Facultad de Ciencias Sociales
Escuela de Psicología

THE EFFECT OF DISCURSIVE AND NON-DISCURSIVE FACTORS ON THE UNDERSTANDING AND AESTHETIC APPRECIATION OF SCIENTIFIC TEXTS

By
JAVIERA PAREDES MAYOR

Thesis presented to the Psychology Department at the Pontificia Universidad Católica de Chile to qualify for the degree of Doctor of Psychology

Thesis Director
Carlos Cornejo Alarcón

Thesis Committee
David Carré[1], Ana Gheorghiu[2], and David Torres[1]

[1] Pontificia Universidad Católica de Chile

[2] University of Portsmouth

January 2022

Santiago de Chile

This thesis has been elaborated according to the three-article format, according to the Pontificia Universidad Católica de Chile Psychology Doctoral Program Rulebook. The present document includes an introduction, theoretical framework, objectives, hypotheses, and a general discussion of the results of the three articles as a whole.

1. Introduction

How we carry out, understand, and communicate scientific work has long been thought of as a logical and rational endeavor by individual scientists, somewhat disconnected from other social factors. A typical image of this is a STEM scientist working alone for long periods inside a laboratory. However, in the last decades, new research from different disciplines has blurred this image, situating the scientist as a person who integrates her subjectivity into her work, interacts within a scientific community, and is integrated into society.

From the social studies of science outlook, the scientist or researcher is embedded in a social system that influences the type of work that she carries out. For example, the social environment influences the configuration and selection of certain types of research projects above others. Social studies of science focus not only on the influence that certain social factors have on scientific work but also on how social dynamics configure scientific work and the researcher's identity. In this sense, research on social factors' impact on science has covered themes from financing research projects to institutional micro-politics (Carré, 2018).

1.1.Social and psychological factors that influence scientific work

Social studies of science have explored the role of variables associated with the social structure of scientific work has on the making of science. For example, a study examined creativity in technicians' work in a paleontology laboratory (Wylie, 2015). Even though technicians are usually described as manual laborers following researcher's lead standardized procedures, technicians depict their work along with artistic and creative elements. In their workdays, technicians choose, apply, and create new ways of preparing fossils in some opportunities. Technicians value art training, the aesthetics of fossils, and the problem-solving aspect of their job.

Wylie (2015) argues that the association of artistic and creative elements in technicians' work questions the usual understanding of this type of job as a low-level, technical one. The key finding of this study is the centrality of creativity in the technicians' work and how this subjective element that goes beyond standardized procedures is crucial to the success of the preparation of fossils. Other studies have explored the impact of professionals' subjectivity on evaluating disciplines such as winemaking. The wine industry in the United States worked as a community to create objective parameters for the description and evaluation of wine in the United States (Shapin, 2016).

Social studies of science have also found that science understanding depends on the personal characteristics of laypeople. For example, a study showed the effect of numeric abilities of laypeople on the perception of climate change messages (Hart, 2013). Participants of lesser numeric ability show greater concern for the victims of climate change and a better disposition for making a donation when the descriptors of the message are modified to numeric data. Instead, people that score high on the numeric ability test do not change their concern or intention to donate if messages are provided through numeric data or not (Hart, 2013). There is also research on the bias in different areas produced by neurophysiological images, the effect of which is currently highly contested. While the original studies provided evidence that pointed to a broader impact of these images on laypeople, later studies prove that this effect is highly complex and contextual to the experimental setting (Baker et al., 2017).

There is also research about the impact of psychological variables of the researcher on scientific work or the personal characteristics of researchers that relate to scientific success (Feist, 2012a; Leahey & Cain, 2013; Simonton, 2003). Psychology of science has addressed the development of researchers within their social and broader context (Feist, 2012b). For example, a study determined that some aspects of people’s personality and the variable need for cognition

influenced the level of interest in and attitudes towards science. Specifically, three Big Five personality variables correlate with an interest in science: openness to experience, introversion, and conscientiousness. Scientific ability in this study is related to individual differences in introversion, openness to experience, conscientiousness, and the need for cognition (Feist, 2012a). Other studies have shown that abilities not related to scientific work influence scientists' research. Hobbies related to visual or musical arts allow scientists to make unexpected connections in their work (Root-Bernstein et al., 1995). Through interviews, the study suggests that scientists are keener to explore unknown or distant scientific problems when they have diverse skills. The integrated action approach allows researchers to make more successful connections than similar peers (Root-Bernstein et al., 1995).

Less attention has been dedicated to the affective elements of the scientist's activities and scientific work itself. This tendency is based on the classic view that emotion interferes with the rational processes of people. This proposal is enhanced by the development of artificial intelligence, where reason is defined as the manipulation of abstract symbols. These symbols do not have any meaning by themselves until they have been united to a reference to actual or possible things (Osbeck et al., 2011). If we consider that rationality and feeling are complementary processes of scientific thought, it would not be strange to think that successful theories or models are evaluated by their logical soundness and beauty or elegance. These aesthetic values are typically accompanied by emotional content. Additionally, it has been proposed that the context of justification of the research process could have an affective component (Thagard, 2002).

Thagard (2002) suggests that one of the reasons for Watson and Crick's conviction of ADN's helicoidal structure was the aesthetic and emotional attractiveness of the final model. Following traditional cognitive models, such beauty would have no place in the justification of the

theory proposed. Affective factors also appear in laboratory practices. In a qualitative study, some descriptions made by researchers regarding scientific protocols and techniques do not fit traditional cognitive categories, mainly because these descriptions have a clear affective or emotional tone (Osbeck et al., 2011). An affective tone is observed in the interview segments that deal with the researcher's experience in the laboratory and the description of laboratory equipment or the instruments that compose it. The phenomenon of the aesthetic and emotional experience of scientific theories within the general parameters of scientific values has been recognized by several theorists. These affective aspects of scientific practice have been primarily discussed in contemporary science studies as autobiographical studies of particular scientists or the analysis of the scientists' personalities (Feist, 2006, 2012a; Osbeck et al., 2011; Simonton, 2003).

Researchers in rhetoric and the history of science have also found that subjectivity permeates the practice and the outcomes of scientific research. Next, I will review some of these disciplines' main ideas and contributions to the present discussion.

1.2.Contributions of rhetoric and history of science to understanding scientific practice

Several authors propose that the person of the researcher is critical in the configuration of scientific practice, outcomes, and how these are articulated in a scientific text. The produced text aims to communicate a particular state of affairs and convince the scientific community that the results presented in the text are accurate and relevant beyond the particular study (Bazerman, 1988; Gross, 2006).

Scientific writing is considered a kind of writing, structured as a unique code compared to other writing styles. The fact that this code is ideally supported by mathematical language makes scientific writing be considered as the most precise one (Bazerman, 1988). One of the main

achievements of this type of writing is that it seems to hide itself. When science is written, it is thought that there is no interpretation at all. This phenomenon occurs because the main idea behind scientific writing is to give an account of the natural world just as it is. Scientific writing would be a faithful representation of the experiment or the observation offered to the reader (Bazerman, 1988).

Nevertheless, the assumed objectivity of scientific prose would be a rhetoric invention carefully created (Gross, 2006). It is an invention because the purpose of the scientific text rests on a non-rational appeal to the authority of reason. The scientific reports are a product of verbal choices designed to capitalize on the most attractive features of a project. These choices embody the myth according to which reason subjugates passions (Gross, 2006). Through varied stylistic techniques, scientific prose creates the feeling that science describes an independent reality, free from any linguistic formulation (Gross, 2006). It is necessary to think that writing, in whatever form, comes up from writers living complex worlds. In this sense, the main takeaway is that language is a situated practice that is sensitive to temporal elements, which configure it as a highly variable system (Bazerman, 1988).

The experimental report as a genre is considered a social reality that helps form discourse in many scientific disciplines. Although writing and reading are activities carried out in private contexts, these activities are social and highly contextualized (Bazerman, 1988). These actions directly speak to their social context and have social aims. In this sense, the social intent of social communities that construct science is to produce a knowledge claim. The project of science is built upon nature reports, and the development of the scientific discourse can be considered the development in which the reports are presented (Bazerman, 1988).

The ideal aim of the empirical report involves articulating language to create a precise, clear, and non-ambiguous symbol system that creates a one-on-one correspondence between objects and symbols. Bazerman (1988) points out that language is a human tool. Therefore, reports of science – like any other written piece – are opaque because they incorporate assumptions about nature and the specific social structure of the scientist's discipline. As Daston & Gallison (2008) suggest, language was created before science, and it is usually inappropriate to express exact measurement or conclusive relations.

Therefore, facts in science would be perceived and represented through an argumentative style actively built by the person who writes. In its beginnings, the scientific discourse was built on the accountability of empirical fact (Bazerman, 1988). The motto of scientific objectivity was precisely “let nature speak for itself” (Daston & Galison, 2008). The scientific article and the representations associated with these publications – images or illustrations – were structured as observations and reports of natural events (Bazerman, 1988; Daston & Galison, 2008). However, through time the scientific article suffered transformations according to the needs and approaches of different scientific communities. As a result, in the 17th century, it became a completely different literary form. Since then, the scientific article compels the author to recognize that she reports the evident truth of natural events and that the article tells a story that can be questioned and has a meaning that can be an object of discussion (Bazerman, 1988).

Accordingly, in the contemporary version of an experimental report, a natural event is not an empirical fact unless the fact is seen and identified as bearing scientific meaning (Bazerman, 1988). Persuasion depends not on the presentation of the actual events as the researcher selected them but on the display of the symbolic representation of these facts in the published report. The objectivity of this symbolic representation is supported in the attempt to eliminate the mediation

played by the observer/researcher between the empirical fact and the reader (Daston & Galison, 2008). Even if the symbolic representation of the facts replaces the direct empirical experience for everyone but the researcher, this representation must appear to be an experience reachable for the audience (Bazerman, 1988).

Since the second half of the 19th century, the scientific facts expressed through language or images were articulated to avoid the temptations from the researcher's subjectivity (Daston & Galison, 2008). These temptations take the form of, for example, interpretations of observations, artistic ability to represent images of observations, or the professional judgment needed to detect and make sense of the natural world. Instead of confronting these factors or making them explicit, the 19th-century researcher transferred the judgment needs to the public. In doing this, the purity of the symbolic representation of an empirical fact was maintained, also recognizing the cost that entails: the role of the public's response, which is the judge regarding the legitimacy of the empirical fact presented to them (Daston & Galison, 2008). Nevertheless, the researchers are particularly interested in persuading readers about the validity and relevance of their work, using the necessary literary elements to achieve this (Bazerman, 1988).

The structure of written scientific work depends on researchers keeping the "purity" of the representations and empirical facts expressed in each report. At the same time, the report contains within its written structure the author's subjectivity that must present and dissuade others about that same object as a way to confirm her theory. Whether text or image, the final product delivered to the broader community represents an empirical fact with some emphases inherent to human labor, which are unavoidable. Contributions from the philosophy of science suggest that all scientific endeavor has affective, aesthetic, and pre-reflexive components. These are part of how scientific judgment works, and it is not possible to ignore it when we talk about the scientist's persona or the products of her work.

1.3. Philosophy of science and the researcher subjectivity

From the philosophy of science, Root-Bernstein (2002) sustains that the scientist's work is similar to the artist's in several aspects, rooted in their motivational source and pre-logical feelings that the author denotes as 'aesthetic cognition.' Identifying scientific problems and their solutions would imply an emotional involvement like the observed in artistic work.

Aesthetic cognition combines knowledge and feeling into an intuition that would be "meta-logical" (Root-Bernstein, 2002). This concept can be characterized as intuition, the base for creative scientific thought. The resolution of research problems or the development of projects is usually solved through non-verbal symbolization, for example, images. The flexibility through which science can be displayed, from mathematical or logical terms to images and diagrams, proves that science is not bounded only to rationality to be articulated or understood (Root-Bernstein, 2002). The author suggests that this pre-logical understanding associated with insight is one of the reasons why scientists value characteristics like elegance. At the same time, the elegance of its presentation allows making multiple connections to other results in the same field (Root-Bernstein, 2002). In sum, the main idea of aesthetic cognition is centered on formulating a brief idea with a significant explanatory power that facilitates understanding of phenomena.

According to Root-Bernstein (2002), each researcher must address how to move from the private sense-feeling that moves their thinking to public languages that they share with others to communicate. Only when the tools of thinking are explicitly distinguished from the tools to communicate can we understand the intimate connection between thought and language or logic and imagination. Researchers often describe that their best discoveries come from combining both factors. Thus, it is necessary to redefine the nature of scientific endeavors not as proofing the validity of specific ideas but rather as proof of the possibility of clear communication of ideas (Root-Bernstein, 2002).

The possibility that the origins of a scientific work come from an intuition based on sensorial and symbolic elements that exceed language implies a new approach to the experience of scientific work (Einstein, 1949; Poincaré, 1946). Although the idea of an aesthetic cognition such as Root-Bernstein's (2002) proposal considers contemporary evidence on the issue, his proposal of a metalogic as an answer to the researcher's capacity of transferring intuition into a conceptual articulation seems to be less articulated. Mainly, his model puts on the same level discursive and non-discursive symbolization. If both types of symbolization work at the same level, the scientist could and would "translate" these structures to communicate her findings to others. The author indicates that this translation is a poor substitute compared with the researcher's understanding of the phenomenon or the answer that she achieved. If the translation through the meta-logical procedure proposed by Root-Bernstein (2002) is carried out exactly as is described, Thagard's (2000) account regarding Watson and Crick's conviction on the helicoidal nature of DNA structure would not be described as it is. How did these researchers persuade the scientific community of their proposal's truthfulness and elegance?

The issue raised by Root-Bernstein's aesthetic cognition is that if non-discursive symbols are equated to discursive symbols (*viz.*, language), an alternative psychological system of human thought becomes necessary. This new system would need to model semantic representations that, to be equated to non-discursive elements of thought, are not built upon on computational or logical rules in the standard sense proposed by current psycholinguistics. Therefore, the suggested equivalence between discursive and non-discursive symbolism would constitute a contradiction when giving an experiential account of cognition's nature and affect in any vital sphere (Shanon, 2013). If translation between the non-discursive and the discursive is so poor, how can any researcher achieve their goal of persuasion and explaining their ideas? Why is the "translation" of

the insights attained by the researcher articulated by other means along with the content of their research (e.g., rhetoric, elegance)?

The present thesis explores non-discursive elements in the scientist's work, specifically in their writings. Also, the thesis enquires into the effect that non-discursive and discursive elements have on the appreciation of scientific texts. Integrating proposals from philosophy, psychology, and experimental aesthetics, I propose that a scientific text is structured in such a way that merges rational argumentation with non-discursive elements. These features aid the comprehension and appreciation of the texts. The presentational would be articulated in discursive and non-discursive factors allowing an understanding of scientific knowledge and its contents slightly closer as thought by the scientists themselves. If the hypothesis is confirmed, it will contribute to a better understanding of the impact of aesthetic and affective aspects upon the development and communication of science. Furthermore, the empirical confirmation of this phenomenon would allow identifying which discursive and non-discursive elements are involved in the understanding, appreciation, or acceptance of the texts and the ideas presented in them. The evidence could eventually show why some scientific texts communicate better or effectively expand a particular research area.

In what follows, I will review theoretical and empirical developments regarding non-discursive symbolic elements, the main conventions of scientific writing, and developments in experimental aesthetics.

2. Theoretical framework

2.1.Presentational symbolization

While human experience is usually articulated by language and propositional thought, people inhabit the world in ways that go beyond discursive symbolism. Polanyi (1962) proposed that the focus of the symbolic articulation in human beings is not centered on language but conceptualization. By articulating symbols, conceptualization would form a reasonable concept for the person's experiences. The reasonableness of thought does not necessarily mean that this concept is rational (or logical). Scientific discovery, for example, allows the researcher to get direct access to the experience of new knowledge without exerting control over this with any pre-established interpretation modes (Polanyi, 1962). If we consider that human thought is structured beyond the logical-propositional symbolism, the human experience is enlarged since language is no longer the only possible semantic of knowledge (Langer, 1942; Langer, 1970). The aim of representing the process of knowing in terms of formal inductive logic would only describe in a pale likeness the original objective, which is to give an account of the human experience of knowledge (Polanyi, 1962).

Another consequence of the traditional definition of conceptualization is the relegation of feeling, defined as any human experience that discursive thought cannot express. Feeling, in this framework, is considered to be a symptom of inner human life, without any intellective or representative function (Langer, 1942; Stern, 2010; Vygotsky, 2008). Both Langer (1942, 1970) and Polanyi (1962) argue that other symbolic forms exist, and they belong to the human experience composed of affective and intellectual elements. These affective and discursive experiences do not necessarily fit into the propositional schemes of language because they are organized differently and are what Langer (1942, 1970) denominates as a *presentational symbol*.

The presentational symbol has been previously described as intuition, feeling, or emotion (Langer, 1942). However, one of its main attributes is centered on articulating emotional, perceptual, and cognitive elements of the human experience (Langer, 1942, 1970). In contrast to discursive symbolization, the presentational occurs empathetically and is sensitive to the person that perceives it (Polanyi, 1962). The meaning associated with these symbols can be renewed each time the person experiments it, and it is generally affective. In the presentational symbol, what is expressed is conveyed through its form and displays a frame of mind and its tensions (Langer, 1970).

In Langer's proposal (1970), presentational symbols are articulated as significant forms. The concept of significant form implies that this symbolization process can be understood and felt outside of the primary field of language. Therefore, the presentational can express ambivalence in content that is not common with discursive symbolization (Langer, 1942, 1970; Stern, 2010). The appreciation and comprehension of these symbols are different from discursive symbols because they are configured in the association between meaning and form, which tends to be described as a 'flash' of understanding (Langer, 1942, 1970).

Polanyi (1962) denominated the phenomena of presentational understanding as learning, either in presentational or discursive contexts. This understanding is not entirely objectified knowledge, but it is part of a physiognomic appreciation, as discussed below (Polanyi, 1962).

2.1.1. Presentational symbolization and its place on consciousness

The meaning associated with the presentational symbol usually unfolds before the person consciously knows what it is seeing or perceiving (Langer, 1970). Congruent with Langer's proposal, recent evidence shows mappings between form and sound (i.e., bouba-kiki effect) can

ensue automatically before achieving full consciousness of the visual form (Hung et al., 2017). The bouba-kiki effect refers to the tendency of people to pair linguistic sounds with particular forms. These forms and sounds do not have a fixed meaning, and forms tend to be completely abstract. The study results show that the preconscious mapping of congruent sensory modalities (e.g., "kiki" belonging to an angular shape) lowered the threshold of conscious perception of the stimuli compared to the perception threshold when they showed incongruent modalities. This phenomenon keeps appearing even when participants were trained to pair different letters to the auditory forms of the word. Therefore, the researchers conclude that the effect is based on phonology and not on the visual characteristics of the letters (Hung et al., 2017). These results support the proposal that some perceptive acts are being shaped before conscious thought or before we have any control of them, and they directly affect our life experiences.

Langer (1942, 1970) and Polanyi (1962) also propose that consciousness is articulated in a focal-subsidiary contrast. A subsidiary model of consciousness means that human experience is not focused on objects but on understanding. Therefore, consciousness and pre-consciousness are phenomena articulated to allow, for example, an account of the human experience of the symbolism in art (Langer, 1970; Vygotsky, 2008). A proposal like the subsidiary model of consciousness has a great affinity with the model of the thematic field of consciousness by Aaron Gurwitsch. Gurwitsch (2009) suggests that an object in the perceptual act or any other act of consciousness offers itself to whom experiences it in a specific and particular way. Then, to think (*noesis*), the person needs to experience the act of making the object conscious from a fixed point of emergence (e.g., memory or perception), and it is also necessary to have the object in view. This object in experience appears under a particular light, as the object-as-it-is (Gurwitsch, 2009).

When thought is articulated, the object that appears to consciousness contains references to all the sides of the object that are not given to immediate perception. These are different elements from which the object offers itself, modes of presentation, and appearance that update themselves after being considered by the act of consciousness. Therefore, facts that are not visible can be perceived, according to Gurwitsch (2009), by way of indication or implication. Perception without vision is achieved through references contained in what is perceived to those facts. However, no mode of presentation in consciousness is ever complete in terms of what in mind is seen. Every act of consciousness – perception, memory, or others – contains in this proposal more than it is given to the direct and immediate view.

In a subsidiary model of consciousness, there is nothing essentially conceptual or explicit about knowledge, and neither is the result of reasoning capabilities or conclusions about the object. The object is presented as an intrinsic element in the act of consciousness. As the act of consciousness and its object are articulated through *Gestalten*, which implies a radical modification of the perceptual act, this allows the model to alter in any manner the facts implicated around the perceptual object (Gurwitsch, 2009). The anticipation of tactile, visual, and other experiences that could be had of the implicated object are confirmed or invalidated in this process. Eventually, the undetermined references are transformed into determined ones.

What is given in a specific act of consciousness about an object, through direct intuition, indwells in a halo of implications and references to the object's aspects that are not perceived (Gurwitsch, 2009). There are potential aspects that could be shown to different modes of appearing of those acts which are being actualized. Implications are acquired through experience, and this acquisition is achieved through gestalt structures or specific perceptual schemas that are applied to each perceptual act until they stabilize through repeated experiences. The stabilization of

implications does not mean that memory is activated each time an object and its implications are perceived. In this model, the stabilization does not significantly contribute to constructing meaning in actual perception.

Therefore, the perception of an object, perceived or used, is embedded with know-how. This know-how does not compare to remembering how to do something but instead refers to managing an act. Gurwitsch (2009) suggests that the "reiterative memory" concept can help to understand this process (Bergson, 1959). Reiterative memory refers to the fact that while experience is used and contributes, it is not evoked or represented in the current act of perception. This concept also echoes Polanyi's (1962) idea of personal knowledge. Personal knowledge is articulated in action, but it cannot be represented or put in discursive terms. Following Gurwitsch (2009), implications could belong within the personal knowledge repertoire proposed by Polanyi (1962). When experience is framed as a reorganization of the perceptual function and restructuration of perceived objects, it is possible to understand how the past can influence the acts of conscience in the present without appearing in front of them (Gurwitsch, 2009).

The environment in which acts of consciousness are located is named by Gurwitsch (2009) the *thematic field*, defined as a relevance dominion. At the same time, relevance is defined as when two objects or themes co-present in the field of consciousness are considered "intrinsically related" given their content (Yoshimi & Vinson, 2015). Therefore, a theme can only be given when located in the center of a thematic field.

This framework has been empirically tested in some experiments that confirm that objects in the periphery of consciousness affect the mode of presentation of the object on the attentional focus (Proffitt et al., 2003; Williams & Bargh, 2008). Thus, the thematic field – as the objects of perception – seems never to be wholly articulated. The incompleteness of the thematic field is

conceptualized by Gurwitsch (2009) as fragmentation. Recent developments in this framework also suggest that the field of consciousness possesses a focus of varying size in which attention is centered. This focus would be surrounded by a peripheral content structure that is not fixed by the attentional process (Yoshimi & Vinson, 2015).

However, no matter how indeterminate a thematic field is, it always possesses a specific character, a distinctive coloring that allows distinguishing a thought trajectory from others (Gurwitsch, 2009). This particular character enables us to understand why the theme is not entirely indifferent to the changes in the thematic field. The changes allow perceiving the theme in a different light, hence allowing the theme to appear from a different perspective.

If presentational symbols are thought to be articulated within a subsidiary account of consciousness, the person can indwell in it and display its effects in the perceptual, cognitive, and emotional aspects (Gurwitsch, 2009; Polanyi, 1962). The presentational is configured in the totality of experience, which speaks of how the person experiences themselves and the world (Langer, 1970). The relationships created or discovered in this space are displayed inside structures composed of statements that do not necessarily denote tangible objects. Nevertheless, these structures refer to the pre-conceptual and contemplative capabilities of the human being (Polanyi, 1962). Rather than comprehending a proposition, something new is discovered compared to the previous lived experiences (Langer, 1942, 1970). As a subsidiary articulation, the presentational appears in our attention related to elements that are present in the thematic field or interwoven in the focal object of the perceptual action as possible factors of non-discursive nature.

The subsidiary localization in the thematic field in the consciousness of the presentational could define the nature of the presented theme or object. It is a way to differentiate, for example, between a poetic and a musical piece. In the poem, discursive, expressive, and presentational

elements are articulated within written language, with the expressive elements farther away from the focus of consciousness. In the first approximation, the focal point is centered in the text, and the expressive or physiognomic characteristics add depth but may not be the central part of the experience. Instead, the musical piece is completely articulated through expressive media, allowing the presentational elements to come before the perceptual action. The greater salience of expression in music is afforded because there are no explicit discursive elements. Therefore, the focus of consciousness is centered on experiencing the presentational aspects in their entirety.

2.1.2. Characterization of presentational symbols

Presentational symbolic activity begets its own knowledge of experience and feelings, especially when it comes to phenomena like tension, expectation, and the feeling of rightness that closes a process of thought (Langer, 1970; Polanyi, 1962). The experience of the perceptual revelation of the presentational symbol allows the person an implicit understanding of these concepts, configuring a new type of knowledge. The framework of symbolic operations is articulated through all symbolic processes of human beings, even though they cannot articulate fully through spoken or written language (Polanyi, 1962).

2.1.2.1.Presentational symbolization as a Gestalt articulation

The presentational has also been characterized as a Gestalt. One characteristic of this Gestalt is that the symbol is immediately captured by perception and its understanding is not dependent on discursive formalization (Langer, 1942, 1970; Stern, 2010). A piece of artwork, for example, is part of a presentational symbolization of an idea. The articulation of very subtle aspects of human experience is commonly labeled as “what one cannot speak about” (Wittgenstein, 1961). The

concepts exposed in the presentational are displayed as a whole, global and indivisible form of understanding.

Vygotsky (2008) proposes that in the research process of artistic works, the methodological approximation to the artwork must not destroy the Gestalt of the phenomenon. This argument is also suggested by other authors (Langer, 1970; Polanyi, 1962; Vygotsky, 2008) because it emphasizes that, to be able to have a comprehensive view of the phenomenon, we must study not only the object on the attentional focus of the perceiver but most importantly, what is in the subsidiary consciousness of perception. That is to say; the lived experience must be considered in its totality. The characterization of the presentational symbol as a Gestalt and as a significant form is part of the same process. When the symbol is presented as a totality, the person directly understands without discursive forms.

2.1.2.2.Presentational symbolization as feeling

Presentational symbolization takes physical form as feeling. Langer (1942, 1970) suggests that feeling does not reflect a state of things, as is the case with the content of propositions, but it is closer to being defined as a process. The feeling-as-process is opposed to the classical approximations of psychology, which regard feeling as a perceptive act closely related or equivalent to emotion (Vygotsky, 2008). In a presentational framework, feeling cannot be objectified or conceptualized from a proposition since an approach through language would stop the process which has created that feeling (Langer, 1970; Polanyi, 1962). This interruption of symbolization also happens when people are asked to reflect on a presentational symbol discursively. The symbolic articulation that allows the experience only occurs when the person is indwelling the symbolism.

In a reflexive state, the person goes through indwelling instead of being guided by it (Polanyi, 1962). Considering the proposal of Gurwitsch (2009), the reflection would involve moving symbolic elements from a subsidiary to a focal point of consciousness, an action that could be achieved with different levels of success. Given these difficulties for studying the presentational symbol, it is assumed that logical conviction is very different the non-discursive, even if they are two processes considered essential to the human being (Langer, 1970). Reason and feeling are regarded as opposite and incompatible pairs. Langer (1970) suggests that logic is part of human thought, but it is not a way of thinking, and therefore, logic and feeling can be part of human experience.

In everyday experience, it is not necessary to define the nature of feeling, more than can be expressed through language, voice, physiognomy, or gesture. "Once the abstraction is made by the eye or the ear, and spontaneously received as something with feeling, it tends to govern the impulse of every stroke or uttered sound or bodily motion" (Langer, 1970, p. 213). However, Langer (1942, 1970) suggests that we need to learn more about the feeling-as-process to know the mind. The dynamic forms of experienced human feeling are primary examples of mental life's source, integration, and rhythms. Feeling as a personal process is characterized then as something constant and systematic that reflects what is going on with a person in levels below the sentient ones. It is a process that happens all over in the organism, and that uses perception and the nervous system.

2.1.2.3.Presentational symbolization as subjectivation of form

The appearance of feeling as a direct way to intuition provokes the sensation of perception of feeling in the person. It is not the perceived object that holds feeling but the presence of a subtle form congruent with the vital experience of the perceiver. Expression through this vital form conveys a sense of reasonableness, a tacit explanation of its display. The abstract recognition of

vital experiences similar to feeling arises from comparing previous experiences with actual ones. Symbolic projection and interpretation of these experiences are spontaneous answers that the person can make conscious or cultivate the ability to make them conscious (Langer, 1970).

Although the presentational, and therefore, feeling, goes beyond complete discursive articulation, this does not mean that this process is ineffable. Some characteristics of this process that several authors have identified are its physiognomic character and its perception as a totality (Langer, 1970; Polanyi, 1962). When the perceptive act is given to the person, the physiognomic character is so predominant that it overcomes the structural qualities of the symbol (Werner & Kaplan, 1964). Feeling is usually the most visible attribute when a complete gestalt-like object is perceived. Furthermore, the articulation process of presentational symbols is located within the matrix of acts of the person, deploying itself in different levels of complexity, from preconscious activity to behavior or speech (Langer, 1970; Polanyi, 1962).

Werner & Kaplan's studies (1964) are consistent with the idea of presentational symbols. Their work adds to previous theory, emphasizing that the dynamics and subjectivation of forms are displayed in activities such as art and discursive activities, i.e., language (Langer, 1970; Werner & Kaplan, 1964). Many reports of participants of these studies point out the intimate relationship between physiognomization of certain words-forms and specific posture-corporal organizations. These studies also added evidence in terms of the embodied character of this symbolization process (Werner & Kaplan, 1964).

The physiognomization of language supposes a return to a mode of symbolization where the internal dynamics of words are shown in the material forms of the discursive symbols. The process of "denaturalization" of the connection between the referent and meaning in language does not mean that these symbols at an advanced level are completely objective phenomena, divorced

from the formative activity of who symbolizes or listens. Werner & Kaplan (1964) suggest that even the most conventional forms of speech – words and phrases – are still an integral part of an articulatory process, a postural corporal activity in which, through its dynamic characteristics, connects these conventional units to their referents.

Given this characterization of language, people can easily give physiognomic characteristics to conventional forms, bringing to the front this activity. In any other way, the authors argue that these typical conventional forms would lose their meaning (Werner & Kaplan, 1964). In regular speech, several stages of physiognomization coexist, from physiognomic patterns to conventional forms. These articulations depend on the nature of the communicative situation and the type of recipient. The conventional forms tend to be selected and shaped with a higher externalization of its dynamic forms, for example, in poetic discourse than in the communication of an impersonal statement of facts.

The physiognomization of words differs from the physiognomization of other non-conventional patterns because there is a restriction imposed on who symbolizes. The liberty of giving form to the pattern – potential vehicle – for the meaning and referent to fit the properties of the vehicle is limited because the relationships of the phonemic configurations to their referents are in large part, already preset in conventional language, and they must be followed (Werner & Kaplan, 1964). This phenomenon happens in the everyday language since articulations such as poetry, the phonemic configurations with their referents are usually in a less rigid relationship, which allows novel forms of this same language.

In contemporary research, a study shows that one of the significant characteristics of the early organismic organization of verbal forms is the apprehension of a general sphere of meaning previous to a more specific determination (Aldunate et al., 2009). In this case, this would recognize

the link between word-meaning (Werner, 1956). If a sound has a particular expressive value, this depends on the dynamic structuration of the word as a whole. In some cases, sound patterns from a technical-geometrical point of view are different, through physiognomization could be apprehended as similar ones (Werner & Kaplan, 1964).

The dynamics that articulate the physiognomy of language are based on modulated properties of the sounds and also in visual, auditory, tactile, and kinesthetic qualities that accompany the declaration or understanding of the word. Werner & Kaplan (1964) highlight that physiognomization of conventional linguistic forms is no different from the process of physiognomization of on non-verbal media for the "natural" representation of meaning. However, this process is different in the case of non-verbal media, such as, for example, expressive line patterns. While expressive lines are not within a sign system whose peculiarity and stability must be maintained, the speech forms must participate in such a system. This proposal of physiognomization of speech forms can be made equivalent in other grammatical categories such as verbs or adjectives.

One of the main differences between the physiognomy of language and non-verbal symbols is that the majority of the non-verbal symbols are formed by not more than two vehicles. Furthermore, while sometimes there are references to an agent, action, or object in the description of these symbols, there is no discrete articulation about the reference in the lineal representation level. Instead, many non-verbal symbols are composed of only one relatively undifferentiated vehicle (Werner & Kaplan, 1964).

In interactive settings, the greater the interpersonal distance between individuals involved in a communicational situation, the more autonomous the symbolic vehicles are to be understood. However, Werner & Kaplan (1964) do not think that consensus regarding the meaning of verbal

symbols must require connotation identities between the diverse participants of this communicational situation. The only fundamental requirement is that those evoked connotations in both individuals are in a comparable position in each person's individual network of meanings. That means both people share a social world, network, or shared meanings.

Although it is not possible to ignore the complexity of the study of the essential conditions or factors that affect the aesthetic response, the reviewed proposals (Gurwitsch, 2009; Langer, 1970; Polanyi, 1962; Werner & Kaplan, 1964) throw light on the meaning of the aesthetic response. Not only should we care about the unity between the aesthetic experience at the individual level and the social function that this could have, but also it is necessary to ensure that the different individual levels –reader, community, and researcher in the present thesis – are not in contradiction. Considering the antecedents presented until now, the study of presentational elements would not only be delimited to the artistic domain – though being the most 'pure' presentation – but also could extend into other expressions of human creativity like scientific texts.

2.2.Approaching the scientific text from a presentational perspective

Considering the previous review is possible to generate a proposal to understand how a researcher articulates her intuitions, questions, and answers from non-discursive elements to discursive media shareable with others. Although the communicative act is aimed to deliver new knowledge, other presentational components have a role in the text's understanding and appreciation (Langer, 1970; Polanyi, 1962; Root-Bernstein, 2002; Werner & Kaplan, 1964).

The dynamic articulation of the presentational in scientific texts can be apprehended through the text's structure, focusing on the physiognomic aspects or the organismic level to which the text alludes to and echoes in the reader (Langer, 1970; Werner & Kaplan, 1964). The

presentational dynamics in science are especially relevant to study given the research and development of Werner and Kaplan's (1964) theory about physiognomization of language and non-verbal forms.

Considering the dynamic articulation of formal language, it is essential to notice that some aspects of writing work as gestures. Sometimes people choose a book because of its cover, an attractive print, or the illustrated text in its interior. Marcuschi (2005) argues that it is a mistake to think of textual formats as informational code without considering all aspects articulated as non-discursive elements in this thesis. These aspects produce meaning effects, and as a secondary medium, would be equivalent to the para-linguistic effects of orality (Marcuschi, 2005). The contextualization given by these non-discursive elements is necessary for the production and reception of the texts. In other words, the articulation of the presentational is needed for human speech to work fully.

An important question is the role of the writer and reader's speech in the construction of the scientific text. For example, we need to consider which aspects of language are prioritized in scientific practices. Discursive or non-discursive forms allow us to get closer to the discovery experience, such as the researcher and lay public went through it. These elements are essential to know how a good scientific text is articulated.

If these presentational elements are thought of as an evident phenomenon, independent of other textual elements – such as metaphor – then it is possible to isolate the scientific text from the writer's subjectivity. However, if the description of the presentational is correct (Langer, 1970), it makes impossible to think of language or scientific formats as free of expressive elements or human subjectivity.

Current research in science communication and impression making gives us some clues for the path to come. A study indicates that first impressions of scientists' faces have an effect on the perception of credibility and the quality of science of the research presented in magazine articles (Gheorghiu et al., 2017). Emerging evidence on similar phenomena makes more research necessary to discover best practices regarding scientific communication and its understanding.

Presentational symbolism is usually outside of the focus of consciousness of persons (Gurwitsch, 2009; Langer, 1970; Yoshimi & Vinson, 2015). In most cases, the presentational is unfolded in the thematic field of consciousness of whom is in front of the text. Consequently, the research on presentational aspects of science should focus on variables that are not necessarily conscious but rather the subsidiary elements that support the text's appreciation and comprehension.

If the study of scientific texts is addressed beyond the explicit content of the text, a new window for research is opened. The study of non-discursive elements of scientific texts must be considered when analyzing conventions in academic writing. Every discipline has standards that orientate authors on creating a good text. Since researchers are guided by manuals and rules from journals and publishers, it is interesting to review how these conventions in academic writing handle extra-discursive elements of texts. On the other hand, experimental aesthetics, even though it does not have the same study object of this thesis, has developed diverse methodologies to explore aesthetic appreciation. Aesthetics is relevant when scientists present their developments, results, and scientific modeling (Root-Bernstein, 2002; Thagard, 2002). Next, I will briefly review the main rules and conventions of academic writing and a state of the art of research in experimental aesthetics.

2.3.Some conventions on the writing of scientific texts

Every scientific text is outlined following some basic rules usually stipulated in writing manuals, in both natural and social sciences. In this section, some of the main conventions in social and natural science writing used by researchers will be described. I will later argue that these conventions also have a role in understanding and appreciating the scientific text.

One of the first requests for a scientific text in most manuals and articles is the use of clear language. This rule implies using language that avoids twofold connotations and aims for accurate and non-emotional writing (Katz, 2006). This directive is similar to the recommendation for researchers to write as clearly as possible. To write clearly means avoiding the style used in prose writing and preferring a plainer style. For example, it is recommended to prevent colloquial expressions that may muddle the meaning of the discussion. Also, it is recommended to avoid specific jargon specific to the researcher's field, comparisons that may be confusing to the reader, or other linguistic devices, such as poetic expressions or metaphors (American Psychological Association, 2020; Day & Gastel, 2011; Katz, 2006). The use of figurative expressions is not recommended since it risks generating confusion or, through using these elements, forcing something that does not sound good in a scientific text (American Psychological Association, 2020; Day & Gastel, 2011).

The use of numbers is also considered a feature of scientific writing. Katz (2006) states that quantifiable words are natural to science, and therefore all facts related to research should be able to be measurable in numbers. Thus, critical adjectives in research should be transformed into indicators expressed through numbers. In the case of variables related to subjective experiences, such as pain, the researcher must be able to define how painful is the event that the participant reports (Katz, 2006, p. X). Also, the reader should be capable of confirming the fundamental

analyses and constructing some estimators of effect size beyond the elements provided by the text (American Psychological Association, 2020; ICMJE, 2016).

Concerning the grammar of the text, the aim is to create a plain text devoid of ambiguity that does not obstruct the communication between writer and reader. One of the grammar characteristics in scientific texts that have changed through time is the use of active and passive voices. Nowadays, several manuals recommend using the active voice in writing instead of the passive voice (American Psychological Association, 2020; Day & Gastel, 2011; Katz, 2006). The passive voice is acceptable when the writer aims to create an expository text or when the writer wishes to focus on the research object or the receptor of the action. These elements are generally described in the methods section of the scientific text (American Psychological Association, 2020). While this is the current consensus of some of the most cited manuals of style, there is still debate within the scientific community regarding using these voices in writing. The advocates of the passive voice emphasize the objectivity and clarity that this tool gives them, while the advocates for the active voice argue, among several other reasons, for a greater level of agency and transparency in writing (Bostian, 1983; Leather, 1996; Ruben, 2012). The use of these voices is also linked to the need to align the sentence's subject. That is, it is necessary to shorten the distance between the subject and object to its maximum, to be able to address one idea in a phrase (American Psychological Association, 2020; Gopen & Swan, 1990). The grammar recommendations are not only restricted to the use of voice and regulate the use of pronouns, adverbs, and other rules associated with the sentence construction in the text (American Psychological Association, 2020; Katz, 2006).

These are some of the stylistic conventions that present greater consensus in academic writing manuals. Although these rules do not exhaust the entirety of conventions mentioned in the

manuals, numbers and grammar are the most visible characteristics and recommendations of academic writing in general. These characteristics are part of the formal organization of scientific texts, but they are not part of the content that the researcher wishes to communicate.

Stylistic conventions are discursive elements in the thematic field of consciousness when the person approaches the text (Gurwitsch, 2009). Therefore, the modification of any of such characteristics would not modify the meaning of the text, but it could alter the clarity of the exposition of the ideas, as some manuals point out (American Psychological Association, 2020; Day & Gastel, 2011; Katz, 2006). Even if they are read as facts in manuals, these recommendations have not been empirically proven. The journal *Nature*, in its writing recommendations, states: "Nature journals prefer authors to write in the active voice ("we performed the experiment...") as *experience* [emphasis added] has shown that readers find concepts and results to be conveyed more clearly if written directly." (*Writing for a Nature Journal*, 2017). These are formal modifications that have a reported effect on the appreciation and understanding of the text, and they need to be studied to transform the experience into empirical evidence.

2.4. Aesthetic appreciation of a scientific text: Current research

Given the current state of the conceptual approaches in empirical aesthetics, the next section will go over a broad corpus of empirical studies on aesthetic experiences, organized through three themes: studies on aesthetic reactions towards varied stimuli; studies that explore the relationship between structural elements of artworks; and finally, the role of contextual variables in the aesthetic experience.

2.4.1. Aesthetic reactions towards diverse stimuli

Most of the studies in this area focus on determining phenomena associated with people's reactions to art stimuli. A descriptive study carried out by Gabrielsson & Lindström (2003)

explored strong experiences in music with Swedish participants. The main categories that arose from the participant's answers were general characteristics, reactions and physical behaviors, perception, cognition, emotions/feelings, existential elements, and personal or social elements. Some of these elements stand out, especially the "double" character of strong experiences in music. This feature allows the experiences regarding music to be classified into two different conventional psychological categories. An example of this phenomenon is the "loss of control" category that was related to "emotions," "cognition and emotions," "new possibilities, insight or necessities," and "positive emotions." Also, in the category "general characteristics" and in the factorial analysis carried out, the researchers found some descriptions related to the inability to put the musical experience in words. This data is relevant since the researchers report that some participants refused to answer the questionnaire as the nature of the experience was untranslatable to them into a closed answer (Gabrielsson & Lindström, 2003).

Other studies explored the possible correspondence between neural mechanisms and phenomenological reports connected to the experience of watching dance (Jola et al., 2012). The results show that empathic abilities in participants heighten cortical excitability for everyday actions that mimic hand movements. In contrast, the visual experience amplifies the same cortical excitability but for formal, stylized movements, such as ballet movements. Concerning the experience of watching dance, participants that reported more enjoyment through the dance interpretation also reported increased levels of emphatic and kinesthetic involvement (Jola et al., 2012; Reason & Reynolds, 2010). Another study explored the coupling effects in spectators that perceive a non-rhythmic and prolonged dance performance. The results suggest a positive relationship between the psychological coupling with the participant's attention to the breathing (of the dancers or themselves). There is also a positive relationship between cognitive coupling,

conceived as the underestimation of the performance's time, and the attention to the breathing and muscular activity of the dancers (Bachrach et al., 2015).

Grosbras, Tan & Pollick (2012), using offline transcranial magnetic stimulation (TMS) in regions involved in the subjective emotional evaluation, explored the causal relationship between cerebral activity and participants' emotional reactions when showed a long segment of dance. The control condition's cerebral activity records through an fMRI negatively correlate with subjective emotional judgment and activity in the right posterior parietal cortex. This cerebral region is known for its involvement in cognitive tasks. Application of the offline TMS in the parietal area of participants' brains in the experimental condition did not affect the general affective response. However, it enhanced the ratings of the moments that produced the more favorable judgments of the dance segment (Grosbras et al., 2012). Additional research has explored the positive relation between neurophysiological reactions (i.e., chills) and the feeling of being moved by artwork (Wassiliwizky et al., 2015) and the relationship between the activation of areas of the brain related to social cognition, predictive inferences regarding the emotional experience of suspense (Lehne et al., 2015). Other studies have also explored the description of peak moments associated with aesthetic appreciation defined by neurophysiological reactions. These reactions appear in greater frequency when the stimuli are related to the display of pro-social behavior (Wassiliwizky et al., 2017).

Schindler et al. (2017) developed an Aesthetic Emotions Scale (AESTHEMOS). Its theoretical framework proposes that previous attempts to measure aesthetic appreciation are not valid enough since they use bipolar dimensions, like valence or arousal. Aesthetic experiences, on the contrary, would usually involve a mix of positive and negative valences and can be experimented as relaxing or activating events. From a twenty-four-theoretical category proposal,

an exploratory factorial analysis retained twenty-one categories. These categories were organized into a broader five groups: prototypical aesthetic emotions, pleasing emotions, epistemic emotions, and negative emotions (Schindler et al., 2017). While a newer development, the AESTHEMOS scale is one of the few scales aiming to measure aesthetic experience beyond artistic stimuli. The construction process involved an exhaustive literature review that added evidence from various disciplines working with aesthetic stimuli, including the arts, marketing, and environmental aesthetics.

2.4.2. Structural elements of artworks and aesthetic reaction

A study explored metrics and rhyme's effect on aesthetic liking, emotional involvement, and positive valence attributions. When the metric and rhyme of the stanza were modified to ensure regularity in both of them, the participants reported greater aesthetic appreciation, processing intensity, and more positive emotions (Obermeier et al., 2013). Similar studies have explored the neurophysiological correlates of the effects of rhyme and metrics (Obermeier et al., 2016). Some results suggest that the stanzas structured with metrics and rhyme produce lower N400/P600 responses than those not structured with these characteristics. The researchers propose that the effects of the rhyme and metrics for N400/P600 show a facilitation process for poetry, and these effects are positively correlated with aesthetic pleasing (Obermeier et al., 2016). Other studies identify that tonal contrast in poetic fragments is associated with the emotional classification of the fragments (Kraxenberger & Menninghaus, 2016a). Also, there is evidence that words with alliteration, assonance, and consonance are identified as happier poetry. In the same study, the words that possessed a stress peak were classified as more emotional when this peak was in more advanced positions in the text (Kraxenberger & Menninghaus, 2017).

2.4.3. The role of contextual variables in aesthetic appreciation.

An experimental study tried to dissociate the effect of the authenticity of the artwork and the physical context of the experience of art (Brieber et al., 2015). The artwork appreciation experience was measured through scales related to liking, interest, valence, and comprehension. Using a within-subject model 2x2 with authenticity and context as independent variables, the researchers concluded that neither the authenticity of the artwork nor the context had an effect on the evaluation of paintings by the participants. A hypothesis for these findings is that the lack of personal involvement and meaning of the artwork added to the context provided to the disparity of results compared with previous studies (Brieber et al., 2015).

The review concerning contemporary empirical aesthetics allows us to understand that aesthetic appreciation comprises several different aesthetic-psychological phenomena. Most research is focused on the effects that artwork has upon aesthetic reactions. Here, the results show the mixture of emotional and cognitive responses in different organismic levels when reacting to an artwork (Gabrielsson & Lindström, 2003; M. H. Grosbras et al., 2012). This mixture has important implications regarding the present doctoral thesis. These conclusions corroborate the idea that perception, comprehension, and appreciation of human work – be this poetry, an artwork, or a scientific text – are mutually articulated between discursive and non-discursive aspects in different proportions (Langer, 1970; Polanyi, 1962; Werner & Kaplan, 1964). Other research in line with the presentational and the articulation of these elements, whether vital or artistic forms show that appreciation and emotional responses are modified when the metric is manipulated (Kraxenberger & Menninghaus, 2016a, 2016b). These results allow us to tentatively propose that similar modifications of these studies could modify the perception, appreciation, and comprehension of any text, including scientific ones. The conceptual contributions on the

subsidiary/thematic field of consciousness (Gurwitsch, 2009; Langer, 1970; Polanyi, 1962) could be manipulated experimentally to capture the delicate presentational articulation and its role outside of the focus of conscience.

Considering all the contributions previously discussed on the theoretical framework, the research question of the present thesis is this: Is there an influence of non-discursive variables in the comprehension and appreciation of a scientific text?

Hypotheses

1. The modification of formal discursive elements will improve the comprehension and appreciation of scientific texts.
    a. The modification of the grammatical person towards the first person will improve the appreciation and comprehension of scientific texts.
    b. The addition of numerical data will improve the appreciation and comprehension of scientific texts.
2. The modification of the non-discursive elements of the texts will improve the comprehension and appreciation of scientific texts.
    a. The modification of titles through the addition of cliché content will improve the comprehension and appreciation of scientific texts.
    b. The modification of the format of the text to a simpler format will improve the comprehension and appreciation of scientific texts.

Objectives

4.1. General objective

To explore the effect of the modification of discursive and non-discursive variables upon the appreciation and comprehension of scientific texts.

4.2. Specific objectives

4.2.1. To determine the effect of the modification of formal discursive elements upon the appreciation and comprehension of the scientific text.

4.2.2. To determine the effect of the modification of non-discursive elements upon the appreciation and comprehension of the scientific text.

5. Article's Outline

In the next section, I will present three articles that show the results of my thesis work. A summary of each article is included, specifying the thesis objective accomplished through each paper.

Article 1 – Effects of discursive presentational variables on the comprehension of scientific texts

The article aimed to determine whether discursive elements in scientific texts influence the reader's comprehension. In an experimental design, the effects were elicited by manipulating the grammatical person and the statistical data of an article's abstract. The results suggest that the experimental manipulations did modify the perception of comprehension of the abstract. The absence of statistical data in abstracts belonging to the natural sciences rose the perceived understanding of the text. Conversely, the addition of statistical data increased the perception of comprehension when in social sciences participants. Nevertheless, the experimental modifications did not affect more complex comprehension measures, such as ratings of comprehension by external judges. The differential impact of the manipulation on the different comprehension measures points out disparities between perceived expertise and skills that could lead to a greater depth of processing in some readers. This article partially answers the first specific objective of the present doctoral thesis.

The article has been submitted to Public Understanding of Science and it is currently under review.

Article 2 – Understanding through content and form: The role of presentational variables on the perception of scientific texts.

In this article, the purpose was to determine whether non-discursive elements in a scientific text influence the comprehension of its content. In an experimental design, the variables manipulated were the style of the title and format of presentation of the text, in this case, an article's abstract. The study results show that the experimental manipulations modified the perception of comprehension of the abstract. Specifically, the inclusion of clichéd titles increased the perception of comprehension in participants. The presentation of the text to a simpler one increased the perception of comprehension reported in the survey. The combination of these presentational characteristics likewise allowed for improved perception of understanding among the study participants. Nevertheless, the experimental modifications did not impact more complex measures, such as ratings of comprehension by external judges. These differences between comprehension measures could indicate differences between readers' levels of expertise and skills that could lead to a greater depth of comprehension in some cases. This article answers partially to the second specific objective of this thesis.

The article has been submitted to Psychological Reports and it is currently under review.

Article 3 – The challenges of measuring aesthetic experiences

This article aimed to reveal some of the challenges associated with developing quantitative measures of aesthetic experiences. After reviewing thirteen articles, we identified three main challenges: a conceptual articulation of aesthetic experiences, the determination of observables, and the examination and documentation of instrument quality. We identified opportunities to define a more robust concept through clarification and differentiation to similar notions regarding the first challenge. Also, there is a need for a more significant link between the attributes of interest

articulated in the theoretical frameworks and the proposed observations intended to gather information about them to determine observables. The link is necessary to create instruments that more accurately characterize the theoretical proposals of the researchers. Finally, the reports of the analyses about the validity of the scales must provide greater clarity regarding the design choices made in the development of the measure. We suggest some avenues of improvement that could lead to a clearer and better-supported interpretation of the scores of these measures. This article partially answers the first and second specific objectives of this thesis.

Regarding the effect of discursive and non-discursive elements on the aesthetic appreciation of scientific texts, these results will be integrated into the discussion section of the present document since an article presenting those results is currently in development.

6. Articles

Effects of discursive presentational variables on the comprehension of scientific texts

Javiera Paredes, David Torres & Carlos Cornejo

Escuela de Psicología, Pontificia Universidad Católica de Chile

Corresponding author

Javiera Paredes, Escuela de Psicología, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile.

Email: jnparede@uc.cl

Abstract

Recently, there has been interest in the effect that presentational (i.e., non-content-related) characteristics of media has on science communication and the role that it plays on understanding. The aim of this study was to determine whether discursive presentational aspects of scientific texts affect comprehension of science content, specifically in texts. The effects were elicited by manipulating the grammatical person and statistical data in which an abstract was written in an experimental design. The results suggest that the experimental manipulations in fact modified the perception of comprehension of the abstract. More specifically, the absence of statistical data in abstracts belonging to the natural sciences rose the perceived understanding of the text. In social sciences, the presence of statistical data rose the perception of understanding reported by participants. However, the modifications did not have an effect on more complex measures of comprehension. The differential effect on measures of comprehension could point out to

differences between levels of expertise and skills that could lead to greater depth of processing in some readers.

Keywords

Comprehension, presentational, discursive, scientific texts, abstracts

Introduction

Scientific texts aim to make statements on nature. But they also look to persuade the scientific community – and the lay public that the assertions presented are true (Bazerman, 1988; Finkler and Leon, 2019; Gross, 2006). Studies have shown that researchers use various tools and criteria beyond accuracy to evaluate these theories (de Regt, 2004; de Regt and Dieks, 2005). Some of these criteria or tools rely on aesthetic cannons determined by the scientific community to which the researcher belongs to assess the information presented to them (de Regt, 2004; McAllister, 1999). These criteria or skills of evaluation allow the scientist to understand and gauge if the theory is true. Scientific understanding then is based not only on the person's content or knowledge on the topic but also the skills that the person has to use this knowledge correctly (de Regt, 2004; de Regt and Dieks, 2005). Therefore, the efficacy of persuasion on understanding science depends not only on the findings presented by the text but also on its symbolic representation and how this elements are used by the reader (Gheorghiu et al., 2017; Hartley, 2003; Kendeou and van den Broek, 2007).

Recently, there has been an increased interest in the effect that presentational, i.e., non-content-related, characteristics of media has on science communication and its understanding. These components have been named in various ways. Some researchers have labeled it non-verbal or peripheral cues (Michalovich and Hershkovitz, 2020; Tal and Wansink, 2016), factors that influence perception (Jamali and Nikzad, 2011), or elements that affect the perception of

comprehensibility of a scientific text (Crick and Hartling, 2015; Garritty et al., 2020), among others. Empirical research has found that fMRI images impact the layperson's ability to judge the accuracy of the neuroscience presented to them (McCabe and Castel, 2008); and there is also evidence that pre-existing highlighting interferes with performance in comprehension when reading textbooks (Silvers and Kreiner, 1997). Yet, the role of presentational elements in scientific communication is far from being fully understood.

Recent empirical evidence shows that presentational elements of discursive nature have a significant effect on the comprehension of scientific texts. For example, Hartley (2003) demonstrated that section-structured abstracts were rated as more clear, informative, and readable than the original abstracts written by academic authors. Similarly, a study exploring the understanding of astronomical pictures by experts and non-experts concluded that variations in type of explanation affected the comprehension of the imagery presented (Smith et al., 2011). Specifically, in non-experts, engaging narrative language was linked to increased understanding of the image accompanying the text.

However, the persuasive role of scientific texts and the importance that the symbolic presentation of the results has on understanding is at odds with several recommendations made by manuals of style in diverse disciplines. Manuals of this kind usually recommend writing practices favoring the use of a language as clear and transparent as possible (Bazerman, 1988; Robbins, 2010). Clear language means avoiding colloquial language, jargon, confusing comparisons, and other linguistic devices, such as poetic expressions or metaphors (American Psychological Association, 2020; Day and Gastel, 2011; Katz, 2006). The fundamental principle behind this is to: "use precise words and, whenever possible, numbers." (Katz, 2006: X).

Following the principle stated above, the idea that research has to be expressed as clearly as possible has, in some cases, moved article writing from description in words to description in numbers. The idea that numbers convey objectivity that words cannot achieve would allow the researchers to reach greater precision and clarity in their writing (Vieira, 2011). Thus manuals of style recommend including descriptive statistics, as they are especially helpful to give the reader a feel of the data presented in the article (Katz, 2006). A consequence of these recommendations is the tendency to convert critical descriptors in research that are of qualitative nature into indicators that can be expressed by numbers. Then, it is the reader who confirms with the information given – statistical analysis, equations, or numerical results – the accuracy of the analyses performed by the researchers (American Psychological Association, 2020). Therefore, considering the high value given in some style manuals to the description of results in numeric form, statistics could play an essential role as a persuasive element compared to the same information provided in an exclusively discursive form.

Another feature in writing that has changed in recent times is the grammatical person in which the text is written. Contrary to previous versions of manuals (e.g., American Psychological Association, 1957) and some style guides geared to students (e.g., McMillan & Weyers, 2007; White, 2000), most contemporary style manuals suggest the use of the first person in detriment to the third person (American Medical Association, 2007; American Psychological Association, 2020; Day and Gastel, 2011; Freeling et al., 2019). Third-person use on an article would be acceptable only when the text's goal is descriptive or needs to focus on the study object or the recipient of an action, which typically describes the article's method section (American Psychological Association, 2020). These non-content-related features of a scientific text, as well as others such as the use of English as the lingua franca of scientific publication (Egger et al., 1997;

Meneghini et al., 2008) or the use of specific pronouns or adverbs (American Psychological Association, 2020; Katz, 2006), could impact on the understanding and perception of the scientific article or text presented to readers.

Marcuschi (2005) suggests that it is inaccurate to think about textual formats as propositional content without considering all its multidimensional aspects, such as the presentational aspects of writing. These presentational aspects of scientific texts create meaning effects and would be equivalent to paralinguistic elements of speech. However, empirical studies that demonstrate that the effect of presentational characteristics of a scientific text on its comprehensibility are scarce. Furthermore, hypotheses tested in previous studies (e.g., Hartley, 2003; Smith et al., 2011) have mainly focused on experts (i.e., scientists, academics). In sum, there is a lack of studies that study how non-content-related features of scientific texts affect their understanding.

The present study aimed to determine whether and how presentational aspects of scientific texts (i.e., grammatical person, presence of statistical data) affect comprehension. For this purpose, we conducted a randomized experimental design using images of the article's abstracts on computer screens, just like an article would be displayed if the participant would have searched for it. Thus, the research question for this study was: Do grammatical person and statistical data in the abstract of a scientific publication affect its comprehension?

Method

The study aimed to explore the effects of discursive modifications on the comprehension of abstracts of scientific texts. The effects were elicited by manipulating the grammatical person in

which the abstract was written and the addition/deletion of statistical data in the abstract of the article presented in an experimental 2 x 2 design.

Participants

The study's participants were 330 undergraduates of the Pontificia Universidad Católica de Chile (age: M = 21.12, SD = 3.19; 179 women). The study aimed to answer its questions in two broad disciplines: natural and social sciences. Therefore, the authors constructed stimuli for the two areas. One hundred and sixty-five students belonging to the area of the natural sciences participated in the study. The participants voluntarily signed up for participating and were allowed to sign up in a raffle for a gift card as compensation for their participation. Ethical approval was granted by the Ethical Committee of Social Sciences at the Pontificia Universidad Católica de Chile following the Declaration of Helsinki. Participants gave informed consent and, since the study was carried out online, were given a choice to download the informed consent from the platform.

Data from 19 participants were excluded from further analyses: 2 of them were under 18 years old; 3 participants registered themselves in a study area that did not correspond to their major; and 14 participants completed the questionnaire at a much faster rate (i.e., two standard deviations from the average time, M = 2.14, SD = 0.24) than the typical participant, suggesting random answering.

Materials

The study was generated from two abstracts of articles belonging to the natural sciences (Delaney et al., 2018; Lagabrielle et al., 2018) and two others belonging to the social sciences and humanities (McDevitt et al., 2018; Street et al., 2018) from a well-known journal. All abstracts in

the same area were selected to have the same baseline on the variables modified and were experimentally changed following the same guidelines (grammatical person and addition/deletion of statistical data).

All stimuli were created through a 962 × 660-pixel screenshot PDF of the original article. The experimental manipulations were made with the Snagit 2018 software that allows the edition of PDF texts. Because most browsers do not display PDF files directly on their web pages, PNG versions of the stimuli were created to better show the stimuli on the platform. After some interviews with undergraduate students before the study's launch, it was clear that a Spanish version of all abstracts was necessary. Therefore, the researchers carried out the Spanish translation of the texts with the assistance of experts in each area. The translated abstracts were added to the stimulus page through a link (i.e., "Para versión en español, click aquí"/"For the Spanish version, click here").

The natural sciences abstracts were modified to be read in first or third person neuter, and the statistical data included in the abstract was kept or deleted and transformed to additional text keeping in the same style and information of the original for the science abstracts. In the case of the social sciences abstracts, the abstract was modified to be read in first or third person neuter, and the statistical data was included in the abstract by adding data from the body of the article to the result sentences in the abstract (See more in Appendix).

In Table 1 below, there is a summary of the experimental conditions and their characteristics.

Table 1. Summary of experimental conditions

			Baseline	Grammatical person	Presence/absence of statistics	Interaction between variables
Natural Sciences	**Lagabrielle et al. (2018)** **Delaney et al. (2018)**	**Grammatical person**	First	Third	First	Third
		Statistics	Yes	Yes	No	No
Social Sciences and Humanities	**Street et al. (2018)** **McDevitt et al. (2018)**	**Grammatical person**	First	Third	First	Third
		Statistics	No	No	Yes	Yes

Each version of the study displayed two abstracts to participants randomized to be assigned to different conditions and shown as rotating blocks to avoid the primacy effect (Glenberg et al., 1980). Block order, stimulus, and participants' assignment to conditions were randomized through the Sosci Survey software. The assignment of the experimental stimuli was counterbalanced.

Procedure

The invitation to the study sent to participants consisted of a brief description of the research and an invitation to sign in with their email address to be sent the link to the survey. Students signed up and the email with the invitation that contained a link to the study in the SoSci platform (http://soscisurvey.de) was sent the same day. A reminder email was sent after four days of the initial invitation.

Once in the platform, participants were asked for their consent to take part in the study. As a first step, participants were asked to fill in some demographic information about themselves, such as gender, age, area of undergraduate study, and level of advancement of undergraduate studies. This section was placed in the survey to allow the platform to assign the students in the natural science or social science/humanities version of the study. Next, an instruction page was added to guide the participants. The command stated: "On the next page, you will find an abstract from a scientific paper. It may take a few seconds to appear. Please read it as many times as you want to.

Underneath the abstract, you will find a link to the Spanish translation if needed. To start, click on the Next button". Afterward, the first abstract was presented on a sole page with a black background and the Spanish translation link.

The following pages in the questionnaire after the display of any abstract inquired about the abstract's perception of comprehension. Checks regarding perceived expertise on the article's subject and language on which the abstracts were added after the questionnaires.

The perceived comprehension of the abstract was assessed first. Participants judged on a five-point scale (1= "Very hard to understand," 5 = "Very easy to understand") the question "What did you think about the abstract that you just read?". Additionally, participants answered an open question worded: "What do you think is the main idea or discovery that the authors wished to communicate through the abstract?". This item allowed for as many answers as the participant wanted to provide regarding each abstract.

Considering the evidence that shows that the skill reading scientific research literature varies depending on the stage of the academic career of the reader (Hubbard and Dunbar, 2017), the last section of the study added a question to measure the self-perception of skill regarding the abstract presented in that particular block (i.e., "How much of an expert do you think you are regarding the article's subject?)". This question was answered on a five-point scale (1 = Not at all, 5 = Very much). Finally, an item on the language in which the abstract was read was added (i.e., "In which language did you read the abstract?"). This question was added to check the possibility that the language in which the abstract was read interfered with the comprehension or appreciation of the abstract.

Closing the study, there was a form to allow the participants to sign themselves up for a raffle as compensation for their participation. After this, they were thanked for their contribution in the exit section of the questionnaire.

Analysis

The analysis of the survey data was carried out using a generalized linear model, controlling for non-independence of the observations linked to the two abstracts read by each participant. The model was run with the Mplus software (Muthén & Muthén, 2017 version 8.4). In the model, the non-independence of the observations – abstracts – was corrected using the TYPE = COMPLEX option of the analysis command in conjunction with the CLUSTER option in the variable command (Muthén & Muthén, 2017). The cluster variable used in this analysis was the case variable that grouped the abstracts read by each participant in the study. In this regression, the distribution is centered in men in the social sciences & humanities area, in the second abstract in participants with a low perception of expertise that read the abstract in Spanish in its more accessible version.

The regression was run with the item "Perception of comprehension" as the dependent variable. The predictor variables were: the study area of participants (sciences/social sciences or humanities), the abstract presented to the participant, the experimental condition within each abstract, the participant's gender, their perception of expertise regarding the subject of the abstract, the language on which the abstract is written and the Flesch reading ease level.

For the analysis of the open-ended question regarding the comprehension of the abstract (i.e., "What do you think that is the main idea or discovery that the authors wished to communicate through the abstract?"), the answers were evaluated by part of the Detailed Rubric for Reading Assessment (Northern Essex Community College, n.d.). The evaluation was based on the rubric

section dedicated to assessing reading comprehension (from 1= Novice to 4 = Internalizing Highest Level), which was rated by a member of the research team and later was triangulated with the author. Also, all of these open-ended answers were coded concerning which of the three main ideas of each abstract were present on the answer. These three ideas in each abstract were related to the study's aims, their main results, and the conclusions and recommendations given by the researchers.

Results

Table 1 below presents the regression coefficients, standard errors for each parameter, and p-values for the regression model on the perception of comprehension.

Table 1. Regression coefficients, standard errors for the regression, and p-values of the item "Perception of comprehension" using as predictor variables the study area of participants, the abstract presented to the participant, the experimental condition within each abstract, the participant's gender, their perception of expertise regarding the subject of each abstract, language of the abstract and the Flesch reading ease level.

	Perception of comprehension		
	Coefficient	S.E.	p-value
Study area	0.006	0.087	0.948
Abstract	0.385	0.117	0.001
Grammatical person	0.110	0.117	0.345
Statistical data	0.428	0.132	0.001
Grammatical person x statistical data	0.518	0.131	0.000
Gender	0.035	0.087	0.690
Expertise perception	0.509	0.099	0.000
Language	-0.090	0.093	0.331

Flesch reading ease	0.366	0.115	0.019

There are significant differences in the scores of the perception of comprehension depending on the following variables: the abstract presented, the presence or absence of statistical data, the interaction between the variables grammatical person and presence/absence of statistical data, the perception of their own expertise and the Flesch reading ease of the study conditions. There are significant differences in the perception of comprehension scores reliant on the abstract shown to participants (R = 0.385, DE = 0.117, 95% CI [0.156, 0.614]). These results indicate that when the Delaney et al. (2018) and the McDevitt et al. (2018) abstracts were presented to participants, the reported perceived comprehension of the abstract rose by 0.385 compared to the other abstracts (Lagabrielle et al., 2018; Street et al., 2018). Significant statistical differences also appear in the perception of comprehension when the variable absence/presence of statistics in the abstract is considered (R = 0.428, DE = 0.132, 95% CI [0.169, 0.687]). Therefore, the absence of statistical data in the science abstract rose the scores of perceived comprehension by 0.43. In contrast, statistical data in the abstract of the social sciences and humanities rose the scores of perceived comprehension of the abstract by 0.43.

The interaction between the variables grammatical person and the presence/absence of statistical data in the abstract also shows significant differences in the scores of the perception of comprehension (R = 0.518, DE = 0.131, 95% CI [0.261, 0.774]). The perception of comprehension rises by 0.52 when the abstract is written in the third person neuter and statistical data is absent in the natural science abstracts. For the social sciences, the perception of comprehension of the abstract rises 0.52 when the abstract is written in the third person neuter and statistical data is present. Significant differences in the perception of comprehension are also present when we

consider as a covariate the participant's self-perceived expertise on the abstract's subject (R = 0.509, DE = 0.099, 95% CI [0.315, 0.703]). When a participant considers herself knowledgeable about the abstract's subject matter, the scores of the perceived comprehension rise by 0.51. Finally, the perception of comprehension scores is also significantly different when we add the Flesch reading level as a covariate in the regression (R = 0.366, DE = 0.132, 95% CI [0.141, 0.591]). As the Flesch reading level of the abstract rises, meaning a more accessible reading, the perceived comprehension of the abstract rises 0.37.

Regarding the analysis of the open-ended questions of the abstracts, in Table 2 below, we present the averages of the rubric scores assessing reading comprehension for each condition in both study areas.

Table 2. Averages the rubric scores assessing reading comprehension for each study condition in the abstracts of natural sciences and social sciences, and humanities.

Natural Sciences		Social Sciences and Humanities	
Condition	Average	Condition	Average
Baseline	2.24	Baseline	2.46
Grammatical person	2.25	Grammatical person	2.28
Presence/absence of statistics	2.28	Presence/absence of statistics	2.34
Gramatical person x statistical data	2.14	Gramatical person x statistical data	2.32

The table shows that the scores do not show a significant difference between comprehension scores evaluated by the rubric (F (7,614) = 0.951, p = 0.466). Regarding the frequencies of the main ideas of the text given by participants in both abstracts in its study area, we present Table 3.

Table 3. Abstract's open-ended questions of comprehension frequencies for the main ideas in the text according to each experimental condition in the area of the natural and social sciences.

	Baseline	Grammatical person	Presence/Absence of statistical data	Grammatical person x statistical data
No main ideas present	50	48	35	54
Idea 1: Aims of the study	64	59	65	67
Idea 2: Results of the study	69	62	58	48
Idea 3: Conclusions and recommendations	63	38	47	55

In the frequencies of the abstracts, the responses related to each main idea presented in the texts are distributed more or less evenly across experimental conditions. A frequency could be of interest, related to the increased amount of responses related to the study results compared to other ideas in the grammatical person condition. Another point of data that could be singled out is the augmented frequency of responses focused on the aims of the study in the interaction condition that mixed the grammatical person and the statistical data

Discussion

Our goal in this study was to explore if presentational cues of scientific texts in the form of stylistic alterations had an effect on the understanding of the text, whether in the perception of the understanding of participants as well as on understanding rated by an external judge. We additionally explored the possible impact of abstract effects, the self-perceived expertise, and the readability of the abstract on the same dependent variables.

An important finding of the study shows that the absence (in the natural science) or the presence (in social science or humanities) of statistical data in the abstracts led to higher ratings of the perception of the understanding of the text by participants. This result would point to the diverse levels of importance that some areas give to statistical data as part of a scientific text's standard presentation. In the natural sciences, such non-content related elements would not be strictly necessary to generate an adequate perception of the understanding of the abstract. On the contrary, the presentation of statistical data for the social sciences and humanities participants seems to be an essential factor in their perception of the understanding of the text. Consequently, social sciences or humanities participants would place greater importance on quantitative indicators as a sign of an intelligible text. This view is consistently supported by some manuals in scientific writing that argue that statistical descriptors grant a precision and clarity that words are not able to (Katz, 2006; Vieira, 2011). Thus, the belief in the power of numbers and statistics as presentational features of scientific texts could be why the scores of the perception of understanding are higher in social sciences and humanities students.

Another finding of the study is that the interaction between the presentation of the abstract in third-person and the absence/presence of statistical data in the presented abstracts also improves the scores of perceived understanding of the abstract. Although the experimental manipulation of exchanging the grammatical person from the first to the third person did not yield statistically significant differences by itself, it does generate significant differences between scores of perceived understanding when paired with the other manipulation. In this case, for natural science training participants, the interaction between the third person and the absence of statistical data rose the scores of perceived understanding of the text. On the other hand, for social sciences and humanities training participants, the interaction between the third person and the presence of statistical data

also rose the scores of perceived understanding of the abstract. An explanation for these results could be found in the discrepancy between the recommendations in style manuals of particular disciplines and style manuals geared to university students. Since the sample of this study comprised undergraduates, the advice that they are more likely to recall is still rooted in the belief that third-person allows an "objective" approach to research and its results (McMillan and Weyers, 2007; Robbins, 2010; White, 2000). Considering this approach in writing styles, in combination with the belief that statistical descriptors give greater clarity in regards to the information presented in the abstract (Katz, 2006; Vieira, 2011), we observed that overall, this sample favored a more conservative stance towards the abstract's presentational features, especially in participants of social sciences and humanities.

Considering the analysis of the open-ended questions, an important issue raised by the results of this study is the disparity of the effects of the experimental manipulation on the scores of the perception of understanding compared to the scores of understanding rated by an external judge. While we detect differences in the scores of the perceived understanding of the abstracts by the participants, these results do not replicate in their open-ended answers when rated for understanding by an external judge. These results are akin to other reports made by Werner da Rosa & Otero's (2018) study, where the experimental manipulation consisted of modifying the document's source shown to secondary students. Here, the students rated the comprehensibility of the paragraphs of different sources. The scores of comprehensibility rose when the texts were supposedly taken from science textbooks compared to peer reports. However, there was no effect of the manipulation on other measures of the study, such as reporting comprehension obstacles that could point to a more in-depth processing and understanding of reading. Altogether, these results point out the differential effect that presentational elements in scientific texts have on the

perception of understanding and the understanding rated by external judges or other measures or in-depth processing in samples of participants in formative stages in their education. Understanding a scientific text does not only implies the assessment of the accuracy or correct recall of the information presented but also entails the use of particular skills and criteria established by specific communities that allow the person to manipulate knowledge correctly (de Regt, 2004; de Regt and Dieks, 2005). It could be hypothesized that in participants of these characteristics, the perception of the understanding of scientific texts can be manipulated through these presentational elements, but these elements are not criteria that participants use to achieve greater or differential understanding in other measures.

Other results need more exploration in further research. One of them is the strong effect that the perception of expertise had on the perception of the understanding of the participants. While there is evidence that natural science-training students in different career stages value sections of articles differently and that their reading skill develop throughout their career (Hubbard and Dunbar, 2017), lesser is known about the reasons why a greater perception of expertise in the discipline would also raise the scores of perception of understanding. One of them could point to the possible relationship between the perception of expertise and the belief of participants that they are able to use their academic skills to approach and understand the text. Also, differences between abstracts were not expected. Specifically, there are significant differences by abstract in scores of the perception of the understanding of the texts. These effects occur across disciplines, and they cannot be accounted for primacy/recency effects since the presentation of the abstracts was counterbalanced. Since the abstracts were authentic published texts, there are some differences between them regarding their overall style and writing, which could cause these differences.

There are also some limitations to this study. First, all participants in this study had undergraduate university training. Further studies should explore the effect of these non-content-related elements in a more diverse sample. Finally, while the modifications shown in this article affected the perception of the understanding of the abstracts, the experimental manipulations were very subtle and directly interwoven in the content of the text. Research should explore the effects of bolder presentational modifications to the texts to understand further the impact of presentational elements in the context of scientific publication and communication.

References

American Medical Association (2007) *American Medical Association Manual of Style*. New York: Oxford University Press.

American Psychological Association (1957) *Publication Manual of the American Psychological Association*. Washington, DC: The American Psychological Association INC. Available at: https://archive.org/details/isbn_9789562912662.

American Psychological Association (2020) *Publication Manual of the American Psychological Association*. 7th ed. Washington, DC: American Psychological Association.

Bazerman C (1988) *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science.* Madison: The University of Winsconsin Press.

Crick K and Hartling L (2015) Preferences of Knowledge Users for Two Formats of Summarizing Results from Systematic Reviews: Infographics and Critical Appraisals. *PLOS ONE* Scherer RW (ed.) 10(10): e0140029. DOI: 10.1371/journal.pone.0140029.

Day RA and Gastel B (2011) *How to Write and Publish a Scientific Paper*. Santa Barbara: California: Greenwood.

de Regt HW (2004) Discussion Note: Making Sense of Understanding*. *Philosophy of Science* 71(1): 98–109. DOI: 10.1086/381415.

de Regt HW and Dieks D (2005) A Contextual Approach to Scientific Understanding. *Synthese* 144(1): 137–170. DOI: 10.1007/s11229-005-5000-4.

Delaney KZ, Vanstone CA, Weiler HA, et al. (2018) Regional adiposity and markers of inflammation in pre-school age children. *Scientific Reports* 8(1): 15204. DOI: 10.1038/s41598-018-33054-1.

Egger M, Zellweger-Zähner T, Schneider M, et al. (1997) Language bias in randomised controlled trials published in English and German. *Lancet* 350(9074): 326–329. DOI: 10.1016/S0140-6736(97)02419-7.

Finkler W and Leon B (2019) The power of storytelling and video: a visual rhetoric for science communication. *Journal of Science Communication* 18(05). DOI: 10.22323/2.18050202.

Freeling B, Doubleday ZA and Connell SD (2019) Opinion: How can we boost the impact of publications? Try better writing. *Proceedings of the National Academy of Sciences* 116(2): 341–343. DOI: 10.1073/pnas.1819937116.

Garritty C, Hersi M, Hamel C, et al. (2020) Assessing the format and content of journal published and non-journal published rapid review reports: A comparative study. *PLOS ONE* Useche SA (ed.) 15(8): e0238025. DOI: 10.1371/journal.pone.0238025.

Gheorghiu AI, Callan MJ and Skylark WJ (2017) Facial appearance affects science communication. *Proceedings of the National Academy of Sciences* 114(23): 5970–5975. DOI: 10.1073/pnas.1620542114.

Gross A (2006) *Starring the Text: The Place of Rhetoric in Science Studies*. SIU Press.

Hartley J (2003) Improving the clarity of journal abstracts in psychology: The case for structure. *Science Communication* 24(3): 366–379. DOI: 10.1177/1075547002250301.

Hubbard KE and Dunbar SD (2017) Perceptions of scientific research literature and strategies for reading papers depend on academic career stage. *PLOS ONE* Ito E (ed.) 12(12): e0189753. DOI: 10.1371/journal.pone.0189753.

Jamali HR and Nikzad M (2011) Article title type and its relation with the number of downloads and citations. *Scientometrics* 88(2): 653–661. DOI: 10.1007/s11192-011-0412-z.

Katz MJ (2006) *From Research to Manuscript: A Guide to Scientific Writing*. DOI: 10.1007/1-4020-4071-7.

Kendeou P and van den Broek P (2007) The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition* 35(7): 1567–1577. DOI: 10.3758/BF03193491.

Lagabrielle E, Allibert A, Kiszka JJ, et al. (2018) Environmental and anthropogenic factors affecting the increasing occurrence of shark-human interactions around a fast-developing Indian Ocean island. *Scientific Reports* 8(1). DOI: 10.1038/s41598-018-21553-0.

Marcuschi LA (2005) A oralidade no contexto dos usos lingüísticos: caracterizando a fala. In: Marcuschi LA and Paiva Dionisio A (eds) *Fala e Escrita*. Belo Horizonte: Autentica, pp. 57–84.

McAllister JW (1999) *Beauty and Revolution in Science*. Cornell University Press.

McCabe DP and Castel AD (2008) Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition* 107(1): 343–352. DOI: 10.1016/j.cognition.2007.07.017.

McDevitt EA, Sattari N, Duggan KA, et al. (2018) The impact of frequent napping and nap practice on sleep-dependent memory in humans. *Scientific Reports* 8(1): 15053. DOI: 10.1038/s41598-018-33209-0.

McMillan K and Weyers J (2007) *How to Write Essays and Assignments*. Pearson Education.

Meneghini R, Packer AL and Nassi-Calò L (2008) Articles by Latin American authors in prestigious journals have fewer citations. *PLoS ONE* 3(11): 1–4. DOI: 10.1371/journal.pone.0003804.

Michalovich A and Hershkovitz A (2020) Assessing YouTube science news' credibility: The impact of web-search on the role of video, source, and user attributes. *Public Understanding of Science* 29(4): 376–391. DOI: 10.1177/0963662520905466.

Northern Essex Community College (n.d.) Detailed Rubric for Reading Assessmnent. Available at: https://facstaff.necc.mass.edu.

Robbins S (2010) Ethics and aesthetics of style: The import of stylistic choices in academic prose. *Filologija* 15. Available at:

http://www.su.lt/bylos/mokslo_leidiniai/filologija/2010_15/robbins.pdf (accessed 13 October 2020).

Silvers VL and Kreiner DS (1997) The effects of pre-existing inappropriate highlighting on reading comprehension. *Reading Research and Instruction* 36(3): 217–223. DOI: 10.1080/19388079709558240.

Smith LF, Smith JK, Arcand KK, et al. (2011) Aesthetics and Astronomy: Studying the Public's Perception and Understanding of Imagery From Space. *Science Communication* 33(2): 201–238. DOI: 10.1177/1075547010379579.

Street SE, Morgan TJH, Thornton A, et al. (2018) Human mate-choice copying is domain-general social learning. *Scientific Reports* 8(1): 1715. DOI: 10.1038/s41598-018-19770-8.

Tal A and Wansink B (2016) Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science* 25(1): 117–125. DOI: 10.1177/0963662514549688.

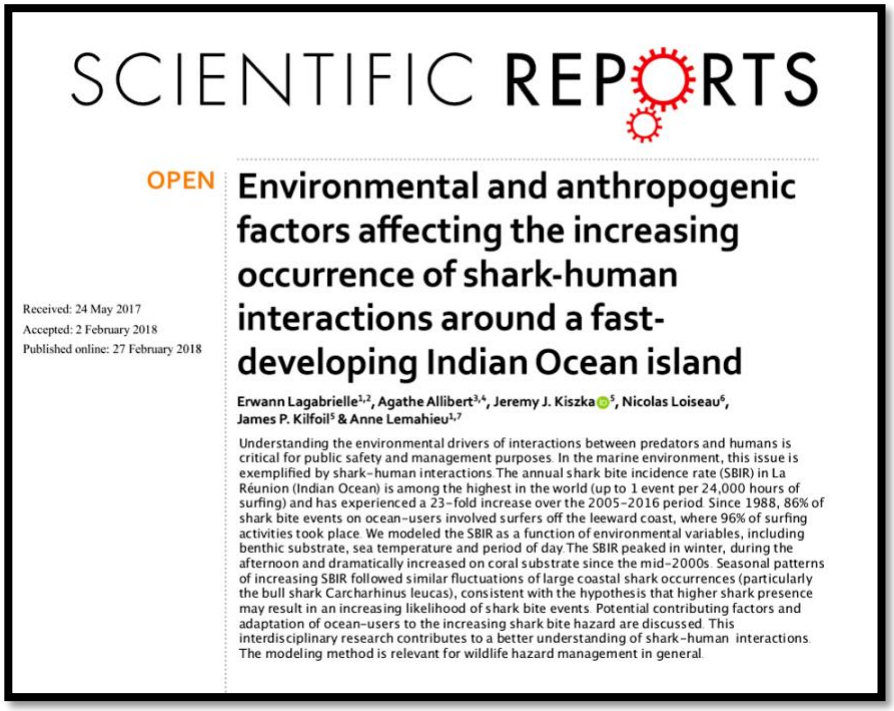
Vieira A (2011) The Art of Technical and Scientific Writing. *Revista de Sistemas de Informacao da FSMA* 8: 22–30.

White B (2000) *Dissertation Skills for Business and Management Students*. London: Thomson.

Appendix

Examples of the experimental manipulations of the study.

Baseline – Natural sciences



SCIENTIFIC REPORTS

OPEN

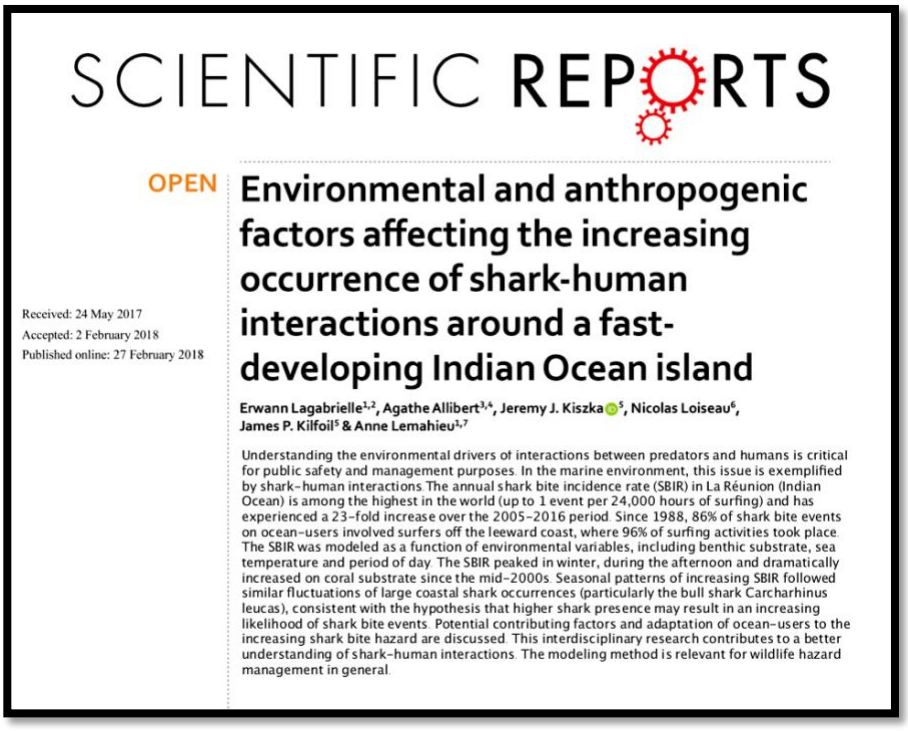
Environmental and anthropogenic factors affecting the increasing occurrence of shark-human interactions around a fast-developing Indian Ocean island

Received: 24 May 2017
Accepted: 2 February 2018
Published online: 27 February 2018

Erwann Lagabrielle[1,2], Agathe Allibert[3,4], Jeremy J. Kiszka[5], Nicolas Loiseau[6], James P. Kilfoil[5] & Anne Lemahieu[1,7]

Understanding the environmental drivers of interactions between predators and humans is critical for public safety and management purposes. In the marine environment, this issue is exemplified by shark–human interactions. The annual shark bite incidence rate (SBIR) in La Réunion (Indian Ocean) is among the highest in the world (up to 1 event per 24,000 hours of surfing) and has experienced a 23-fold increase over the 2005–2016 period. Since 1988, 86% of shark bite events on ocean-users involved surfers off the leeward coast, where 96% of surfing activities took place. We modeled the SBIR as a function of environmental variables, including benthic substrate, sea temperature and period of day. The SBIR peaked in winter, during the afternoon and dramatically increased on coral substrate since the mid-2000s. Seasonal patterns of increasing SBIR followed similar fluctuations of large coastal shark occurrences (particularly the bull shark Carcharhinus leucas), consistent with the hypothesis that higher shark presence may result in an increasing likelihood of shark bite events. Potential contributing factors and adaptation of ocean-users to the increasing shark bite hazard are discussed. This interdisciplinary research contributes to a better understanding of shark–human interactions. The modeling method is relevant for wildlife hazard management in general.

Grammatical person – Natural sciences



SCIENTIFIC REPORTS

OPEN

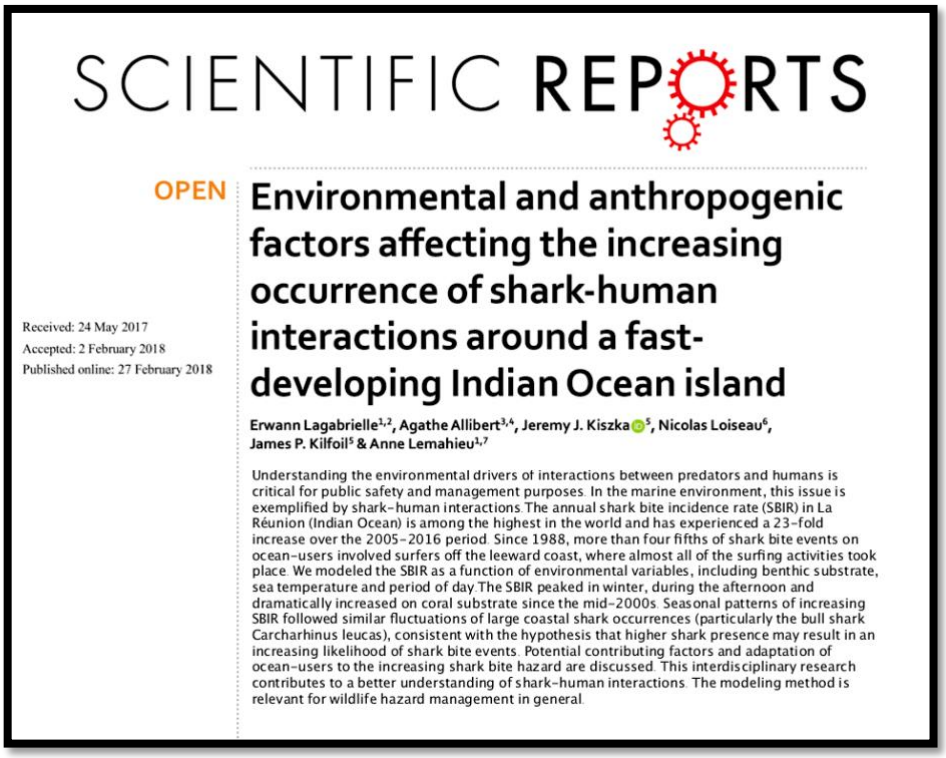
Environmental and anthropogenic factors affecting the increasing occurrence of shark-human interactions around a fast-developing Indian Ocean island

Received: 24 May 2017
Accepted: 2 February 2018
Published online: 27 February 2018

Erwann Lagabrielle[1,2], Agathe Allibert[3,4], Jeremy J. Kiszka[5], Nicolas Loiseau[6], James P. Kilfoil[5] & Anne Lemahieu[1,7]

Understanding the environmental drivers of interactions between predators and humans is critical for public safety and management purposes. In the marine environment, this issue is exemplified by shark–human interactions. The annual shark bite incidence rate (SBIR) in La Réunion (Indian Ocean) is among the highest in the world (up to 1 event per 24,000 hours of surfing) and has experienced a 23-fold increase over the 2005–2016 period. Since 1988, 86% of shark bite events on ocean-users involved surfers off the leeward coast, where 96% of surfing activities took place. The SBIR was modeled as a function of environmental variables, including benthic substrate, sea temperature and period of day. The SBIR peaked in winter, during the afternoon and dramatically increased on coral substrate since the mid-2000s. Seasonal patterns of increasing SBIR followed similar fluctuations of large coastal shark occurrences (particularly the bull shark Carcharhinus leucas), consistent with the hypothesis that higher shark presence may result in an increasing likelihood of shark bite events. Potential contributing factors and adaptation of ocean-users to the increasing shark bite hazard are discussed. This interdisciplinary research contributes to a better understanding of shark–human interactions. The modeling method is relevant for wildlife hazard management in general.

Statistical data – Natural sciences



SCIENTIFIC REPORTS

OPEN

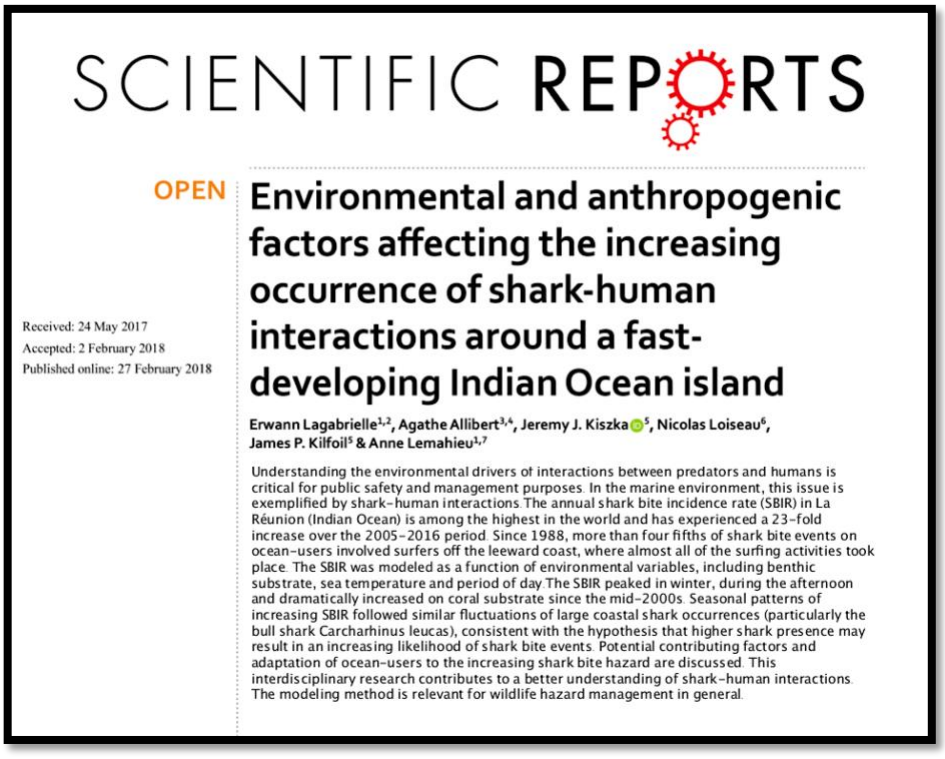
Environmental and anthropogenic factors affecting the increasing occurrence of shark-human interactions around a fast-developing Indian Ocean island

Received: 24 May 2017
Accepted: 2 February 2018
Published online: 27 February 2018

Erwann Lagabrielle[1,2], Agathe Allibert[3,4], Jeremy J. Kiszka[5], Nicolas Loiseau[6], James P. Kilfoil[5] & Anne Lemahieu[1,7]

Understanding the environmental drivers of interactions between predators and humans is critical for public safety and management purposes. In the marine environment, this issue is exemplified by shark–human interactions. The annual shark bite incidence rate (SBIR) in La Réunion (Indian Ocean) is among the highest in the world and has experienced a 23-fold increase over the 2005–2016 period. Since 1988, more than four fifths of shark bite events on ocean-users involved surfers off the leeward coast, where almost all of the surfing activities took place. We modeled the SBIR as a function of environmental variables, including benthic substrate, sea temperature and period of day. The SBIR peaked in winter, during the afternoon and dramatically increased on coral substrate since the mid-2000s. Seasonal patterns of increasing SBIR followed similar fluctuations of large coastal shark occurrences (particularly the bull shark Carcharhinus leucas), consistent with the hypothesis that higher shark presence may result in an increasing likelihood of shark bite events. Potential contributing factors and adaptation of ocean-users to the increasing shark bite hazard are discussed. This interdisciplinary research contributes to a better understanding of shark–human interactions. The modeling method is relevant for wildlife hazard management in general.

Grammatical person and statistical data – Natural sciences



SCIENTIFIC REPORTS

OPEN

Environmental and anthropogenic factors affecting the increasing occurrence of shark-human interactions around a fast-developing Indian Ocean island

Received: 24 May 2017
Accepted: 2 February 2018
Published online: 27 February 2018

Erwann Lagabrielle[1,2], Agathe Allibert[3,4], Jeremy J. Kiszka[5], Nicolas Loiseau[6], James P. Kilfoil[5] & Anne Lemahieu[1,7]

Understanding the environmental drivers of interactions between predators and humans is critical for public safety and management purposes. In the marine environment, this issue is exemplified by shark–human interactions. The annual shark bite incidence rate (SBIR) in La Réunion (Indian Ocean) is among the highest in the world and has experienced a 23-fold increase over the 2005–2016 period. Since 1988, more than four fifths of shark bite events on ocean-users involved surfers off the leeward coast, where almost all of the surfing activities took place. The SBIR was modeled as a function of environmental variables, including benthic substrate, sea temperature and period of day. The SBIR peaked in winter, during the afternoon and dramatically increased on coral substrate since the mid-2000s. Seasonal patterns of increasing SBIR followed similar fluctuations of large coastal shark occurrences (particularly the bull shark Carcharhinus leucas), consistent with the hypothesis that higher shark presence may result in an increasing likelihood of shark bite events. Potential contributing factors and adaptation of ocean-users to the increasing shark bite hazard are discussed. This interdisciplinary research contributes to a better understanding of shark–human interactions. The modeling method is relevant for wildlife hazard management in general.

Understanding through content and form: The role of presentational variables on the perception of scientific texts

Javiera Paredes, David Torres & Carlos Cornejo

Escuela de Psicología, Pontificia Universidad Católica de Chile

Corresponding author

Javiera Paredes, Escuela de Psicología, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile.

Email: jnparede@uc.cl

Abstract

There is emerging evidence that suggest that presentational variables (i.e., non-content-related) have an effect on the perception of science presented to laypeople and experts alike. The aim of this study was to determine whether presentational aspects of scientific texts affect comprehension of science content, specifically in scientific abstracts. The effects were elicited by manipulating the style of the title and the format of the abstract presented in an experimental design. The results suggest that the experimental manipulations did in fact modified the perception of comprehension of the abstract. More specifically, the inclusion of clichéd titles increased the perception of comprehension in participants. The modification of the format of the text to a simpler one also increased the comprehension reported in the survey. The combination of these presentational characteristics likewise allowed for an improved understanding perception among the participants of the study. Nevertheless, the modifications did not have an effect on more complex measures of comprehension. The differential effect on measures of comprehension could point out to

differences between levels of expertise and skills that could lead to greater depth of processing in some readers.

Keywords

Understanding, presentational, scientific texts, abstracts

Scientific communication and discourse are articulated by several different participants of the scientific community, one of the most important being researchers and scientists (Burns et al., 2003; Gheorghiu et al., 2017). Some of the aims of such communication are to allow understanding of the information presented in each piece of work (Burns et al., 2003; Kappel & Holmen, 2019) to persuade the readers that the facts and observations presented in such media are true (Bazerman, 1988; Finkler & Leon, 2019; Gross, 2006). The most used element of communication and dissemination is the scientific article. Studies have explored how some elements of these texts, not related to the content of an article, have an effect on the perception and understanding of scientific work (Crick & Hartling, 2015; Jamali & Nikzad, 2011; McCabe & Castel, 2008; Smith et al., 2011).

For the purposes of this study, we will consider all of the factors and elements not related to the content of scientific texts as *presentational variables*. In whichever form they adopt, presentational variables affect the perception and the understanding of scientific texts. Following Susanne Langer's work (1942, 1976) we suggest that all the elements mentioned contemporary research could be part of this denomination. This concept has also been named as "extraneous presentational factors" (Gheorghiu et al., 2017). The *presentational* as a concept denotes that these elements are not related to the content of the text but, nonetheless, they impact the understanding of the document (Langer, 1976). These formal elements, working together with the content of the text, allow understanding and persuade the community of the claims of the article. The

presentational features of scientific texts have an effect on the content displayed and work similarly to the paralinguistic elements in oral communication. In this sense, the contextualization function that these presentational elements provide is necessary for the full operation of language in everyday use, including scientific texts (Marcuschi, 2005)

While there is a wide array of studies on this phenomenon, most of them do not define it through a single concept or theoretical framework. For example, it has been reported the influence of clarity, comprehensibility, and aesthetic appeal of specific information displays upon perception of scientific texts (Crick & Hartling, 2015; Egger et al., 1997; Jamali & Nikzad, 2011) as well as the impact of non-verbal elements or peripheral cues on scientific texts (Michalovich & Hershkovitz, 2020; Tal & Wansink, 2016). Also, studies that inquiry into presentational features typically focus on the modification of the contents of the article and measure the effects of these modifications on the perception of the readers. Studies about public understanding of science have found that content-related presentational variables affect the understanding of science media. For instance, graphs and formulas positively impact the public's belief in product efficacy (Tal & Wansink, 2016). Another study evaluated people's satisfaction with explanations on whether these explanations contained irrelevant neuroscientific information. The results suggested that lengthy justifications that included neuroscientific information were judged more satisfying (Weisberg et al., 2015).

However, less attention has been given to the effect of non-content related, presentational variables on the communication and understanding of science. In the health sciences, studies have explored the preference of clinicians for different displays of information regarding systematic reviews (Crick & Hartling, 2015). Bibliometric evidence suggests that question titles or titles with colons are less cited than other types of titles (Jamali & Nikzad, 2011) and that articles with short

titles are cited more often than different types of titles (Paiva, 2012). In psychology, evidence supports that the researchers' attractiveness positively affects the interest regarding the research presented. On the other hand, the appearance of morality and competence are the factors that weigh on the quality and interest judgments of readers (Gheorghiu et al., 2017).

There is incipient empirical evidence that shows that presentational variables significantly affect the perception and understanding of the science presented in texts or textual displays. Hartley (2003) demonstrated that structuring the content of articles' abstracts was rated as more clear and informative than the original abstracts. Smith et al. (2011) explored the role of color, scale, and type of explanation on the understanding of astronomical pictures by experts and non-experts. Their results suggest that augmentation was a key feature for appreciating astronomical images, especially in a non-expert sample. The inclusion of illustrative scales that allow people to get a sense of the pictures' size was also helpful to all participants, irrespective of their expertise level. The interpretation of these colors differs in terms of the expertise of who is exposed to the picture. In this sense, non-experts interpret the colors differently from experts (e.g., blue to represent hotter regions of space by experts).

The study of non-image related presentational elements on the understanding of scientific texts is also important, given the considerable variation in how these texts are presented nowadays. The present study will focus on two presentational variables: the title of the text and the format on which the text is presented. Previous evidence suggests that formatting and titles are important parts of the process of perception and selection of scientific texts (Hartley, 2003; Jamali & Nikzad, 2011; Paiva, 2012; Whissell, 2012). Also, through recommendations and guidelines, authors and editors have recognized the value of these variables on science communication success (e.g., Annesley, 2010; Day & Gastel, 2011; Knight & Ingersoll, 1996). Furthermore, the creation of

cliché titles to draw audiences to the latest research and the format of presentation of scientific texts (from the structuration of abstracts to stylistic emphases) are presentational elements that are commonly used but seldom studied.

Title is the first and most-read element of articles (Spotti Lopes Fujita et al., 2018). The author draws from other sections of the article to create a descriptor for the content of the whole text that at the same time plays the role of attention-grabber (Annesley, 2010). The main recommendations regarding style are to inform the readers in the article's title of the independent and dependent variables, the observed effect, and the population studied in 10 words or less (Annesley, 2010; Knight & Ingersoll, 1996; Paiva, 2012). Studies have found that titles summarizing results were correlated with higher citation rates (Paiva, 2012) compared with question titles or titles with colons that were less cited (Jamali & Nikzad, 2011). However, recent evidence from educational psychology challenges those findings. A survey presented to psychology students found that participants preferred longer titles, titles that contained colons, and titles that were phrased as questions. The researchers also found that participants did not favor result or method-focused titles over each other or titles that used clichés (Hallock & Bennett, 2021). Clichés in the case of titles are characterized as stereotyped expressions that are widely known (Goodman, 2012). Considering current research and the tensions between the ideas of attractiveness and succinct information (Whissell, 2012), the ideal title would be attractive enough to draw more readers and informative of the critical elements of the article in a relatively concise manner.

The presentation format also could be a significant variable with an effect on the understanding of scientific texts. A study in behavioral sciences and design explored the use of signaling devices as a mechanism to make the reading of texts more productive (i.e., improve recall

of the text's content). The study suggests a statistically significant difference in text recall comparing the standard text and the text using signaling devices, with content recall being higher in the latter (van der Meij et al., 2013). In science education, researchers have found that secondary students who are shown texts supposedly taken from students' reports were rated as less comprehensible than reports taken from textbooks (Werner da Rosa & Otero, 2018). Considering complementary formats of scholarly communication such as the pre-print, the structuration of such texts and their study could be crucial to understand the comprehension of the science presented in them. One definition presents pre-prints and related documents as bulkier and less legible documents informing other scholars about the latest scientific developments (McKiernan, 2000). The possibility that a plainer format, such as one found in a pre-print format, could affect the comprehension of scientific texts is worth exploring, considering the increasing accessibility of these documents by researchers outside of the community and members of the public. Most studies reviewed modify the content presented to the participants of the studies, which could obscure the effect of the sole modification of the format on the understanding of the science.

The present study explores whether presentational modifications of scientific texts – cliché titles and a simple text format – affect the reader's understanding of the text. To this purpose, we conducted a randomized experimental design using images of the article's abstracts on computer screens. This display allowed to show this section of the article, as shown if the participant searched for it.

We hypothesized that cliché titles would ensue better scores on the study participants' overall understanding of the texts. In addition, we hypothesized that the traditional format of presentation of the abstract would result in better scores on the understanding of the texts by the

participants. We did not anticipate any interaction effects since there are no reports of this happening for this specific combination of presentational variables.

Method

The study aimed to explore the effects of presentational modifications on comprehension of abstracts of scientific texts. The effects were elicited by manipulating the style of the title in which the abstract was written, and the format of the article presented in an experimental 2 x 2 design.

Participants

This study's participants were 234 members of the Pontificia Universidad Católica de Chile (age: M = 21.36, SD = 1.9; 134 women). The study aimed to answer its questions in two broad disciplines: natural and social sciences. Therefore, the authors constructed stimuli for the two areas. One hundred and thirty-three students belonging to the area of the natural sciences participated in the study. The participants voluntarily signed up for participating and were allowed to sign up in a raffle for a gift card as compensation for their participation. Ethical approval was granted by the Ethical Committee of Social Sciences at the Pontificia Universidad Católica de Chile following the Declaration of Helsinki. Participants gave informed consent and, since the study was carried out online, were given a choice to download the informed consent from the platform.

Data from 11 participants were excluded from further analyses. These participants completed the questionnaire at a much faster rate (i.e., two standard deviations above the average response time, M = 2.15, SD = 0.07) than the typical participant of the study, suggesting random answering.

Materials

The study was generated from two abstracts of articles belonging to the natural sciences (Sitterlé et al., 2019; Skinner et al., 2019) and two abstracts belonging to the social sciences and humanities (Baus et al., 2019; Bressan, 2018) from the same highly cited multidisciplinary journal (See Appendix). The survey presented them with two abstracts to read and rate depending on the participant's study area. All abstracts were selected to have the same baseline on the variables modified: type of title and format of presentation. The type of title manipulation changed the original title to a clichéd version and the format of presentation varied from the original format to a simple MS word format. The baseline of all abstracts were the original title and format of presentation.

All stimuli were created through a high-quality screenshot PDF of the original article. The experimental changes were made with the Snagit 2018 software that allows the edition of PDF texts. Because most browsers do not display PDF files directly on their web pages, PNG versions of the stimuli were created to better show the images on the platform. After some interviews with undergraduate students before the study's launch, it was clear that a Spanish version of all abstracts was necessary. Therefore, the researchers carried out the Spanish translation of the texts with the assistance of experts in each area. Translated abstracts were added to the stimulus page through a link (i.e., "Para versión en español, click aquí"/"For the Spanish version, click here").

In all abstracts, the baseline condition kept the journal's original title and original format. The second condition modified the title of the abstracts for a cliché version. The format of presentation of the journal was kept in the same style as the original. In the third condition, in which the abstract format is changed, the abstract was presented in a word processor format (i.e., white background, Times New Roman 12-point font) instead of the journal format. In this case, the title

of the abstract was kept in the same manner as the original. Finally, in the fourth condition both variables were simultaneously modified. Thus, the abstract was altered in its title and presentation format, having a cliché title and being presented as simple word processor file.

In Table 1, a summary of the experimental conditions and their characteristics is presented.

Table 1. Summary of experimental conditions

			Baseline	Title	Format	Interaction
Natural Sciences	**Sitterlé et al. (2019)**	**Title**	No	Yes	No	Yes
	Skinner et al. (2019)	**Format**	No	No	Yes	Yes
Social Sciences	**Baus et al. (2019)**	**Title**	No	Yes	No	Yes
	Bressan et al. (2018)	**Format**	No	No	Yes	Yes

Each version of the study displayed two abstracts to participants randomized to be assigned to different conditions and shown as rotating blocks to avoid primacy effect issues (Glenberg et al., 1980). Block order, stimulus, and participants' assignment to conditions were randomized through the Sosci Survey software. The selection of the experimental stimuli was counterbalanced.

Procedure

Participants were invited in person by means of a brief description of the study and, if they were interested, asked for writing their email address to be contacted for participating in an online research survey. Students who signed up received the same day an email with the invitation containing a link to the study in the SoSci platform (http://soscisurvey.de). A reminder email was sent after four days of the initial invitation.

Once in the platform, participants were asked for their consent to take part in the study. As a first step, participants were asked to fill in some demographic information about themselves, such as gender, age, area of undergraduate study, and level of advancement of undergraduate studies.

This section was placed in the survey to allow the platform to assign the students in the natural science or social science/humanities version of the survey.

Next, an instruction page was added to guide the participants. The command stated: “On the next page, you will find an abstract from a scientific paper. It may take a few seconds to appear. Please read it as many times as you want. Below the abstract, you will find a link to the Spanish translation if needed. To start, click on the Next button”. Afterwards, the first abstract was presented on a sole page with a black background and the Spanish translation link.

The perceived comprehension of the abstract was assessed first. Participants judged on a five-point scale (1= “Very hard to understand,” 5 = “Very easy to understand”) the question “What did you think about the abstract that you just read?”. Additionally, participants answered an open question worded: “What do you think that is the main idea or discovery that the authors wished to communicate through the abstract?”. This item allowed for as many answers as the participant wanted to provide regarding each abstract.

Considering the evidence that scientific research reading skills vary depending on the stage of the reader’s academic career (Hubbard & Dunbar, 2017), the study included a question to measure the expertise self-perception (i.e., “How much of an expert do you think you are regarding the article’s subject?)”. This question was thought to be answered on a five-point scale (1 = Not at all, 5 = Very much). Additionally, an item on the language in which the abstract was read was added (i.e., “In which language did you read the abstract?”). This question was added to check the possibility that the language in which the abstract was read interfered with the comprehension or appreciation of the abstract.

Finally, there was a form to allow the participants to sign themselves up for a raffle as compensation for their participation. After this, they were thanked for their contribution in the exit section of the questionnaire.

Analysis

The analysis of the survey data was carried out using a generalized linear model, controlling for non-independence of the observations linked to the two abstracts read by each participant. The model was run with the Mplus software (Muthén & Muthén, 2017 version 8.4). In the model, the non-independence of the observations – abstracts – was corrected using the TYPE = COMPLEX option of the analysis command in conjunction with the CLUSTER option in the variable command (Muthén & Muthén, 2017). The cluster variable used in this analysis was the case variable that grouped the abstracts read by each participant in the study. In this regression, the distribution is centered in men in the social sciences & humanities area, in the second abstract in participants with a low perception of expertise who read the Spanish abstract.

The regression was run with the item "Perception of comprehension" as the dependent variable. The predictor variables were the study area of participants (sciences/social sciences or humanities), the abstract presented to the participant, the experimental conditions within each abstract, the participant's gender, their perception of expertise regarding the subject of the abstract, and the language on which the abstract is written.

For the analysis of the open-ended question regarding the comprehension of the abstract (i.e., "What do you think is the main idea or discovery that the authors wanted to communicate through the abstract?"). The answers were evaluated by means of the Detailed Rubric for Reading Assessment (Northern Essex Community College, n.d.). The evaluation was based on the rubric

section dedicated to assessing reading comprehension (from 1= Novice to 4 = Internalizing Highest Level), which was rated by a research team member and later was triangulated with the authors. Also, all these open-ended answers were coded concerning which of the three main ideas of each abstract were present on the answer. These three ideas in each abstract were related to the study's aims, the central hypothesis of their study, the results and the conclusions stated by the researchers.

Results

Table 1 below presents the regression coefficients, standard errors for each parameter, and p-values for the regression model.

Table 1. Regression coefficients, standard errors for the regression, and p-values of the item "Perception of comprehension" using as predictor variables the study area of participants, the abstract presented to the participant, the experimental condition within each abstract, the participant's gender, their perception of expertise regarding the subject of each abstract, and the language of the abstract.

	Perception of comprehension		
	Coefficient	S.E.	p-value
Study area	-0.473	0.109	< 0.001
Abstract	0.277	0.106	0.009
Style of the title	0.417	0.135	0.002
Format	0.514	0.151	0.001
Title's style x Format	0.329	0.150	0.028
Gender	-0.167	0.105	0.111
Expertise perception	0.846	0.108	< 0.001
Language	-0.454	0.107	< 0.001

Significant differences in the perception of comprehension in almost all variables except the gender of the participant were found. We can observe significant differences in the perception of comprehension scores reliant on the study area to which the participant belonged (R = -0.473, DE = 0.109, 95% CI [-0.687, -0.259]). These results indicate that when the participant was a natural sciences student, the perceived comprehension on any of the abstracts read fell by 0.47 compared to the responses of social sciences or humanities students. There are also significant differences in the perception of comprehension scores reliant on the abstract shown to participants (R = 0.277, DE = 0.106, 95% CI [0.069, 0.485]). These results indicate that when the Skinner et al. (2018) and the Bressan (2018) abstracts were presented to participants, the perceived comprehension of the text rose by 0.28 compared to the other two abstracts (Baus et al., 2019; Sitterlé et al., 2019).

Significant differences are similarly detected in the scores of the perception of comprehension of the abstract depending on title of the abstract (R = 0.417, DE = 0.135, 95% CI [0.152, 0.682]). When the abstract title was written in its cliché version, the scores of the perceived comprehension of the text rose by 0.42. We observe significant differences between perceived comprehension scores when considering the format covariate in the regression (R = 0.514, DE = 0.151, 95% CI [0.218, 0.810]). The perceived comprehension of the abstract rises by 0.51 when the text shown to participants is the MS Word format. The interaction between the cliché title and presentation format in the abstract also shows significant differences in the scores of the perceived comprehension of the texts (R = 0.329, DE = 0.150, 95% CI [0.035, 0.623]). The perceived comprehension of the abstract rises by 0.33 when the abstract has a cliché title and its format of presentation changes to a simpler one, such as a MS Word format. Significant differences in the perceived comprehension scores are also present when we consider as a covariate the participant's self-perceived expertise on the abstract's subject (R = 0.846, DE = 0.108, 95% CI [0.634, 1.058]).

When a participant considers herself knowledgeable about the abstract's subject matter, the scores of the perception of comprehension of the text rise by 0.85. Finally, the perception of comprehension scores is also significantly different when we consider the language on which the participants read the abstract as a covariate in the regression (R = -0.454, DE = 0.107, 95% CI [-0.664, -0.244]). When the abstract is read in Spanish, the perceived comprehension of the texts falls by 0.45.

Regarding the analysis of the open-ended questions of the abstracts, in Table 2 below, we present the averages of the rubric scores assessing reading comprehension for each condition in both study areas.

Table 2. Averages the rubric scores assessing reading comprehension for each study condition in the abstracts of natural sciences and social sciences, and humanities.

Natural Sciences		Social Sciences and Humanities	
Condition	Average	Condition	Average
Baseline	2.70	Baseline	2.72
Style of title	2.90	Style of title	2.90
Format	2.58	Format	2.52
Style of title x format	2.58	Style of title x format	2.54

The average scores of each condition for reading comprehension rated by an external judge were not significantly different from each other (F (7,468 = 1.49, p = 0.17). Regarding the frequencies of the text's main ideas in both abstracts in its study area, we present Table 3.

Table 3. Abstract's open-ended questions of comprehension frequencies for the main ideas in the text according to each experimental condition in both areas of study (natural sciences and social sciences and humanities).

	Baseline	Style of title	Format	Style of title x format
No main ideas present	17	21	24	37
Idea 1: Aims of the study	47	24	36	32
Idea 2: Hypothesis/results of the study	28	34	44	34
Idea 3: Results and conclusions of the study	65	76	53	53

The responses related to each main idea are distributed in a clear trend, with no differential distribution in each condition for the two areas. In general, answers to open-ended questions focused on the results and conclusions of the studies. Another trend that appears shows that the experimental manipulations may have generated less focus on the study's aims as the main idea.

Discussion

This study aimed to explore if presentational modifications of scientific texts, in the form of a clichéd title and a simple format, affected the comprehension of the science presented in each abstract. We additionally explored the possible impact of each abstract as a unique text and self-perceived expertise, gender, and language of the text on the same dependent variables.

Overall, the results of this study suggest that presentational variables --title and format modifications-- affected the perception of comprehension of the abstracts across all participants. In the case of the title modification, the addition of a clichéd title to the abstract rose the scores of

the perceived comprehension of the abstract across all participants. A recent study on the appreciation of titles in a sample of psychology students found that participants preferred long titles and titles phrased as a question, while no significant preference for clichéd titles was found (Hallock & Bennett, 2021). However, it should be considered that Hallock & Bennett's study (2021) and ours measured different dependent variables, i.e., preference between two types of article titles and perception of comprehension of a specific text. While in the current study participants did not indicate any specific preference for clichéd titles, they may perceive that a better understanding of the text is achieved when this type of title is used.

The results of this study also do not follow style recommendation in several disciplines, where the use of cliché titles is disincentivized (Day & Gastel, 2011; DeBakey & DeBakey, 1983; Goodman, 2012). The conventional approach towards this presentational variable seems to be more flexible when it is considered by people still on a formative process; such is the case with graduate students. A more colloquial title does seem to improve the perception of the understanding of the abstract, even though it does not fit an objective approach to scientific communication and the presentation of its results.

Regarding the paper format, the change from the traditional article format to a simpler format improved the perception of comprehension of the abstract across all participants. A hypothesis that could explain these results is that there might be a negative emotional response to the standard article formatting. On one hand, the participants may have perceived the original format of presentation as an element of the study that was fabricated. Therefore, a plainer layout may have improved the perception of comprehension since it did not have any visual distractors. Also, the fact that participants were undergraduate students with some experience in academic activities such as reading scientific articles might have a part on these results as well. Since not all

participants were well versed with this practice, the mere fact of reading an article abstract might be the ground for a negative affective reaction to the activity. Consequently, the negative disposition to the article from the formatting diminishes the perception of understanding of the text. To the authors' knowledge, this is the first time that the effect of the format of presentation on the perception of understanding has been tested on scientific texts without content modifications like previous studies did (Hartley, 2003; van der Meij et al., 2013; Werner da Rosa & Otero, 2018). Contrary to what was thought about simple formats (McKiernan, 2000), presentations such as these are perceived by readers with some university training as less bulky and, maybe, more legible.

An important finding is the significant effect of the interaction of the format of presentation and the modified title. Participants perceived a greater understanding of the text when presented in a plainer layout and had a clichéd title. This configuration is counterintuitive considering the recommendations made by academic writing (Day & Gastel, 2011; Goodman, 2012; McKiernan, 2000; McMillan & Weyers, 2007). However, since the presentational features of scientific information have been changing more rapidly in recent years, it can be hypothesized that a simpler text configuration combined with a memorable title could change the perception of the text's difficulty and, therefore, increase the perception of comprehension. This is related to Werner da Rosa & Otero's (2018) results, where the understanding of science in secondary students was linked to the authorship of the text. In our study, the perception of a scientific text that could be in the process of writing or a pre-print stage makes the abstract appear more approachable to participants and, consequently, easier to understand.

Although the experimental manipulations of this study modify the perception of comprehension, the analysis of the open-ended questions regarding the understanding of the main idea of the texts is not in accordance with the previous results. The disparity between the perception

of understanding and other measures of depth of processing is also identified in Werner da Rosa & Otero's study (2018) and a previous study by the authors (Paredes, Torres & Cornejo, 2021). Some accounts of philosophy of science state that understanding in science comprises several elements. Achieving comprehension in these proposals is the assessment of the accuracy and truth of the statements presented by the text and the use of particular skills and criteria developed by specific scientific communities that allow the person to manipulate this knowledge properly (de Regt, 2004; de Regt & Dieks, 2005). A better understanding of the mismatch between these two types of measures could be reached when understanding of science is thought through this definition. A hypothesis of the differential effect between the perception of comprehension and other measures is that while participants reached a better perception of their understanding through the introduced presentational elements, these elements did not lead to a greater depth of processing. Participants in the early stages of formation, like undergraduates, may have some training in determining the accuracy of the statements in a text but may have less training or knowledge regarding the specific skills and criteria necessary to manage correctly the knowledge presented. Further studies are needed to see if, in other more expert samples, these effects disappear.

Other results of the study point out to the difference of scores of the perception of understanding depending on the language on which the abstract was read. Higher scores on the perception of understanding related to the English version of the abstract suggest that the belief that this language is the lingua franca of science (Meneghini et al., 2008) also affects how the abstract is perceived and understood. Evidence shows that studies published in English are ever-increasing and that those studies report more significant results than studies published in other languages (Egger et al., 1997; Ramírez-Castañeda, 2020; Solovova et al., 2018). Reading

publications in this language could be linked to the perception in laypeople that science written in English has better quality compared to the science published in other languages.

The results of this study also suggest that participants trained in natural sciences rate their perceived understanding of the abstract significantly lower than the social sciences and humanities-trained participants. An explanation for this difference could be that the effect of the presentational elements is not strong enough to bring the science-trained students to the same level of perception of their comprehension, given that they have an increased familiarity with academic journals, their reputation, and format of presentation. More research is needed to understand the underlying factors that could be contributing to this difference.

In the present study, the perception of understanding also varied depending on which abstract was presented to the participants, regardless of their area of expertise. In the design of the experiment, the authors favored an ecological approach where four different genuine abstracts were considered. The abstract selection process opens the door to many other explanations for this unexpected difference. One hypothesis is that the topics of the Skinner et al. (2019) and Bressan (2018) abstracts were more attractive to the sample of this study than the other texts. There is evidence that the public is particularly interested in climate change topics (Baram-Tsabari & Segev, 2011; European Commission, Directorate-General for Communication, 2020; Falchetti et al., 2007), and in other cases, the preference could be guided by previous experiences and personal interests (Falk et al., 2007). Other possibilities could explain these results, and more research on this is needed. Another unforeseen result was the strong effect of the perception of expertise of the participants on the perceived understanding of the abstract across all areas. Previous studies suggest that undergraduate students value different sections of articles depending on their discipline and that reading skills do develop throughout their careers (Hubbard & Dunbar, 2017; Vosniadou,

2019), but more knowledge is needed to understand why a greater perception of expertise would also affect the perception of the understanding of texts that relate broadly to their discipline. One hypothesis could point to the possible relationship between the perception of expertise and the belief of participants that they are able to use their academic skills to approach and understand the text. Therefore, a student who expresses more confidence in their knowledge of the subject presented in the abstract may perceive that they better understood the text presented to them in this study.

Some limitations of this study should be mentioned. All participants in this study had undergraduate university training. Further studies should explore the effect of these presentational elements in scientific texts in a more diverse sample. Also, future research should include additional measures for the perception of understanding and reading comprehension, which could give better clues to figure out the difference between the measures of this study.

The results presented in the present study provide evidence of how presentational variables in texts affect the understanding of scientific texts. Therefore, reading and understanding scientific texts also depend on the document's content and how it is displayed. This evidence follows other studies that suggest that presentational features have a role in articulating and communicating science and its results (Gheorghiu et al., 2017; Hartley, 2003; Smith et al., 2011). We need to consider presentational variables as significant elements when the presentation of the results of scientific research is outlined and carried out. A previous study suggests that there is an important effect of presentational elements embedded in the content of scientific texts in the understanding of science (Paredes, Torres Iribarra & Cornejo, 2021), and this study confirms the relevance of the presentational aspects that are not linked to content as a similarly significant element in these psychological processes. These results allow us to sustain that science texts work as a whole

meaning-making unit that integrates content and presentational aspects to the reader's benefit. Considering the results of this study, some configurations of these elements allow for a better perception understanding in participants that are being trained in their disciplines. In this respect, the presentational elements of the texts work as paralinguistic elements of oral discourse (Marcuschi, 2005). These elements frame the experience of the text, not only in artistic contexts but also in scientific practice and communication.

References

Annesley, T. M. (2010). The Title Says It All. *Clinical Chemistry*, *56*(3), 357–360. https://doi.org/10.1373/clinchem.2009.141523

Baus, C., McAleer, P., Marcoux, K., Belin, P., & Costa, A. (2019). Forming social impressions from voices in native and foreign languages. *Scientific Reports*, *9*(1), 414. https://doi.org/10.1038/s41598-018-36518-6

Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science.* The University of Winsconsin Press.

Bressan, P. (2018). Systemisers are better at maths. *Scientific Reports*, *8*(1), 11636. https://doi.org/10.1038/s41598-018-30013-8

Burns, T. W., O'Connor, D. J., & Stocklmayer, S. M. (2003). Science Communication: A Contemporary Definition. *Public Understanding of Science*, *12*(2), 183–202. https://doi.org/10.1177/09636625030122004

Crick, K., & Hartling, L. (2015). Preferences of Knowledge Users for Two Formats of Summarizing Results from Systematic Reviews: Infographics and Critical Appraisals. *PLOS ONE*, *10*(10), e0140029. https://doi.org/10.1371/journal.pone.0140029

Day, R. A., & Gastel, B. (2011). *How to write and publish a scientific paper*. Greenwood.

de Regt, H. W. (2004). Discussion Note: Making Sense of Understanding*. *Philosophy of Science*, *71*(1), 98–109. https://doi.org/10.1086/381415

de Regt, H. W., & Dieks, D. (2005). A Contextual Approach to Scientific Understanding. *Synthese*, *144*(1), 137–170. https://doi.org/10.1007/s11229-005-5000-4

DeBakey, S., & DeBakey, L. (1983). The title. What's in a name? *International Journal of Cardiology*, *2*(3–4), 401–406. https://doi.org/10.1016/0167-5273(83)90012-8

Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *Lancet*, *350*(9074), 326–329. https://doi.org/10.1016/S0140-6736(97)02419-7

Finkler, W., & Leon, B. (2019). The power of storytelling and video: A visual rhetoric for science communication. *Journal of Science Communication*, *18*(05). https://doi.org/10.22323/2.18050202

Gheorghiu, A. I., Callan, M. J., & Skylark, W. J. (2017). Facial appearance affects science communication. *Proceedings of the National Academy of Sciences*, *114*(23), 5970–5975. https://doi.org/10.1073/pnas.1620542114

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., Gretz, A. L., Fish, J. H., & Turpin, B. M. (1980). A two-series account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(4), 355–369. https://doi.org/10.1037/0278-7393.6.4.355

Goodman, N. (2012). Familiarity breeds: Clichés in article titles. *British Journal of General*

Practice, *62*(605), 656–657. https://doi.org/10.3399/bjgp12X659420

Gross, A. (2006). *Starring the text: The place of rhetoric in science studies*. SIU Press.

Hallock, R. M., & Bennett, T. N. (2021). I'll Read That!: What Title Elements Attract Readers to an Article? *Teaching of Psychology*, *48*(1), 26–31. https://doi.org/10.1177/0098628320959948

Hartley, J. (2003). Improving the clarity of journal abstracts in psychology: The case for structure. *Science Communication*, *24*(3), 366–379. https://doi.org/10.1177/1075547002250301

Hubbard, K. E., & Dunbar, S. D. (2017). Perceptions of scientific research literature and strategies for reading papers depend on academic career stage. *PLOS ONE*, *12*(12), e0189753. https://doi.org/10.1371/journal.pone.0189753

Jamali, H. R., & Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics*, *88*(2), 653–661. https://doi.org/10.1007/s11192-011-0412-z

Kappel, K., & Holmen, S. J. (2019). Why Science Communication, and Does It Work? A Taxonomy of Science Communication Aims and a Survey of the Empirical Evidence. *Frontiers in Communication*, *4*, 55. https://doi.org/10.3389/fcomm.2019.00055

Knight, K. L., & Ingersoll, C. D. (1996). Structure of a Scholarly Manuscript: 66 Tips for What Goes Where. *Journal of Athletic Training*, *31*(3), 201–206.

Langer, S. K. (1942). *Philosophy in a new key*. Harvard University Press.

Langer, S. K. (1976). The Dynamic Image: Some Philosophical Reflections on Dance on JSTOR.

Salmagundi, *33*, 76–82. https://doi.org/10.2307/40546920

Marcuschi, L. A. (2005). A oralidade no contexto dos usos lingüísticos: Caracterizando a fala. In L. A. Marcuschi & A. Paiva Dionisio (Eds.), *Fala e Escrita* (pp. 57–84). Autentica.

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, *107*(1), 343–352. https://doi.org/10.1016/j.cognition.2007.07.017

McKiernan, G. (2000). arXiv.org: The Los Alamos National Laboratory e-print server. *The International Journal on Grey Literature*, *1*(3), 127–138.

McMillan, K., & Weyers, J. (2007). *How to write essays and assignments*. Pearson Education.

Meneghini, R., Packer, A. L., & Nassi-Calò, L. (2008). Articles by Latin American authors in prestigious journals have fewer citations. *PLoS ONE*, *3*(11), 1–4. https://doi.org/10.1371/journal.pone.0003804

Michalovich, A., & Hershkovitz, A. (2020). Assessing YouTube science news' credibility: The impact of web-search on the role of video, source, and user attributes. *Public Understanding of Science*, *29*(4), 376–391. https://doi.org/10.1177/0963662520905466

Northern Essex Community College. (n.d.). *Detailed Rubric for Reading Assessmnent*. https://facstaff.necc.mass.edu

Paiva, C. E. (2012). Articles with short titles describing the results are cited more often. *CLINICS*, *67*(5), 5. https://doi.org/10.6061/clinics/2012(05)17

Ramírez-Castañeda, V. (2020). Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian

researchers in biological sciences. *PLOS ONE, 15*(9), e0238372. https://doi.org/10.1371/journal.pone.0238372

Sitterlé, E., Maufrais, C., Sertour, N., Palayret, M., d'Enfert, C., & Bougnoux, M.-E. (2019). Within-Host Genomic Diversity of Candida albicans in Healthy Carriers. *Scientific Reports*, *9*(1), 2563. https://doi.org/10.1038/s41598-019-38768-4

Skinner, C., Gattinger, A., Krauss, M., Krause, H.-M., Mayer, J., van der Heijden, M. G. A., & Mäder, P. (2019). The impact of long-term organic farming on soil-derived greenhouse gas emissions. *Scientific Reports*, *9*(1), 1702. https://doi.org/10.1038/s41598-018-38207-w

Smith, L. F., Smith, J. K., Arcand, K. K., Smith, R. K., Bookbinder, J., & Keach, K. (2011). Aesthetics and Astronomy: Studying the Public's Perception and Understanding of Imagery From Space. *Science Communication*, *33*(2), 201–238. https://doi.org/10.1177/1075547010379579

Solovova, O., Santos, J., & Veríssimo, J. (2018). Publish in English or Perish in Portuguese: Struggles and Constraints on the Semiperiphery. *Publications*, *6*(2), 25. https://doi.org/10.3390/publications6020025

Spotti Lopes Fujita, M., Agustín-Lacruz, M.-C., & Terra, A. L. (2018). Journals' guidelines about title, abstract and keywords: An overview of Information Science and Communication Science areas. *European Science Editing*. https://doi.org/10.20316/ESE.2018.44.18015

Tal, A., & Wansink, B. (2016). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*, *25*(1), 117–125. https://doi.org/10.1177/0963662514549688

van der Meij, H., van der Meij, J., & Farkas, D. K. (2013). QuikScan formatting as a means to improve text recall. *Journal of Documentation*, *69*(1), 81–97. https://doi.org/10.1108/00220411311295333

Vosniadou, S. (2019). The Development of Students' Understanding of Science. *Frontiers in Education*, *4*, 32. https://doi.org/10.3389/feduc.2019.00032

Weisberg, D. S., Taylor, J. C. V., & Hopkins, E. J. (2015). Deconstructing the Seductive Allure of Neuroscience Explanations. *Judgment and Decision Making*, *10*(5), 15.

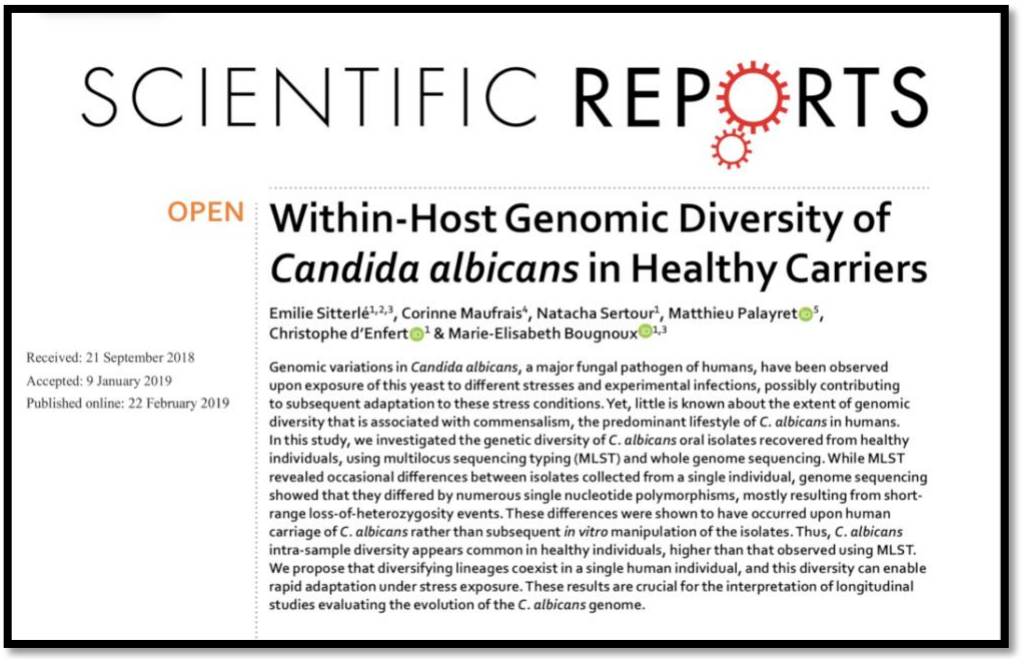
Werner da Rosa, C., & Otero, J. (2018). Influence of source credibility on students' noticing and assessing comprehension obstacles in science texts. *International Journal of Science Education*, *40*(13), 1653–1668. https://doi.org/10.1080/09500693.2018.1501168

Whissell, C. (2012). The Trend toward more Attractive and Informative Titles: American Psychologist 1946–2010. *Psychological Reports*, *110*(2), 427–444. https://doi.org/10.2466/17.28.PR0.110.2.427-444

Appendix

Examples of the experimental manipulations of the study.

Baseline – Natural sciences



SCIENTIFIC REPORTS

OPEN

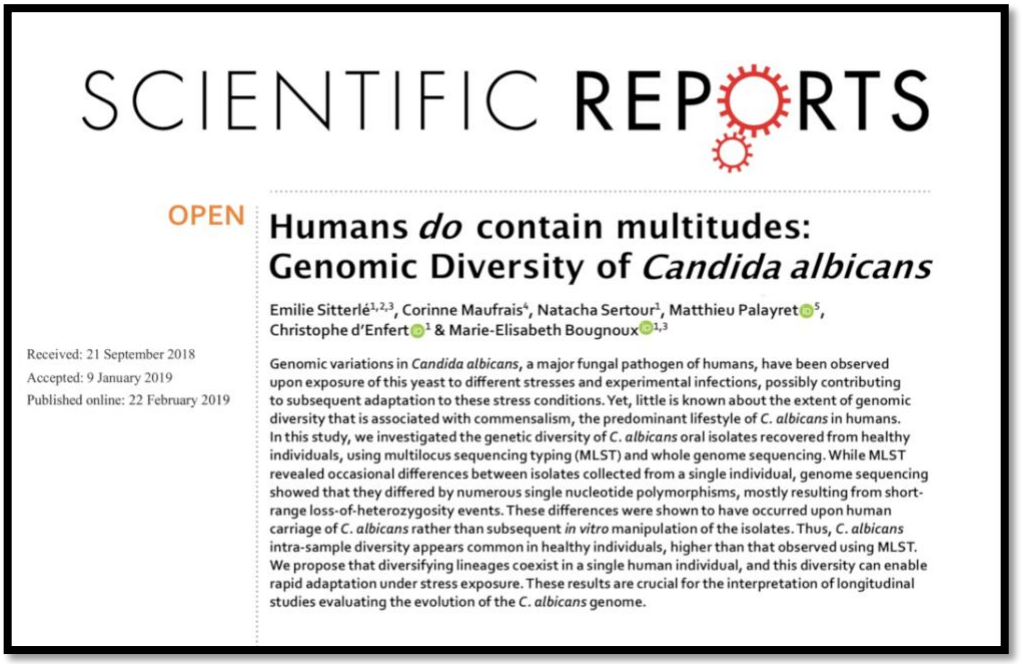
Within-Host Genomic Diversity of *Candida albicans* in Healthy Carriers

Emilie Sitterlé[1,2,3], Corinne Maufrais[4], Natacha Sertour[1], Matthieu Palayret[5], Christophe d'Enfert[1] & Marie-Elisabeth Bougnoux[1,3]

Received: 21 September 2018
Accepted: 9 January 2019
Published online: 22 February 2019

Genomic variations in *Candida albicans*, a major fungal pathogen of humans, have been observed upon exposure of this yeast to different stresses and experimental infections, possibly contributing to subsequent adaptation to these stress conditions. Yet, little is known about the extent of genomic diversity that is associated with commensalism, the predominant lifestyle of *C. albicans* in humans. In this study, we investigated the genetic diversity of *C. albicans* oral isolates recovered from healthy individuals, using multilocus sequencing typing (MLST) and whole genome sequencing. While MLST revealed occasional differences between isolates collected from a single individual, genome sequencing showed that they differed by numerous single nucleotide polymorphisms, mostly resulting from short-range loss-of-heterozygosity events. These differences were shown to have occurred upon human carriage of *C. albicans* rather than subsequent *in vitro* manipulation of the isolates. Thus, *C. albicans* intra-sample diversity appears common in healthy individuals, higher than that observed using MLST. We propose that diversifying lineages coexist in a single human individual, and this diversity can enable rapid adaptation under stress exposure. These results are crucial for the interpretation of longitudinal studies evaluating the evolution of the *C. albicans* genome.

Title manipulation – Natural Sciences



SCIENTIFIC REPORTS

OPEN

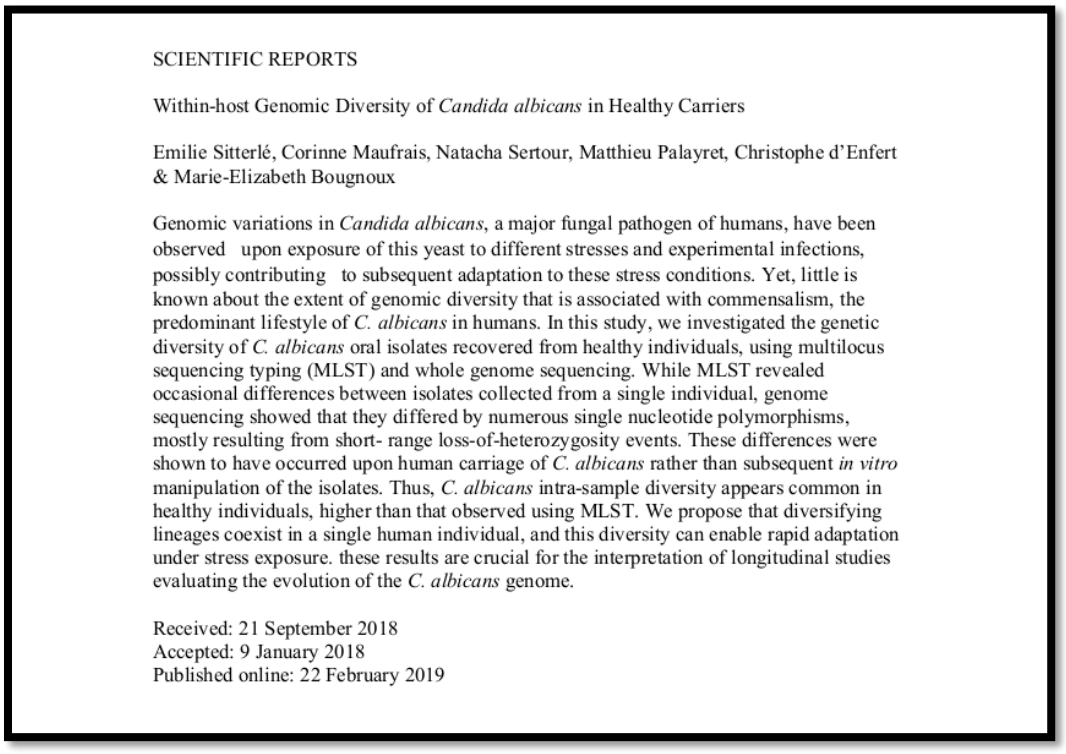
Humans *do* contain multitudes: Genomic Diversity of *Candida albicans*

Emilie Sitterlé[1,2,3], Corinne Maufrais[4], Natacha Sertour[1], Matthieu Palayret[5], Christophe d'Enfert[1] & Marie-Elisabeth Bougnoux[1,3]

Received: 21 September 2018
Accepted: 9 January 2019
Published online: 22 February 2019

Genomic variations in *Candida albicans*, a major fungal pathogen of humans, have been observed upon exposure of this yeast to different stresses and experimental infections, possibly contributing to subsequent adaptation to these stress conditions. Yet, little is known about the extent of genomic diversity that is associated with commensalism, the predominant lifestyle of *C. albicans* in humans. In this study, we investigated the genetic diversity of *C. albicans* oral isolates recovered from healthy individuals, using multilocus sequencing typing (MLST) and whole genome sequencing. While MLST revealed occasional differences between isolates collected from a single individual, genome sequencing showed that they differed by numerous single nucleotide polymorphisms, mostly resulting from short-range loss-of-heterozygosity events. These differences were shown to have occurred upon human carriage of *C. albicans* rather than subsequent *in vitro* manipulation of the isolates. Thus, *C. albicans* intra-sample diversity appears common in healthy individuals, higher than that observed using MLST. We propose that diversifying lineages coexist in a single human individual, and this diversity can enable rapid adaptation under stress exposure. These results are crucial for the interpretation of longitudinal studies evaluating the evolution of the *C. albicans* genome.

Format manipulation – Natural sciences



SCIENTIFIC REPORTS

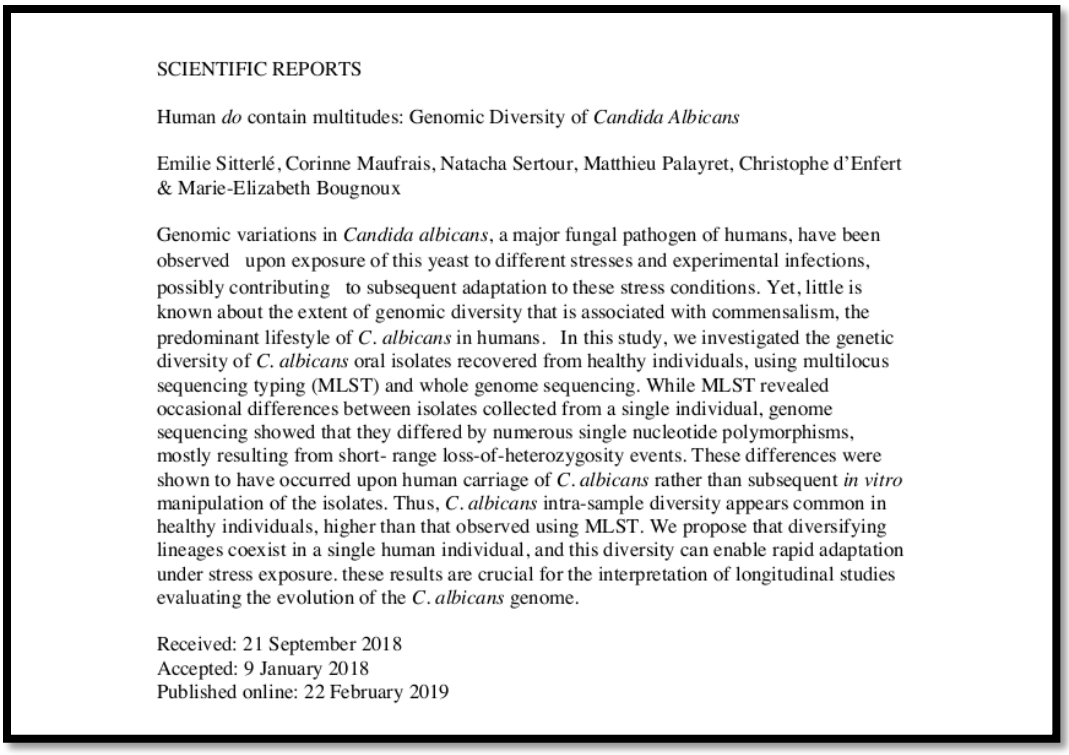
Within-host Genomic Diversity of *Candida albicans* in Healthy Carriers

Emilie Sitterlé, Corinne Maufrais, Natacha Sertour, Matthieu Palayret, Christophe d'Enfert & Marie-Elizabeth Bougnoux

Genomic variations in *Candida albicans*, a major fungal pathogen of humans, have been observed upon exposure of this yeast to different stresses and experimental infections, possibly contributing to subsequent adaptation to these stress conditions. Yet, little is known about the extent of genomic diversity that is associated with commensalism, the predominant lifestyle of *C. albicans* in humans. In this study, we investigated the genetic diversity of *C. albicans* oral isolates recovered from healthy individuals, using multilocus sequencing typing (MLST) and whole genome sequencing. While MLST revealed occasional differences between isolates collected from a single individual, genome sequencing showed that they differed by numerous single nucleotide polymorphisms, mostly resulting from short- range loss-of-heterozygosity events. These differences were shown to have occurred upon human carriage of *C. albicans* rather than subsequent *in vitro* manipulation of the isolates. Thus, *C. albicans* intra-sample diversity appears common in healthy individuals, higher than that observed using MLST. We propose that diversifying lineages coexist in a single human individual, and this diversity can enable rapid adaptation under stress exposure. these results are crucial for the interpretation of longitudinal studies evaluating the evolution of the *C. albicans* genome.

Received: 21 September 2018
Accepted: 9 January 2018
Published online: 22 February 2019

Title and format manipulation – Natural sciences



SCIENTIFIC REPORTS

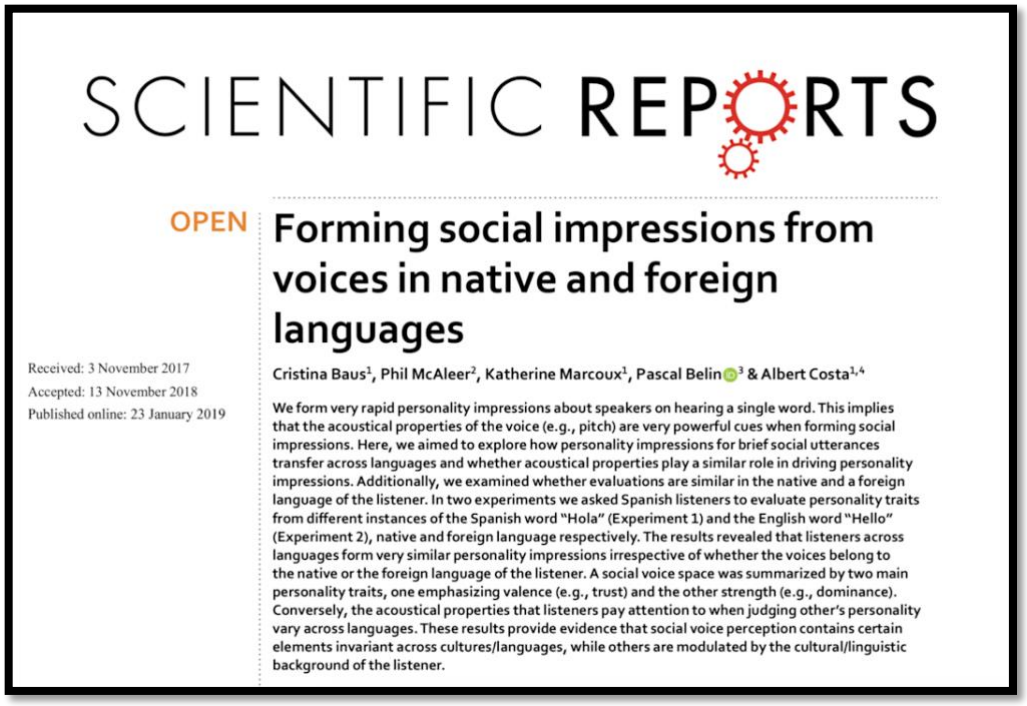
Human *do* contain multitudes: Genomic Diversity of *Candida Albicans*

Emilie Sitterlé, Corinne Maufrais, Natacha Sertour, Matthieu Palayret, Christophe d'Enfert & Marie-Elizabeth Bougnoux

Genomic variations in *Candida albicans*, a major fungal pathogen of humans, have been observed upon exposure of this yeast to different stresses and experimental infections, possibly contributing to subsequent adaptation to these stress conditions. Yet, little is known about the extent of genomic diversity that is associated with commensalism, the predominant lifestyle of *C. albicans* in humans. In this study, we investigated the genetic diversity of *C. albicans* oral isolates recovered from healthy individuals, using multilocus sequencing typing (MLST) and whole genome sequencing. While MLST revealed occasional differences between isolates collected from a single individual, genome sequencing showed that they differed by numerous single nucleotide polymorphisms, mostly resulting from short- range loss-of-heterozygosity events. These differences were shown to have occurred upon human carriage of *C. albicans* rather than subsequent *in vitro* manipulation of the isolates. Thus, *C. albicans* intra-sample diversity appears common in healthy individuals, higher than that observed using MLST. We propose that diversifying lineages coexist in a single human individual, and this diversity can enable rapid adaptation under stress exposure. these results are crucial for the interpretation of longitudinal studies evaluating the evolution of the *C. albicans* genome.

Received: 21 September 2018
Accepted: 9 January 2018
Published online: 22 February 2019

Baseline – Social Sciences



SCIENTIFIC REPORTS

OPEN

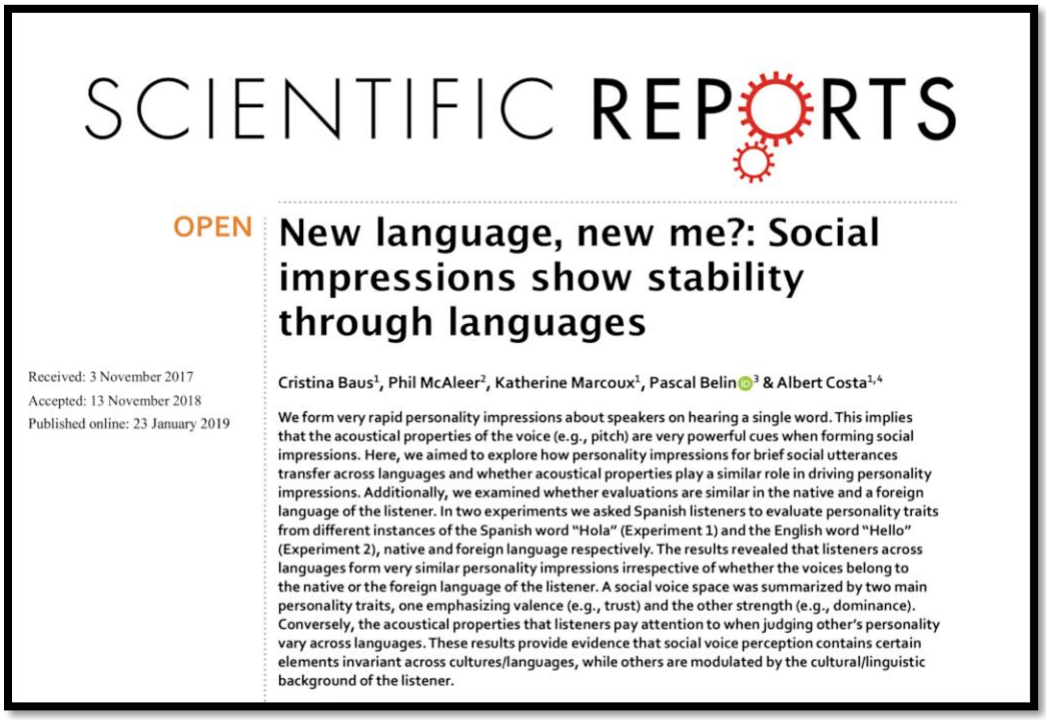
Forming social impressions from voices in native and foreign languages

Received: 3 November 2017
Accepted: 13 November 2018
Published online: 23 January 2019

Cristina Baus[1], Phil McAleer[2], Katherine Marcoux[1], Pascal Belin[3] & Albert Costa[1,4]

We form very rapid personality impressions about speakers on hearing a single word. This implies that the acoustical properties of the voice (e.g., pitch) are very powerful cues when forming social impressions. Here, we aimed to explore how personality impressions for brief social utterances transfer across languages and whether acoustical properties play a similar role in driving personality impressions. Additionally, we examined whether evaluations are similar in the native and a foreign language of the listener. In two experiments we asked Spanish listeners to evaluate personality traits from different instances of the Spanish word "Hola" (Experiment 1) and the English word "Hello" (Experiment 2), native and foreign language respectively. The results revealed that listeners across languages form very similar personality impressions irrespective of whether the voices belong to the native or the foreign language of the listener. A social voice space was summarized by two main personality traits, one emphasizing valence (e.g., trust) and the other strength (e.g., dominance). Conversely, the acoustical properties that listeners pay attention to when judging other's personality vary across languages. These results provide evidence that social voice perception contains certain elements invariant across cultures/languages, while others are modulated by the cultural/linguistic background of the listener.

Title manipulation – Social Sciences



SCIENTIFIC REPORTS

OPEN

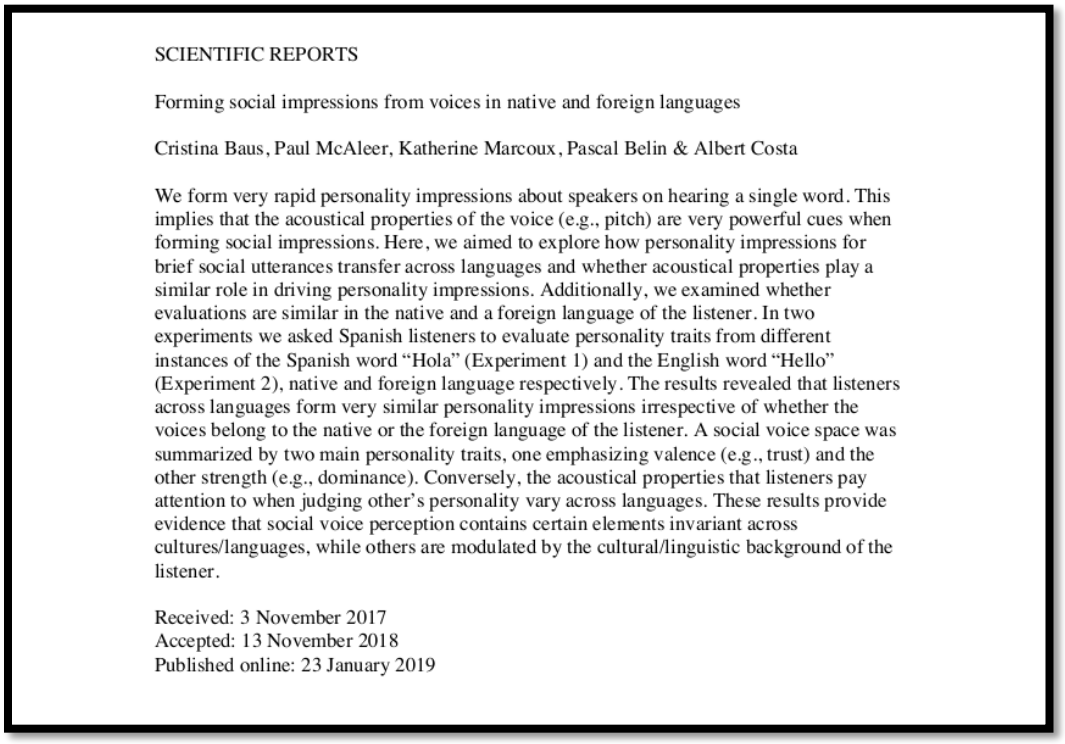
New language, new me?: Social impressions show stability through languages

Received: 3 November 2017
Accepted: 13 November 2018
Published online: 23 January 2019

Cristina Baus[1], Phil McAleer[2], Katherine Marcoux[1], Pascal Belin[3] & Albert Costa[1,4]

We form very rapid personality impressions about speakers on hearing a single word. This implies that the acoustical properties of the voice (e.g., pitch) are very powerful cues when forming social impressions. Here, we aimed to explore how personality impressions for brief social utterances transfer across languages and whether acoustical properties play a similar role in driving personality impressions. Additionally, we examined whether evaluations are similar in the native and a foreign language of the listener. In two experiments we asked Spanish listeners to evaluate personality traits from different instances of the Spanish word "Hola" (Experiment 1) and the English word "Hello" (Experiment 2), native and foreign language respectively. The results revealed that listeners across languages form very similar personality impressions irrespective of whether the voices belong to the native or the foreign language of the listener. A social voice space was summarized by two main personality traits, one emphasizing valence (e.g., trust) and the other strength (e.g., dominance). Conversely, the acoustical properties that listeners pay attention to when judging other's personality vary across languages. These results provide evidence that social voice perception contains certain elements invariant across cultures/languages, while others are modulated by the cultural/linguistic background of the listener.

Format manipulation – Social sciences



SCIENTIFIC REPORTS

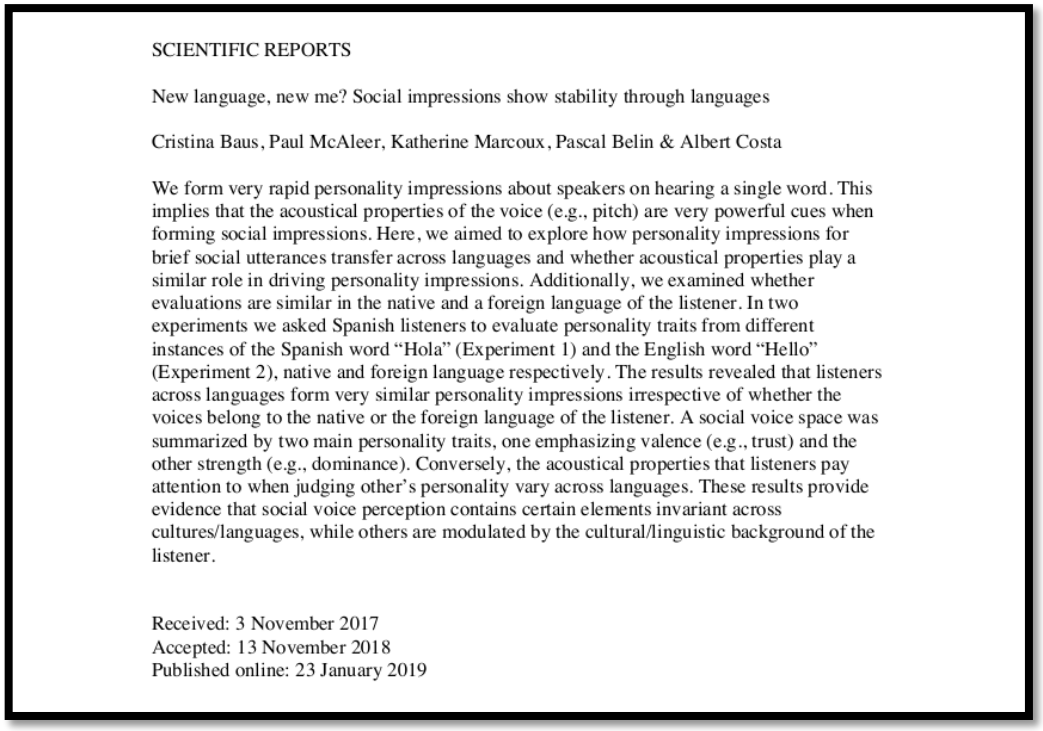
Forming social impressions from voices in native and foreign languages

Cristina Baus, Paul McAleer, Katherine Marcoux, Pascal Belin & Albert Costa

We form very rapid personality impressions about speakers on hearing a single word. This implies that the acoustical properties of the voice (e.g., pitch) are very powerful cues when forming social impressions. Here, we aimed to explore how personality impressions for brief social utterances transfer across languages and whether acoustical properties play a similar role in driving personality impressions. Additionally, we examined whether evaluations are similar in the native and a foreign language of the listener. In two experiments we asked Spanish listeners to evaluate personality traits from different instances of the Spanish word "Hola" (Experiment 1) and the English word "Hello" (Experiment 2), native and foreign language respectively. The results revealed that listeners across languages form very similar personality impressions irrespective of whether the voices belong to the native or the foreign language of the listener. A social voice space was summarized by two main personality traits, one emphasizing valence (e.g., trust) and the other strength (e.g., dominance). Conversely, the acoustical properties that listeners pay attention to when judging other's personality vary across languages. These results provide evidence that social voice perception contains certain elements invariant across cultures/languages, while others are modulated by the cultural/linguistic background of the listener.

Received: 3 November 2017
Accepted: 13 November 2018
Published online: 23 January 2019

Title and format manipulation – Social Sciences



SCIENTIFIC REPORTS

New language, new me? Social impressions show stability through languages

Cristina Baus, Paul McAleer, Katherine Marcoux, Pascal Belin & Albert Costa

We form very rapid personality impressions about speakers on hearing a single word. This implies that the acoustical properties of the voice (e.g., pitch) are very powerful cues when forming social impressions. Here, we aimed to explore how personality impressions for brief social utterances transfer across languages and whether acoustical properties play a similar role in driving personality impressions. Additionally, we examined whether evaluations are similar in the native and a foreign language of the listener. In two experiments we asked Spanish listeners to evaluate personality traits from different instances of the Spanish word "Hola" (Experiment 1) and the English word "Hello" (Experiment 2), native and foreign language respectively. The results revealed that listeners across languages form very similar personality impressions irrespective of whether the voices belong to the native or the foreign language of the listener. A social voice space was summarized by two main personality traits, one emphasizing valence (e.g., trust) and the other strength (e.g., dominance). Conversely, the acoustical properties that listeners pay attention to when judging other's personality vary across languages. These results provide evidence that social voice perception contains certain elements invariant across cultures/languages, while others are modulated by the cultural/linguistic background of the listener.

Received: 3 November 2017
Accepted: 13 November 2018
Published online: 23 January 2019

The challenges of measuring aesthetic experiences

Javiera Paredes* and David Torres Irribarra

Escuela de Psicología, Pontificia Universidad Católica de Chile

Author note

Javiera Paredes https://orcid.org/0000-0003-1828-3005

Correspondence concerning this article should be addressed to Javiera Paredes, e-mail: jnparede@uc.cl

Abstract

Researchers in the last decades have been actively pursuing the development of quantitative measures of aesthetic experiences. Focusing on concepts within aesthetic experience in the arts, we reviewed 13 articles that reported the creation of an instrument to measure constructs such as aesthetic appreciation, aesthetic evaluation, among others. We identify three challenges related to the construction and validation of these instruments: a conceptual articulation of aesthetic experiences, the determination of observables, and the examination and documentation of instrument quality. Regarding the first challenge, we identified opportunities to define the concept to be measured more clearly through the process of clarification and differentiation from other concepts within the discipline. The second challenge relates to the need for a stronger link between the attributes of interest articulated in the theoretical frameworks and the proposed observations intended to gather evidence about them. This link is necessary to create instruments that more accurately characterize the theoretical proposals of the authors. Finally, to surmount parts of the third challenge, reporting of analyses about validity must provide greater clarity regarding the design choices made by the researchers. Ideally, these choices should be informed by literature on their discipline and current perspectives on assessment practices. We suggest avenues of improvement that could lead to clearer and better-supported interpretations of the scores, basing them on a better-defined characterization of the attributes, a clearer specification of the uses of the instrument, and a more explicit discussion of the design decisions that shaped the process of instrument development.

Keywords: aesthetics, measurement, challenges, quantitative

The challenges of measuring aesthetic experience

Quantifying the subjective experience of art is undoubtedly an ambitious goal and creating an assessment instrument that can meaningfully summarize these experiences quantitatively raises both theoretical and methodological difficulties. Despite these conceptual and practical complexities, empirical researchers on aesthetic appreciation have been prolific in developing assessment instruments aimed precisely at the production of quantitative estimates of this, at least to some extent, ineffable attribute.

Until the end of the nineteenth century, although aesthetics was a well-developed subject in philosophy (Baumgartner et al., 2006; Kant, 1914), there was no interest or necessity to measure how beautiful or pleasurable the artistic object was for the people who experienced it. Some philosophers and scholars classified the aesthetic experience (AE) as tough to grasp and conceptualize since art and its expression does not necessarily follow representational or logical rules (Wittgenstein, 1961/2002).

With the arrival of psychology as a field of study independent of philosophy, the interest in studying the aesthetic experience continued. This field focused on solving the problem of approaching the conceptualization and assessment of aesthetic experiences at the psychological level. The attempt to characterize AE using questionnaires to conduct a quantitative analysis of the results is predicated on the idea that AE is either a quantitative attribute—a non-trivial assumption to make or support (Michell, 2009, 2012) – or at the very least, that it is sensible to attempt to model it in such a way (Author, 2021).

One of the first examples of the attempt to measure AE as a quantitative attribute is Israeli's (1928) article on the psychological affective reactions to visual art. In this study, participants were subject to a free emotional reaction experiment to color reproductions of paintings. Since then,

there have been many studies in several disciplines that aim to measure AE in a quantitative form (e.g., Andringa, 1996; Asmus, 1985; Wanzer et al., 2018). The method through which this quantification is attempted usually takes the form of questionnaires presented to the participant after participating or watching some sort of artistic expression. Visual art tends to be the predominant object chosen to display to participants.

Even though the instruments that have been developed to assess AE present differences, the authors of each instrument necessarily had to, either tacitly or explicitly, deal with a common set of conceptual and technical challenges in order to guide the process of instrument development and the interpretation of its results.

A first challenge relates to the definition of the attribute of interest, or stated more concretely: what is it precisely that the authors aim to measure? This question is, of course, central to the process of assessment development, as it defines the scope of the attribute of interest, its structure (i.e., does it make sense to consider it a quantity?), and its dimensionality (i.e., is it a single, unitary attribute or a myriad of different, specific attributes that somehow are integrated into a whole?). Questions about the structure of the attribute are challenging not only due to the diverse variety of theoretical approaches to the topic but also due to the unobservable or latent nature of the attributes of interest.

A second challenge relates then to the latter point: where or how can we observe AE? Once a definition of AE has been selected, and therefore its structure and dimensionality, the development of an assessment demands the creation or selection of a set of items that will be considered as providing the observable elements that will serve as the basis for our inference regarding the latent attribute of AE. Depending on the kind of instrument being developed, these items may be designed to collect self-reports or potentially standardized observations by raters. In

any case, they must have defined which behaviors or declared attitudes and beliefs are seen as informative to specific dimension or dimensions posited by the definition of AE that has been adopted.

A third and final challenge relates to the empirical assessment of the developed instrument. Once the definitional aspects are settled, and they have guided the selection of specific items deemed observable to produce information about a person's AE, then we face the challenge of documenting the quality of the results of such an instrument in order to support our claims that the produced scores are interpretable as quantitative measures of the attribute that we have defined as the measure of our instrument.

First Challenge: Conceptual articulation of Aesthetic Experiences (AE)

Conceptual clarity regarding what we want to measure is a necessary condition for a meaningful construction of interpretation of a measurement result (Wilson, 2004). The interpretation of quantitative results is predicated on a coherent alignment between theoretical definitions and methodological choices. If quantitative assessments of AE are to be meaningfully discussed in the literature, then clarity regarding their different theoretical foundations becomes a necessity to correctly understand the extent to which results, potentially sharing the same label, could be addressing very different attributes.

Although there is a healthy amount of work regarding the measurement of AE, the conceptualization of the phenomena is diverse. Any effort aimed at measuring and quantifying AE will necessarily be rooted in theoretical concepts developed in psychology and philosophy such as aesthetic perception and evaluation (Andringa, 1996; Hagtvedt et al., 2008), aesthetic experiences

(Silvia & Nusbaum, 2011; Vukadinovic & Markovic, 2012), aesthetic appreciation (Hager et al., 2012; Rowold, 2008), among others.

An illustration of the challenges faced when attempting to measure constructs in this area of the social sciences can be exemplified with the concept of emotion. The definitions made by different scholars are varied and do not have the same theoretical or philosophical foundations. For example, on one hand, there is an evaluative outlook on emotion that emphasizes the notion of appraisal, a concept that would explain how emotions are elicited. Appraisal is defined as the process in which the significance and importance of a situation for a person are made manifest. This process gives way to attraction or aversion. Emotions then would be a felt tendency toward any case that the person appraises as beneficial (or harmful) (Scarantino & de Sousa, 2018). Another approach to emotion is emotion as a feeling. In this tradition, emotion is thought of as a class of feelings, a specific kind of subjective experience. The classical proposal for this idea is the James-Lange theory. In this theory, emotions are feelings articulated by perceptions of changes in physiological conditions relating to the body's autonomic and motor functions (James, 1884; Scarantino & de Sousa, 2018). Considering only these brief definitions, we can identify some critical distinctions between these two approaches to emotion. To a non-expert reader or consumer of results from the different assessment instruments rooted in these different theoretical perspectives, this is not apparent if these differences are not made explicit by other authors that use those concepts for their proposals. The lack of differentiation between approaches to emotion, in this particular case, could cause miscommunication between researchers that may be approaching emotion in different ways but think that they are comparable measurement results.

Second Challenge: Determination of observables

Ideally, how and attributes are characterized should not only be internally coherent but also should relate to other attributes and phenomena within the larger area of study. Historically, this network of relations was understood as a nomological network, following the foundational work of Cronbach and Meehl (1955) on validity. The construction of this ideal nomological network—with clear and differentiated concepts and their relations—is still a goal far from achieved within psychology. However, the expectations that attributes under study that are treated as theoretically distinct should be specified in terms of their differences (or their potential relation) is a more modest goal that historically underpins treatments of predictive and concurrent validity (Cronbach & Meehl, 1955) and current approaches to convergent and discriminant validity evidence (Maul, 2013). This exercise would help study different and distinctive attributes in any discipline, allowing researchers to distinguish concepts with the same name but have distinct functional attributes (Wilson, 2004). A consequence of tackling the challenges of measuring an attribute is that researchers need to identify theoretically identify ways of empirically observing it and, therefore, that can serve as a basis for making inferences about it.

This ideal path between concepts, constructs, and inferences from measurement is not achieved successfully in all situations. For example, if two different attributes share the same observations, indetermination problems quickly appear (Salzberger, 2013). More precisely, two possible issues are related to indetermination, one associated with the existence of two different concepts that have the same observables and therefore yield the same inferences. Another is related to having one concept that has several observables that do not predict the phenomena consistently. These two possible problems and their consequences are of particular importance for science in general, especially in the social sciences. Going back to the issues introduced in the case of

aesthetic experiences, where conceptualization is diverse and measurement opaque, how does psychology create a robust research program when constructs that are defined as theoretically distinct are occasionally associated with the same set of observables? In order to do so, we need to tackle both the challenge of developing a shared understanding of the construct of AE (or at least clarity regarding its different characterizations) and the challenge of producing instruments that reflect these definitions.

This article will focus on five concepts presented in the literature: aesthetic experiences, aesthetic appreciation, aesthetic emotions, aesthetic evaluation, and aesthetic perception/judgment. We propose that a central problem related to this challenge is the disconnect between the conceptual articulation of AE concepts and the determination of its observables, which has led to a proliferation of competing definitions of AE that, despite their differences, are measured using similar questionnaires. A glance of the item content of different articles that developed measures for distinct phenomena, such as aesthetic emotions or aesthetic appreciation, provides related item categories and items. A study that constructed a measure of aesthetic perception and evaluation (Hagtvedt et al., 2008) has an item regarding the creativity of the artwork (e.g., How well do these attributes describe this artwork?: "Creative") that is similar (e.g., "This painting features a high level of creativity") to another study that constructed an instrument aimed at measuring aesthetic appreciation (Hager et al., 2012). While there is some reference to instruments in each researcher's specific discipline, such as emotion surveys in psychological approaches measuring AE (Rowold, 2008; Schindler et al., 2017), there is little cross-disciplinary citation in these studies.

Third Challenge: Examining and documenting instrument quality

Even in the cases where the two previous challenges have been successfully addressed, clearly characterizing the construct of AE and establishing a clear connection between it and a set of specific observables, a third issue relates to empirically documenting the performance of an instrument. The quality of an assessment instrument has traditionally been judged both in terms of the evidence to support the validity of the inferences that are made based on the instrument as well as the evidence that its precision or reliability can adequately support such inferences (American Educational Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 1999, 2014). More recently, the extent to which the instrument performs fairly across all potential persons in a population has also been added as a criterion to assess the instrument's quality (AERA et al., 2014).

The current *Standards for Educational and Psychological Testing* (henceforth the Standards; AERA et al., 2014) present a framework for documenting the quality of the inferences that can be supported by an assessment instrument based on an examination of validity, precision/reliability, and fairness. The Standards present a framework with a unified conception of validity, emphasizing the need to document the theoretical rationale and empirical evidence used to support the intended uses of any assessment instrument. The support for the intended uses can be gathered by examining five different evidence sources: based on content, based on response processes, based on internal structure, based on relation to other variables, and finally based on consequences. A second key element that must be documented is the degree of precision of the instrument, known usually as reliability, which requires the determination of the amount of measurement error associated with the instruments' results. Finally, the latest version of the Standards included for the first-time fairness as a third essential aspect to be documented, explicitly

promoting the active documentation of lack of bias in the instrument that could help or hinder relevant subgroups.

A literature review reveals that the level of documentation regarding the empirical evidence collected to inform validity, reliability, and fairness varies wildly, with some areas, and specific kinds of evidence within those areas, being better represented than others. When results are documented, the evidence points to problems in the instrument's performance, with cases where, for instance, the measures of the reliability of the instruments were not optimal or final solutions of factor analysis were disregarded.

The role of measurement in AE research

As it is probably clear to the reader, these three challenges are not independent of each other, as any good measurement instrument must successfully articulate them to produce results meaningfully and reliably. Based on these three interconnected challenges, we are interested in evaluating the extent to which instruments aimed at quantifying aesthetic experiences can support a meaningful interpretation of their results, a necessary precondition to fruitfully conduct empirical research based on said results. To achieve this, we review currently available questionnaires related to aesthetic experiences in the fields of experimental aesthetics, psychology, and philosophy and examine their similarities, differences, and psychometric properties. Specifically, we analyze the conceptual definitions in which they are based, the dimensional structure that their developers posited, the observable indicators that were selected to create them, and, finally, what steps were followed to document their validity, reliability, and fairness (AERA et al., 2014).

This article reviewed 13 questionnaires related to aesthetic experiences, focusing on the conceptual definition, instrument development, and how the theoretical framework presented in

each study is or is not congruent with the developed measurement instruments. The revision will focus on measures that have their objects exclusively on artistic creation. Based on this analysis, we argue that a central problem in the aesthetic experiences field is the disconnect between theoretical developments and their measures. This disconnect has led to a proliferation of competing definitions of AE that, despite their differences, are measured using similar questionnaires. The evidence makes the case that improving how AE is measured is a crucial step to generate consistent and more fruitful empirical research in this area.

Methodology

To select the articles reviewed in this study, we searched websites such as Google Scholar and the library's online service of a large South American university. The keywords used to perform this search were "measure," "instrument," "aesthetic appreciation," "aesthetics." Results were selected with one single criterion; the articles must have explicitly reported the construction of a measure related to aesthetic experiences. Empirical research articles and measurement development articles were both selected when this criterion was fulfilled. As the articles reporting the construction of aesthetic appreciation measures were few, the search criterion was extended to concepts such as "aesthetic experiences," "aesthetic emotions," "aesthetic perception." The selection criterion was maintained, and the search stopped when we considered that the different concepts were represented in at least two separate articles.

Results

Conceptual articulation of Aesthetic Experiences (AE)

The conceptual articulation of aesthetic experiences in the articles reviewed share some

similarities that group them in three different approaches: (a) articles that do not provide an explicit definition or theoretical framework of aesthetic experiences, (b) articles that by its nature – experimental – give a brief definition of aesthetic experiences and, (c) articles that provide a theoretical framework of AE in a variety of degrees.

Approaches to characterizing AE

Of the thirteen articles, two did not provide any definition or provided an implicit definition of the concept utilized for the construction or use of the AE measurement (Asmus, 1985; Israeli, 1928). In the oldest article, the lack of definition and theoretical framework could be attributed to a lack of description of aesthetic experiences within psychology (Israeli, 1928). Therefore, this article could be viewed in psychological science as the empirical starting point of a construct definition for the concept of aesthetic/affective reactions to art. However, there were ample theoretical developments in other areas such as philosophy or aesthetics that could have been used to provide a framework to understand the researcher's work. The source of the absence of a theoretical definition of concepts and constructs could also be traced back to the structure of scientific articles at the time, where the focus used to be on the results of an experimental procedure instead of structuring a sound theoretical background to the empirical work (Daston & Galison, 2008).

The earliest contemporary article that deals with the construction of measures of aesthetic emotions tried to assess the affective-emotional states that participants developed in response to musical stimuli (Asmus, 1985). The author does not explicitly define what an affective-emotional state is, nor does it distinguish it from other similar concepts. Although one could argue that the interpretation for these proposed states could be inferred through previous work that he mentions as the basis for this research, the reader will not know whether the author entirely agrees with the

epistemological stance that the cited work backs. The article also refers to a duality of concepts; the "feelingful/emotional" emotional states (Asmus, 1985, p. 19). In contrast with the other definitions of similar concepts, feelings are conceptualized as the experiential or subjective component of emotions (Gross, 2010; Scherer, 2005). On the other hand, emotions are a complex articulation of different elements, such as bodily experiences and feelings (Scherer, 2005). From the reading of the article, it's not possible to disentangle the difference – if it exists – between these concepts.

The second category presented is the articles that provide a brief definition of aesthetic experiences by their empirical focus. The earliest of them is a study on the effects of narrative distance on readers' emotional involvement and response (Andringa, 1996). This study was one of the first research articles that involved the evaluation of the cognitive perception of the artwork and emotional variables related to empathy. In this proposal, there are two types of emotion; emotions relating to the participant involvement and empathetic response processes (F-emotions), and emotions that relate to the fiction as an artistic object (A-emotions) (Tan, 1994). A similar case is found in a study that deals with physiological aesthetic responses and self-reported aesthetic perception in a museum setting (Tröndle & Tschacher, 2012; Tschacher et al., 2012). Tschacher et al. (2012) discuss the theoretical and empirical need to assess and consider art perception as a "bottom-up" process and how physiological data could help achieve this aim. In these studies, there is a focus on the embodied approach of aesthetic experience that entails the cognitive aspect of this phenomenon and the emotional and bodily components that would make up a complex psychological process (Tröndle & Tschacher, 2012; Tschacher et al., 2012).

Another similar article is Silvia and Nusbaum's (2011). The authors describe the link between aesthetic chills and other unusual aesthetic states, self-reported by their participants. The

article distinguishes briefly between three strands of aesthetic experience studies: studies in aesthetic emotions, experimental aesthetics, and unusual aesthetic states. These unique aesthetic states are experientially described as chills, absorption, or feeling touched, for example (Silvia & Nusbaum, 2011). The focus of the study is to extend the literature on individual differences regarding these aesthetic states. While there is some reference to studies of the discipline that focuses on individual differences or experimental approaches, the article does not detail the theoretical basis of the phenomena or the empirical developments of the field.

Table 1

List of article authors and their study objects.

Authors	Study object
Israeli (1928)	Affective reactions to paintings
Asmus (1985)	Affective responses to music
Andringa (1996)	Appreciation and perception of emotional states
Konijn (1999)	Aesthetic emotions
Stamatopoulou (2004)	Aesthetic experiences
Hagtvedt, Hagtvedt & Patrick (2008)	Perception and evaluation of visual art
Zentner, Grandjean & Scherer (2008)	Aesthetic emotions
Rowold (2008)	Aesthetic appreciation
Silvia & Nusbaum (2011)	Aesthetic experience
Vukadinović & Markovic (2011)	Aesthetic experience of dance
Tschacher, Greenwood, Kirchberg, Wintzertith, van der Berg & Trondle (2012)	Aesthetic perception
Hager, Hagemann, Danner & Schankin (2012)	Aesthetic appreciation
Schindler et al. (2017)	Aesthetic emotions
Wanzer, Procter Finley, Zarian & Cortez	Aesthetic appreciation

Finally, the third category encompasses the articles that provide a theoretical framework of AE, which we will discuss below. These articles distinguish two or more components of aesthetic experiences and describe them in various levels of detail.

The broad structural components of AE

We have identified three broad components that comprise aesthetic experience in most articles that have a detailed definition of the phenomenon: cognitive, affective, and supplementary components. It is essential to notice that sometimes these components are labeled something else by researchers (e.g., epistemic emotions vs. cognitive components). While recognizing this, we decided to label them in one of the three components since they share broadly the same aim.

In the case of cognitive components of aesthetic experience, most researchers aim to access the evaluative, rational, and thought-related aspects of the aesthetic experience. A common aim of this component is to describe the evaluative process that a person experiences when observing a painting or another artwork (e.g., Hager et al., 2012). Two proposals within the cognitive elements of AE can be distinguished by comparing the theoretical frameworks of the articles reviewed. First, the approximation to cognitive elements of AE thought as the evaluation of the artwork based on the perception of complexity and also the level of surprise/curiosity/novelty caused by the artwork and reported by the participant (Hagtvedt et al., 2008; Konijn, 1999; Schindler et al., 2017; Vukadinovic & Markovic, 2012; Wanzer et al., 2018). These components of the cognitive response in AE are usually related to other psychological phenomena like a higher information processing rate (Konijn, 1999), the highlight of the role of expertise or prior experience (Konijn, 1999; Wanzer et al., 2018), and their role in the final process of cognitive evaluation of the artwork.

The other framework used to articulate the cognitive components of aesthetic experience is

led by scholars working with concepts such as aesthetic experience (Stamatopoulou, 2004) and aesthetic appreciation (Hager et al., 2012; Rowold, 2008). The framework presented in these three articles is very similar and proposes sequential cognitive processing with several stages. Usually, this process starts with perceptual or sensory appraisals followed by diverse processes, such as the cognitive restructuring of the information that allows the explicit elaboration of the meaning of the artwork in a loop process (Stamatopoulou, 2004). In other cases, it is proposed that after the perceptual and sensorial stage, elements such as personal interests or knowledge lead the evaluative process and judgment of an artwork (Leder et al., 2004; Rowold, 2008). Hager et al. (2012) propose that this cognitive processing is conscious and that the perceptive and sensory stage is followed by a process of explicit classification where the artwork's style and context are processed. Following Leder et al. (2004), this model emphasizes the importance of expertise and personal interests for the final extraction of the meaning of the artwork and its evaluation.

There are some differences between these two proposals. Considering the first grouping, which emphasizes the evaluative aspect of AE, Konijn (1999) approaches the cognitive aspect of aesthetic emotions as one directly related to the sociological status of the person as the reason why they partake in artistic experiences. On the other hand, Hagtvedt et al. (2008) highlight that this evaluation and the characteristics of aesthetic experiences may belong to the object itself or the emotions elicited by the object assessed. These processes, they argue, are interlinked in experience. Although considering elements in the evaluation of the artworks such as novelty and surprise could lead the reader to think that the cognitive component is related to a significant hedonistic part of aesthetic experiences, Vukadinovic and Markovic (2012) reject such stance. Another proposal (Schindler et al., 2017) argues that the components involved in the cognitive articulation of aesthetic emotions are the basis of the search for meaning, especially the elements related to

curiosity, interest, and surprise. These elements could lead to a feeling of insight or knowing when experiencing aesthetic objects.

Regarding the second grouping that focuses on sequential cognitive processing, Stamatopoulou (2004) suggests that the perceptual and sensory cues, along with the cognitive elements in the following information stage, act like a synthesized dynamic complex. The features presented in this model interact in something akin to a loop, where the process works as a whole. The author also states that the extraction of meaning in the primary processes, likely after the perceptual and sensory cues, cannot be easily verbalized. Verbalizing the possible meaning of the artwork is only possible after elaborating the cognitive restructuring of the whole evaluative process. Rowold's (2008) and Hager et al.'s (2012) theoretical background regarding the cognitive components of AE comes from the same initial framework proposed by Leder et al. (2004). The main differences between the articles are the detail in which each team of authors describes the model and the emphasis made by Hager et al. (2012) on the conscious nature of the cognitive processes related to the evaluation of the artwork.

A more concise description can be found in Table 2. As we will see in the descriptions of the theoretical emotional and supplementary components of AE, the cognitive components are the elements that are the most developed theoretically and, therefore, offer more details about the way that the conceptual models are articulated to produce the aesthetic experience.

Table 2

Summary of the models of cognitive elements of aesthetic experiences.

Models	Similarities	Differences
1. Evaluation of the artwork based on the perception of	Associated with prior experience (Konijn,	Related to the sociological status of the person. Also

complexity or level of surprise/novelty (Konijn, 1999; Schindler et al., 2017; Vukadinovic & Markovic, 2012; Wanzer et al., 2018)	1999; Wanzer et al., 2018)	associated with a higher processing rate (Konijn, 1999). Evaluation may belong to art object or the emotions elicited by the object (Hagtvedt et al., 2008) Novelty and surprise are not related to a hedonistic element of AE (Vukadinovic & Markovic, 2012) Evaluation key to finding meaning. This process could lead to insight experiences (Schindler et al., 2017)
2. Sequential cognitive processing (Hager et al., 2012; Rowold, 2008; Stamatopoulou, 2004)	The process starts with perceptual or sensorial appraisals followed by different processes (Hager et al., 2012; Rowold, 2008; Stamatopoulou, 2004). Some models emphasize the role of personal experience (Hager et al., 2012; Rowold, 2008)	Perceptual and sensorial cues, along with cognitive processing, form a dynamic complex loop. Discursive meaning appears very late in processing (Stamatopoulou, 2004) The sequential cognitive processing is of conscious nature (Hager et al., 2012)

Regarding the emotional components of AE, three theoretical proposals are presented in the reviewed articles. One relates to the conception of emotion in AE as involvement or empathy when given the art object or play (Andringa, 1996; Konijn, 1999; Stamatopoulou, 2004). Usually, a brief explanation of the meaning of empathy is presented, with some differences. In the case of

the articles authored by Andringa (1996) and Konijn (1999), they follow Tan (1994), distinguishing emotions triggered by the participant's involvement and empathetic processes and emotions that relate to the fiction or object of art. In Konijn (1999), the emphasis is on the task or personal associations concerning aesthetic emotions. Stamatopoulou (2004) elaborates the concept of empathy with the idea of emotional distance in aesthetic experiences. Emotional closeness is generated through empathy and the consequent distance when we are not experiencing empathy, leading to a first emotional state of relief or catharsis. This process seems to depend on the motivational framing in which the participant is immersed.

The second theoretical proposal regarding emotional components in AE is articulated through concepts such as valence or arousal (Hagtvedt et al., 2008; Rowold, 2008; Vukadinovic & Markovic, 2012; Wanzer et al., 2018). The valence of emotion usually refers to how positive or negative the reaction to the artwork is. On the other hand, arousal is defined as a feeling state of activation of the participant, which can range from drowsiness to frantic activity (Hagtvedt et al., 2008). The differences between the articles on the formulation of the valence/arousal emotional component are of varied nature. In one study, valence and arousal are considered to influence the artwork's information processing, which affects the final evaluation made by the participant (Hagtvedt et al., 2008). The valence/arousal component in the theoretical articulation of Rowold (2008) also focuses on the link that emotions and pleasure have in aesthetic appreciation processes. It is also vital for the author to consider the personal associations that the object may have to the participant, which could modify the emotions related to the artwork. This focus on the personal experience or emotional recognition of something experienced beforehand as vital to the aesthetic emotion is also highlighted by Vukadinovic & Markovic (2012). A different approach is taken by Wanzer et al. (2018) that proposes that aesthetic emotions are the result of aesthetic appreciation

but not part of its process.

Finally, the last proposal on the emotional components of AE comes from the articles of Zentner, Grandjean & Scherer (2008), and Schindler et al. (2017). Both studies advocate for a complex view of emotions that goes beyond valence and arousal. Emotion in Zentner et al.'s (2008) article is thought of as the subjective component of emotion; this is how the emotion feels to each person. There is a distinction between coarse and refined emotions. The latter term is used to label emotions that are not manifest in behavior and refer to complex events or subtle events. Standard emotion labels do not correctly describe these occurrences, art being one of them. Although Schindler et al. (2017) developed its theoretical framework of emotion with references to the concept of valence, they also advocate for a complex view of emotion. In this case, aesthetic emotions are emotions that people feel rather than emotions represented by the artwork. This idea rings close to the views portrayed by Zentner et al. (2018). The main difference between both proposals is a much broader scope of the latter, where Schindler et al. (2017) consider emotion as a subjective feeling and a myriad of other concepts, such as valence and epistemic emotions, for example. This mix between emotional components and other components, depending on the authors' theoretical framework, is crucial in the next element of AE, labeled here as supplementary concepts. Although these ideas are not the main focus of the conceptualization of aesthetic experiences in the articles reviewed, they usually play a role in adding elements identified in the phenomenology of aesthetic experience but are not articulated or described before.

Two elements can be identified in the supplementary components of AE. The first has to do with psychophysiological phenomena, motivational and emotional concepts related to the phenomenal experience of AE (Schindler et al., 2017; Silvia & Nusbaum, 2011; Stamatopoulou, 2004). In Stamatopoulou (2004), these components are grouped in a concept called paratelic

motivational mode. This mode includes concepts such as activity orientation towards the artwork, the feeling of prolonged suspension, and high arousal reported when someone is *absorbed* with an artwork. The conceptualization of the psychophysical experience of AE is related to phenomena such as chills, being touched, or awed by the art that is presented (Silvia & Nusbaum, 2011). Some emotions are also classified as prototypical, like fascination or being moved by the artwork (Schindler et al., 2017). All of these concepts point to the phenomenal aspects of aesthetic experience and come from varied theoretical origins. However, other elements are also supplementary to the cognitive and emotional components of AE. Some authors differentiate perception processes from the different parts of aesthetic experience (Hager et al., 2012; Stamatopoulou, 2004; Wanzer et al., 2018). Some authors emphasize connecting processes, such as the case with Stamatopoulou (2004) and the concept of expressive perception introduced in the article. This concept refers to a convergence of feeling patterns between perception and empathy. In other cases, authors use perceptual processes to refer to the analyses that allow memory integration between the new information given by the artwork and previous knowledge (Hager et al., 2012). Some other elements mentioned by other articles do not appear in others. That is the case with the concept of communicative dimensions related to the interpretation of the artwork proposed by Wanzer et al. (2018).

Considering the conceptual articulation of the articles reviewed, we can identify that most definitions presented seem to share a large portion of the theoretical grounding regarding aesthetic experiences. In summary, most definitions of AE consider two key aspects: cognitive and emotional elements. Later, these aspects are complemented by some other AE concepts that help to give more depth to the definition. Usually, these complementary concepts try to include or articulate into the theoretical models embodied accounts of AE or psychophysical aspects that aim

to account for the phenomenal experience of aesthetics. How do these conceptualizations are different since they share so much of their broad conceptual framework? A short answer would be that most AE concepts share most of their theoretical framework, even though authors label them as distinctive concepts. The differences between these concepts are sometimes scant and, since there is little to no cross-reference between areas of study, some distinctions tend to be shared across concepts. A question arises then regarding the extent to which the conceptual similarities that can be seen across the articles will be reflected in the construction of the assessment instruments, or if the theoretical distinctions that different authors present will be revealed more clearly in the way in which they seek to collect empirical data to inform their inferences about AE.

Selecting observables to characterize AE

An analysis of the types of items that each developed instrument selected for their final form shows that although the researchers aim to measure a specific construct defined in a theoretically distinctive way, most of the instruments are very much alike in structure. A summary of the formal characteristics of the instruments that were reviewed is presented in Table 3.

Table 3

Original and final count of items, final count of dimensions, and format of items proposed by twelve articles reviewed.

Article	N Original items	N Final items	N Final dimensions	Format
Asmus (1985)	296	41	9	Likert 1-4
Andringa (1996)	NA	18	5	Likert 1-4
Konijn (1999)	63	63	3	Likert 1-4

Stamatopoulou (2004)	72	28	5	Likert 1-5
Hagtvedt et al. (2008)	NR	36	8	Likert 1-9
Rowold (2008)	76	29	3	Likert 1-5
Zentner et al. (2008)	515	40	9	Likert 1-4
Silvia & Nusbaum (2011)	12	8	3	Likert 1-7
Vukadinović & Marković (2011)	35	35	3	Semantic differential 1-7
Tschacher et al. (2012)	19	19	5	Likert 1-5
Hager et al. (2012)	76	29	6	Likert 1-5
Schindler at al. (2017)	122	42	21	Likert 1-5
Wanzer et al. (2018)	28	22	5	Likert 1-7

NR: Not reported

Comparing the items selected across instruments

Instruments that aimed to measure aesthetic emotions (Konijn, 1999; Schindler et al., 2017), aesthetic appreciation (Hager et al., 2012; Rowold, 2008; Stamatopoulou, 2004), and aesthetic perception/evaluation (Andringa, 1996; Hagtvedt et al., 2008; Tröndle & Tschacher, 2012; Tschacher et al., 2012) have in their midst items that measure cognitive aspects of the phenomena. There are items, for example, that aim to measure if the stimulus presented caused curiosity or interest in the participant: “Made me curious” (Aesthetic emotions, Schindler et al., 2017), “This painting makes me curious” (Aesthetic appreciation, Hager et al., 2012), “Arousal of curiosity” (Aesthetic perception, Hagtvedt et al., 2008). This similarity is not only restricted to cognitive-related items, but it is also present in affect/emotion items. For example, negative emotions: “Made me sad” (Schindler et al., 2017), “This painting makes me feel sad” (Hager et al., 2012), “Sadness” (Hagtvedt et al., 2008) or “Sad” (Asmus, 1985). The only category that appears solely in a construct is self-reference related to instruments measuring aesthetic experiences. “Relate the artwork to my personal situation” (Rowold, 2008) or “This painting makes me think

about my own life history" (Hager et al., 2012). A supplementary fourth category deals with items that seek to measure the "strong" aesthetic experiences such as the feeling of flow, absorption or sublime (Csíkszentmihályi, 1990; Makin, 2017; Tellegen & Atkinson, 1974), the physiological correlates of this experience (Silvia & Nusbaum, 2011; Tröndle & Tschacher, 2012; Tschacher et al., 2012) or the expressive characteristics of the stimuli (Arnheim, 1966; Stamatopoulou, 2004; Vukadinovic & Markovic, 2012).

The process of making this comparison across instruments is when the reader most notices the need for construct clarification. Next is an example of what the readers of articles are missing when this clarification process is not considered. In the study carried out by Hagtvedt et al. (2008), working on aesthetic perception and evaluation, arousal and valence influence the artwork's information processing. Rowold (2008) instead focuses on the role that these affective elements (i.e., emotions and pleasurable feelings) have in the process of aesthetic appreciation. In the case of Rowold (2008), the affective elements work in tandem with personal experiences and associations that could modify the emotions that each person feels when confronted with an artwork. The synergy between affect and personal experience is an approach also taken by the work of Vukadinovic & Markovic (2012) on aesthetic experiences. Instead, Wanzer et al. (2018) propose that emotions result from the aesthetic appreciation process and not part of it. With just these three approaches, we could see that while the basic definition of the emotional element is the same, the role that emotional elements have on the process is quite different. The distinction regarding observables for assessment is difficult to make when we consider definitions such as the ones made by Hagtvedt et al. (2008), Rowold (2008), and Vukadinovic & Markovic (2012). The differences between these definitions rest on how complex the relationship between arousal/valence and other variables is in the aesthetic experience. How can this be expressed in the form of clear observables

for people to self-report? An intuitive response would be to produce questions that deal with arousal/valence and questions that specify a relationship between these concepts and personal experiences. In the study of Rowold (2008), while there are observables for both concepts, they are not related in any way. Although Vukadinovic & Markovic (2012) state the importance of personal experiences in the aesthetic experience of dance, they do not add any observable that could account for that relationship specified in their theoretical framework. The result of these decisions is three sets of emotion related observables that are similar in style.

Some reasons for this consistent similarity between measures may be a high degree of reliance by researchers on previous measures that may belong to different conceptualizations of aesthetic experiences, the lack of a detailed clarification exercise at the moment of articulating the attribute, or an analytical approach that did not always fit with the theoretical approach proposed by the researchers (Wilson, 2004). Another potentially significant alternative is that all of these different conceptual proposals are ultimately underlined by a single overarching attribute. Examining this alternative would require additional empirical studies to assess, for instance, the convergent or divergent nature of the different constructs through the use of methods such as the multi-trait multi-method approach (Campbell & Fiske, 1959). The possibility that the distinctions between these constructs could ultimately be both theoretically and empirically indistinct is a relevant hypothesis to consider along with the others previously outlined. It is necessary to understand why we find this divergence between distinctive theoretical constructs on the one hand and very similar structures in measurements on the other.

Comparing the instrument development process

When it comes to the development of the instruments, researchers use three main techniques to create items: proposals by non-experts that describe the experience, suggestions by judges or

experts in the field, and items constructed from other studies or by a literature review. Regarding including proposals by participants in the instrument construction, these activities take several different forms. For example, in the oldest article, the study participants were given the experiment and were asked to note any emotional reaction to the painting presented (Israeli, 1928). In a contemporary study, participants' suggestions for emotional items regarding AE were taken into account in the pilot study process (Andringa, 1996). The most common case is researchers asking for proposals from lay people or undergraduate students on the type of experience researched (Hagtvedt et al., 2008; Stamatopoulou, 2004; Vukadinovic & Markovic, 2012). For example, in Stamatopoulou (2004), 45 people between regular people and students were interviewed to identify salient characteristics of aesthetic experiences. This process was later refined by searching for items that indicated agreement between subjects. Frequently, this item construction technique was mixed with the other two to some degree (Asmus, 1985; Hagtvedt et al., 2008; Stamatopoulou, 2004; Vukadinovic & Markovic, 2012). Often, item proposals by non-experts were balanced by later refinement by experts or suggestions by the experts themselves. The latter case is illustrated in Asmus (1985), where the first stage of the item construction involved musically experienced undergraduate and graduate students giving a list of adjectives related to affective responses to music. After creating a pool with these proposals and adding items through a literature review, which is not cited in the article, the list was further refined by 14 musically expert judges that defined the final pool of items. This process is close to the process undertaken by Hagtvedt et al. (2008), where experts reviewed an initial pool made by non-experts.

Lastly, most of the articles reviewed included a literature review as inspiration or source for the item and instrument construction (Asmus, 1985; Hager et al., 2012; Hagtvedt et al., 2008; Konijn, 1999; Rowold, 2008; Schindler et al., 2017; Silvia & Nusbaum, 2011; Tschacher et al.,

2012; Vukadinovic & Markovic, 2012; Wanzer et al., 2018; Zentner et al., 2008). Some articles use this single technique for the creation/selection of items and the creation of the instruments (Hager et al., 2012; Konijn, 1999; Rowold, 2008; Schindler et al., 2017; Silvia & Nusbaum, 2011; Tschacher et al., 2012; Wanzer et al., 2018), while other articles have a hybrid approach that includes the literature review in some part of the process of the instrument construction. For example, Vukadinovic & Markovic (2011) gathered proposals of adjectives for the aesthetic experience of dance from college undergraduates and experts in the first round. Afterward, the item pool was reduced by the level of agreement between subjects in some items. A literature review aided the selection of items. The construction process is taken a step further by Zentner et al. (2008), where the instrument construction process implied three stages. First, a pool was created from existing items from literature and others developed by the research team based on a literature review. Later, this list was rated by participants, and the pool was reduced through the level of agreement of the subjects on the descriptors of musical experience. Finally, the list was further refined by the judgments of experts on the field.

Some differences appear between articles in the process of item/instrument construction. The main difference between the articles reviewed and the article by Zentner et al. (2008) is that the latter is the only one that formalizes this item construction process through an independent study. Most of the other processes are not described in detail and involve a small sample from where the instrument is developed. Vukadinovic and Markovic (2011) and Tschacher et al. (2012) also explicitly cite some of their previous work as the foundations of the instruments constructed. The instrument construction in the article in Schindler et al. (2017) is also unique because it incorporates ideas for items from a broad disciplinary base, such as design and environmental studies, not restricting itself to art aesthetics.

The consolidation of the item pool and its structure varies depending on the period on which the study was carried out and the aims of each study reviewed. Five studies do not report a pilot review (or the study was not labeled as such) for the construction of its AE measurement instrument (Asmus, 1985; Hager et al., 2012; Konijn, 1999; Silvia & Nusbaum, 2011; Wanzer et al., 2018). Some cases include empirical studies in which item and instrument construction works as a side note (Konijn, 1999; Silvia & Nusbaum, 2011), but some of these studies are validation efforts that include just one study for identification and validation of an instrument (Asmus, 1985; Hager et al., 2012; Wanzer et al., 2018). The rest of the studies reported one pilot study, and only the article of Hagtvedt et al. (2008) reported two pilots. In the earlier studies, such as Israeli (1928), the number of participants of the pilot was small (N = 6), but the rest of the studies report between 100 (Rowold, 2008; Schindler et al., 2017; Vukadinovic & Markovic, 2012) and 3000 (Tschacher et al., 2012) participants as part of their pilot efforts.

Most of the articles reviewed that added a pilot process to create the measurement instrument used statistical methods such as principal components analysis or factor analysis to determine the instrument's structure and functioning. One main issue in this section is the inconsistency in the reporting of the results of these analyses. Some articles name the process but do not give any detail about the analysis results that would help the reader understand the researcher's choices (Andringa, 1996; Hagtvedt et al., 2008; Tschacher et al., 2012). In most reports, the researchers detail some elements of the analysis, such as the type of factor analysis and the criteria for selecting the best structural solution. Still, the detail in the reporting is inconsistent. For example, researchers define solutions by high loadings but do not define what is considered a high loading. Or, the article reports the item characteristics of the solution that was selected but did not describe the items that did not fit the final solution. These issues are problematic because

experts have identified shortcomings in the use of factor analysis, like the use of arbitrary cutoff standards for high factor loadings or the hiding of problematic solutions through the reporting of only final solutions in the article (Preacher & MacCallum, 2003).

The variety of approaches regarding the pilot studies and the use of statistical methods has some consequences in terms of reporting the process of instrument consolidation, mainly an opacity in the decision process for the reader of the articles. In some cases, there is an explicit reference to the specific items that the authors decided to delete from the final form (Konijn, 1999; Schindler et al., 2017; Wanzer et al., 2018). In other articles, while they did not detail the items deleted, the rationale for eliminating items is straightforward and details the criteria on which decisions were made (Hager et al., 2012; Stamatopoulou, 2004; Vukadinovic & Markovic, 2012). It also happens that some articles do not explicitly detail the reason for the selection, deletion, or adjustment of items even though they identify them (Hagtvedt et al., 2008; Konijn, 1999; Zentner et al., 2008). The depth of the explanation for these decisions also varies within the articles.

The instrument development process across the articles reviewed is well documented in terms of how the items were created or selected to articulate the first form of the scales proposed. Afterward, the decision chain described in the texts could be improved in clarity and detail for the understanding of readers. The challenges for future attempts regarding the process of instrument development can be outlined in two parts: the creation of greater consistency and clarity in the reports of the process and the need for greater integration of the broader literature into the construction of the item/instrument. Considering the latter challenge, while the process of item construction in most articles, especially the ones dedicated to development is very comprehensive, the lack of cross-referencing between AE literature has as a consequence some similarities between the constructs and instruments. These similarities could derive in two instruments intended to

measure different constructs, but that in practice are indistinguishable. Regarding the first issue, it would be important that the reports regarding the instrument's development are detailed just as much as the rest of the analyses presented in the article. Greater detail in reporting would allow the reader to understand the researcher's choices and have the tools to evaluate the quality of the development process.

As we will examine, these issues continue and move over from the instrument's conceptual articulation to the documentation of instrument quality and the presentation of evidence of validity and reliability of the instruments' final forms.

Documenting instrument quality

With the Standards framework in mind, the evidence presented across the articles when documenting the instruments could be improved significantly. In general, it can be argued that some of the articles consider to some extent content-based evidence. However, as discussed before, an explicit characterization of the construct and the relation to the observables are not always present. Aside from content-based evidence, the main focus regarding the documentation of validity is on evidence-based based on internal structure through the use of factor analytic techniques. Lastly, in a few of the articles, there is consideration of evidence of relations with other variables. However, though validity is addressed to some extent, very little information is presented regarding reliability, and no evidence is presented regarding fairness.

In general, a positive aspect across most articles was the inclusion of more than one study and, in particular, the inclusion of a validation study to examine the performance of the instrument. Only two of the thirteen articles did not report a follow-up study of their first study (Tschacher et al., 2012; Vukadinovic & Markovic, 2012). Most of the other articles reported results of one

validation study, except for Andringa (1996) and Hagtvedt et al. (2008) that reported two studies. The number of participants of these studies ranges from 150 (Hagtvedt et al., 2008) in their first study to 2057 (Asmus, 1985).

Evidence of validity

As mentioned, the main focus when documenting empirical evidence about validity is on analysis from factor analytic models; however, the reporting of both how the analysis was conducted and its results is in general incomplete, and when it is documented it exhibits practices that could be improved.

The first issue is the presentation of the type of analysis conducted with sufficient details to understand and reproduce the methodology clearly. Except for one article (Israeli, 1928), all the articles have a validation study report the analysis carried out in their studies. Two of them (Andringa, 1996; Konijn, 1999) report that they carried out scale/factor analysis on the study data, but they do not specify the type of analysis used. The remaining articles do report their choice of analysis: principal components analysis (Asmus, 1985; Hager et al., 2012; Stamatopoulou, 2004), exploratory factor analysis (Schindler et al., 2017; Silvia & Nusbaum, 2011; Stamatopoulou, 2004; Wanzer et al., 2018) and confirmatory factor analysis (Hagtvedt et al., 2008; Rowold, 2008; Schindler et al., 2017; Zentner et al., 2008).

It is worth noting that some methodologically problematic practices were detected among the papers that reported its factor analytical methods with enough details. One of them is the choice of the model to be used throughout the data analyses. Most studies reviewed use principal components analysis (Asmus, 1985; Hager et al., 2012; Rowold, 2008; Stamatopoulou, 2004; Tschacher et al., 2012; Vukadinovic & Markovic, 2012) or factor analysis (Schindler et al., 2017;

Silvia & Nusbaum, 2011; Wanzer et al., 2018; Zentner et al., 2008) as their model to analyze the underlying structure of the data. Some of these studies use two types of models to analyze their data depending on the stage of the item development process. While no use of any of the models is inherently incorrect, the choice between them has important implications. One of the consequences of using principal component analysis (henceforth PCA) is that this approach does not explicitly model error variance. When error variance is not accounted for, the possibility of a substantive interpretation of components is problematic. PCA usually overestimates loadings and underestimates correlations between factors (Preacher & MacCallum, 2003). The possibility of interpretation of the structural components of the data is key when constructing instruments of measurement, and the use of this type of analysis would difficult the interpretation of results in any step of the process. If the researcher aims to identify interpretable constructs that explain correlations between measured variables, factor analysis in any of its forms is the choice recommended (Furr, 2013; Preacher & MacCallum, 2003). The reasoning behind this recommendation stems from the traditional aims in the process of instrument development, that is, to make generalizable inferences about the attribute. In this case, the inferences are based on the sources of common variation underlying the observed data.

Another questionable practice related to the use of factor analysis is the use of several factor analyses on the same data set (Stamatopoulou, 2004; Vukadinovic & Markovic, 2012) and, in particular, fitting both EFA and CFA data sequentially to the same dataset (Hager et al., 2012; Schindler et al., 2017), a practice that is commonly discouraged for its tendency to produce overfitting of models to the data (Fokkema & Greiff, 2017). Similarly, other articles presented unconventional practices that are not clearly justified, as in the case of Zentner et al. (2008). In this case, the authors clearly report their analysis, indicating how they selected a 9-factor solution based

on a CFA model, only to later run an EFA to examine for second-order factors. In addition to being strange in terms of moving towards an exploratory model after settling on a confirmatory solution, this practice could have been avoided by directly modeling the second-order factors in CFA and comparing their fit indices with the previously modeled fit indices.

A final issue that arises around the use of factor analysis in the development of AE instruments is the criteria to retain or discard factors. The selection of factors and "naming" of factors raises both conceptual and methodological issues, with the former being deeply rooted in the issues discussed in the first two challenges. After all, the measurement of attributes is traditionally expected to be theoretically guided and not merely a data-mining exercise (Wilson, 2004; Author, 2021). Regarding the methodological challenges, current measurement literature promotes best practices in the use of factor analytic techniques, such as the examination of the scree plot for the number of eigenvalues that fall before the drop, the comparison of this analysis with a parallel analysis of the same data or the use of fit indices in case the use of SEM (Furr, 2013; Preacher & MacCallum, 2003). The reporting of the criteria on which the researchers base their decisions is not always clear. In some cases, the choices made regarding the retention or discard of factors are in line with the recommendations (Schindler et al., 2017; Zentner et al., 2008). In the majority of articles, the researchers selected one of the recommendations followed by other criteria that have been discouraged as best practices in measurement, for example, the use of eigenvalues larger or equal than one to determine retention (Asmus, 1985; Hager et al., 2012; Hagtvedt et al., 2008; Rowold, 2008; Stamatopoulou, 2004; Vukadinovic & Markovic, 2012; Wanzer et al., 2018). Finally, other result reports do not report these criteria, or the criteria are opaque (Andringa, 1996; Konijn, 1999; Silvia & Nusbaum, 2011; Tschacher et al., 2012). The importance of communicating these best practices and encouraging their use in the development process of AE instruments can

increase the interpretability of the resulting solutions and increase the potential reproduction of the results, consolidating the instrument's use among its users.

A second issue is the presentation of the results of the selected analysis, where we can see that relevant summary statistics are not presented, with only three studies reporting the amount of the total variance explained by the initial set of items, ranging from 37% to 59.75% (Rowold, 2008; Stamatopoulou, 2004; Vukadinovic & Markovic, 2012). And only three articles report the total variance explained by their final instruments, ranging from 53% to 71% (Hager et al., 2012; Hagtvedt et al., 2008; Wanzer et al., 2018).

A third issue is the clarification of the rationale guiding the interpretation of results and decision-making when making decisions based on the factor analytical results. By way of illustration, we will consider the instrument's final structure presented by Schindler et al. (2017) in their article. The first analysis carried out on the data was an EFA for ordinal responses, specifying 24 factors. The results allowed the authors to conclude that a 7-factor solution could also be a possible factor structure, which they also carried out. Although the model fit was excellent in the 24-factor EFA, only 20 of the 24 factors were interpretable from the previously established categories. The 7-factor EFA showed acceptable indices of model fit. The authors selected 42 items with these results and grouped them in 21 sub-scales with two items each. The criteria used by the authors to assign the items were: common loading of the items in the same factor in 24-factor EFA, similar loading profile in the 7-factor EFA. These criteria were effective for 17 of the 21 factors of the instrument. The other items were constructed into sub-scales, considering previous theoretical and empirical findings. Here, although the analytical choices of the authors are elaborated in the article, some reservations arise. As the solution (mix between the 24- and 7- factor solution) relies on this ad hoc procedure, questions of the stability of the solution to other contexts or over time are

more prominent.

Aside from the evidence of internal structure using factor analysis, the second most prevalent source of validity evidence based on relation to other constructs, with some of the studies have considered evidence in relation to other variables, attempting for instance to articulate a construct differentiation or correlation from other constructs (Hager et al., 2012; Stamatopoulou, 2004; Wanzer et al., 2018; Zentner et al., 2008). For example, in the article of Stamatopoulou (2004), the Aesthetic Experience Scale was correlated with constructs such as familiarity with diverse aesthetic experiences (r = .30) and self-perception of aesthetic competence (r = .50). In the case of Wanzer et al. (2008), the evidence compiled from the development of the scale was compared with scores of the Curiosity and Exploration Inventory (CEI-II; r = .28), Inspiration Scale (r = .46), and the Openness to Experience scale (r = .45). The researchers in this article also added evidence regarding the difference in scores between individuals with more or less artistic experience. The results show that having more art training or having a job related to the area did not relate to aesthetic experiences. However, women reported more intense aesthetic experiences in some instrument's subscales, such as emotional, perceptual, and flow experience. Also, participants with lower education levels reported higher intensity levels regarding emotional and flow experiences in the scale (Wanzer et al., 2018).

Evidence of reliability

Most of the articles document some evidence of reliability, except for the articles by Israeli (1928), Tschacher et al. (2012), and Vukadinović & Marković (2012). Most studies (9 out of 13) rely only on Cronbach's alpha to assess the internal consistency of the instruments as their only source of reliability evidence. Just one of the articles, Rowold (2008), reported on the temporal stability of the assessment results by using a test-retest study of the final form of the instrument a

month later.

Some studies report high reliability coefficients through all of their subscales (Hager et al., 2012; Hagtvedt et al., 2008), while other studies have reported mixed results, ranging from .80 to .60 (Andringa, 1996; Asmus, 1985; Rowold, 2008; Schindler et al., 2017; Zentner et al., 2008). A few studies report lower reliability coefficients, ranging from .70 to a lower .30 (Konijn, 1999; Stamatopoulou, 2004; Wanzer et al., 2018).

In general, we can see that results for at least one type of reliability are reported. The quality of instrument development would benefit from multiple sources of errors as considered by alternate methods of reliability, such as test-retest, alternate forms, and internal consistency.

Discussion

Considering the diversity of proposals regarding the measurement of aesthetic experiences, we identified three main challenges: (a) defining the aesthetic experience, (b) the determination of observables, and (c) examining and documenting instrument quality.

Regarding the first challenge, which is concerned with defining the aesthetic experience, while there is a good amount of theoretical work in psychology related to aesthetic experiences from which researchers can build their conceptual foundations, less work is done on comparing and differentiating the theory proposed with other similar concepts. Only three of the thirteen articles reviewed made a distinction between their theoretical proposal and the current developments within the discipline (Hager et al., 2012; Schindler et al., 2017; Zentner et al., 2008). One outcome related to this lack of comparison between frameworks is research within a discipline that employs different concepts (e.g., epistemic emotions vs. cognitive elements of AE) but ends articulating a construct that resembles others in different frameworks. An emphasis on the

clarification, differentiation (and potentially of amalgamation) of constructs would allow the measurement developers to distinguish between their conceptual framework and the theories of fellow researchers to disentangle possible points of convergence (Wilson, 2004). Or, if the evidence does not support these differentiations, to question the theoretical articulation of these concepts and review if the empirical evidence supports a sole conceptualization of the phenomena. Although this is not always feasible in empirical or experimental articles, the conceptual clarification and its implications on assessment development are critical, as these instruments are used as the basis to conduct and interpret experimental studies given that they are used to, for instance, quantify the effects of any manipulation on the dependent variables. Differentiation not only would need to be restricted to the articulation of the theoretical framework and the constructs developed from it, but it would also be helpful in consideration of the following two challenges related to AE (Furr, 2013; Wilson, 2004).

Concerning the determination of observables, only a few articles reviewed clearly linked their theoretical framework with their proposed measurement construct. Going through the reports of these processes, there were some examples where a clear connection between the theoretical framework was present, its concepts, and the construct developed to measure aesthetic experiences (e.g., Hager et al., 2012). As could be expected, this happened when the article was entirely focused on the development of measures. But, in some cases, there was no distinct connection between the article's conceptual proposal and the measures used or developed in later stages. This lack of connection is easier to verify in empirical papers (e.g., Silvia & Nusbaum, 2011; Tschacher et al., 2012). While there is some theoretical presentation of the concepts and measurements used, there is no direct link or citation of this relationship. In other cases, there is a clear link between some concepts and constructs, while other measures presented in the articles do not have a theoretical

counterpart (e.g., Andringa, 1996). Therefore, it is tough for readers to check if the authors' proposal is consistent in theoretical terms and makes replication of results more difficult. As with the previous challenge, there is the need for an explicit clarification between the conceptual developments and frameworks within a concept or discipline and the link of these theoretical elaborations with the final construct definition and proposed observable. In empirical articles, an extensive clarification may not be called for, but the presentation of a direct relation between the theoretical claims made in it and the measurements utilized in the experimental setting would be a significant and necessary improvement. At the very least, a direct citation or brief explanation could allow the reader to research and make their own minds regarding the relationship between concept and construct. In articles dealing with construct development of AE measures, clarification is critical for understanding the process and the overall consistency between conceptual developments and their corresponding constructs. Without performing this step, an opacity between theory and measurement procedures is installed very early in the development process.

A further challenge detected on the determination of observables has to do with the measurement construction reports. The result section in most of the articles reviewed was brief. This characteristic meant, for example, that there was little mention of whether the items or instruments cited are used in their entirety or a few selected items were selected. A consequence of choosing one option over the other is that it could affect the psychometric characteristics of the scales used (Furr, 2013). In general, the most common steps reported in measurement construction in these articles involve the selection of techniques used to confirm or explore the structure of the questionnaire, the final form of the scales with their psychometric characteristics, and some reference to reliability in the form of Cronbach's Alpha. Opacity in reporting is common regarding the rationale of researchers for selecting/discarding items and the choices made by researchers

regarding analytical procedures. While it is not expected that the authors give a meticulous account of their decision process in terms of item selection in the first stages, for example, it should be important to mention the characteristics of the items deleted and the reasons – empirical or theoretical – why the items are being deleted or replaced. In more comprehensive result sections of instrument development, such as Schindler et al. (2017) or Zentner et al. (2008), such descriptions are reported by the authors, especially at the in-between stages of factor structure research. These descriptions allow the reader to understand the rationale under the decisions made by the researchers and how the elimination or replacement of items conforms the instrument under a particular theoretical proposal.

Other issues stem from the search of the final form of the instrument. While the articles reviewed always name and specify the analytical method used to identify the instrument's structure, their usage is not always in line with the recommended practices in statistics or psychometrics (Furr, 2013; Preacher & MacCallum, 2003). This worrying trend of misuse of statistical analyses should be a call of attention to authors to detail their construction and analysis process explicitly and to the readers to critically assess the reports and statements made in these articles. One suggestion for the improvement of this issue would be a greater alignment between the analytical methods used to characterize the structure of AE instruments with the knowledge already existent in the statistical/measurement disciplines. These guidelines would allow researchers to build upon a solid theoretical and empirical base of all their analytical choices. For example, the usage of confirmatory factor analysis for an instrument that is grounded in a well-developed theoretical framework—or avoiding using the same data for two or more different factor analyses (Fokkema & Greiff, 2017).

Another critical issue that needs to be addressed is the need for further examination and

instrument quality documentation. Standards in measurement regarding reliability and validity have long since moved away from relying solely on the use of Cronbach's alpha as a measure of reliability (AERA et al., 2014). The consequence of the lack of evidence for the reliability or validity of these scales is that the researcher does not have a substantive claim regarding the consistency of the scores through the testing process. In some cases, in the articles reviewed, this is especially worrying since the only report of reliability shows low scores of the coefficients. One of the suggestions for the researchers that wish to move towards updated forms of validity and reliability is the necessity to provide, first and foremost, an interpretation of the scores and an explicit statement on the intended uses of the instrument (Wilson, 2004). A validity argument that adequately supports inferences based on an instrument does not depend solely on the data provided by a single article; it also is dependent on the evidence gathered by other studies using the same instrument that share a common understanding of the construct, and ideally, that share the specific use or interpretations of the instrument's scores. Thus, the need for an interpretation of the scores based on a common understanding of the attribute and the specification of the uses of the instrument is vital. Without these specifications, there is no way to truly advance in establishing a solid and ongoing evidentiary base for the validity of these instruments (AERA et al., 2014).

While this article's objective was to identify some points of tension of the construction and development of measures of aesthetic experiences in general, the goal was not to make this review a census of the developed measures of aesthetic phenomena. Considering this antecedent, a limitation of the current study is the small quantity of reviewed articles compared to the total amount of publications that develop measures related to aesthetic experiences, or for that matter, other measures of aesthetic phenomena outside of the artistic domain. While there has been significant development of theory and measurements of topics such as the aesthetic experience of

the environment, we focused on the field of aesthetic phenomena within artistic experiences because it is one of the precursors of these studies. Further studies should explore if measurement efforts in these related study domains replicate some of the same issues exposed in this article.

All the challenges pointed out in this article are surmountable. The field of measurement of aesthetic experiences is just developing, and it would benefit from reconsidering the practices used to create instruments, promoting a solid conceptual alignment in the development process and the generation of substantial evidence for their validity and reliability. More cooperation and knowledge exchange between the aesthetic experiences research and experts in psychometrics and statistics is needed. This article hopes to point out some of the challenges associated with the construction of measures in aesthetic experiences so that future teams of researchers can acknowledge and tackle them, highlighting that measurement issues are not "just" a methodological problem, as they are deeply interconnected with issues of theory development and the interpretation of experimental studies.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Author. (2021). [Title omitted for masked review].

Andringa, E. (1996). Effects of 'narrative distance' on readers' emotional involvement and response. *Poetics*, *23*(6), 431–452. https://doi.org/10.1016/0304-422X(95)00009-9

Arnheim, R. (1966). *Toward a psychology of art: Collected Essays*. University of California.

Asmus, E. P. (1985). The Development of a Multidimensional Instrument for the Measurement

of Affective Responses to Music. *Psychology of Music*, *13*(1), 19–30. https://doi.org/10.1177/0305735685131002

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology*, *60*(1), 34–43. https://doi.org/10.1016/j.ijpsycho.2005.04.007

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin*, *52*, 281–302.

Csíkszentmihályi, M. (1990). *Flow: The psychology of optimal experience.* Harper & Row.

Daston, L., & Galison, P. (2008). *Objectivity* (Third). Zone Books.

Fokkema, M., & Greiff, S. (2017). How Performing PCA and CFA on the Same Data Equals Trouble: Overfitting in the Assessment of Internal Structure and Some Editorial Thoughts on It. *European Journal of Psychological Assessment*, *33*(6), 399–402. https://doi.org/10.1027/1015-5759/a000460

Furr, M. (2013). *Scale Construction and Psychometrics for Social and Personality Psychology*. SAGE.

Gross, J. J. (2010). The Future's So Bright, I Gotta Wear Shades. *Emotion Review*, *2*(3), 212–216. https://doi.org/10.1177/1754073910361982

Hager, M., Hagemann, D., Danner, D., & Schankin, A. (2012). Assessing aesthetic appreciation

of visual artworks—The construction of the Art Reception Survey (ARS). *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 320–333. https://doi.org/10.1037/a0028776

Hagtvedt, H., Patrick, V. M., & Hagtvedt, R. (2008). The Perception and Evaluation of Visual Art. *Empirical Studies of the Arts*, *26*(2), 197–218. https://doi.org/10.2190/EM.26.2.d

Israeli, N. (1928). Affective reactions to painting reproductions: A study in the psychology of esthetics. *Journal of Applied Psychology*, *12*(1), 125–139. https://doi.org/10.1037/h0070445

James, W. (1884). What is an emotion? *Mind*, *9*(2). https://doi.org/10.1093/mind/os-IX.34.188

Kant, I. (1914). *Critique of Judgement*. The MacMillan Co. of Canada.

Konijn, E. A. (1999). Spotlight on spectators: Emotions in the theater. *Discourse Processes*, *28*(2), 169–194. https://doi.org/10.1080/01638539909545079

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508. https://doi.org/10.1348/0007126042369811

Makin, A. D. J. (2017). The Gap Between Aesthetic Science and Aesthetic Experience. *Journal of Consiousness Studies*, *24*(1–2), 30.

Maul, A. (2013). Method Effects and the Meaning of Measurement. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00169

Michell, J. (2009). The psychometricians' fallacy: Too clever by half? *British Journal of Mathematical and Statistical Psychology*, *62*(1), 41–55. https://doi.org/10.1348/000711007X243582

Michell, J. (2012). Alfred Binet and the concept of heterogeneous orders. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00261

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's Electric Factor Analysis Machine. *Understanding Statistics*, *2*(1), 13–43. https://doi.org/10.1207/S15328031US0201_02

Rowold, J. (2008). Instrument Development for Esthetic Perception Assessment. *Journal of Media Psychology*, *20*(1), 35–40. https://doi.org/10.1027/1864-1105.20.1.35

Salzberger, T. (2013). Attempting Measurement of Psychological Attributes. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00075

Scarantino, A., & de Sousa, R. (2018). Emotion. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2018/entries/emotion/

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., & Scherer, K. R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLOS ONE*, *12*(6), e0178899. https://doi.org/10.1371/journal.pone.0178899

Silvia, P. J., & Nusbaum, E. C. (2011). On personality and piloerection: Individual differences in aesthetic chills and other unusual aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts*, *5*(3), 208–214. https://doi.org/10.1037/a0021914

Stamatopoulou, D. (2004). Integrating the Philosophy and Psychology of Aesthetic Experience: Development of the Aesthetic Experience Scale. *Psychological Reports*, *95*, 673–695. https://doi.org/10.2466/pr0.95.2.673-695

Tan, E. (1994). Film induced affect as a witness emotion. *Poetics*, *23*, 7–32.

Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ('absorption'), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology*, *83*(3), 268–277. https://doi.org/10.1037/h0036681

Tröndle, M., & Tschacher, W. (2012). The Physiology of Phenomenology: The Effects of Artworks. *Empirical Studies of the Arts*, *30*(1), 75–113. https://doi.org/10.2190/EM.30.1.g

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K., & Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96–103. https://doi.org/10.1037/a0023845

Vukadinovic, M., & Markovic, S. (2012). Aesthetic experience of dance performances. *Psihologija*, *45*(1), 23–41. https://doi.org/10.2298/PSI1201023V

Wanzer, D. L., Finley, K. P., Zarian, S., & Cortez, N. (2018). Experiencing flow while viewing art: Development of the aesthetic experience questionnaire. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000203

Wilson, M. (2004). *Constructing Measures: An Item Response Modeling Approach*. Routledge. https://doi.org/10.4324/9781410611697

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. Routledge.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Contributions

Contributed to conception and design: JP, DTI

Contributed to acquisition, analysis and interpretation of data: JP

Drafted and/or revised the article: JP, DTI

Approved the submitted version for publication: JP, DTI

Funding information

This work was supported by the Chilean National Commission for Scientific and Technological Research (CONICYT-PCHA/Doctorado Nacional/2018).

Competing interests

We have no known conflict of interest to disclose.

7. General discussion

This thesis aimed to explore whether the discursive and non-discursive modifications of scientific texts produced changes in the comprehension or appreciation of the texts. Using an experimental design, I explored through two studies whether the modifications of presentational variables influenced the perception of the comprehension and appreciation of a short scientific text.

7.1. On the understanding of the scientific text

Regarding the effects of the modifications of discursive and non-discursive variables on the comprehension of scientific texts, the experimental manipulations did modify the participants' perception of understanding. In the first study, the absence of statistical data in natural science abstracts led to higher ratings of the perception of comprehension of the text. On the contrary, in the case of social sciences, statistical data in the abstract leads to higher ratings of self-reported understanding. These results highlight the diverse importance that those areas give to statistics as part of a scientific text's presentation. In the natural sciences, statistical data would not be strictly necessary to generate an adequate perception of the understanding of the abstract. In the social sciences, however, the addition of statistical data seems essential in the perception of the understanding by participants. The participants with some training in social sciences seem to place greater importance on quantitative indicators as a sign of a comprehensible text. The latter results are consistent with prescriptions from several manuals of scientific writing that statistical descriptors in texts grant precision and clarity that words cannot give (Katz, 2006; Vieira, 2011). The belief in the power of numbers and statistics as discursive elements of scientific texts could be why the scores of the perception of understanding are higher in social sciences undergraduates.

Results from the first study show that the interaction between the display of the abstract in third person plural and the absence/presence of statistical data also improved the scores of

perceived comprehension of the abstract in both areas of study. Even though the manipulation of the grammatical person of the abstract did not yield statistically significant differences alone, when paired with the manipulation of statistical data, it generates substantial differences in the perceived comprehension of the text. For participants from natural sciences formation programs, the interaction between third person plural and the absence of statistical data rose the scores of the perceived understanding of the text. In the case of social sciences and humanities students, the interaction between third person plural and the presence of statistical data also rose the scores of the perceived understanding of the abstracts. This interaction can be explained through the discrepancy between the recommendations in style manuals of particular disciplines and the style manuals orientated towards university students. The advice that students are more likely to recall is still rooted in the belief that the third person plural reflects an "objective" approach to research and its divulgation (McMillan & Weyers, 2007; Robbins, 2010; White, 2000). This recommendation, together with the belief that statistical descriptors give greater clarity to abstracts, allows for a more conservative stance towards the abstract's presentational features, especially in participants trained in the social sciences.

In the second study, where presentational non-discursive features of the abstracts were manipulated, there were also significant results about the perceived comprehension of the scientific text. Here, the modification of both variables caused changes in the ratings of understanding reported by participants. When the title was modified to a clichéd version, the ratings of perceived comprehension rose across all study participants. These results contrast with a recent study that shows that psychology undergraduates prefer longer titles and titles phrased as a question but did not show a significant preference for clichéd titles (Hallock & Bennett, 2021). These contrasting results can be explained considering that the dependent variable measured in both studies is

different; in the case of Hallock & Bennett's (2021) study, the variable was labeled as the preference for one of two titles. The participants of the second study of the present thesis did not have to indicate any specific preference for the titles of each abstract. However, they reported a better understanding of the text when the clichéd title was used. More research is needed to disentangle differences between preference and comprehension in science to account for both sets of results. The results of this particular study also do not follow recommendations and research in other disciplines, where the use of clichéd titles is discouraged (Day & Gastel, 2011; DeBakey & DeBakey, 1983; Goodman, 2012). The conventional approach suggested by the scientific community seems to be more flexible when it is rated by people still on a formative process; such is the case with undergraduate students. A more colloquial title does seem to improve the perception of the understanding of the abstract, even though it does not fit an objective cannon to scientific communication and the presentation of its results.

When the paper format was modified experimentally from a traditional article format to a simpler one, there was an improvement in the perception of comprehension of the abstract across all study participants. These results could be explained by a negative holistic-affective response to the standard article format. The participants may have perceived the original design of the article's presentation as an element fabricated by the study's researchers since it is different from most traditional article layouts. Therefore, a plainer layout may have improved the perception of understanding since it did not have any other visual characteristics besides the text itself. The fact that the study participants were undergraduates with some experience in academic activities such as reading scientific articles might resonate with these results. Since not all participants had training and were well versed with this practice, the mere fact of reading the abstract might be the ground for an adverse affective reaction to the activity. Therefore, the negative disposition to the article's

formatting produces a diminished perception of the understanding of the text contents. To the author's knowledge, this is the first time that the effect of the format of presentation on the perception of understanding has been tested on scientific texts without content modifications like previous studies have done (Hartley, 2003; van der Meij et al., 2013; Werner da Rosa & Otero, 2018). Contrary to what has been thought about simpler formats (McKiernan, 2000), presentations such as the implemented one in this study are perceived as less bulky and, maybe, more intelligible, at least by readers with some university training.

Finally, another important finding of the second study is the significant effect of the format of presentation and the modified title. The results show that participants rated their perceived understanding higher when the abstract was presented in a plainer layout and had a clichéd title. This configuration is counterintuitive considering the recommendations on academic writing (Day & Gastel, 2011; Goodman, 2012; McMillan & Weyers, 2007). However, since presentational features of scientific information in texts have been changing more rapidly in recent years, it could be hypothesized that a more straightforward text configuration combined with a more memorable title could change the perception of the text's difficulty. Therefore, this modification would increase the perception of comprehension. The perception of a scientific text in pre-writing or a pre-print stage makes the abstract appear more approachable to participants and, consequently, easier to understand.

In both studies, the analysis of the open-ended questions on understanding raises an important issue since these results show no differences between any experimental manipulation. Therefore, there is a disparity between the effects of the experimental manipulation on the perception of understanding and understanding through an external judge's ratings of the open-ended questions. Although intriguing, these results are akin to other reports made by Werner da

Rosa & Otero's (2018) study, where the experimental manipulation consisted of modifying the document's source shown to secondary students. In the study, the participants rated the comprehensibility of the paragraphs that came from different sources. The scores rose when the texts were supposedly taken from science textbooks compared to peer reports. However, this manipulation had no effect on other study measures, such as reporting comprehension obstacles that could point to a more in-depth processing and understanding of reading. Altogether, the results of this thesis studies on the difference of perception of comprehension versus comprehension point out the differential effect that overall presentational (i.e., discursive or non-discursive) elements in scientific texts have on the perception of understanding and the understanding of the text rated by external judges or other measures of in-depth processing in samples of participants in formative stages in their education. Understanding a scientific text does not only implies the assessment of the accuracy or correct recall of the information presented in the text but also entails the use of particular skills and criteria established by specific communities that allow the person to manipulate this knowledge correctly (de Regt, 2004; de Regt & Dieks, 2005). A hypothesis explaining these results could be that the perception of the understanding of scientific texts can be manipulated through these presentational elements. However, these elements are not criteria that participants use to achieve greater or differential understanding in other measures.

I additionally explored the impact of abstract effects, the self-perceived expertise, and the readability of the abstract on the perceived understanding of the abstracts by the participants. In both studies, we identified a strong effect of the perception of expertise on the reported understanding of the abstract read. While there is evidence that students' reading skills develop throughout their career (Hubbard & Dunbar, 2017), less is known about why a greater perception of expertise would also raise the scores of perception of the understanding of a text. It is necessary

to highlight that the article did not always belong to the domain reported by the participant. One of the reasons could point to the possible relationship between the perception of expertise and their belief that they can transfer the skills of their discipline to approach and understand the text.

Also, differences between abstracts were not expected in any study. In the first study, significant differences by abstract in scores of the perception of the understanding of the texts appear. These effects occur across disciplines, and they cannot be accounted for primacy or recency effects since the presentation of the abstracts was counterbalanced. The abstracts were authentic published texts, and we favored an ecological approach where four different genuine abstracts were considered. The abstract selection process opens many other explanations for this unexpected difference. In the case of the second study, one hypothesis is that the topics of the Skinner et al. (2019) – farming and its effects on climate change – and Bressan (2018) – personal characteristics and mathematical ability – abstracts were more attractive to the sample of this study than the other texts. There is evidence that the public is particularly interested in climate change topics (Baram-Tsabari & Segev, 2011; European Commission, Directorate-General for Communication, 2020; Falchetti et al., 2007). In other cases, the preference could be guided by previous experiences and personal interests (Falk et al., 2007). Many other possibilities could explain these results, and more research is needed.

We obtained unexpected results in the second study concerning the significant differences between scores of perceived understanding of the abstract depending on the language on which participants read the abstract. Higher scores on the perception of understanding related to the English version of the abstract suggest that participants believe English works as the lingua franca of science (Meneghini et al., 2008) and also impacts how the abstract is perceived and understood. Evidence shows that studies published in English are ever-increasing and that those studies report

more significant results than studies in other languages (Egger et al., 1997; Solovova et al., 2018). Reading scientific texts in English could be linked to the perception in laypeople that science is of better quality than the science published in other languages, given that researchers have migrated towards publication in English to reach more prestigious journals and publishers.

Finally, in the second study, we obtained an unexpected gap in the scores of perceived understanding between the areas considered. The results suggest that participants trained in the natural sciences rate their perceived understanding of the abstract significantly lower than those trained in social sciences. A possible explanation for this difference could be that the effect of the presentational elements is not strong enough to bring the science-trained students to the same level of perception of their comprehension, given that they have an increased familiarity with academic journals. A higher-level knowledge of the journal system and its networks, reputation, and presentation format allows the natural science-trained students to identify a high-ranking journal rapidly. Therefore, the perception that the abstract contents may be more difficult to understand since the source is genuine academic writing may be more salient to these students.

7.2. On the aesthetic appreciation of the scientific text

Regarding the effects of the experimental manipulations in the aesthetic appreciation, no significant differences were detected between manipulations in the studies of this thesis. While the experimental manipulations did not yield the expected effect on the aesthetic appreciation of a scientific text, some other control variables did.

In the first study, a strong effect of expertise on the aesthetic appreciation of the text was detected. Here, a higher perception of expertise across all areas of knowledge resulted in higher ratings of aesthetic appreciation. The second study replicated these results, but only for the positive and epistemic emotions considered in the aesthetic appreciation measure. Overall, these results

follow some established and new evidence suggesting that experts in several disciplines value the stimuli presented differently than laypeople (Era et al., 2019; Leder et al., 2014; Pihko et al., 2011). The current studies results' are similar to those from a study by Hayn-Leichsenring et al. (2021). Their results indicate that mathematics undergraduates report a higher agreement regarding the aesthetic judgment of equations compared to other participants. Also, understanding the meaning of an equation turned out to be an essential factor contributing to aesthetic appeal among experts. These results are thought to be brought by their specialized and advanced training in their specific domain (Hayn-Leichsenring et al., 2021). A greater perception of expertise could have contributed to an increased appreciation of the texts in both studies. A plausible reason is that a higher level of training of the undergraduates has allowed them to integrate themselves into the knowledge community and, therefore, to the knowledge aesthetics of their discipline. Those skills incorporated into their reading practice will enable them to appreciate the text (de Regt & Dieks, 2005; Hayn-Leichsenring et al., 2021). To the best of my knowledge, this is the first time this effect of expertise in the aesthetic appreciation of science texts across domains has been empirically shown.

There are also effects on the aesthetic appreciation of the text depending on the gender in the first study. When the study participant was male, there were higher ratings of aesthetic appreciation regarding positive and prototypical aesthetic emotions linked to the aesthetic appreciation of the abstract presented in the experiment. These results are uncommon since few studies explore the differences between gender in the aesthetic appreciation of any subject, especially in the natural sciences. Some studies focus on individual differences in aesthetic appreciation but do not focus on gender, and other studies only focus on a male or female sample (Marin & Leder, 2018; Palmer et al., 2016; Silvia & Nusbaum, 2011).

The area of training of the participants also had an unexpected effect on the aesthetic appreciation of the text. In the first study, the natural science trained participants reported increased negative emotions regarding the aesthetic appreciation of the text compared to their social science and humanities counterparts. In the second study, the natural science trained participants scored significantly lower on epistemic emotions regarding the aesthetic appreciation of the text compared to their social science and humanities counterparts. The results suggest that natural science students could be less focused on the rhetorical or presentational elements that make a text more appealing to the reader. Therefore, they rate their aesthetic appreciation of the abstracts in general as lower than other participants.

Contrary to the hypothesis proposed in the thesis, the presentational manipulations – whether discursive or non-discursive – did not influence the aesthetic appreciation of the texts. Besides the fact that most people do not consider science texts as aesthetic objects, there might be other reasons for which these experimental manipulations did not work in both studies. Following the theoretical framework of this thesis, consciousness itself is articulated in a subsidiary mode which allows understanding of the world (Langer, 1970; Polanyi, 1962). Therefore, the act of perceiving and reading the abstract presented is situated in a background that allows the co-presence of multiple objects or themes around the acts of consciousness (Gurwitsch, 2009). Only themes or objects that belong to the thematic field of the text would have the capacity by implication to alter in any way the perceptual field of the participants. While the presentational modifications of the abstracts may dwell in the thematic field by reading each abstract, it was not one of the elements closest to the focus of consciousness (Gurwitsch, 2009). Being the presentational modifications outside of the primary perceptual field of consciousness when reading

the abstracts, the participants did not explicitly acknowledge the effect of these modifications or could not articulate it with the pre-made answers of the survey.

The aesthetic effects of the presentational modifications could also be mediated by expertise. Here, the tacit knowledge of participants trained in natural science allowed them to have presentational elements as more prominent themes in their thematic field when reading the abstracts (Polanyi, 1962). Since asking the participants about aesthetic properties may be difficult given the subsidiary aspect of these properties on consciousness, an indirect or qualitative approach to studying the phenomena may be preferable. Some research currently uses eye-tracking to analyze attentional patterns when viewing art (e.g., Massaro et al., 2012; Pelowski et al., 2018). Therefore, measuring other aesthetic reactions with this instrument would not be a stretch, especially when the object has similar visual characteristics.

Given the difficulties of measuring aesthetic appreciation, the thesis also focused on reviewing the processes of quantitative measurement in aesthetics and the repercussion these processes have on the outcomes. In this article, we identified three main challenges regarding the measurement of aesthetic experiences: defining the aesthetic experience itself, the determination of observables, and examining and documenting instrument quality. Regarding the first challenge, while there is considerable work in psychology related to aesthetic experiences, less work has been done about comparing and differentiating the theory proposed with other similar concepts or frameworks. An emphasis on clarification and differentiation (or possible amalgamation) of constructs would allow measurement developers to distinguish between their conceptual framework and the theories of fellow researchers to disentangle possible points of convergence (Wilson, 2004). If the evidence does not support these differentiations, it would be possible to question the theoretical articulations of these concepts. Also, it would allow reviewing if the

evidence supports a sole conceptualization of the phenomena. Concept clarification and differentiation are critical processes since the measurements based on these concepts are used in experimental settings to quantify the effects of any manipulation on the dependent variables.

On the challenge of the determination of observables, only a few articles noticeably linked their theoretical proposal to their measurement construct. In some cases, there was no particular connection between the article's conceptual proposal and the measures used or developed in the later stages of the article. In other articles, there is a clear link between some concepts and constructs, while other measures do not have a clear theoretical counterpart. Therefore, it is not easy for readers to check if the developers' proposal is consistent in theoretical terms, making replication more difficult. As with the previous challenge, there is the need for an explicit clarification between the conceptual developments and these theoretical elaborations to the final construct definitions and the proposed observables. Clarification is critical for understanding the process and overall consistency between concepts and their corresponding constructs. If this step is not performed, an opacity between theory and measurement procedures is installed early in the development process.

Another challenge detected regarding the determination of observables is the measurement construction reports. The procedure or result sections in most of the articles reviewed are brief. In general, the most common steps reported in measurement construction involve selecting techniques used to confirm or explore the structure of the questionnaire, the final form of the scales with their psychometric characteristics, and some reference to reliability in the form of Cronbach's alpha. Opacity in reporting is typical regarding researchers' rationale for selecting or discarding items and the choices made regarding analytical procedures. These descriptions allow the reader to understand the motivation under the decisions made and how the elimination or replacement of

items conforms to the instrument under a particular theoretical proposal. Other issues stem from the search for the final form of the instrument. While the articles reviewed always name and specify the analytical method used to identify the instrument's structure, their usage is not always in line with the recommended practices in statistics or psychometrics (Furr, 2013; Preacher & MacCallum, 2003). One suggestion for improving this issue would be to focus on a better alignment between the analytical methods used to characterize the structure of the instrument and the knowledge already existent in statistics or measurement disciplines. Following these guidelines would allow researchers to build upon a solid base for all their analytical choices.

Finally, there is still a challenge concerning further examination and instrument quality documentation. Standards in measurement regarding reliability and validity have long since moved away from relying on Cronbach's alpha as a measure of reliability (AERA et al., 2014). No evidence for reliability or validity results in a lack of a substantive claim of the consistency of the scores through the testing process. A suggestion for future researchers is to provide, first and foremost, an interpretation of the scores and an explicit statement on the intended uses of the instrument (Wilson, 2004). A validity argument that adequately supports inferences based on an instrument does not depend solely on the data provided by a single article; it also depends on the evidence gathered by other studies using an instrument that shares a common understanding of the construct. Ideally, these studies have the same specific use or interpretations of the instrument's score. Without these specifications, there is no way to advance in establishing a solid base for the validity of these instruments (AERA et al., 2014).

7.3. On the effect of control variables on credibility and perceived quality of the scientific text

In both studies, there is evidence of the effect of control variables (i.e., credibility and perceived quality of the science) on the texts. The natural science trained participants rated the abstracts as more credible and of better quality than their social sciences counterparts in both studies. A variety of possible reasons could be behind these results. One of them could be the greater familiarity of the students of natural sciences disciplines with the scientific journals and their presentation format. These abstracts all came from the same reputable high-impact source. Knowing that information could have tipped the credibility above the level compared to other participants not in scientific communities familiar with specific scientific journals.

In the first study, there is also a significant effect of perception of expertise on the text's credibility and scientific quality. When the perception of expertise of the participant was higher, the ratings of credibility and quality of the science rose. While there is evidence that students' reading skills develop throughout their career (Hubbard & Dunbar, 2017), less is known about why a greater perception of expertise in the discipline raises scores on aesthetic appreciation, credibility, and perception of the quality of the texts.

7.4. Conclusion

The results of both empirical studies of this thesis suggest that presentational elements influence the perception and understanding of scientific texts. These results contribute to the vision that science and its outcomes, such as the scientific paper, are part of a social practice beyond discursive thought (Gross, 2006; Root-Bernstein, 2002; Thagard, 2002). Even though science and the practice of scientists are thought to be of objective nature, the evidence found in our studies suggests that such practices make use of other elements besides the propositional content of what

is communicated. Namely, there are presentational elements that frame the content in such ways that allow distinctions regarding the understanding or credibility of the text to be made just based on these changes. It is important to note that these changes do not need to be of a grand scale to have an impact. Previous research has mainly focused on structural presentational changes to prove that these elements influence how science is perceived (McCabe & Castel, 2008; van der Meij et al., 2013; Vidal-Abarca & Sanjosé, 1998). However, as the current studies have suggested, even a slight modification of the wording of a text can change readers' perception of its scientific quality and comprehensibility.

In our case, the presentational elements modified were close to the thematic field of the studies' participants. The modifications of these elements triggered changes in the perceptual field effectively regarding the comprehension of the abstract. This effect did not replicate in the appreciation of the texts. This lack of impact probably is due to the difficulty of assessing aesthetic appreciation, considering its presentational features and the nature of the instrument used to measure it. Future research might consider a more nuanced measurement of the perception of the text to determine if the experimental manipulations did influence the overall appreciation of the texts.

Following the theoretical framework proposed in this thesis, the effect of the presentational elements on the perception of scientific texts is co-concurrent with the effect of many other variables that the studies in the present thesis are measured as control variables. The effect of perceived expertise, for example, tells us that many other elements are co-present in the perceptual field of consciousness (Gurwitsch, 2009; Yoshimi & Vinson, 2015). These elements are part of how we perceive the stimuli part of our attentional focus. The place of each component on the thematic field of consciousness could affect the prominence in which the person considers the

presentational as a vital element of the perception and understanding of scientific texts. The prominence of this phenomenon in the arts may be due to the closeness of presentational elements to the focus of consciousness. Therefore, we could expect that the role of the presentational in science, although recognized, does not have such importance to scientists. A minor preponderance does not mean that presentational elements do not play an important role in science communication. It just means that these elements are further away from the focus of the attention of the people that work in this area of human knowledge.

A thought-provoking implication of this thesis is to think of all types of knowledge as categories permeated with symbolic elements that go beyond those represented by language. Doubtless, language is a crucial element to communicate and establish consensus in human practices – such as science – but it is not the unique way in which human knowledge is articulated. We need to recognize that while language is the form that allows us to structure a specific type of information, there are other types of experience articulated through non-discursive symbols. Usually, this type of symbolization has been identified in practices such as art and craftsmanship (Era et al., 2019; Polanyi, 1962; Stamatopoulou & Cupchik, 2017), but the increasing amount of evidence up until now suggest that it is part of all human practices in a greater or lesser degree. In the sciences, the presentational is configured through a set of aesthetic or conventionally determined practices by the scientific community to which the researchers belong. However, the effect of these elements on the conscious perception of people is very delicate, and it needs to be measured in more refined ways to provide a more robust knowledge about its workings.

These results are also significant in the field of public understanding of science. Our results suggest that laypeople and people who have some training in the discipline are subject to these effects in the presentation of scientific content or texts. The distinction between the perception of

laypeople vs. experts regarding science does not seem as blunt when we consider that the current evidence for the effect of presentational features on understanding permeates all levels of experience. A distinction that could help delineate the differences between levels of expertise is the specific presentational features that non-experts and experts use to determine the understanding or comprehensibility of science media in general. Another relevant finding is that presentational elements have a differential effect depending on the participants' area of expertise. While there have been some developments in both disciplines, more research has been dedicated to the understanding of science in STEM-related disciplines (e.g., Crick & Hartling, 2015; Hart, 2013; Hayn-Leichsenring et al., 2021). The current results prove that more studies are needed to reveal how presentational elements are perceived through different scientific practices.

More research to confirm these hypotheses and other previous ideas is needed. Considering the findings and the variety of areas related to this topic, a greater cross-reference between disciplines is required to strengthen the findings on presentational elements. The current distance between philosophy of science, the studies of public understanding of science, psychology, and rhetoric studies allows considerable research inside specific fields. However, it does not benefit from evidence from other disciplines.

There are some limitations of this thesis worth mentioning. As the participants were only undergraduates, we cannot state that these modifications would have the same effect in a sample such as experts in each field or laypeople without college training. Different samples would be needed to extend the claims of the thesis to other groups of people that participate in science communities. Likewise, further studies would benefit from including a greater variety of measurements of the dependent variables contemplated in our studies, especially regarding the understanding of scientific texts. As shown by previous and current studies, the distinction between

the perception of understanding and deep measures of understanding is an important issue. Considering the current results and other studies, the two concepts correspond to different phenomena.

Overall, these results support the hypothesis that scientific texts are understood as a complex meaning-making unit that integrates content and presentational aspects to the reader's benefit. Some configurations of these elements allow for a better perception understanding in participants trained in sciences. In this respect, non-discursive elements of the texts work as paralinguistic elements of oral discourse (Marcuschi, 2005). These elements frame the whole experience of the text, not only in artistic contexts but also, as our data reveal, in scientific practice and communication.

8. References

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Aldunate, N., Infante, J., Carré, D., & Cornejo, C. (2009). Saber -cómo sin saber- qué. Estudio microgenético de la percepción de caras. *Avances en Psicología Latinoamericana*, *27*(2), 311–326.

American Psychological Association. (2020). *Publication Manual of the American Psychological Association* (7th ed.). American Psychological Association.

Bachrach, A., Fontbonne, Y., Joufflineau, C., & Ulloa, J. L. (2015). Audience entrainment during live contemporary dance performance: Physiological and cognitive measures. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00179

Baker, D. A., Ware, J. M., Schweitzer, N. J., & Risko, E. F. (2017). Making sense of research on the neuroimage bias. *Public Understanding of Science*, *26*(2), 251–258.

Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science.* The University of Winsconsin Press.

Bergson, H. (1959). *Matter and Memory*. Double Day/Anchor Books.

Bostian, L. R. (1983). How Active, Passive, and Nominal Styles Affect Readability of Science Writing. *Journalism Quarterly*, *60*(4 (Winter)), 635–640. https://doi.org/10.1177/107769908306000408

Brieber, D., Leder, H., & Nadal, M. (2015). The Experience of Art in Museums: An Attempt to Dissociate the Role of Physical Context and Genuineness. *Empirical Studies of the Arts*,

33(1), 95–105. https://doi.org/10.1177/0276237415570000

Carré, D. (2018). Towards a Cultural Psychology of Science. *Culture & Psychology*, 1–30. https://doi.org/10.1177/1354067X18790000

Crick, K., & Hartling, L. (2015). Preferences of Knowledge Users for Two Formats of Summarizing Results from Systematic Reviews: Infographics and Critical Appraisals. *PLOS ONE*, *10*(10), e0140029. https://doi.org/10.1371/journal.pone.0140029

Daston, L., & Galison, P. (2008). *Objectivity* (Third). Zone Books.

Day, R. A., & Gastel, B. (2011). *How to write and publish a scientific paper*. Greenwood.

de Regt, H. W. (2004). Discussion Note: Making Sense of Understanding*. *Philosophy of Science*, *71*(1), 98–109. https://doi.org/10.1086/381415

de Regt, H. W., & Dieks, D. (2005). A Contextual Approach to Scientific Understanding. *Synthese*, *144*(1), 137–170. https://doi.org/10.1007/s11229-005-5000-4

DeBakey, S., & DeBakey, L. (1983). The title. What's in a name? *International Journal of Cardiology*, *2*(3–4), 401–406. https://doi.org/10.1016/0167-5273(83)90012-8

Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *Lancet*, *350*(9074), 326–329. https://doi.org/10.1016/S0140-6736(97)02419-7

Einstein, A. (1949). *Autobiographical Notes*. Open Court.

Era, V., Candidi, M., & Aglioti, S. M. (2019). Contextual and social variables modulate aesthetic appreciation of bodily and abstract art stimuli. *Acta Psychologica*, *199*, 102881. https://doi.org/10.1016/j.actpsy.2019.102881

Feist, G. J. (2006). How Development and Personality Influence Scientific Thought, Interest, and Achievement. *Review of General Psychology*, *10*(2), 163–182. https://doi.org/10.1037/1089-2680.10.2.163

Feist, G. J. (2012a). Predicting interest in and attitudes toward science from personality and need for cognition. *Personality and Individual Differences*, *52*(7), 771–775. https://doi.org/10.1016/j.paid.2012.01.005

Feist, G. J. (2012b). The psychology of science is off and running but where do we go from here? In R. W. Proctor & E. J. Capaldi (Eds.), *Psychology of Science: Implicit and Explicit Processes* (pp. 45–66). https://doi.org/10.1093/acprof:oso/9780199753628.001.0001

Furr, M. (2013). *Scale Construction and Psychometrics for Social and Personality Psychology*. SAGE.

Gabrielsson, A., & Lindström, S. (2003). Strong Experiences Related to Music: A descriptive System. *Musicae Scientiae*, *7*(2), 157–217. https://doi.org/10.1177/102986490300700201

Gheorghiu, A. I., Callan, M. J., & Skylark, W. J. (2017). Facial appearance affects science communication. *Proceedings of the National Academy of Sciences*, *114*(23), 5970–5975. https://doi.org/10.1073/pnas.1620542114

Goodman, N. (2012). Familiarity breeds: Clichés in article titles. *British Journal of General Practice*, *62*(605), 656–657. https://doi.org/10.3399/bjgp12X659420

Gopen, G. D., & Swan, J. A. (1990). The Science of Scientific Writing. *American Scientist*, *78*(6), 550–558.

Grosbras, M. H., Tan, H., & Pollick, F. (2012). Dance and emotion in posterior parietal cortex: A

low-frequency rTMS study. *Brain Stimulation*, *5*(2), 130–136. https://doi.org/10.1016/j.brs.2012.03.013

Grosbras, M.-H., Tan, H., & Pollick, F. (2012). Dance and emotion in posterior parietal cortex: A low-frequency rTMS study. *Brain Stimulation*, *5*(2), 130–136. https://doi.org/10.1016/j.brs.2012.03.013

Gross, A. (2006). *Starring the text: The place of rhetoric in science studies*. SIU Press.

Gurwitsch, A. (2009). *Volume I. Constitutive Phenomenology in Historical Perspective*. Springer.

Hallock, R. M., & Bennett, T. N. (2021). I'll Read That!: What Title Elements Attract Readers to an Article? *Teaching of Psychology*, *48*(1), 26–31. https://doi.org/10.1177/0098628320959948

Hart, P. S. (2013). The role of numeracy in moderating the influence of statistics in climate change messages. *Public Understanding of Science*, *22*(7), 785–798. https://doi.org/10.1177/0963662513482268

Hartley, J. (2003). Improving the clarity of journal abstracts in psychology: The case for structure. *Science Communication*, *24*(3), 366–379. https://doi.org/10.1177/1075547002250301

Hayn-Leichsenring, G. U., Vartanian, O., & Chatterjee, A. (2021). The role of expertise in the aesthetic evaluation of mathematical equations. *Psychological Research*. https://doi.org/10.1007/s00426-021-01592-5

Hubbard, K. E., & Dunbar, S. D. (2017). Perceptions of scientific research literature and strategies for reading papers depend on academic career stage. *PLOS ONE*, *12*(12),

e0189753. https://doi.org/10.1371/journal.pone.0189753

Hung, S.-M., Styles, S. J., & Hsieh, P.-J. (2017). Can a Word Sound Like a Shape Before You Have Seen It? Sound-Shape Mapping Prior to Conscious Awareness. *Psychological Science*, *28*(3), 263–275. https://doi.org/10.1177/0956797616677313

ICMJE. (2016). Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. *Citeseer*, *December*, 1–17. https://doi.org/10.1080/08941920.2016.1150542

Jola, C., Ehrenberg, S., & Reynolds, D. (2012). The experience of watching dance: Phenomenological – neuroscience duets. *Phenomenology and the Cognitive Sciences*, *11*(1), 17–37. https://doi.org/10.1007/s11097-010-9191-x

Katz, M. J. (2006). *From research to manuscript: A guide to scientific writing*. https://doi.org/10.1007/1-4020-4071-7

Kraxenberger, M., & Menninghaus, W. (2016a). Emotional effects of poetic phonology, word positioning and dominant stress peaks in poetry reading. *Scientific Study of Literature*, *6*(2), 298–313. https://doi.org/10.1075/ssol.6.2.06kra

Kraxenberger, M., & Menninghaus, W. (2016b). Mimological reveries? Disconfirming the hypothesis of phono-emotional iconicity in poetry. *Frontiers in Psychology*, *7*(NOV), 1–9. https://doi.org/10.3389/fpsyg.2016.01779

Kraxenberger, M., & Menninghaus, W. (2017). Affinity for Poetry and Aesthetic Appreciation of Joyful and Sad Poems. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.02051

Langer, S. K. (1942). *Philosophy in a new key*. Harvard University Press.

Langer, S. K. (1970). *Mind: An essay on human feeling: Vol. I*. The Johns Hopkins Press.

Leahey, E., & Cain, C. L. (2013). Straight from the source: Accounting for scientific success. *Social Studies of Science*, *43*(6), 927–951. https://doi.org/10.1177/0306312713484820

Leather, S. R. (1996). The case for the passive voice. *Nature*, *381*(6582), 467–467. https://doi.org/10.1038/381467a0

Leder, H., Gerger, G., Brieber, D., & Schwarz, N. (2014). What makes an art expert? Emotion and evaluation in art appreciation. *Cognition and Emotion*, *28*(6), 1137–1147. https://doi.org/10.1080/02699931.2013.870132

Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, W., Jacobs, A. M., & Koelsch, S. (2015). Reading a Suspenseful Literary Text Activates Brain Areas Related to Social Cognition and Predictive Inference. *PLOS ONE*, *10*(5), e0124550. https://doi.org/10.1371/journal.pone.0124550

Marcuschi, L. A. (2005). A oralidade no contexto dos usos lingüísticos: Caracterizando a fala. In L. A. Marcuschi & A. Paiva Dionisio (Eds.), *Fala e Escrita* (pp. 57–84). Autentica.

Marin, M. M., & Leder, H. (2018). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. *Personality and Individual Differences*, *125*, 80–90. https://doi.org/10.1016/j.paid.2017.12.027

Massaro, D., Savazzi, F., Di Dio, C., Freedberg, D., Gallese, V., Gilli, G., & Marchetti, A. (2012). When Art Moves the Eyes: A Behavioral and Eye-Tracking Study. *PLoS ONE*,

7(5), e37285. https://doi.org/10.1371/journal.pone.0037285

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, *107*(1), 343–352. https://doi.org/10.1016/j.cognition.2007.07.017

McKiernan, G. (2000). arXiv.org: The Los Alamos National Laboratory e-print server. *The International Journal on Grey Literature*, *1*(3), 127–138.

McMillan, K., & Weyers, J. (2007). *How to write essays and assignments*. Pearson Education.

Meneghini, R., Packer, A. L., & Nassi-Calò, L. (2008). Articles by Latin American authors in prestigious journals have fewer citations. *PLoS ONE*, *3*(11), 1–4. https://doi.org/10.1371/journal.pone.0003804

Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., von Koppenfels, M., & Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(2), 362–373. https://doi.org/10.3758/s13415-015-0396-x

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. A. (2013). Aesthetic and Emotional Effects of Meter and Rhyme in Poetry. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00010

Osbeck, L. M., Nersessian, N. J., Malone, K. R., & Newstetter, W. C. (2011). *Science as Psychology* (First). Cambridge Press University.

Palmer, S. E., Schloss, K. B., & Griscom, W. S. (2016). Individual Differences in Perceptual Preference. *Electronic Imaging*, *2016*(16), 1–6. https://doi.org/10.2352/ISSN.2470-

1173.2016.16.HVEI-113

Pelowski, M., Leder, H., Mitschke, V., Specker, E., Gerger, G., Tinio, P. P. L., Vaporova, E., Bieg, T., & Husslein-Arco, A. (2018). Capturing Aesthetic Experiences With Installation Art: An Empirical Assessment of Emotion, Evaluations, and Mobile Eye Tracking in Olafur Eliasson's "Baroque, Baroque!" *Frontiers in Psychology*, *9*, 1255. https://doi.org/10.3389/fpsyg.2018.01255

Pihko, E., Virtanen, A., Saarinen, V.-M., Pannasch, S., Hirvenkari, L., Tossavainen, T., Haapala, A., & Hari, R. (2011). Experiencing Art: The Influence of Expertise and Painting Abstraction Level. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00094

Poincaré, H. (1946). *The Foundations of Science*. Science Press.

Polanyi, M. (1962). *Personal Knowledge*. The University of Chicago Press.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's Electric Factor Analysis Machine. *Understanding Statistics*, *2*(1), 13–43. https://doi.org/10.1207/S15328031US0201_02

Proffitt, D. R., Stefanucci, J., Banton, T., & Epstein, W. (2003). The role of effort in perceiving distance. *Psychological Science*, *14*(March 2003), 106–112. https://doi.org/10.1002/jlcr

Reason, M., & Reynolds, D. (2010). Kinesthesia, Empathy, and Related Pleasures: An Inquiry into Audience Experiences of Watching Dance. *Dance Research Journal*, *42*(2), 49–75. https://doi.org/10.1017/S0149767700001030

Robbins, S. (2010). Ethics and aesthetics of style: The import of stylistic choices in academic

prose. *Filologija*, *15*. http://www.su.lt/bylos/mokslo_leidiniai/filologija/2010_15/robbins.pdf

Root-Bernstein, R. S. (2002). Aesthetic cognition. *International Studies in the Philosophy of Science*, *16*(1), 61–77. https://doi.org/10.1080/02698590120118837

Root-Bernstein, R. S., Bernstein, M., & Garnier, H. (1995). Correlations between avocations, scientific style, work habits, and professional impact of scientists. *Creativity Research Journal*, *8*(2), 115–137.

Ruben, A. (2012). *How to write like a scientist*. http://www.sciencemag.org/careers/2012/03/how-write-scientist

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., & Scherer, K. R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLOS ONE*, *12*(6), e0178899. https://doi.org/10.1371/journal.pone.0178899

Shanon, B. (2013). *The representational and the presentational*. Andrews UK Limited.

Silvia, P. J., & Nusbaum, E. C. (2011). On personality and piloerection: Individual differences in aesthetic chills and other unusual aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts*, *5*(3), 208–214. https://doi.org/10.1037/a0021914

Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin*, *129*(4), 475–494. https://doi.org/10.1037/0033-2909.129.4.475

Solovova, O., Santos, J., & Veríssimo, J. (2018). Publish in English or Perish in Portuguese: Struggles and Constraints on the Semiperiphery. *Publications*, *6*(2), 25.

https://doi.org/10.3390/publications6020025

Stamatopoulou, D., & Cupchik, G. C. (2017). The Feeling of the Form: Style as Dynamic 'Textured' Expression. *Art & Perception*, *5*(3), 262–298. https://doi.org/10.1163/22134913-00002066

Stern, D. (2010). *Forms of vitality*. Oxford University Press.

Thagard, P. (2002). The passionate scientist: Emotion in scientific cognition. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 235–250). Cambridge University Press.

van der Meij, H., van der Meij, J., & Farkas, D. K. (2013). QuikScan formatting as a means to improve text recall. *Journal of Documentation*, *69*(1), 81–97. https://doi.org/10.1108/00220411311295333

Vidal-Abarca, E., & Sanjosé, V. (1998). Levels of comprehension of scientific prose: The role of text variables. *Learning and Instruction*, *8*(3), 215–233. https://doi.org/10.1016/S0959-4752(97)00020-0

Vieira, A. (2011). The Art of Technical and Scientific Writing. *Revista de Sistemas de Informacao Da FSMA*, *8*, 22–30.

Vygotsky, L. (2008). *Psicología del Arte*. Paidós.

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017). The emotional power of poetry: Neural circuitry, psychophysiology, compositional principles. *Social Cognitive and Affective Neuroscience*, *2017*, 1–12. https://doi.org/10.1093/scan/nsx069

Wassiliwizky, E., Wagner, V., Jacobsen, T., & Menninghaus, W. (2015). Art-elicited chills indicate states of being moved. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 405–416. https://doi.org/10.1037/aca0000023

Werner da Rosa, C., & Otero, J. (2018). Influence of source credibility on students' noticing and assessing comprehension obstacles in science texts. *International Journal of Science Education*, *40*(13), 1653–1668. https://doi.org/10.1080/09500693.2018.1501168

Werner, H. (1956). Microgenesis and aphasia. *The Journal of Abnormal and Social Psychology*, *52*(3), 347–353. https://doi.org/10.1037/h0048896

Werner, H., & Kaplan, B. (1964). *Symbol Formation* (2nd ed.). John Wiley & Sons.

White, B. (2000). *Dissertation skills for business and management students*. Thomson.

Williams, L. E., & Bargh, J. A. (2008). Experiencing Physical Warmth Promotes Interpersonal Warmth. *Science*, *322*(5901), 606–607. https://doi.org/10.1126/science.1162548

Wilson, M. (2004). *Constructing Measures: An Item Response Modeling Approach*. Routledge. https://doi.org/10.4324/9781410611697

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. Routledge.

Writing for a Nature journal. (2017). http://www.nature.com/authors/author_resources/how_write.html

Yoshimi, J., & Vinson, D. W. (2015). Extending Gurwitsch's field theory of consciousness. *Consciousness and Cognition*, *34*, 104–123. https://doi.org/10.1016/j.concog.2015.03.017