IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 31, 2021, accepted May 3, 2021, date of publication May 10, 2021, date of current version May 19, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3078429

Visual Information Literacy: Definition, Construct Modeling and Assessment

ANGELA LOCORO[1], **WILLIAM P. FISHER, JR.**,[2], **AND LUCA MARI**[1]
[1]Università Carlo Cattaneo–LIUC, 21053 Castellanza (VA), Italy
[2]BEAR Center, Graduate School of Education, University of California, Berkeley, Berkeley, CA 94720, USA

Corresponding author: Angela Locoro (alocoro@liuc.it)

This work was supported by Università Carlo Cattaneo - LIUC, funding the 2019 research project with title ''Walking the last mile from data to decision-making: the role of data validation and data visualization.''

ABSTRACT A major problem in education and visual information design is that, while tools to measure people's reading and writing ability with texts and numbers are ripe, the ability to properly process information from data graphics – an ability that can be called Visual Information Literacy – is still off the radar, and even less interest is apparently devoted to its evaluation. The purpose of this research is that of presenting an exploration of methods and tools towards the measurement of data graphics effectiveness and efficiency, and of proposing a definition of 'Visual Information Literacy', together with the design of a model characterizing it as a developmental skills progression that covers the cognitive abilities activated when dealing with data graphics. A final goal of this paper is to report a first round of results assessing the validity of the model designed, by bringing statistical evidence that data graphics comprehension depends on the matching of users' ability and data graphics difficulty. The contribution of this paper is twofold: comparing the current research on Visual Information Literacy and advancing it by designing a model for its characterization to allow the design of a Visual Information Literacy measurement scale standard.

INDEX TERMS Computers and information processing, information and communication technology, information management, visual communication, visualization.

NOMENCLATURE

ATVC	Attitude Towards Visual Communication.
BAS	Bear Assessment System.
DG	Data Graphics.
DWI − FW	Legacy Data Visualizaion Literacy Framework.
AM	Abstract Mappings.
AS	Abstract Systems.
C	Compute.
E	Explain.
ELA	English Language Arts Literacy.
GD	Goal Description.
I	Infer.
M	Map.
O	Observe.
P	Principles.
PISA	Program For International Student Assessment.
QSC	Quantile Skill and Concept.
R	Reason.
RM	Representational Mappings.
RS	Representational Systems.
SA	Single Abstractions.
SR	Single Representations.
VLAT	Visual Literacy Assessment Test.

The associate editor coordinating the review of this manuscript and approving it for publication was Ali Kashif Bashir.

I. INTRODUCTION AND MOTIVATIONS

The ability to understand and reproduce human written signs systems, usually called literacy,[1] may rely on many codification tools: words, numbers, graphics, and the like. Cognitive models help hypothesize, analyze, and assess how individuals understand complex concepts (declarative knowledge) and apply skills (procedural knowledge), also by means of their

[1]Some authors (cf for example Taylor [1]) argue that this is an improper or rather metaphorical extension of the meaning of the word, which is claimed to have to do only with sign systems of words, as the etymology indicates from the Latin introduction of this word, defined as ''[…] having knowledge of letters'' (see at https://www.etymonline.com/word/literate.). However, we advocate the evolution that the cultural meaning of the word literacy has had through studies in Information science, Visual arts, and Educational studies, where often what it is called literacy goes together with text comprehension, or rather goes beyond it (assuming it for acquired and required), and includes the capacity of managing numbers (that some also call ''numeracy'') as well as understanding pictures (that some also call ''graphicacy'').

literacy level, throughout their active lifetime [2]. Many literacy assessment tests, based on cognitive tasks, became a standard measuring instrument of student's literacy with texts,[2] with numbers,[3] and even with "domain-based literacies".[4] Despite these well established efforts to provide solid tools for educational assessment, the initiatives to envision cognitive models from which to start investigating how people process information are rarer. This lack of vision limits the pace of interpretability in ever faster shifting cultural horizons like that of the knowledge society we are living in [3]. Information is a key asset of both production and communication [4], and information management and processing is one of the crucial aids towards ethical and sustainable decision making (see for example Floridi in [5]).

Fact saliency is ever more supported by visual media. The reason is that the supportive power of visual language allows a faster comprehension and selection of relevant patterns in information. This capability may be enabled by its analogical nature, when compared to the digital nature of words [6]. The basic intuition of this distinction regards the fact that information is encoded in visual representation by means of analogies (e.g., in a bar chart, the height of bars represents not only the quantitative intensities of values of a property but enables their immediate comparison — a higher bar immediately recalls a greater quantity). This representational encoding more naturally (and "innately") adheres to the way individuals observe, perceive and understand the world they are immersed in. On the contrary, digital languages are the result of a conventional agreement among individuals, hence less "spontaneous" and more cognitively demanding (their codification does not recall any perceptual attitude and does not allow any analogical, "at a glance" form of decoding).

Visual Information Literacy, in short the ability to properly process information related to data graphics, i.e., encoding information into data graphics and decoding information from data graphics, is one of the conceptual and practical skills that are now deemed as important as textual literacy and numeracy.[5] The growing importance of understanding data that become bigger, faster and even more complex renders "digital citizenship" key competences such as Visual Information Literacy an inevitable wealth of skills for good, responsible and sustainable information interpretation and information-based decision making.

Nevertheless, little attention has been paid to it so far in the field of educational assessment tests [7] and, in general, in the design of students' curricula [8]. Likewise, in the fields of Information and Data Visualization, recent attempts to model Visual Information Literacy to assess the ability of individuals to deal with data graphics have shown a lack of rigour and of a methodology able to lead to a standard measurement scale like the ones designed for text literacy and numeracy. Lacking this perspective, the field of Information Systems is also affected [9].

Arguably, this short-sightedness for Visual Information Literacy models, assessments and synergies in educational and professional learning have severe repercussions on peoples' lives. Likewise, the unavailability of models and the limited applicability of the above assessment tools to adults is also a sign that the consequences of illiteracy, which may dramatically reduce the chance of making good decisions in critical contingencies and future scenarios, do not urge stakeholders to provide models and tools.

This paper aims at bringing to the forefront the problem of the lack of a model characterizing Visual Information Literacy, of a standard measurement scale for assessing the literacy level of individuals, and of a methodology for the design of both the model and the measurement scale standard. In so doing, the paper develops along the following directions to: make an overview of the main terms, models and assessment methods in education and data visualization domains, and propose a definition of Visual Information Literacy for a specific domain and class of visualizations, that of data graphics (Section II); introduce a cross domain conceptual framework of developmental cognition to design a Visual Information Literacy model based on this framework but referencing peculiarity of data graphics, and design an initial set of items to assess the validity of a construct based on the model (Section III); report the results of an initial survey administering those items, for the analysis (Section IV) and discussion (Section V) of the construct assessment and the items reliability, together with the limitations of what has been proposed so far (Section VI). This paper contributes to make a point about the need to provide standard tools for defining and measuring Visual Information Literacy, and to propose a definition and a model characterizing it, in the direction of designing a standard measurement scale for taking Visual Information Literacy as seriously as text literacy and numeracy for an informed citizenship.

II. BACKGROUND

A. WHAT IS VISUAL INFORMATION LITERACY

Our overview of the literature does not claim to be exhaustive. Nonetheless we are well aware of the main current studies on Visual Information Literacy assessment and we refer in the following to these specific sources: the main definitions of what we have termed "Visual Information" are taken from Education and Literacy Studies [1], [10] and Data Visualization and Visual Information fields [7], [11], [12]; the mainstream researches about what we have termed "Visual Information Literacy" assessment are taken from the above

[2]E.g., see the "Program for International Student Assessment" or PISA, available at https://www.oecd.org/pisa/, or the Lexile framework, available at https://lexile.com/.

[3]E.g., see the PISA and the Quantile framework, available at https://www.quantiles.com/.

[4]E.g., see again the definition of the PISA test whose aim is to "measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges".

[5]See, e.g., the Association of College and Research Libraries (ACRL) introductory statement: "The importance of images and visual media in contemporary culture is changing what it means to be literate in the 21st century." available at www.ala.org/acrl/standards/visualliteracy

sources. In particular, we considered the literature related to the current edition of the most important conference in the Visualization field (IEEE Vis[6]), together with some foundational studies (e.g., in the Journal of Visual Literacy [13]) and most recent advances (e.g., in the IEEE Transactions on Visualization and Computer Graphics [7], [12]) about assessing Visual Information Literacy[7] We also referred to the Educational and Literacy Studies field.[8] We excluded from our literature overview domain-based visual literacy (e.g., in Biochemistry or Medicine [15]) and too broad studies (e.g., those encompassing multimodal or visual art media and languages).

The literature examined proposes several key terms to designate Visual Information Literacy or some of its aspects, depending on the kind of visual language and the strand of research dealing with each of them. In the following paragraphs, each definition is presented from the narrowest to the broadest scope, whereas in the next Section (II-B) we start from a generic view to delve into specific models.

This topic has been dubbed in similar ways, but defined very differently, for example:

1) VISUALIZATION LITERACY

As proposed by [7], [11], [12], visualization literacy regards the comprehension of images designed to visualize data, where visualization literacy means ''the ability to confidently use a given data visualization to translate questions specified in the data domain into visual queries in the visual domain, as well as interpreting visual patterns in the visual domain as properties in the data domain'' [7];

2) VISUAL LITERACY

This notion, in the words of [1] relates to all of the ''formally-defined symbols and visual conventions'' such as ''diagrams, maps, charts, graphs, [and] explanatory pictorial representations''. For Taylor, visual literacy is ''a proficiency in reading and writing systems for encoding ideas in visual form [with the aim of] becoming fluent in the skills of reading and understanding communications, and constructing new communications''. This term is again proposed with similar nuance in [13], [16], [17], where it includes the broadest domain of natural image representation (e.g., photographs), and is defined as ''the abilities to understand, interpret, and evaluate visual messages'' [18];

3) GRAPHICACY

This term was coined by [10], where it refers to the ''critical use of inscriptions, the knowledgeability relative to sketches, photographs, diagrams, maps, plans, charts, graphs and other non-textual, two dimensional formats''.

The definitions and usage of the terms are all different and, as suggested, they hint at scopes of different size: for example, the above notion of Visual Literacy is understood to have too wide a scope for our purposes, as it may include also art work and natural objects. On the other hand, the above notion of Visualization Literacy is too narrowly defined, as it seems to only take partial note of the translation mechanisms between data graphics and their interpretation. In search of a notion encompassing as many ''information literacy of data graphics'' aspects as possible, maintaining that visual literacy refers to the ''ability to manage a conventional language of formally-defined symbols'' [1], we adopted the term ''Visual Information Literacy'' to better denote our scope and objects of investigation. Also identified so far in the literature across different domains, Visual Information Literacy is a multi-faceted concept understood to borrow aspects from visual, media and information literacy at the same time. Referring again to Taylor's paper, the ''visual'' aspect points to ''the ability to read and write visual systems of signs such as diagrams, maps, charts, graphs, and explanatory pictorial representations'', which are ''formally-defined symbols and visual conventions''. The ''information'' aspect points to the ''abilities to work well with modern information sources'', by means of strategies to ''locate, compare, evaluate, organize, apply and synthesize'' data, with the aim to ''communicate, explain and instruct people about abstract concepts and ideas'' towards ''construct[ing] new knowledge, understanding and [refining] communication skills'' [ibid.]. This definition includes the kind of graphical artifacts that are the objects of application of Visual Information Literacy, as well as some crucial aspects of literacy (human capabilities) in processing them. One of those aspects concerns the semiotic process of interpretation of signs, which leads from data to information and stands behind any communication process (cf. Eco's semiotics of language), and the related aspect of ''critical understanding'' of information (cf. Pierce's pragmatics). Another aspect concerns the knowledgeable ''creation of dimensionally new information'' that emerges from considering the ''literacy''-sided concept of ''gateway skill'', i.e., the generative property embodied in and enacted by any language of ''opening up access to a whole wide world of learning and understanding'' [ibid.].

As said above, this definition has an object of application: graphical artifacts. In this regard, there is no agreement about the most generic concept among 'chart', 'graph' and 'plot' (provided that they are not coextensive but only partially overlapping concepts), that may subsume all the visual information which we refer to: diagrams, maps, charts, graphs, and explanatory pictorial representations. Some authors subsume them under ''charts'' [19]; Wikipedia also refers to them as ''charts'', intended as ''representations of data''[9]; older but foundational references of semiotics and

[6]See the last edition at http://ieeevis.org/year/2020/welcome

[7]In particular we referred to two main recent studies, one European [7] and two American [12], [14] about Visual Literacy assessment, and to the literature referenced in them.

[8]We referred in particular to one of the standard educational test in the US, developed at MetaMetrics for Math skills and concepts, available at https://www.quantiles.com.

[9]Although there is no reference on their page to a prominent and more general visual information device – that of infographics, which not only represents data but also help contextualize and interpret them with complementary graphical elements.

grammatical aspects subsume them under the term "graphics" ([20], [21]); Edward Tufte, a prominent scholar in the field ([22], [23]), recently subsumed them under the notion of "data graphics". The subtle differences between the words "chart", "graph" and "plot" each emphasize a function of their being visual: the more symbolic, metaphorical, and iconic aspect of "data representations" (like being maps for the territory), the "visual rendering of mathematical objects", and the "representation of points as visual marks on a plane", respectively. None of them seems to properly subsume the others and we prefer to rely on a term that may subsume all of them, that is "data graphics". For this reason, we will refer to all of them by this term in the rest of this paper.

A comparison of the main definitions of "visual literacy", approaches to the topic, their underlying models and proposed assessments are reported in the synthetic view of Table 9 in Appendix A. As the focus of this paper is on a specific kind of visual language (that of visual information) and on the characterization of its cognitive elaboration and comprehension, the digression about the many notions of visual literacy ends here; the interested reader is invited to refer to the works mentioned in the above referenced table for further inquiry.

B. THEORETICAL MODELS FOR VISUAL INFORMATION LITERACY

As also recently noted by [14], existing models for Visual Information Literacy assessment focus on some aspects of the human cognition spectrum, and a unifying framework encompassing all of them has so far been disregarded. Some types alternatively characterizing the current models are:

1) Type 1: syntactic aspects of data graphics and their involvement with tasks [24];
2) Type 2: query answering skills with data graphics [25];
3) Type 3: cognitive tasks to be executed with and through data graphics [7].

Attempts have been made in the literature to assess Visual Information Literacy in reference to models of the three above kinds:

1) CURCIO

The most cited model underlying the concept of Visual Information Literacy assessment is the one by Curcio [25]. This model was devised in reference to data graphics and the basic cognitive operations associated with them. Regarding the above list, this is a Type 2 model, and depicts a three-steps interaction with data graphics, where each step is characterized by a level of "graph comprehension" (ibidem), namely:

1) a level where data are extracted;
2) an interpretation level where data are integrated with previous knowledge;
3) and a final step, where new knowledge is generated.

The more abstract structure of the model is a taxonomy of skills that people should use for finding answers by querying data graphics:

1) the elementary skill of information extraction;
2) the intermediate skill of finding relationships;
3) the advanced skill of synthesizing both.

However, this model is too generic to achieve either a sufficiently broader or deeper view of the levels of a cognitive underlying model for Visual Information Literacy.

2) BLOOM

Another model of recent exploitation is Bloom's taxonomy of educational objectives [26], which is a Type 3 model, referring to the above list. This model was not designed having Visual Information Literacy in mind nor it has been adapted to it in the study about Visual Information Literacy assessment where it was exploited.

Six levels of subsequent understanding of data graphics are described by mapping Bloom's abstract levels to more concrete levels [14], namely:

1) the knowledge level of recalling basic analogies with other objects and examples;
2) the comprehension level of understanding information;
3) the application level of solving new problems with the acquired understanding;
4) the analysis level of thorough dissection of conceptual features;
5) the synthesis level of creating new knowledge;
6) the evaluation level of judgement of what has been internalized.

Although it is sufficiently well structured and broad in scope to hold as a conceptualization of literacy, the order of some of its levels, as they were mapped in the same order as Bloom's original model, is disputable (e.g., may the comprehension and application of a phenomenon be antecedent to its analysis?). Furthermore, when applied to the field of data visualizations, its progression levels and their order could not be validated for data visualization tasks [14].

a: PINKER

A third model, the one exploited by [7], is also a Type 3 model, and borrows its cognitive levels and conceptual ground from Pinker's [27] graphs comprehension theory and its refinement by Trickett [28]. Boy and colleagues identified common steps for this family of graphs comprehension models, in sequential order:

1) a user's goal;
2) the process of data graphics schema elaboration with gestalt principles;
3) the encoding of salient features;
4) the identification of cognitive and interpretative strategies to be used with the data graphics at hand;
5) the information extraction;
6) the comparison of visual chunks;
7) the extraction of the relevant information to satisfy the initial goal.

In this model, the process of understanding comes from two opposite directions: from prior knowledge to interpretation, and from perception of the current features to interpretation. However, this model does not explicitly conceptualize how progression of "graphs comprehension" (ibidem) takes place, nor what the progression order is in which cognitive skills should apply and what supporting the assessment of Visual Information Literacy really means.

The model that mostly resembles the idea of developmental levels of Visual Information Literacy (Bloom's model) was not specifically designed nor adapted for exploring and measuring Visual Information Literacy and, hence, did not provide terrain for foundational notions about it. The rest of the models examined above were supportive of the Visual Information Literacy empirical assessment, rather than being the object and motivation of empirical assessment.

Moreover, they do not provide at all or provide only in ła limited way a conceptual framework for interpreting the results towards the establishment of a standard measurement scale of Visual Information Literacy.

After this overview of the background literature, many research questions remain unanswered: for example, what makes data graphics easy or difficult to interpret? What are the cognitive levels characterizing Visual Information Literacy? Are those levels the same or in the same order or importance or in magnitude similar to the ones devised for texts and numbers comprehension?

These questions all have to do with the foundational question: is it possible to construct a standard measurement scale for Visual Information Literacy? Furthermore, they have driven our search for a theoretical model, the provision of a developmental construct and the design of a test for the validation of the model.

C. ASSESSMENT TOOLS

Most of the studies mentioned in Section II-A came with methods for visual literacy assessment with data graphics (relative names are reported in Table 9 of Appendix A and abbreviations are explained in the Nomenclature). With the exception of DVI-FW [11], whose educational material included a part that was not freely accessible, all the others were available for scrutiny. These assessment tools were all based on visual items design and administration [7], [12], [14]. For VLAT, Lee *et al.* [12] conceived 12 data graphics and 53 multiple-choice test items covering 8 tasks with data graphics. Their assessment method followed a rigorous protocol borrowed from the American Educational Standard [29]. Analyses of responses were summarized in a table where each data graphics and corresponding items were classified according to a difficulty index and a discriminating index related to how well a task discriminated the difficulty of a data graphics depending on fixed thresholds in the respondents' performance. Boy *et al.* [7] assessed 12 tasks (including min and max computations, comparisons, averages, and the like) for each of the 4 kinds of data graphics (two line charts, one bar chart, and one scatterplot) designed for their experiment. Their analysis was based on Item Response Theory (IRT). Some items could not be validated and were discarded. A 2-parameter logistic curve seemed to better fit people's ability vs items discrimination. However, in this study a full measurement model was not developed nor completely assessed (for example, the logistic curve has quantitative domain and codomain; but how could we assess the fact that the properties it represents are in fact quantitative?). Burns *et al.* [14] ran 6 questions, one for each of the 2 versions (one badly designed, and one conventionally designed) of the 3 data graphics tested against the six levels of their model and the 18 tasks related to them, three tasks for each level. They elaborated the results with regression analysis for dichotomous items, and chi-squared analysis for responses distribution of Likert-scale and descriptive items. As also recalled in Section II-B, these authors could not validate the order of levels of Bloom's model for data visualization tasks and items.

III. METHOD

After this thorough analysis of existing models for Visual Information Literacy and the detailed report of their assessment methods, we may conclude that none of the models designed for data graphics and for Visual Information Literacy assessment were shown to follow any rigorous methodology in their design, nor were existing models (e.g., Bloom) either revisited for fitting or found to reach a validation from their assessment with data graphics (see the declaration of Burns at al. in the most recent work (2020) about Visual Information Literacy [14]).

Our model of human learning progression is intended to describe how skills are developed towards the highest literacy achievements, i.e., the ability to master the declarative and procedural knowledge of, as well as the creativity, around the topic of interest. Such models are called "cognitive developmental models" from Piaget's theories about stages of development characterizing them in terms of hierarchical levels of skills [31]. In these terms, each developmental level is a conceptualization of the kind of knowledge and the related skills activated by such knowledge. Besides their descriptive utility, models of this kind also serve as theoretical reference for guiding in the design of measuring instruments for assessing the learning progression of individuals with the topic of interest, along the progression steps of a model that can be built upon. Constructing a model of this kind is part of a rigorous methodology to construct standard measurement scales in the educational field to measure individuals' progression with a certain topic [32]. We followed this rigorous methodology to design and implement our Visual Literacy model, and items for its assessment. In particular, we adhered to the BAS workflow[10] as an "integrated approach to developing assessments that provide meaningful interpretations of student work relative to the cognitive and developmental goals of a curriculum" [33].[11] This methodology is aimed at validating a model quantitatively, and taking it as the base for constructing a measurement tool of the level of literacy of individuals on a measurement scale [34] [35].

[10] Available at https://bearcenter.berkeley.edu/page/about-bear.

[11] In the present work, we present the implementation of the first two steps of the BAS: that of the construct modelling and that of the item preparation and administration.

A model for Visual Information Literacy development should satisfy a minimum set of requirements to achieve the ambitious goal of becoming a standard reference model for describing Visual Information Literacy, and for measuring how much of this literacy individuals attain during their lifetime. With this aim, we hypothesize that Visual Information Literacy is a unidimensional property and that it could be described with total order levels. Furthermore, our implementation design is rigorously developed along the following steps, borrowed from the methodology above:

- identify the levels of Visual Information Literacy that define the developmental progress of individuals in processing data graphics;
- describe those levels in terms of what kind of descriptive and procedural knowledge is necessary to activate the Visual Information Literacy skills required to achieve each level;
- prepare items for an assessment test aimed at verifying the validity of its levels, as they were defined and described, and the reliability of these levels to constitute a good tool to assess the level of Visual Information Literacy of individuals with data graphics.

The above method is the one we followed to implement the design of our Visual Information Literacy model. Our model is descriptive, explanatory and predictive. A Visual Information Literacy model may also be prescriptive for educational and data visualization design practices when associated with a standard measurement scale.

In the next sections we apply the above steps to implement the design of a Visual Information Literacy model, to the items preparation and to the administering of items for construct validity assessment.

A. A MODEL FOR VISUAL INFORMATION LITERACY

In [30], Dawson-Tunik *et al.* made some reflections about how to characterize a cognitive framework to describe human cognitive development from concrete objects to abstract concept processing through the identification of a common pattern among several models. Their hypotheses were that human reasoning grows progressively with a structure based on the sequential nesting of subsequent levels of development. After having achieved one level, humans internalize it through an act of abstraction from the concrete level into a higher level of knowledge [30, p.165]. Furthermore, she argued that progressive attainment of judgement capacity could be measured along a variable that may identify the patterns of performance of indviduals along stages of development [30, p.170]. Her framework aligned several cognitive models: Fischer's model [31] [36], which emphasizes a "structure of hierarchical skills"; Commons and colleagues' model [37], which describes a "hierarchy of cognitive tasks"; Case [38] and Demetriou's and Valanides' model [39], which outlines a "hierarchy of processing [cognitive] functions" [30, p.165]. Dawnson-Tunik integrated all of the aspects of each of the above models into her cognitive framework by describing seven levels of abstractions in sequential nesting order that individuals are deemed to manage in relation to their cognitive maturity.

Starting from Dawson-Tunik's framework, we designed a Visual Information Literacy model, by mapping developmental levels of Visual Information Literacy into the sequentially nested conceptual structure resulting from the application of Piaget's theory of cognitive development upon which the above framework lies. In particular, a descriptive view of the levels conceived is reported in the list below, whereas a more formal schema is reported in Appendix B. Table 1 conveys a detailed yet descriptive view of the six levels of the model, making it a blueprint for the design of items based on it.

The following descriptive mapping takes into consideration all the levels of the framework proposed by Dawson-Tunik, along with the peculiarities of data graphics for characterizing a Visual Information Literacy property with the following developmental levels:

1) MAP: single representations (SR) level, which is supposed to conceptualize visual syntax elements and their organization into data graphics (e.g., knowing what the titles, the legends, the scales, the kind and the features depicted in a data graphics are); and single abstractions (SA) level, which was supposed to conceptualize the knowledge of referring single elements to single properties of represented entities (e.g., knowing what the title, the legends and the features are representing) [20];
2) INTEGRATE: representational mappings (RM) level, which is supposed to conceptualize the knowledge of having observed the whole data graphic by having related single structures into an integrated structure of the entity represented in it (e.g., observing a pie chart and mapping it to the representation of the proportion of values of a characteristic of a sample population) [10];
3) COMPUTE: representational systems (RS) level, which is supposed to conceptualize the mapping from the knowledge of visual systems to the knowledge of other systems such as the quantitative one. For example, taking the above mentioned pie chart, at this level quantification of the proportions of value takes place (e.g., resulting that one proportion value is twice another proportion value). At this level, computational tasks are supposed to be supported by visual clues [12], [14], [40];
4) REASON: abstract mappings (AM) level, which is supposed to conceptualize the knowledge acquired with data graphics and transform it into conclusions about the phenomena represented in it (e.g., conclude from the integration and computation on the above pie chart that there is a disproportion between two strata of the sample population represented) [13];
5) INFER: abstract systems (AS) level, which is supposed to conceptualize the control of systematic knowledge for the acquisition of new knowledge, within a certain confidence interval. This new knowledge is inferred from the knowledge of previous conclusions drawn

TABLE 1. Progress levels [30] for a developmental cognitive model of Visual Information Literacy. "Data graphics" is abbreviated to "DG" in the Table.

Level	Examples	Progress and order structure
1 - Map		1st order - Single representations and abstractions (SR & SA)
Know what information is encoded in DG	Identify titles, annotations, and legends.	
Know how information is encoded in DG	Identify scales (categorical, numerical, ordinal), and aggregation level of data.	
Know which visual structures encoded information in DG	Identify how many features are selected and displayed; whether what is displayed is static or dynamic; what is the superimposed structure.	
2 - Integrate		1st and 2nd orders Representational mappings (RM)
Identify, compare and extract qualitative information from association patterns in DG	Integrate association patterns such as clustering, and shape, height and width in simple and conjoint distributions, etc.	
Identify, compare and extract qualitative information from relationship patterns in DG	Integrate bivariate patterns such as correlations, regressions, linearity and non linearity patterns in scatterplots, proportional patterns in area charts, etc.	
Identify, compare and extract qualitative information from time dependent patterns in DG	Integrate time dependent linear patterns such as trend shape, orientation, rates of change, etc.	
3 - Compute		2nd order Representational systems (RS)
Extract quantitative information from association patterns in DG	Do computations on association patterns such as clustering, and shape, height and width in simple and conjoint distributions, etc.	
Extract quantitative information from relationship patterns in DG	Do computations on bivariate patterns such as correlations, regressions, linearity and non linearity patterns in scatterplots, proportional patterns in area charts, etc.	
Extract quantitative information from time dependent patterns in DG	Do computations on time dependent linear patterns such as trend shape, orientation, rates of change, etc.	
4 - Reason		
Make interpretations by integrating information in DG	Reason by integrating information and knowledge of visual patterns within or between DG	2nd and 3rd orders
Draw logical conclusions by integrating interpretations in DG	Reason by integrating information, knowledge and interpretations of visual patterns within or between DG	Abstract mappings (AM)
5 - Infer		3rd order Abstract systems (AS)
Make predictions by integrating conclusions drawn from DG	Making predictions by integrating conclusions of visual elements within or between DG	
6 - Explain		3rd order (meta) - Principles (P)
Describe the given answers, intent and technical aspects of DG	Describe technical, graphical and rhetorical elements in DG	
Comment rhetorical failures and graphical limitations in DG	Consider limitations, cognitive biases, distortions, wrong or misleading use of graphical and rhetorical elements in DG	
Propose translations between wrong and right DG	Comparing right and wrong element of DG	

from the data graphics (by referring again to the above example, making predictions about how the above disproportion will most probably evolve over time) [13];

6) EXPLAIN: principles (P) or judgmental level, which is supposed to conceptualize the capacity to explain how visual information has been organized and exploited, and the critical knowledge of explaining how the elaborated visual information can be improved (e.g., arguing that a pie chart is not the most appropriate chart to represent trends over time) [10], [41].

A total of six levels were deemed to fit into the domain of visual information concepts and skills, resulting in the conceptual model depicted in Table 1. In the model proposed in this study, two levels were collapsed into one (SR and SA levels), and the developmental levels were kept conceptually separated from the structure of the 1st, 2nd and 3rd "hierarchical orders" discussed below. The six levels of the Visual Information Literacy model are described in the "Level" column. This label is in line with the requirement of the theory, i.e., that of defining a progression of cognitive abilities that represent progressing skills with visual information. An example of cognitive abilities that each level requires attainment of is described in the "Examples" column. The separation of the six levels from their internal structures (i.e., the hierarchical order of the column "Progress and order structure" was functional to the semiotic lenses described below: that of representational (1st and 2nd orders, that of syntax and semantics), abstract (1st, 2nd and 3rd orders, that of syntax, semantics and pragmatics, respectively), and systemic (2nd and 3rd orders, that of semantics and pragmatics), which are recursively applicable to the design of items.

The choice of collapsing the SR and AR levels came from a consideration about the nature of diagrammatic representations: as stated in [42, p.6], "diagrams are constituted always to some degree by a mode of representation that is also an *immediate* non-linguistic presentation [...] meaning not necessarily the simple identity of sign and entity, but their at least minimal overlap and ontological continuity [...] In this way the *content* of a diagram is already at least partly present directly and immediately in its *form*. Its syntax is already an instance of its semantics [...]". This strict correspondence of the single representation (the visual form) with its meaning (the single abstraction) was the reason for the choice to collapse the two levels in visual language paradigms. Also the subsequent levels should be read by virtue of the nature of the visual information language we are considering in the model, through the theoretical lenses of visual language semiotics and of its three layers: syntax, semantics and pragmatics (see [43], [44], and the foundational study on the semiology of data graphics [20]).

In particular, as said above, the first level of the model, "Map"(resp. "Single Representations" and "Single Abstractions" in Dawson-Tunik's framework), contains syntactic aspects, those of individual visual signs and their identification and recognition. Likewise, the second level, "Integrate" (resp. "Representational Mappings"), is where individual visual signs are mapped into coherent visual relations; at this level also semantic aspects are involved, those of recognition and interpretation of "iconic" structures. The third level, "Compute" (resp. "Representational Systems"), includes the semantic mappings of two systems, in the sense that the system of meaning of visual signs is transferred into the system of meaning of quantitative signs, for further manipulation. These three levels may also be considered at a concrete stage of signs manipulation and interpretation. The next three levels may be considered at the abstract stage of manipulation and interpretation of signs. The fourth level, "Reason" (resp. "Abstract Mapping"), contains both semantic and pragmatic aspects, as far as the meaning conveyed by the visual signs and their further manipulation is properly consolidated into formal interpretations about the entity represented, and the abstractness of formal reasoning converges into the concreteness and practical understanding of the entity represented. The fifth level, "Infer" (resp. "Abstract Systems"), includes pragmatics aspects, in terms of the systematic use of interpretations and conclusions about the entity in order to predict or guess new knowledge from it. The last level, "explain" (resp. "Principles"), is related to the ability to give account of the practical reasons why a visual information is used and even of how it is used for communication and action, hence it fully expresses a meta-layer of visual signs comprehension, which is still related to pragmatics.

TABLE 2. Codification of the construct model of Table 1 into a Visual Information Literacy construct with Data graphics type, Developmental Levels, and the three order structures that can be activated while interacting with data graphics.

Visualization		Levels		Orders	
id group	id kind	id	descr	id	descr
v:	d: dot-plots	m	map	r	representat.
very	i: icon-graphs	it	integrate	a	abstract
simple	di: dot-icon c.	c	compute	s	systemic
	dba: dot bar c.	rr	reason		
s:	ba: bar charts	if	infer		
simple	bu: bubble-charts	e	explain		
	st: treemaps				
	pa:parallel-line				
c:	li: line-charts				
complex	bali: double-chart				
	g: maps				
	a: area-charts				

The substantive model described in Section III-A became a Visual Information Literacy blueprint for designing items, which were aligned with data graphics comprehension tasks along the six theoretical levels of development listed above and in Table 1. As said above, this blueprint was the generative structure for items design, as a combination of skills and concepts derived from the initial database of educational goals selection (see the next Section III-B), the theoretical model and the attempts reported in the literature to measure Visual Information Literacy. Table 2 reports the coding of these elements that, taken in combination with each other, generated a set of items for assessing the construct validity at each level of the developmental construct.

The salient steps of items preparation are depicted in the workflow of Figure 1. For the items preparation, we followed "Principled Assessment Design", which encompasses several methods of assessment design [45], none of which requires automatization of procedures. We strictly followed the recent recommendations of Kosh [46], where she admits the "lack of content-specific processes for implementing automatic items generation". We then performed some manual selections with the aid of experts in the educational domain and items bank design, in order to reach the highest accuracy possible.

B. ITEMS PREPARATION

A preliminary screening activity of the Math and Skills Database at MetaMetrics[12] was carried out, in order to extract the Quantile Skills and Concepts (QSCs) graph,[13] a model of relations of math skills with its prerequisite, supporting, and impending (subsequently reachable) skills, together with information for each QSC such as its code, and a list of associated Goal Descriptions (GDs) based on the "Common

[12]https://hub.lexile.com/math-skills-database.

[13]https://www.quantiles.com/wp-content/uploads/2019/10/11 × 17-Quantile-Map.pdf.

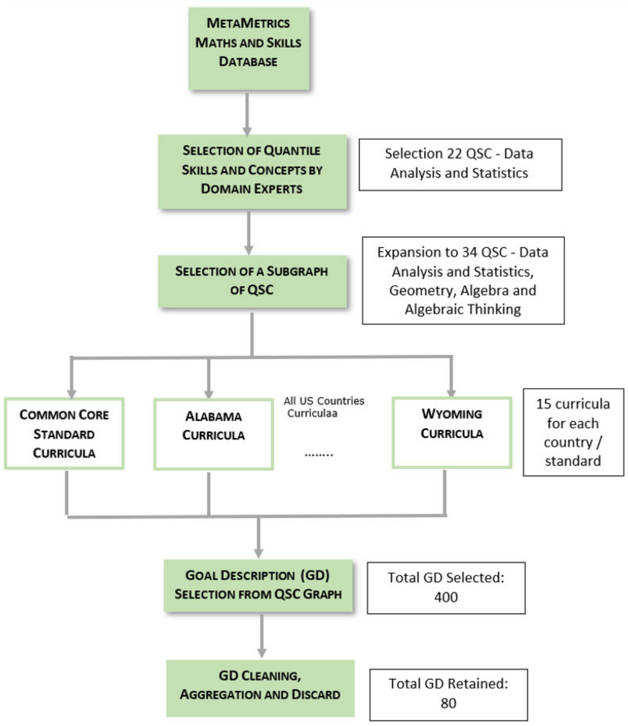


FIGURE 1. Workflow for the GDs selection.

Core State Standards'',[14] a US standard for educational tests from Kindergarten to Grade 8 students, plus a preparatory mathematics curriculum for college students. This standard refers to ''high-quality academic standards in mathematics and English language arts/literacy (ELA)'' devised to ''outline what a student should know and be able to do at the end of each grade''. Stemming from the Math Skills Database and from a bunch of QSCs provided by MetaMetrics educational domain and MetaMetrics items bank experts and designers, we selected by hand a sub-graph of thirty-four QSCs, those related to the strand of ''Data Analysis, Statistics, and Probability'', as well as partly related to the ''Algebra and Algebraic Thinking'' and ''Geometry'' strands, which contain all of the visual information comprehension QSCs and GDs, as a part of the more general Quantile assessment system.[15] An excerpt of the resulting QSCs is available in Table 3, where we also reported the ''Quantile score'', a measure of the level attained for each educational goal.[16]

Starting from this sub-graph, we manually extracted the GDs for each QSC; in order to obtain the most complete, accurate and wide list of GDs, we checked both the ''Common Core Standard'' curricula for each educational

[14]http://www.corestandards.org/.

[15]https://www.quantiles.com/. We asked the same experts to check our manual selection, in order to certify the accuracy of manual selection of GDs. The Math and Skills Database is not freely available, but a trial version of it was available for 30 days, in the period October 2019.

[16]As the computation of this score is not freely available, we did not consider it in our model design; we only considered it in the definition of criteria for data graphics difficulty (see the next paragraphs of this section, where we address the systematization of criteria for classification of data graphics difficulty).

TABLE 3. An excerpt of the resulting QSCs for visual skills for the Strand ''Data Analysis, Statistics, and Probability''.

QSC ID	Quantile score	GD
QSC 20	-110Q	Organize display and interpret information in concrete or picture graphs.
QSC 5	970Q	Answer comparative and quantitative questions about charts and graphs.
QSC 182	490Q	Select the appropriate graph that best displays the given data.

TABLE 4. Examples of Visual Information Literacy GDs.

Visual Information GDs
Collect, record, interpret, represent, and describe data in a graph.
Interprets graphs by applying correct understanding of whole numbers.
Compares two proportional relationships using graphs.
Makes qualitative statements about the rate of change of a function from graphs.
Draw conclusions and make predictions from information in a graph.
Makes correct observations from the graph about the shape of the data.
Recognizes distorsions in the way data are displayed.

grade and the complete list of other standard curricula available for each State. We repeated this operation for all of the US States. We extracted all of the GDs present in each and every one of the curricula examined.. In this way, we gathered almost 400 unique GDs (after deleting by hand the perfect redundancies, where identified). In a second step, very similar GDs were aggregated, and very similar GDs at different educational grades were collapsed by hand to one ''visual'' GD alone. GDs that were definitely not related to ''visual concepts'' were removed. For both the aggregation and final selection steps, two authors ran both tasks separately, and compared their outcomes until a Cohen's kappa substantial agreement of 0.65 was reached [47], [48]. At the end, a final list of GDs amounted to around 80 statements about visual information comprehension abilities. An excerpt of the outcome of this preliminary phase of work is reported in Table 4.

Items were designed stemming from the GDs and a screening of data graphics, partly designed from scratch and partly available at the Infographics Portal ''Beautiful News''.[17] Items were prepared as yes/no, multi-choice and free text questions.[18] Regarding data graphics, they were selected and classified as ''very simple'', ''simple'' and ''complex'' (see the construct codified for items preparation in Table 2),

[17]https://informationisbeautiful.net/beautifulnews/.

[18]Free text items were essentially of the kind ''Write a short paragraph of critical review about this information visualization''. Critical reviews may regard the grammatical elements of the data graphics (e.g., the choice of colors,the font size, the data graphics fit for the data, and the like) or the rethorical management of the story that the data graphics tell (e.g., the emphasis on a part of the data to support an argument by choosing a data graphics rather than a still valid alternative). The responses should be argumented enough to constitute correct critics, and they were coded by summing up correct explanations for each of the aspects examined in the response and dividing them up for all the correct explanations about a data graphics (domain experts in the field of data graphics helped us identify a list of possible critics for each of the data graphics to which we add the free text question). Coding of the free text questions was done according to three possible scores: wrong explanation (0), partial explanation (1), complete explanation (2), which were further dicotomized.

TABLE 5. List of criteria used to classify data graphics into very simple (1), simple (2), and complex (3).

	Criteria			
Sample	Aggreg.	Syntax	Percept.	Concept
(1) Small discrete ($\leq$ 100 obs)	None	Dots	Pointwise	Unidim. details
(2) Big discrete ($\geq$ 100 obs)	Yes	Bars & derivatives (e.g., stepped lines) Slopes Pie Treemaps	Block-wise	Unidim. & multidim. aggregation (proportion)
(3) Continuous	Yes	Lines Areas Geo maps	Line slope Color density Continuous change	Uni-multidim. continuum & aggregation Rate of change

according to a combination of the following criteria (gathered from the Quantile Knowledge Base,[19] from [49], and from [20]):

1) sample size and kind (e.g., small samples of discrete data were supposed to be easier to manage than big samples of continuous data);
2) level of data aggregation (e.g., data graphics where each sample element is represented are easier than data graphics where only aggregated points are represented; data graphics where proportional aggregation refers to discrete elements are easier than data graphics where proportional aggregation refers to continuous elements);
3) data graphics syntax (e.g., data graphics with data codified in dots are easier to process than bars in bar charts, which are easier to process than geographical maps);
4) perceptual rules (e.g., bar charts are perceived as allowing a more precise hence more immediate comparisons with respect to color density);
5) information kind (e.g., pointwise information is easier to understand than the abstract concept of continuous change that the slope of a line conveys).

In Table 5hl, classification criteria are listed and schematized. The number of data graphics selected from the portal was 19, for a total of 22 data graphics and 90 items exploited. Figures 2 and 3 report an absolute frequency distribution of items associated with construct levels, data graphics kind and order structure (see Table 2 for an overview), with items number and responses collected, respectively.

Exemplary questionnaire items related to a data graphics are reported in Table 6.

IV. ASSESSMENT RESULTS

Responses related to the items associated with data graphics were collected by administering a survey through the Limesurvey web platform[20] in the period May 2020 - October 2020. A total of 3,001 responses were collected, ranging from yes/no questions to free text questions of the kind exemplified in Table 6. Each response was coded as binary (right or wrong): for yes/no questions the coding was straightforward; for multi-choice items the right choice was rated 1 and the wrong ones were rated 0; for free text responses, after a careful examination by at least three raters, right answers were coded with 1 and wrong answers with 0 (an inter-rater agreement was computed by adjusting single item controversies until a final agreement of $\alpha = 80\%$ was reached [50]). For all the statistical tests, we applied standard procedures of statistical hypothesis testing by adopting a confidence level of .95 and a significance level of .05. Data were analyzed for statistical significance of the following null hypotheses:

1) $H0_1$: there is no difference in respondents' performance on the same level with data graphics of different kind (very simple, simple and complex);

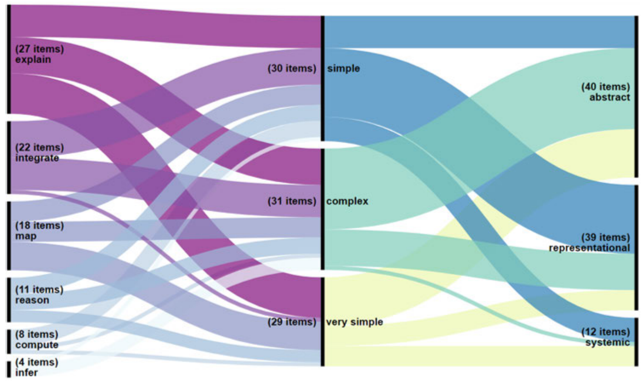


FIGURE 2. Alluvia chart of the distribution of items (absolute frequencies) for the different construct levels (on the left), kind of data graphics (in the middle), and structure order (on the right). All the elements are put in size order from top to bottom, where size is the proportion of items designed for each of them.

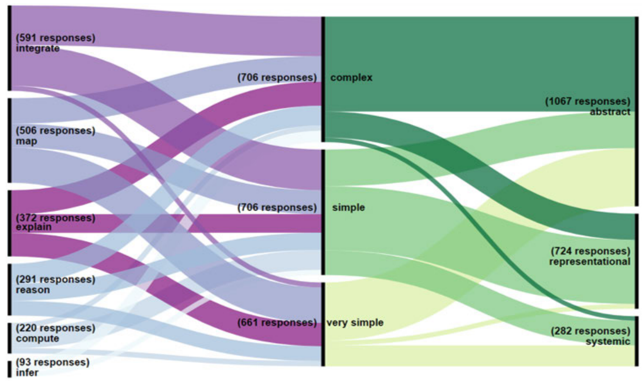


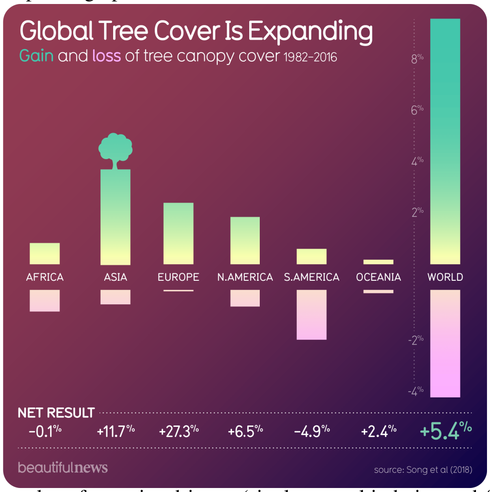
FIGURE 3. Alluvia chart of the distribution of responses (absolute frequencies) related to items for the different construct levels (on the left), kind of data graphics (in the middle), and structure order (on the right). All the elements are put in size order from top to bottom, where size is the proportion of responses obtained for each of them.

[19]Also systematized by us from the quantification in Quantile scores of MetaMetrics, derived from thousands of items administration to students of all the US curricula, from Kindergarden to pre-college Math curricula.

[20]https://www.limesurvey.org/en/.

TABLE 6. An example of questionnaire about a data graphics and its associated items. Data graphics source: https://informationisbeautiful.net/beautifulnews/877-global-tree-cover/.

Example of graph:



Examples of associated items (single or multi-choice and free text):
The bars of the data graphics represent:
- Coverage quantities
- Coverage trends
- Coverage variations

What is the more correct and complete conclusion that can be drawn from the graph?
- No world area increased its tree canopy
- Almost every world area increased its tree canopy
- The tree canopy worldwide balance is positive

What is the continent with the highest tree canopy gain?

____

What is the percentage value corresponding to the highest quantity of tree canopy?
- 11,7%
- 5,4%
- 9,4%

Write a short paragraph of critical review about this information visualization:

____

____

2) $H0_2$: if there is any difference, it is not the case that the performance degrades with the raising of data graphics complexity;
3) $H0_3$: there is no difference in respondents' performance on the same level and different orders;
4) $H0_4$: if there is any difference, it is not the case that this is due to the increasing complexity in the structure order (representational structures are easier to be processed than abstract structures, and the latter are easier than systemic structures);
5) $H0_5$: there is no difference in respondents' performance on the six levels of the Visual Information Literacy construct;
6) $H0_6$: if there is any difference, this is not due to an increasing order of difficulty in cognitive levels reflecting the order of levels of the construct.

Each hypothesis test was designed by considering that the outcome variable (the right or wrong response to each item) was in a functional relationship with variables such as the construct level tested with each item (LEV), the kind of data graphics exploited (CRT), and the kind of order structure (either representational, abstract or systemic) required to give the correct answer to any item (RAS). Assuming that the dependent variable Y is the outcome of the above function for each right or wrong answer (resp. codified as $y_i = 1$ or $y_i = 0$), a binary logistic regression method was chosen to run statistical tests.[21] This method was applied twice: to estimate the expected value of Y from the values of LEV, CRT and RAS in order to predict and assign the most probable value

[21] Data were analyzed in *R*, with *RStudio* 1.3.1093 version, available at https://rstudio.com/products/rstudio/download/. The binary regression algorithm was run on a Windows 10 PC, with Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz, 16GB RAM, with computational times for running the algorithm between 0.60 and 0.68 secs. The function used in *R* was *glm*. The computational complexity of the binary regression algorithm (glm) applied to our data is dependent by the model parameters p and the sample size n, formally resulting in a computational complexity of $O(p^2 * n + p^3)$.

of Y to all the non response items (a total of 928 out of 3,001); to run tests in order to reject the null hypotheses above, with the complete set of responses. Binary logistic regression [51] was formally characterized as follows:

- the expected value of Y is modelled with the equation
$$E[Y|x] = \pi(x); \tag{1}$$
- $\pi(x)$ is the logistic regression function, and it has the form
$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}; \tag{2}$$
- the linearization function of (2), $g(x)$, is the logit transformation of (2) and has the form
$$g(x) = ln[\frac{\pi(x)}{1-\pi(x)}] = \beta_0 + \beta_1 x. \tag{3}$$
The coefficients estimated with (2) are in relation with the odds ratio (OR) given by the model, i.e., how likely is outcome 1 wrt outcome 0;
- given all of the above, an observation y may be expressed with the equation
$$y = E[Y|x] + \epsilon \tag{4}$$
where ϵ is the estimation error, which should be distributed binomially instead of normally, as it would be in linear regression, because of the logistic regression assumed as the current model.

Running binary logistic regression on the RAS variable did not bring any statistically significant coefficient. Hence, null hypotheses $H0_3$ and $H0_4$ could not be discarded; neither age nor gender of respondents bring any statistical significance to the Visual Information Literacy problem.[22] Hence, statistical tests with respondents' strata were not further considered in subsequent analyses.

The binary logistic regression models considered for running tests were the following:

- Model 1:
$$\begin{aligned} E[correct|X_{CRT}] &= \pi(x) = ln[\frac{\pi(x)}{1-\pi(x)}] \\ &= \beta_0 + \beta_1 X_v + \beta_2 X_s + \beta_3 X_c \end{aligned} \tag{5}$$
where v=very simple, s=simple, and c=complex kind of data graphics (see Table 2 for details);
- Model 2:
$$\begin{aligned} E[correct|Z_{LEV}] &= \pi(z) = ln[\frac{\pi(z)}{1-\pi(z)}] \\ &= \beta_0 + \beta_1 Z_m + \beta_2 Z_o + \beta_3 Z_c \\ &\quad + \beta_4 Z_{rr} + \beta_5 Z_{ii} + \beta_6 Z_e \end{aligned} \tag{6}$$
where indices of the variables are the construct levels (m=map, it=integrate, c=compute, rr=reason, if=infer and e=explain), as being modelled in Table 2.

Results are reported according to the APA style [52], and by means of tables and figures, as an aid to better clarify and synthesize our results.

[22] For testing differences among respondents' strata we ran Fisher's test.

A. PREDICTING RESPONSES FOR NON RESPONSE ITEMS

A binary logistic regression was run for each level of the model on each subset of the total 2,073 items with responses, which were supposed to represent the training set of the binary logistic regression method exploited. The predicted probability of the model was assumed to represent the accuracy (resp. the uncertainty) with which the model correctly assigned a value to the missing response. A value of 0 or 1 was then assigned to the test set of non response items, according to the following predicted probability thresholds:

- assign a value of 0 if the predicted probability of correctness was under 70%;
- randomly assign a value that is either 0 or 1 if the predicted probability of correctness was between 70% and 90%;
- assign a value of 1 if the predicted probability of correctness was equal or above 90%.

The test set of 928 non response items was assigned correct response values in 43% of the cases.

The total responses divided by level and kind of data graphics with descriptive statistics is reported in Figure 5.

B. TESTING DATA GRAPHICS DIFFICULTY AT EACH COGNITIVE LEVEL

Results of the binary logistic regression method run on Model 1 equations, one for each level, are reported in Table 7. With the exception of the data graphics of the simple and complex kind at the "reason" level, all levels show statistically significant coefficients. At the "integrate" and "compute" level it is confirmed that varying from very simple to simple data graphics, the odds of giving the correct answer decreases, by 72% and 54%, respectively, all other elements being equal. At the "map", "integrate" and "infer" level the same trend is confirmed when passing from simple to complex data graphics: the odds of giving the correct answer decrease by 51%, 81% and 97%, respectively. An inverse situation arises when passing from very simple to simple data graphics at the "map", "reason" and "explain" level: all of the other elements being equal, the odds on performing more accurately increases instead of decreasing. At the "compute" and "explain" levels, when passing from simple to complex data graphics the odds on performing better also increase, although at a lower rate than when passing from very simple to simple (see Table 7). This outcome may have different explanations: for example, computation task items were concentrated more in simpler data graphics. Surprisingly, passing from very simple data graphics to simple data graphics seems to increase the accuracy of some responses at the easiest and at the more difficult levels, which are "map" and "explain". This anomaly at the extremes be a reason for outer causes to be taken into consideration, such as for example the greater familiarity with data graphics such as bar charts with respect to very simple data graphics such as dot plots. Also a raw counting of frequency of correct responses like the one reported in Figure 4, seems to confirm that simple data graphics were the "easier to work with".

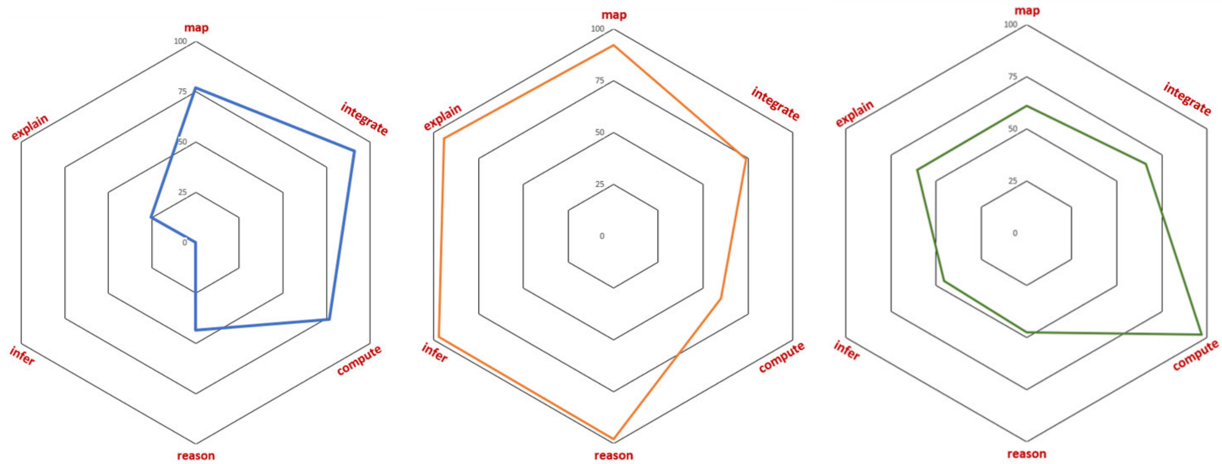


FIGURE 4. Radar charts of the proportion of correct responses for each level of Visual Information Literacy, from left to right: the blue line for "very simple" data graphics; the orange line for "simple" data graphics; the green line for "complex" data graphics.

TABLE 7. Regression results for Model 1, with Models Coefficients and their Significance (Signif. codes: <.001='***', 0.001='**', 0.01='*', 0.05='.', <.1=' '), Fit test (χ^2), Odds ratio (OR) with Confidence intervals (CI), and OR expressed as ratios (of how many times the chance of succeeding with the current data graphics type increases (positive ratio) or decreases (negative ratio) with respect to the previous (an easier) data graphics type.

Levels	Model	Coefs	Fit test	OR	CI 2.5%	CI 97.5%	Ratio
map	(Interc.)	1.21***	χ^2(2,599)	3.35	2.54	4.49	2.35
	type s	1.26***	49.59	3.51	1.92	6.88	2.51
	type c	-0.77***	p<.001	0.46	0.30	0.71	-0.54
integrate	(Interc.)	2.31***	χ^2(2,727)	10.07	6.13	17.86	9.07
	type s	-1.27***	38.32	0.28	0.15	0.49	-0.72
	type c	-1.64***	p<.001	0.19	0.10	0.34	-0.81
compute	(Interc.)	1.18***	χ^2(2,264)	3.25	2.09	5.24	2.25
	type s	-0.78**	24.46	0.46	0.26	0.81	-0.54
	type c	2.29*	P<.001	9.85	1.95	179.70	8.85
reason	(Interc.)	-0.26	χ^2(2,366)	0.77	0.52	1.14	-0.23
	type s	4.03***	126.26	56.14	19.48	238.24	55.14
	type c	0.16	p<.001	1.17	0.70	1.97	0.17
infer	(Interc.)	3.47***	χ^2(1,130)	32.00	6.89	569.40	31.00
	type c	-3.65***	33.92,p<.001	0.03	0.00	0.13	-0.97
explain	(Interc.)	-1.07***	χ^2(2,898)=	0.34	0.27	0.44	-0.66
	type s	3.88***	330	48.64	28.20	89.81	47.64
	type c	1.49***	p<.001	4.46	3.19	6.27	3.46

A ranking of data graphics in order of percentages of correct responses by each is presented in Figure 6, where simple data graphics are top ranked with respect to complex data graphics, and, in a counter intuitive way, very simple data graphics are concentrated on a lower percentage range of correct responses with respect to complex data graphics and simple data graphics. Overall, the results show that the null hypothesis $H0_1$ can be rejected, whereas the null hypothesis $H0_2$ cannot be rejected.

C. TESTING CONSTRUCT LEVELS DIFFICULTY

Results for binary logistic regression on Model 2, for the six levels of the Visual Information Literacy construct, are reported in Table 8. With the exception of the "integrate" level, all the other five levels have significant coefficient (all with p<.001). Hence, the null hypothesis $H0_5$ can be rejected. Furthermore, the resulting odds seem to show a degradation of performance (see the Perc% column of Table 8) when passing from an easier level to a more difficult level. Hence, $H0_6$ null hypotheses can be also rejected. Indeed, it may result as an empirical evidence that Visual Information Literacy has a developmental cognitive structure that is outlined by the model proposed, in the same order as the model proposed: mapping syntax and semantics of data graphics to entities as a first step; integrating data graphics to interpret aspects of such phenomena; making computations to grasp proportional and quantitative patterns; reasoning on qualitative patterns such as statistics, relations, associations, correlations, distributions, and the like; inferring new knowledge based on the knowledge, observation and interpretation acquired in the preceding levels of data graphics comprehension; explaining the aspects of comprehension achieved and suggesting improvements in visualizations.

V. DISCUSSION

A. ADVANCEMENTS IN THE THEORY OF VISUAL INFORMATION LITERACY

This study makes a contribution in raising concern about the lack in both the educational and data graphics literature of a unique definition and of a model for Visual Information

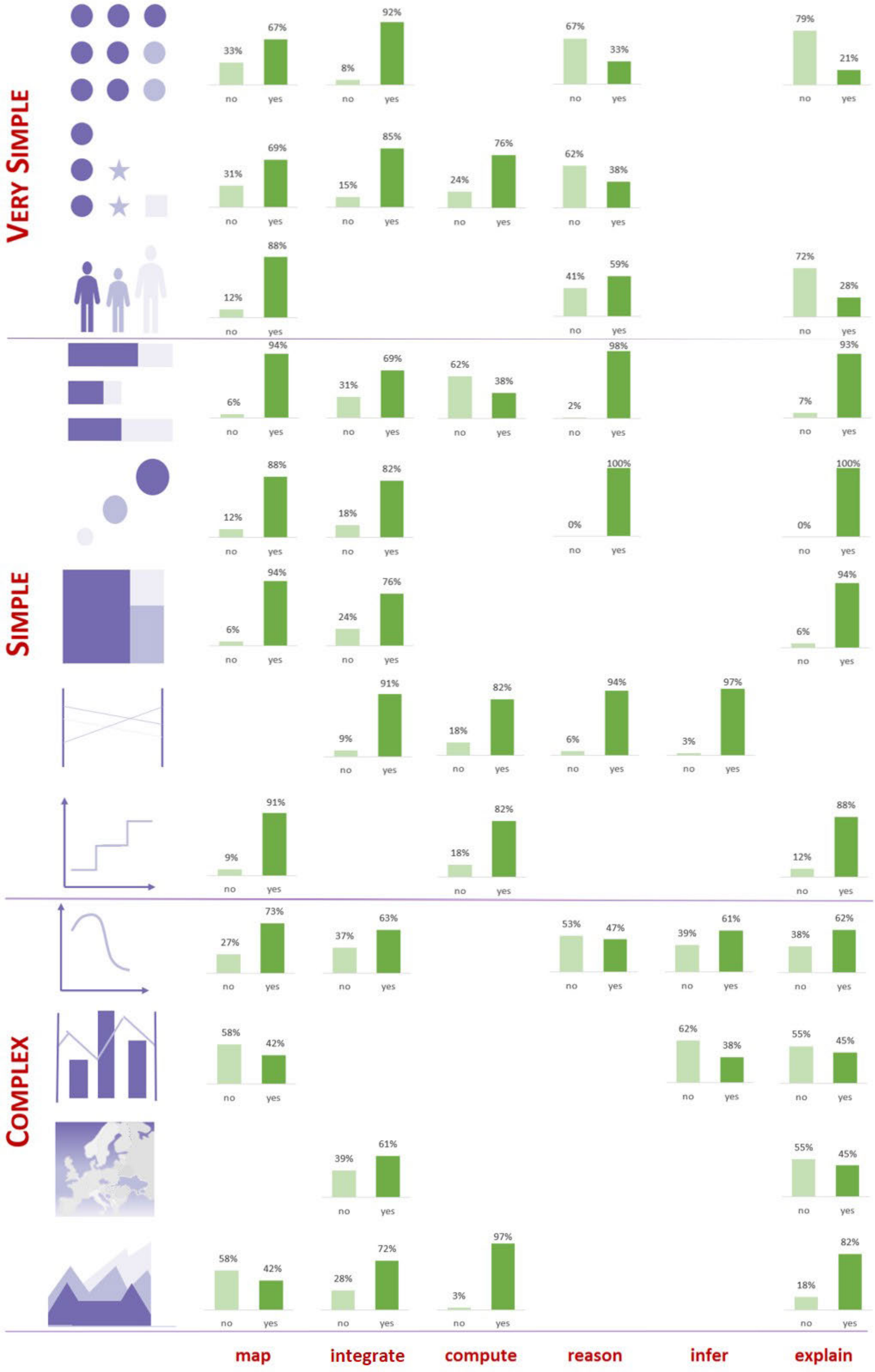


FIGURE 5. Proportion of responses for each kind of data graphics and level. From top to bottom: dot plot, bar plot, icon array, bar chart, bubble chart, treemap, slope chart, stepped line chart, line chart, double axis chart, choropleth map, and area chart.

TABLE 8. Regression results for Model 2, with Models Coefficients and their Significance (Signif. codes: <.001='*', 0.001='**', 0.01='*', 0.05='.', <.1=' '), Fit test (χ^2), Odds ratio (OR) with Confidence intervals (CI), and OR expressed as Percentages.**

Model	Coefs	Fit test	OR	CI 2.5%	CI 97.5%	Perc%
(Interc.)	1.19***	χ^2(5;2,995)	3.30	2.74	4.00	230
level it	-0.10	91.14	0.91	0.70	1.17	-9
level c	-0.31 .	p<.001	0.73	0.53	1.02	-27
level r	-0.60***		0.55	0.41	0.73	-45
level if	-0.86***		0.42	0.29	0.63	-58
level e	-0.89***		0.41	0.32	0.52	-59

Literacy, which could fit data graphics characteristics and provide the foundation for the design of a standard measurement scale for data graphics difficulty on one hand, and for literacy progress of individuals in managing data graphics on the other hand. A theoretical model was proposed within a meta perspective strand of research (discussed in Section III-A), in the direction of yielding generalizability of hypotheses and results, after noting the limits, in this regard, latest results in state-of-the-art literature in the fields of data graphics design (see, again, the latest results yielded in [7] and [14] for Visual Information Literacy assessment). Furthermore, a contemporary empirical scrutiny of existing standardized test materials for Math skills (described in Section III-B) was also properly carried out, to find standardized educational goal descriptions that, although not supported by an explicit cognitive model, could constitute the core of an initial set of visual items that we re-designed with the support of a conceptual construct (our model) as well as domain experience (that of MetaMetrics domain experts in education and item bank design and maintenance). Our major result in the assessment phase was the validation of the global ordering of levels of our Visual Information Literacy model. This is an advancement with respect to the study of [14], where the order of levels of their model could not be confirmed, although they applied the same analysis method as ours. Our choice of analyzing items response with a binary logistic regression method had the intent to quantitatively measure the performance accuracy of individuals at each level of the model and for degrees of data graphics difficulty, in order to validate the Visual Information Literacy construct proposed in this study. Only two null hypotheses could not be rejected: those of the internal structure of data graphics difficulty for each level. For a couple of levels, the order of difficulty of data graphics was not the one expected, although for four levels out of six the order was the one expected. This blurred evidence may be due more to the lack of further data points to feed the statistical model than to wrong assumptions about the order of developmental levels or the difficulty of data graphics. Related to this, the surprising performance on simple data graphics compared to very simple data graphics for some levels of the construct may shed light on the need to better calibrate further dimensions of data graphics difficulty than the ones explored in the literature so far, for example familiarity [53] or the cognitive load of digging into details (see for example [24] in this respect).

B. IMPLICATIONS FOR VISUAL INFORMATION LITERACY ITEMS DESIGN

As already noticed in [54] and [55], information quality and design quality dimensions may impact on people's interaction and performance with visual information. In particular, dimensions such as familiarity with some kind of data graphics entered the design discourse on a par with quantitative dimensions such as redundancy or multi-dimensions (see for example [53] for a whole detailed view of the design dimensions of infographics). The contribution of this study reveals that the more data graphics become complex from the difficulty criteria perspective[23] [20], [49] (e.g., continuous data samples represented along a line vs discrete data samples represented in bar charts; patterns matching such as the rate of change of a line chart vs proportions of a bar chart) the poorer the human performance becomes through dealing with them. Although in our experiment this result was clearer when passing from simple to complex data graphics than from very simple to simple data graphics, this may be considered as a valid result of our assessment phase, and a challenging topic for further assessments.

C. ANSWERS TO THE RESEARCH QUESTIONS

With respect to the initial research questions of Section I, also derived from the lack of answers in the literature so far, which we have discussed by taking into examination the main works on Visual Information Literacy assessment in Section II, we asked first what makes a data graphics easy or difficult. As already outlined in the previous Section, we answered this question by systematizing well known criteria for classifying data graphics difficulty such as the kind of data, the kind of patterns on data and the kind of information represented through data and patterns, with the levels of our Visual Information Literacy model (the model has been detailed in Table 1). The second question was related to what makes a graph-supported task easy or difficult. Our contribution in this direction was the characterization of a property, Visual Information Literacy, through a model of progressive difficulty that may explain the increasing effort needed to execute tasks depending on the data graphics difficulty and the literacy level attained by individuals (see the proposed construct depicted in Table 2 and Figure 7). The third question we asked was related to how many and what levels of literacy with data graphics could characterize Visual Information Literacy. The provision of a theoretical model of Visual Information Literacy with its assessment was provided as an answer. The last question we asked was about the relations to the homomorphism of this model with other kind of literacies. The relevance of levels for Visual Information Literacy may be deemed homomorphic to that of Dawson-Tunik's developmental progression framework, which was validated for cognitive development and moral achievements. Our contribution is to have validated it also for cognitive development in dealing with data graphics (see section IV for the detailed report of results).

[23] See Table 5 of criteria in Section III-B.

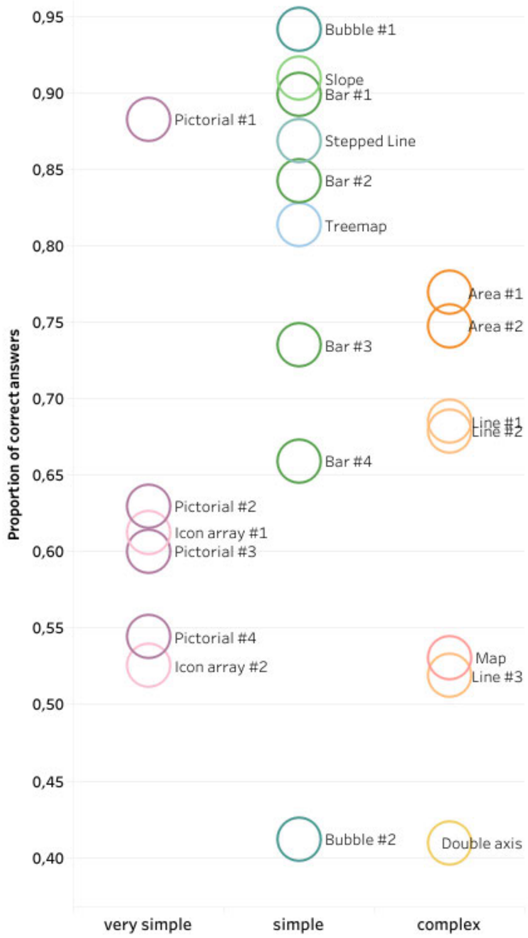


FIGURE 6. Proportion of correct responses for each data graphics exploited in our experiment. From left to right, each kind of data graphics is aligned by column. The baseline was changed to 0,40, which is the minimum proportion of correct responses, for better readability of the data graphics (and not to intentionally violate design guidelines we are well aware of [19]).

VI. LIMITATIONS AND CONCLUSION

The model may have potentials for development into a fully fledged and mature measurement tool like the PISA and the *Quantile* tests mentioned in Section I, towards a Visual Information Literacy standard measurement scale. To achieve this goal, the survey should be extended either vertically, by means of items design to be added to the items repository, or horizontally, by administering items to many targets and worldwide, in order to calibrate items difficulty and use this calibration to measure individuals' ability into a generalized measurement scale. The first goal should consider the complex activity of designing items for the purpose of maintaining the compliance of tasks across different levels and data graphics kind (e.g., the manual computational precision on a continuous sample scale vs a discrete sample scale). This should then be agreed upon by a group of expert educational designers and raters, whose agreement should be assessed by a inter-rater agreement threshold. In addition, the statistical treatment of the wider results should be oriented to measure items difficulty and individuals' ability along a common measurement scale, for example by adopting statistical techniques such as for example Item Response Theory or the Rasch method.[24] These provide a common scale for measuring both items difficulty and individuals' ability and are used to test the same person on items of progressive difficulty to determine his/her ability threshold. These experimental frameworks have already been adopted in reading comprehension frameworks [56] and partly in Visual Information Literacy assessment [7]. The mirror side of reading Visual Information is writing it: the ability to design, propose or draw the right data graphics against a piece of non visual information, to test whether writing data graphics capacity has the same developmental progression as reading data graphics. Proposing items to assess individuals on the task of drawing the correct graph, is something that has not yet received sufficient attention from scholars of any of the fields mentioned in this study, nor could it find a space for development in our study. Nonetheless, this aspect should be considered for future development, being part and parcel of the conceptualization of Visual Information Literacy.

That said, this study has contributed to make an overview of a very current and urgent topic of interest for all the stakeholders interested in the future of visual information processing and management, and to argue for a multidisciplinary approach towards the development of a standard measurement scale for Visual Information Literacy, taking into account educational, information processing and management, and data visualization theories and practical techniques of design. A definition of Visual Information Literacy was provided, together with an overview of current models and assessment systems. A cognitive model based on a theoretical framework of human cognitive and moral developmental progress, grounded in semiotics, was presented, and those aspects were highlighted in their originality against existing approaches that do not take any of them into account. Furthermore, evidence that this developmental theory is applicable to Visual Information Literacy emerged. In this sense, Visual Information Literacy has been shown to follow a model of cognitive developmental process for data graphics comprehension, where developmental levels follow a progression order that is the same as that of the other kind of literacies explored in the literature. Results of a survey based on administering visual items and with the intent to obtain precise responses to the research questions translated into hypotheses were presented and discussed. Further hints about the need to extend the investigation to data graphics writing, and about the potential resonance of this study on the scholarly debates and on the design of applications were also offered in this study.

APPENDIX A

A literature overview of the terms and definitions regarding different notion of Visual Literacy.

[24]https://www.rasch.org/.

TABLE 9. An overview of definitions, conceptual frameworks and assessment for visual literacy.

Expression	Definition	Explicit cognitive model	Assessment method
Visual literacy (Fransecky [17])	"A group of vision-competencies a human being can develop by seeing and at the same time having and integrating other sensory experiences. [...] a visual literate person [is able to] discriminate and interpret the visible actions, objects, and symbols natural or man-made, that he encounters in his environment. Through the creative use of these competencies, he is able to communicate with others."	none	none
Visual literacy (Ausburn [58])	"A group of skills which enable an individual to understand and use visuals for intentionally communicating with others"	none	F-sort and Q-sort tests: categorizing (art) images and identifying them [59]
Visual literacy Index (Avgerinou [13])	"In the context of human, intentional visual communication, visual literacy refers to a group of largely acquired abilities, i.e., the abilities to understand (read) and to use (write) images, as well as to think and learn in terms of images."	Debes' hierarchy for audiovisual literacy [60]	Attitude towards visual communication (ATVC) questionnaire
Graph Literacy (Galesic [61])	"The ability to understand graphically presented information"	Graph comprehension skills three levels model [25]	Graph Literacy Scale for medical information
Visualization literacy (Boy [7])	"The ability to use well-established data visualizations (e.g., line graphs) to handle information in an effective, efficient, and confident manner."	adapted from Pinker's Graph Comprehension [27] and [28]	Visualization Literacy test with 6-tasks, 48-items and 4 kinds of charts
Data visualization literacy (Borner [11])	"The ability to make meaning from and interpret patterns, trends, and correlations in visual representations of data"	Legacy Data Visualization Literacy Framework (DVl-FW)	Legacy 8-levels Data Visualization Literacy tests
Visualization literacy (Lee [12])	"The ability and skill to read and interpret visually represented data in and to extract information from data visualizations."	Graph comprehension skills three levels model [25]	Visual Literacy Assessment Test (VLAT) with 12 data visualizations and 61 items
Visual literacy competency (ACRL [16])	"A set of abilities that enables an individual to effectively find, interpret, evaluate, use, and create images and visual media. Visual literacy skills equip a learner to understand and analyze the contextual, cultural, ethical, aesthetic, intellectual, and technical components involved in the production and use of visual materials. A visually literate individual is both a critical consumer of visual media and a competent contributor to a body of shared knowledge and culture."	N/A	N/A
Data visualization understanding (Burns [14])	"[The ability to] locate and report specific pieces of (visual) information"	Bloom's taxonomy of educational objectives [26]	Visualization understanding test with three charts and six questions for six levels of understanding

APPENDIX B

Formalization of the Visual Information Literacy construct.

In the Figure 7, each level has a functor between objects of a category and objects of a subsequent category [61]. Let $L1$ and $L2$ be such categories, hence:

- let an object of category L_1 be C_1 and an object of category L_2 be C_2
- a functor $F\colon L_1 \to L_2$ is a map between each L_1 object $C1$ and L_2 object $F(C_1) = C_2$ and each arrow of category L_1, $f\colon C_1i \to C_1j$ to an L_2 arrow $F(f)\colon F(C_1i) \to F(C_1j)$

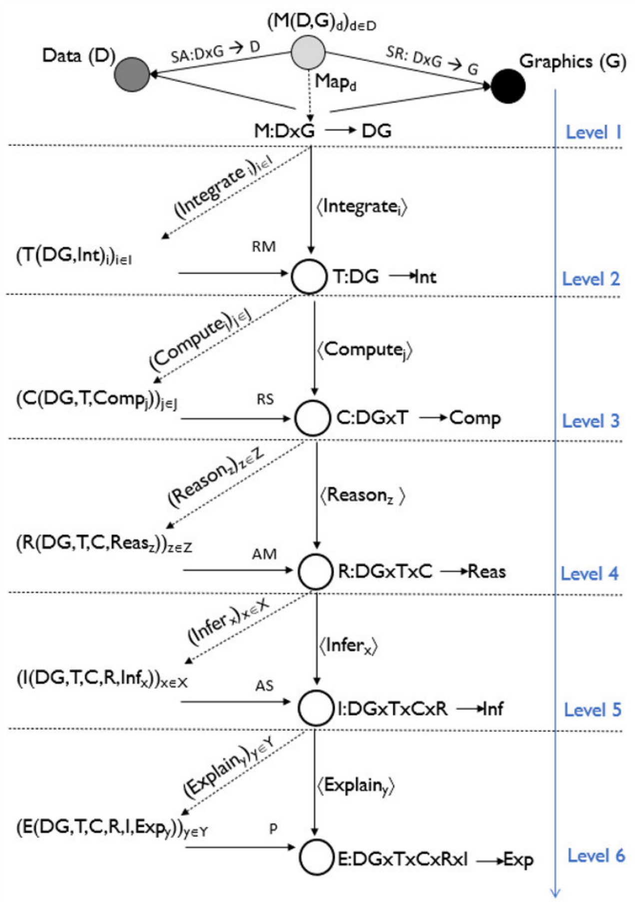


FIGURE 7. The Visual Information Literacy construct model in formal notation. Each Level is denoted by a family of functions (skills) whose domain is formed by objects (concepts) of the previous levels and whose codomain is formed by objects of the current level. The model can be interpreted as a sequence of nested and developmental levels (see the arrows and their direction), in the sense that the skills (arrows) necessary to reach each subsequent level should follow from the correct and complete comprehension and application of concepts and skills of the previous ones.

- $F(id_{C1i}) = id_{F(C1i)}$
- given two arrows $f: C_1i \to C_1j$ of L_1 and $g: C_2x \to C_2y$ of L_2, a composition arrow has the property that $F(g \circ f) = F(g) \circ F(f)$.

For example, Level 2 of Figure 7 is a category whose objects are Int_i and whose arrows are $i: DG_i \to Int_i$. Level 3 of Figure 7 is a category whose objects are $Comp_j$, and whose arrows are $f: DG \times Int_i \to Comp_j$. A functor F is the family of functions from objects $DG \times Int_i$ of Level 2 to objects $DG \times Int \ \times Comp_j$ of Level 3, by $F(i \circ f) : F(i) \to F(f)$.

REFERENCES

[1] C. Taylor, "New kinds of literacy, and the world of visual information," *Literacy*, vol. 13, no. 11, pp. 1–22, Sep. 2013.

[2] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. New York, NY, USA: Basic Books, 2018.

[3] S. Paavola and K. Hakkarainen, "The knowledge creation metaphor–An emergent epistemological approach to learning," *Sci. Educ.*, vol. 14, no. 6, pp. 535–557, 2005.

[4] M. Castells, *The Information Age*, vol. 98. New York, NY, USA: Oxford Blackwell Publishers, 1996.

[5] L. Floridi, *The Ethics of Information*. Oxford, U.K.: Oxford Univ. Press, 2013.

[6] A. Frigerio, A. Giordani, and L. Mari, "On representing information: A characterization of the analog/digital distinction," *Dialéctica*, vol. 67, no. 4, pp. 455–483, 2013.

[7] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete, "A principled way of assessing visualization literacy," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 1963–1972, Dec. 2014.

[8] R. Lowe, "Visual literacy and learning in science. ERIC digest," ERIC Clearinghouse Sci., Math., Environ. Educ., Columbus, OH, USA, Tech. Rep. ED463945, 2000.

[9] P. Beynon-Davies and R. Lederman, "Making sense of visual management through affordance theory," *Prod. Planning Control*, vol. 28, no. 2, pp. 142–157, Jan. 2017.

[10] W.-M. Roth, L. Pozzer-Ardenghi, and J. Y. Han, *Critical Graphicacy: Understanding Visual Representation Practices in School Science*, vol. 26. Dordrecht, The Netherlands: Springer, 2005.

[11] K. Börner, A. Maltese, R. N. Balliet, and J. Heimlich, "Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors," *Inf. Vis.*, vol. 15, no. 3, pp. 198–213, Jul. 2016.

[12] S. Lee, S.-H. Kim, and B. C. Kwon, "VLAT: Development of a visualization literacy assessment test," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 551–560, Jan. 2017.

[13] M. D. Avgerinou, "Towards a visual literacy index," *J. Vis. Literacy*, vol. 27, no. 1, pp. 29–46, Jan. 2007.

[14] A. Burns, C. Xiong, S. Franconeri, A. Cairo, and N. Mahyar, "How to evaluate data visualizations across different levels of understanding," 2020, *arXiv:2009.01747*. [Online]. Available: http://arxiv.org/abs/2009.01747

[15] Y. Okan, M. Galesic, and R. Garcia-Retamero, "How people with low and high graph literacy process health graphs: Evidence from eye-tracking: Graph literacy and health graph processing," *J. Behav. Decis. Making*, vol. 29, nos. 2–3, pp. 271–294, Apr. 2016.

[16] A. of College and Research Libraries (ACRL). (2018). *Visual Literacy Competency Standards for Higher Education*. [Online]. Available: https://www.ala.org/acrl/standards/visualliteracy

[17] R. B. Fransecky and J. L. Debes, "Visual literacy: A way to learn–A way to teach," Assoc. Educ. Commun. Technol., Washington, DC, USA, Tech. Rep. ED064884, 1972.

[18] V. J. Bristor and S. V. Drake, "Linking the language arts and content areas through visual technology," *J. (Technol. Horizons Educ.)*, vol. 22, no. 2, p. 74, 1994.

[19] A. Cairo, *Charts Lie: Getting Smarter About Visual Information*. New York, NY, USA: W.W. Norton & Company, 2019.

[20] J. Bertin, *Semiology of Graphics (WJ Berg, Trans.)*. Madison, WI, USA: Univ. Wisconsin Press, 1983.

[21] L. Wilkinson, "The grammar of graphics," in *Handbook of Computational Statistics*. New York, NY, USA: Springer, 2012, pp. 375–414.

[22] E. R. Tufte, N. H. Goeler, and R. Benson, *Envisioning Information*, vol. 126. Cheshire, CT, USA: Graphics Press, 1990.

[23] E. R. Tufte, *Beautiful Evidence*, vol. 1. Cheshire, CT, USA: Graphics Press, 2006.

[24] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proc. IEEE Symp. Vis. Lang.*, Washington, DC, USA, Sep. 1996, pp. 336–343. [Online]. Available: http://dl.acm.org/citation.cfm?id=832277.834354

[25] S. N. Friel, F. R. Curcio, and G. W. Bright, "Making sense of graphs: Critical factors influencing comprehension and instructional implications," *J. Res. Math. Educ.*, vol. 32, no. 2, pp. 124–158, 2001.

[26] B. S. Bloom, *Taxonomy of Educational Objectives: Cognitive Domain*, vols. 1–20. New York, NY, USA: McKay, p. 24, 1956.

[27] S. Pinker, "A theory of graph comprehension," in *Artificial Intelligence and the Future of Testing*. Hillsdale, NJ, USA: Roy Freedle, Educational Testing Service, 1990, pp. 73–126.

[28] S. B. Trickett and J. G. Trafton, "Toward a comprehensive model of graph comprehension: Making the case for spatial cognition," in *Diagrammatic Represent. Inference*, D. Barker-Plummer, R. Cox, and N. Swoboda, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 286–300.

[29] A. E. R. Association, "Aera," in *American Psychological Association [APA], and National Council on Measurement in Education [NCME]*. Chicago, IL, USA: NCME, 2014.

[30] T. L. Dawson-Tunik, M. Commons, M. Wilson, and K. W. Fischer, "The shape of development," *Eur. J. Developmental Psychol.*, vol. 2, no. 2, pp. 163–195, 2005.

[31] K. W. Fischer, "A theory of cognitive development: The control and construction of hierarchies of skills," *Psychol. Rev.*, vol. 87, no. 6, p. 477, 1980.

[32] M. Wilson, "Measuring progressions: Assessment structures underlying a learning progression," *J. Res. Sci. Teaching*, vol. 46, no. 6, pp. 716–730, Aug. 2009.

[33] M. Wilson and K. Sloane, "From principles to practice: An embedded assessment system," *Appl. Meas. Educ.*, vol. 13, no. 2, pp. 181–208, Apr. 2000.

[34] M. Wilson, *Constructing Measures: An Item Response Modeling Approach*. Evanston, IL, USA: Routledge, 2004.

[35] R. F. DeVellis, *Scale Development: Theory and Applications*, vol. 26. Newbury Park, CA, USA: Sage, 2016.

[36] J. Piaget, *The Equilibration of Cognitive Structures (t. Brown & KJ Thampy, Trans.)*. Cambridge, U.K.: Harvard Univ. Press, 1985.

[37] M. L. Commons and F. Richards, "A general model of stage theory," *Beyond Formal Oper.*, vol. 1, pp. 120–140, Mar. 1984.

[38] R. Case, *The Mind's Staircase: Exploring the Conceptual Underpinnings of Children's Thought and Knowledge*. Hillsdale, NJ, USA: Psychology Press, 2013.

[39] A. Demetriou and N. Valanides, "A three-level theory of the developing mind: Basic principles and implications for instruction and assessment," Lawrence Erlbaum Associates, Mahwah, NJ, USA, Tech. Rep., 1998.

[40] D. Dowding, J. A. Merrill, N. Onorato, Y. Barrón, R. J. Rosati, and D. Russell, "The impact of home care nurses' numeracy and graph literacy on comprehension of visual display information: Implications for dashboard design," *J. Amer. Med. Inform. Assoc.*, vol. 25, no. 2, pp. 175–182, Apr. 2017. [Online]. Available: https://doi.org/10.1093/jamia/ocx042

[41] J. Hullman and N. Diakopoulos, "Visualization rhetoric: Framing effects in narrative visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2231–2240, Dec. 2011.

[42] R. Gangle, *Diagrammatic Immanence*. Edinburgh, U.K.: Edinburgh Univ. Press, 2015.

[43] D. Petri, P. Carbone, and L. Mari, "Quality of measurement information in decision-making," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 1003816.

[44] L. Mari, M. Wilson, and A. Maul, *Measurement Across the Sciences: Developing a Shared Concept System for Measurement* (Springer Series in Measurement Science and Technology). Cham, Switzerland: Springer, 2021.

[45] M. A. Simpson, A. Kosh, L. Bickel, J. Elmore, E. Sanford-Moore, H. Koons, and M. Enoch-Marx, "Theory-based item families and the effect of varying grain size on the exchangeability of item isomorphs," Nat. Council Meas. Educ., Chicago, IL, USA, Tech. Rep. Document1, 2015.

[46] A. E. Kosh, "A process for automatically generating algebra items," *J. Appl. Test. Technol.*, vol. 20, no. 1, pp. 16–33, 2019. [Online]. Available: http://www.jattjournal.com/index.php/atp/article/view/140827

[47] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.

[48] K. A. Hallgren, "Computing inter-rater reliability for observational data: An overview and tutorial," *Tuts. Quant. Methods Psychol.*, vol. 8, no. 1, p. 23, 2012.

[49] W. S. Cleveland and R. McGill, "Graphical perception: Theory, experimentation, and application to the development of graphical methods," *J. Amer. Stat. Assoc.*, vol. 79, no. 387, pp. 531–554, Sep. 1984.

[50] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*. Newbury Park, CA, USA: Sage, 2018.

[51] D. W. Hosmer, Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 398. Hoboken, NJ, USA: Wiley, 2013.

[52] *Publication Manual of the American Psychological Association Washington*, Amer. Psychol. Assoc., Washington, DC, USA, 2010.

[53] A. Cairo, *The Functional Art: An Introduction to Information Graphics and Visualization*. San Francisco, CA, USA: New Riders, 2012.

[54] A. Locoro, F. Cabitza, R. Actis-Grosso, and C. Batini, "Static and interactive infographics in daily tasks: A value-in-use and quality of interaction user study," *Comput. Hum. Behav.*, vol. 71, pp. 240–257, Jun. 2017.

[55] A. Locoro, F. Cabitza, A. Ravarini, and P. Buono, "IGV short scale to assess implicit value of visualizations through explicit interaction," *Appl. Sci.*, vol. 10, no. 18, p. 6189, Sep. 2020.

[56] A. J. Dray, N. J. S. Brown, R. Diakow, Y. Lee, and M. R. Wilson, "A construct modeling approach to the assessment of reading comprehension for adolescent readers," *Reading Psychol.*, vol. 40, no. 2, pp. 191–241, Feb. 2019.

[57] L. J. Ausburn and F. B. Ausburn, "Visual literacy: Background, theory and practice," *Programmed Learn. Educ. Technol.*, vol. 15, no. 4, pp. 291–297, Nov. 1978.

[58] A. Rourke and Z. O'Connor, "Look before you leap: Testing some assumptions on visual literacy and predominant learning modalities of undergraduate design students in Australia and New Zealand," *Int. J. Learn., Annu. Rev.*, vol. 16, no. 8, pp. 33–46, 2009.

[59] J. L. Debes, "The loom of visual literacy–an overview," *Audiovisual Instr*, vol. 14, no. 8, pp. 25–27, Oct. 1969.

[60] M. Galesic and R. Garcia-Retamero, "Graph literacy: A cross-cultural comparison," *Med. Decis. Making*, vol. 31, no. 3, pp. 444–457, 2011.

[61] B. C. Pierce, *Basic Category Theory for Computer Scientists*. Cambridge, MA, USA: MIT Press, 1991.

ANGELA LOCORO received the B.Sc. degree in computer science and the Ph.D. degree in computer science from the University of Genova. She is currently an Assistant Professor with LIUC–Università Cattaneo, Castellanza, Italy, where she teaches statistical data analysis and data visualization. She is developing theories about visual information literacy and data visualization for decision making in organizations. Her research interests include data visualization, human–computer interaction, knowledge management, and information processing. Her publication records include peer-reviewed journals, such as *Computers in Human Behavior*, *Information Systems Frontiers*, *Computer Supported Cooperative Work (CSCW)*, and *Information Technology and Behavior*. She has served as a committee member, the chair, and the track chair for many international conferences about Human-Computer Interaction and Information Systems.

WILLIAM P. FISHER, JR. received the Ph.D. degree from the University of Chicago, where he was mentored by Benjamin D. Wright and supported by a Spencer Foundation Dissertation Research Fellowship. His online course on Rasch measurement applications attracts students from all over the world. His recent work focuses on the role uniform units of measurement play in harmonizing individual decisions and behaviors. His articles have appeared in *Frontiers of Psychology*, *Journal of Applied Measurement*, *Theory & Psychology*, the *International Journal of Educational Research*, the *International Journal of Transpersonal Psychology*, *Human Studies*, the *Archives of Physical Medicine and Rehabilitation*, *Neuropsychology*, *Advances in Nursing Science*, the *Journal of Nursing Measurement*, *Measurement*, and others.

LUCA MARI received the M.S. degree in physics from the University of Milan, Italy, in 1987, and the Ph.D. degree in measurement science from the Politecnico di Torino, Italy, in 1994. Since 2006, he has been a Full Professor of Measurement Science with LIUC–Università Cattaneo, Castellanza, Italy, where he teaches courses on measurement science, statistical data analysis, and system theory. He has been the Chair of the TC7 (Measurement Science) of the International Measurement Confederation (IMEKO). He is currently the Chair of the TC1 (Terminology), the Secretary of the TC25 (Quantities and Units) of the International Electrotechnical Commission (IEC), and an IEC Expert in the WG2 (VIM) of the Joint Committee for Guides in Metrology (JCGM).