL learner decontextualized and codified within the classroom, but by the researcher who is the one who is able to influence the very policies which govern and influence learning and the classroom. It is no wonder then that learners' voices are lost in the corporate shuffle.

Semantic Reliability and New Frames of Reference in ESL/EFL

While scholarship has argued to include giving the learners an opportunity to have their own voices and to have meaningful expression (Hanauer, 2012; Iida, 2017) scholarship has not adequately represented meaningful assessment of ESL/EFL students' learning. Scholarship in second language acquisition SLA is not exempt from this growing trend to question the gold standard of evaluation, as previously hot topics of the Social Turn and the Bilingual Turn were mentioned (Bylund et al., 2013; Ortega, 2009). I propose that the field of literature can assist the educational researcher. Literary research and scholarship are different from both human and hard sciences because we, as readers, have no background knowledge, only previous tools of analysis; put another way, when we read a story, we are propelled forwards because we have little or no background knowledge: "Literary schemata are created internally, within the literature itself. They're not projected from the outside" (Widdowson, 1983, p. 30). Literature is criticized from a variety of lenses: but first by which it is *resonant*. Is it believable? Do you believe it? Did it make you feel something? (It is then criticized by its aesthetics and a variety of other things which I won't get into.)

But what makes true literary criticism of use to educational research(ers), is its manner of interrogation of every very word, every representation, every context. While the author, the context, and the protagonist are all considered unreliable, we count their journey and the emotions we as readers are rendered, collated to be emotionally resonant, or not. We analyze protagonists with subjective lenses and layers. In the same way, we could apply this to add variance and give researchers opportunities to explore possible extraneous variables that are addressed in giving students' and teachers' opportunities to voice their experiences while simultaneously taking on the emotional resonant valence that comes with the problems and their experiences. Or we can include these idiosyncrasies in the presentation of research and data – to connect researchers to learners' stories – for data presentation to be more holistic.

Literary critics and literary criticism are unknown to most scholars and misunderstood across disciplines. A literary critic is meant to be objective and to remove their previous bias and aesthetics aside in the same way a scientist is meant to be objective. An example of a literary critic's removal from a text is where every word is semantically rendered variable as in Widdowson's example of the word "breakfast" in conventional discourse:

In conventional discourse you immediately have a set of expectations: If the heading is "breakfast" (for example on a menu), and you find "egg" you won't be surprised to find "bacon," "coffee" etc. because you are immediately engaged with breakfast frame of reference. You will be surprised to find "embryo" coming up. You'd be surprised to find "chicken" perhaps: you don't have chicken for breakfast in your culture. As soon as you engage in conventional discourse, you immediately have a set of expectations, and when you're talking to somebody else or writing to someone else or reading what somebody else has written, you're seeking to get some kind of convergence of these "schemata," these frames of reference. (Widdowson, 1983, p. 30)

Semantic reliability is necessary to be accurate and consistent in presenting data and to negotiate basic meaning. I believe that semantic reliability should be a standard in ESL/EFL. Scientists should not be sloppy with their language especially after going through all the effort to propose their controls for their experiments. Hermeneutic analyses, common in literary analysis, is valuable in ESL/EFL to negotiate meaning with non-native speakers and also to aid them in expressing their stories while also being present to both their struggles to give voice to their experience to access new language and negotiate culture. More pressing is how hermeneutic analysis is necessary for researchers in ESL/EFL: if what researchers say lacks semantic meaning or consistency, then what is the point in statistical control? Moreover, if we choose the right words but then the data has no resonant value, what does it serve that correlations are demonstrated if they cannot be communicated? It is here that I believe that literary resonance is useful to educational researchers so that the readers of their scholarship can connect with their findings. The next section presents examples of learners' stories in ESL.

Students' Stories in ESL: Motivation, Identity; Outliers and the Rest of Us

In the years that I have studied applied linguistics and second language research, I have been formally instructed to rarely trust learners' personal stories because they are not scientifically objective. Informally, I know other phenomena to be true: that success in the FSL classroom does not lead to success in speaking the language. In a personal story, I spent nine formal years studying French in Ontario to be awarded with a scholarship to France only to realize that all of those years of study did not equip me with enough language for any spoken interaction; I was considered an A student. Later still, I did not have enough French to navigate the stigma against Anglophones which remains in some parts of Montreal. I wanted to learn French to be what I considered a "true" Canadian; instead, I was ostracized over poor grammar, faulty vocabulary, and mispronounced words. I still have trouble identifying as a French speaker and with speaking French. On the other hand, when I went to live in Colombia, I learnt Spanish through immersion. My roommates did not speak English and took me out dancing. I learnt how to dance salsa and bachata faster than I learnt to speak Spanish, but these were opportunities to both engage with language and culture. I have never taken a language class and within two years I was presenting in Spanish nationally. I now consider myself a bilingual English-Spanish speaker.

My personal experience I find harrowing as both a language researcher and educator because what value is there in learning language in a classroom if the learners do not ever feel like they ever accomplish anything beyond becoming a human conjugation system? Collectively, we as language researchers are afraid of qualitative measures because how would we assess the language learners' learning gains in an environment in which we cannot control for? It is too messy. Instead, we as researchers have not been able to explain a variety of phenomena, for example, why some people learn languages instantly, and others spend twenty-years studying the language and do not succeed. Instead, we piecemeal their learning a language into different parts – disciplines – such as phonology, phonetics, morphosyntax, semantics; we further condense these disciplines into smaller pieces – variables – that can be measured and tested, which are so small that it divorces the macroscopic person, for example, the acquisition of morphosyntactic “ed” endings, or the phonological acquisition of /i/ (i.e., in the word: beat) and /I/ (i.e., in the word: bit). This assembly process of building a learner's language through grammar negates the value in their stories. Denzin (2017) urges researchers to not just *interpret* the world, but to *change* it, (p. 9, *his emphasis*). I contend that this involves changing how research is conducted in applied linguistics. I believe L2 learners' stories would be helpful, with the inclusion of their voices, their words, and would benefit educational research in second language acquisition. For example, why did Alice Kaplan succeed in learning French, and why did Richard Watson “fail”? They both were successful scholars, who despite not being linguists, wrote books about their experiences learning French. Their autobiographical self-chronicled experience of L2 acquisition of French informs our understandings of education as they are relatable, resonant stories (Ortega, 2009).

Kaplan grew up in Minnesota in the 1960s and claimed to have a “life-long affiliation with French” (Ortega, 2009, p. 147) and interest with the culture: she writes that “learning French was connected to my father” (Ortega, 2009, p. 147). On the other hand, Watson grew up in Iowa in the 1930s and his lack of ability to speak French came to a tipping point at fifty-five when his validity as a Cartesian philosopher, he felt, was at stake by being able to present in French at a conference. He writes: “All those years of guilt and embarrassment at being a Cartesian scholar who could not speak French...I would learn to speak French, whatever it took, however long” (as cited in Ortega, 2009, p. 65). His experience in learning French was anxiety driven.

Did Kaplan's gender aid her acquisition? Did Watson's extrinsic motivation detract from his ability to acquire the language, although he put in more effort? Must language be reduced to measurable, controllable parts? Or does acquisition have to do with the way in which the speakers identify with using French that marks true linguistic and social success? Is the emotional experience and identity of *speaking* French related to *being French*? Or does it have to do with interest in the culture? Did the fact that Kaplan learned French in France on exchanges, and through an undergraduate program in French literature help her gain the language? Does it come down to immersion versus classroom?

Scholarship in SLA has looked to address some of these questions. To take the last question as an example, scholarship currently aligns in looking at the role of the classroom as having negligible impact on the learner. Bardovi-Harlig (2015) claims that that the role of instruction and second language acquisition itself are one and the same: “When I look at the results of instructional investigations, I see second language acquisition itself (p. 66).” To further this point, authorship in this textbook collectively says that the *role of instruction* is “of limited use” (VanPatten & Williams, 2015, p. 30); “ineffective” (Ellis & Wulff, 2015, p. 88), and to go so far in its concluding chapter to claim that “instruction plays no substantial role in SLA,” (Ortega, 2015, p. 262). Other scholars address, as Ullman (2015) does that “implicit, instructed, immersion-like instruction is more effective” (p. 152) for L2 instruction.

Given the current “turns” in SLA, and the research I have presented it is clear that it could be far more useful to start looking at learners’ stories: these stories are *helpful* in sharing a type of validity that is available because Kaplan and Watson’s experiences are *resonant*. It does not matter for purposes of literary criticism if Kaplan or Watson lived (they did) or if they lied (they claim to not have); what is of value here is the fact that their stories inform and can direct L2 acquisition/instruction to try and better understand, say, the links between intrinsic and extrinsic motivation, or perhaps even identity construction and language learning which is a hot topic.

Other stories beyond biographical accounts are often collected through longitudinal studies; an example of this is accent acquisition. Moyer (2013) interrogates whether successful pronunciation instruction is just about learning a system of foreign sounds, and whether or not fluency is actually a precondition for forming deep social bonds. Such as in the example of “Julie” a British woman who married an Egyptian man and never took classes, and she was validated by her speech community to have done so within two years of residence. Julie was able to successfully “fool” native Arabic speakers of being a native speaker. How come, then, that Laura, who has a master’s degree from an American university, who was also married to an Egyptian man and typically passed off as native speaker in her speech, did not pass all the grammar tests, and yet had much more “education”? Where are their stories? Could their stories perhaps inform accent acquisition? I suggest for ESL scholars to question their “ethical responsibility” (Denzin, 2017, p. 9) and to include narrative research as a method in their research practices.

Agreement and equality regarding the shared value and position of qualitative and quantitative measures must be agreed upon across educational research; this movement must also be applied towards ESL. Applying Khagram and Thomas’ (2020) Platinum Standard to ESL would encourage learners’ learning experience to be shared – something currently viewed as less important than experimental methods in ESL research. Educational researchers in ESL currently lack two things: the voices of their subjects, and an ability to adequately analyze the learners’ experiences because there are too many confounding variables which would be marked as extraneous. Instead of discarding extraneous variables, I recommend including narratives as a method of evaluation so that learners’ stories can be heard and that it may provide opportunities for the educational research to resonate with the intended audience. When educational researchers are able to narratively express their data as resonant, they will gain the authority and reliability to influence educational interventions and policy more readily.

I have proposed that resonance be included as a method of evaluation in educational research specifically for second language education. Contemporary scholarship has lost touch with the learner, words have either lost or shifted in their meaning, or have more than one meaning that complicate scholarship. Emotions are now rated by Likert scales instead of the experience being shared through stories. I have discussed the implications of resonance being silenced or marked as extraneous in the evaluation of the L2 learner, and the teachers of those learners in the classroom and in the propagation of educational researcher’s presentation of objective data. I have attempted to make a case in this essay – to *essai*, to try – to convince the reader of this paper for why a literary evaluation of resonance should be considered for us to connect with the learners’ stories, but also to not to deprive them of their voices through ventriloquism or neglect. I hope that this reader leaves this essay with: an understanding for why connection and resonance should be included in a pluralistic presentation of validity in L2 education to authenticate human learning experience; and that resonance should also drive the presentation of scholarship in educational science, not solely statistical control; and that there is value beyond the statistical gold standard and that value can be found in alternative epistemological research such as literary analytical evaluations of resonance. Lastly, I hope this article encourages an urgency to adhere to standards of semantic reliability lest the

objectivity of research and the purpose of sharing the data become confused with sloppy language, ultimately resulting in the failure to express one's research and the added consequences of failing to resonate with its audience.

References

- Altrichter, H. (1986). Visiting two worlds: An excursion into the methodological jungle including an optional evening's entertainment in the Rigour Club. *Cambridge Journal of Education*, 16(2), 131-143.
- Bardovi-Harlig, K. (2015). One functional approach to SLA: The Concept-Oriented approach. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 54-74). Routledge.
- Block, D. (2003). *The social turn in second language acquisition*. Edinburgh University Press.
- Bruner, J. S. (1986). *Actual minds, possible worlds*. Harvard University Press.
- Bylund, E., Hyltenstam, K., & Abrahamsson, N. (2013). Age of acquisition effects or effects of bilingualism in second language ultimate attainment? In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 69-102). John Benjamins.
- Carter, R., & Long, M. N. (1991). *Teaching literature*. Longman.
- Carter, R. (2007). Literature and language teaching 1986-2006: A review. *International Journal of Applied Linguistics*, 17(1), 3-13.
- Dadds, M. (2008). Empathetic validity in practitioner research. *Education Action Research*, 16(2), 279-290.
- Denzin, N. K. (2017). Critical qualitative inquiry. *Qualitative Inquiry*, 23(1), 8-16. <https://doi.org/10.1177/1077800416681864>
- Dirkx, J. M. (2006). Studying the complicated matter of what works: Evidence-based research and the problem of practice. *Adult Education Quarterly*, 56(4), 273-290.
- Ellis, N., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 75-93). Routledge.
- García, O. 2009. Education, multilingualism and translanguaging in the 21st century. In T. Skutnabb-Kangas, R. Phillipson, A. K. Mohanty & M. Panda (Eds.), *Social justice through multilingual education* (pp. 140-158). Multilingual Matters.
- Hammersley, M. (2015, May). *Against 'gold standards' in research: On the problem of assessment criteria* [Paper presentation]. 'Was heißt hier eigentlich "Evidenz"?', Frühjahrstagung 2015 des AK Methoden in der Evaluation Gesellschaft für Evaluation (DeGEval), Fakultät für Sozialwissenschaften, Hochschule für Technik und Wirtschaft des Saarlandes, Saarbrücken, Germany. http://www.degeval.de/fileadmin/users/Arbeitskreise/AK_Methoden/Hammersley_Saarbruecken.pdf
- Hanauer, D. I. (2012). Meaningful literacy: Writing poetry in the language classroom. *Language Teaching*, 45(1), 105-115. [doi:10.1017/S0261444810000522](https://doi.org/10.1017/S0261444810000522)
- Iida, A. (2008). Poetry writing as expressive pedagogy in an EFL context: Identifying possible assessment tools for Haiku poetry in EFL freshman college writing. *Assessing Writing*, 13(3), 171-179. <https://doi.org/10.1016/j.asw.2008.10.001>
- Iida, A. (2010). Developing voice by composing haiku: A social-expressivist approach for teaching haiku writing in EFL contexts. *English Teaching Forum*, 48(1), 28-34.
- Iida, A. (2017). Expressing voice in a foreign language: Multi-writing haiku pedagogy in the EFL context. *TEFLIN Journal*, 28(2), 260-276.
- Stalin, J. 1879-1953 Soviet Dictator. (2016). In S. Ratcliffe (Ed.), *Oxford essential quotations*

- (4th ed.). Oxford University Press.
<https://www.oxfordreference.com/view/10.1093/acref/9780191826719.001.0001/q-oro-ed4-00010383>
- Kim, J., & Duff, P. (2012). The language socialization and identity negotiations of generation 1.5 Korean Canadian university students. *TESL Canada Journal*, 29(6), 81-102.
- Kuru Gönen, S. (2018). Implementing poetry in the language class: A poetry-teaching framework for prospective English language teachers. *Advances in Language and Literary Studies*, 9(5), 28-42. <http://dx.doi.org/10.7575/aiac.all.v.9n.5p.28>
- Langer, S. K. (1953). *Feeling and form*. Scribner.
- Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge University Press.
- Opler, M. E. (1945). Themes as dynamic forces in culture. *American Journal of Sociology*, 51(3), 198-206.
- Ortega, L. (2009). *Understanding second language acquisition*. Routledge.
- Ortega, L. (2015). Second language learning explained? SLA across 10 contemporary theories. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 245-272). Routledge.
- Oxford University Press. (n.d.-b). Reliability, n. In *Oxford English Dictionary*.
- Oxford University Press. (n.d.-c). Resonance, n. In *Oxford English Dictionary*.
- Oxford University Press. (n.d.-c). Validity, n. In *Oxford English Dictionary*.
- Oxford University Press. (n.d.-d). Ventriloquism, n. In *Oxford English Dictionary*.
- Ryan, G., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, 15(1), 85-109.
- Seilman, U., & Larson, F. (1989). Personal resonance to literature: A study of reminders while reading. *Poetics*, 18(1-2), 165-177.
- Spiro, R. J. (1982). Long-term comprehension: Schema-based versus experiential and evaluative understanding. *Poetics*, 11(1), 77-86.
- Ullman, M. (2015). The Declarative/Procedural Model: A neurobiologically motivated theory of first and second language. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 135-158). Routledge.
- VanPatten, B. (2015). Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 113-134). Routledge.
- Wassiliwizky, E., & Menninghaus, W. (2021). Why and how should cognitive science care about aesthetics? *Trends in Cognitive Sciences*, 25(6), 437-449.
- Widdowson, H. G. (1983). Talking shop: H. G. Widdowson on literature and ELT. *ELT Journal*, 37(1), 30-35.

Author Note

Anastasia Boldireff is currently doing her PhD at Concordia University. Ms. Boldireff has been a language teacher for the last nine years. She has presented at multiple conferences in the last five years, where she has been both a panel and plenary speaker on topics ranging from Bilingualism and Biculturalism, Literacy, Writing, Classroom Management, the Communicative Approach, Public Speaking, and Second Language Pronunciation. She received her Honours Bachelor of Arts in English and Creative Writing from York University, Toronto and then went on to obtain her Master of Arts in Literature from the University of Toronto. In addition, she holds her TESL Diploma in ESL teaching and a diploma in Latin-English Translation from the Toronto School of Theology. She has taught in Northern Ireland, Colombia and Canada. She is an academic, athlete, activist, and poet. Please direct

correspondence to anastasia.boldireff@mail.concordia.ca.

Copyright 2021: Anastasia Boldireff and Nova Southeastern University.

Article Citation

Boldireff, A. (2021). Questioning standards of evaluation in educational research: Do educational researchers ventriloquize learners' voices in L2 education? *The Qualitative Report*, 26(6), 1724-1735. <https://doi.org/10.46743/2160-3715/2021.4691>
