CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

# The Role of Verbal and Visual Text in the Process of Institutionalization

Meyer, Renate; Jancsary, Dennis; Höllerer, Markus A.; Boxenbaum, Eva

*Document Version*
Accepted author manuscript

*Published in:*
Academy of Management Review

*DOI:*
10.5465/amr.2014.0301

*Publication date:*
2018

*License*
Unspecified

*Citation for published version (APA):*
Meyer, R., Jancsary, D., Höllerer, M. A., & Boxenbaum, E. (2018). The Role of Verbal and Visual Text in the Process of Institutionalization. *Academy of Management Review*, *43*(3), 392-418. https://doi.org/10.5465/amr.2014.0301

Link to publication in CBS Research Portal

**General rights**
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**
If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 26. Aug. 2022

CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

# The Role of Verbal and Visual Text in the Process of Institutionalization

Meyer, Renate; Jancsary, Dennis; Höllerer, Markus A.; Boxenbaum, Eva

*Document Version*
Accepted author manuscript

*Published in:*
Academy of Management Review

*DOI:*
10.5465/amr.2014.0301

*Publication date:*
2018

*Creative Commons License*
Unspecified

*Citation for published version (APA):*
Meyer, R., Jancsary, D., Höllerer, M. A., & Boxenbaum, E. (2018). The Role of Verbal and Visual Text in the Process of Institutionalization. *Academy of Management Review*, *43*(3), 392-418. https://doi.org/10.5465/amr.2014.0301

Link to publication in CBS Research Portal

**General rights**
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**
If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 30. Mar. 2021

# THE ROLE OF VERBAL AND VISUAL TEXT IN THE PROCESS OF INSTITUTIONALIZATION

*Renate E. Meyer, WU Vienna & Copenhagen Business School*

renate.meyer@wu.ac.at

*Dennis Jancsary, WU Vienna*

dennis.jancsary@wu.ac.at

*Markus A. Höllerer, WU Vienna & UNSW Sydney Business School*

markus.hoellerer@wu.ac.at

*Eva Boxenbaum, Mines ParisTech & Copenhagen Business School*

eva.boxenbaum@mines-paristech.fr

**Acknowledgements:** The authors acknowledge financial support from the Danish Research Council (DFF-1327-00030). Eva Boxenbaum also acknowledges support from the French National Research Agency (ANR-14-CE29-0008). We wish to thank SCANCOR for providing an excellent working environment during our sabbaticals. Dennis Jancsary also thanks the Weatherhead Center for International Affairs at Harvard University for its support. We are grateful to our editor and the three anonymous reviewers for their in-depth engagement with our ideas and for their valuable suggestions.

# THE ROLE OF VERBAL AND VISUAL TEXT

# IN THE PROCESS OF INSTITUTIONALIZATION

*Abstract.* In this article, we develop novel theory on the differentiated impact of verbal and visual texts on the emergence, rise, establishment, and consolidation of institutions. Integrating key insights from social semiotics into a discursive model of institutionalization, we identify distinct affordances of verbal and visual text based on the constitutive features of these respective semiotic modes. In an effort to extend scholarly inquiry into the relationship of text and institutions, we develop a set of propositions on how and under which conditions verbal and visual text, respectively, facilitate the institutionalization of novel ideas in each stage of the process. Our theory development has implications for research on institutions as communicative accomplishments, contributes to the nascent line of multimodal research, and provides novel insights into institutional emergence.

*Keywords*. Visual text; verbal text; affordances; institutionalization; social semiotics; institutional theory; multimodality

## INTRODUCTION

The central role of communication and language for the emergence of institutions has long been acknowledged (e.g., Berger & Luckmann, 1967). More recent literature has revitalized this agenda (e.g., Ashcraft, Kuhn, & Cooren, 2009; Harmon, Green, & Goodnight, 2015; Ocasio, Loewenstein, & Nigam, 2015; Li, 2016). The core premise is that "speech and other forms of symbolic interactions are not just seen as expressions or reflections of inner thoughts or collective intentions, but as potentially formative of institutional reality" (Cornelissen, Durand, Fiss, Lammers, & Vaara, 2015: 11). In this article, we extend insights in this field of inquiry by developing novel theory on the impact of different semiotic modes (i.e., socially shaped resources of meaning making like verbal and visual sign systems) (Kress, 2010) across specific stages in the process of institutionalization – a question that, as of yet, has been neglected.

Such neglect is surprising. Core work notes that 'texts' are not exclusively verbal (e.g., Cooren, Kuhn, Cornelissen, &, Clark, 2011; Phillips, Lawrence, & Hardy, 2004). In addition, an emerging stream of literature points to the substantial role of visual imagery in the constitution of organizations and institutions (for overviews, see, e.g., Bell, Warren, & Schroeder, 2014; Meyer, Höllerer, Jancsary, & van Leeuwen, 2013; Puyou, Quattrone, McLean, & Thrift, 2012). Still, most literature tends to regard the verbal as the sole, or, at least, the dominant, semiotic mode – and other modes as either an insubstantial add-on or governed by more or less the same rules and mechanisms as the verbal.

We contend that this bias has led to impoverished theories of institutionalization in which, implicitly, no difference regarding the impact of different semiotic modes is assumed. Consequently, more theory development is warranted to establish systematically in which communicative situations differences between these modes actually impact institutionalization, and how. In this article, we contrast the relevance of verbal text and visual text regarding the institutionalization of novel ideas – that is, of 'packages' of ideational and

behavioral, material and symbolic elements that were previously not available in the shared experiences of a particular social community (Berger & Luckmann, 1967). In more detail, we address the question of *how, and under which conditions, the use of verbal and visual text, respectively, has more potential to facilitate the institutionalization of novel ideas.*

To do so, we build on, and extend, a performative approach to communication (e.g., Cornelissen *et al.*, 2015; Green & Li, 2011) which understands communication as a generative force that builds, sustains, and challenges institutions through eliciting reactions in audiences. Our communicative model, accordingly, comprises the following components: Actors ('who?'), in their everyday social interaction, disseminate their ideas through texts ('what?'; e.g., Ashcraft *et al*., 2009; Li, 2016; Phillips *et al.*, 2004), using rhetoric and argument ('how?'; e.g., Green, Li, & Nohria, 2009; Suddaby & Greenwood, 2005) to communicate with audiences ('with whom?') in order to create, or transform, meaning and discourse, whereby they eventually contribute to institutional maintenance or change, and consolidate novel interpretations into full-fledged institutions ('to what effect?'; e.g., Gray, Purdy, & Ansari, 2015; Zilber, 2009). The model implies a mutually constitutive relationship between text (i.e., its production, dissemination, and reception), discourse, and institutions, and acknowledges that communication and institutions are co-constituted through recursive processes. Following Phillips *et al.* (2004), we focus on the vital role of texts, defined as symbolic expressions – spoken, written, or depicted – using a physical medium that permits storage and makes it accessible to others.

Discursive and rhetorical approaches acknowledge that the impact of texts on institutions varies with their content and composition (e.g., Green *et al.*, 2009; Sillince & Barker, 2012). We offer a first extension of the performative model by drawing on social semiotic theory to distinguish between a verbal and a visual *semiotic mode* of communication. We elaborate how the constitutive features of each mode endow text that draws on it with particular *affordances* (e.g., Gibson, 1986; Kress, 2010), that is, enabling and constraining

potentials for action and meaning-making that are realized relationally between a text, its producer(s), and its recipient(s).

As a second extension of the model, we contrast the salience of verbal and visual affordances for institutionalization by unpacking the stage model of institutionalization as proposed by Berger and Luckmann (1967; see also Tolbert and Zucker, 1996). For each stage, we develop specific *communicative requirements* that need to be overcome in order to move on successfully to the next stage (such as, for instance, initial placement, or justification). We outline a set of arguments explaining how the relative relevance of either verbal or visual text for accomplishing these requirements ('how?') is related to the *characteristics of the communicative situation*: the position of the text producer ('who?'; e.g., Phillips *et al.*, 2004), the characteristics of the novel idea ('what?'; e.g., Eisenman, 2013), and the match between the idea and the context into which it is communicated ('to whom?'; e.g., Meyer & Höllerer, 2010). These arguments lead to the development of a set of propositions that match specific verbal and visual affordances with particular requirements in distinct communicative situations.

Overall, our propositions suggest a differentiated impact of both verbal and visual text across all stages of institutionalization, a topic that has so far been ignored in the literature. Our work, consequently, contributes to existing models on the role of communication and text in the process of institutionalization as well as to the nascent line of multimodal research on institutions. We also extend these models by providing novel insights into institutional emergence and innovation.

The remainder of this article proceeds as follows. First, we elaborate on the characteristics of verbal and visual text. We then introduce the stage model of institutionalization, and identify characteristics of the communicative situation and communicative requirements for each stage. At the core of our article, we develop a set of propositions on how, and under which conditions, the affordances of verbal and visual text

facilitate the institutionalization of novel ideas. We conclude with implications for institutional theory, and outline promising avenues for future research.

## MEANING, TEXT, AND MULTIPLE SEMIOTIC MODES OF COMMUNICATION

In order to differentiate between and characterize the multiple modes that texts activate, we mainly build on social semiotics (e.g., Kress & van Leeuwen, 2001, 2006). Its concern with the construction and stabilization of meaning through signs in social contexts makes social semiotics highly compatible with discourse analysis (e.g. van Leeuwen & Wodak, 1999; Wodak & Meyer, 2016) and institutional theory in the tradition of Berger and Luckmann (e.g., van Leeuwen, 2007; Li, 2016). A semiotic mode of communication is "a socially made and culturally given semiotic resource for making meaning" (Kress, 2010: 79). Texts are assemblages of signs (e.g., images, words, sounds) instantiated in physical media. They draw, to varying degrees, on one or more semiotic modes. A social semiotic perspective distinguishes analytically between 'visual texts' and 'verbal texts' according to the semiotic mode that they activate in order to organize meaning, relate to audiences, and compose meaningful internal structures. Visual text includes, for instance, photographs as well as graphs, charts, and cartoons, but excludes any kind of 'mental' image and visual impressions that are not stored on a material support (e.g., what I see when walking through the park, or what I imagine visually) – neither would be considered 'text'. Visual aspects of written text (i.e., typography, color, etc.) and material artefacts are not the focus of our article, and neither are moving visual texts (e.g., videos). Verbal text, on the other hand, includes both written documents and oral recordings.

The basic insights that we derive from a semiotic perspective are further substantiated by knowledge from cognitive psychology. This latter literature clarifies how the different semiotic modes operate cognitively and, as such, sustains the claims by social semiotics.

Cognitive psychology explains the activation of particular sensory equipment, information processing, and storage/retrieval of information, as well as the sensory repositories on which texts rely (Atkinson & Shiffrin, 1968). A sensory repository refers to the initial holder of the information that we receive through a particular sensory function, such as sight, hearing, or touch. Whereas visual text relies exclusively on the iconic repository (i.e., sight), verbal text uses either the echoic repository (if spoken) or the iconic repository (if written) to store the initial information we absorb. This process makes the activation of the visual sensory equipment a necessary, but insufficient criterion for distinguishing visual from verbal text. The articulation of social semiotics with relevant knowledge from cognitive psychology is highly conducive for studying processes of institutionalization. The reason is that signs need to be cognitively available to actors before they can be mobilized for meaning construction.

Below, we distill a number of constitutive features for each of the two semiotic modes as well as particular affordances of verbal and visual text that flow from these insights. Finally, we elaborate on the advantages of each mode for accomplishing certain communicative challenges at key stages of institutionalization.

## Constitutive Features of Semiotic Modes of Communication

Constitutive features of a mode are properties that are characteristic for this mode. These properties permeate and underlie every text that builds on this mode within a certain cultural community. We differentiate three types of constitutive features: *Semiotic* features describe how a particular mode enables meaning construction. *Cognitive* features refer to the way in which a particular mode is perceived and processed through the cognitive apparatus of individuals. *Cultural* features, finally, concern the social organization and regulation of a mode and its use within a specific social setting.

Most texts encountered in and around organizations are multimodal, i.e., they draw on different modes. Multimodality is a common feature in many established genres of organizational communication, such as corporate reports, logos, websites, media reports, digital presentations, intranet, newsletters, or even emails. For instance, PowerPoints (e.g., Kaplan, 2011) combine verbal (e.g., words and sentences) and visual (e.g., layout, color, font) information. In extracting the constitutive features of each mode from existing literature to systematize the differences between them, we use a 'pure' manifestation of the verbal mode (i.e., spoken text without visual characteristics) and the visual mode (i.e., photography without verbal characteristics). We acknowledge that spoken verbal text overlaps with the auditory mode, just as written verbal text overlaps with the visual mode but for purposes of analytical clarity, we leave out the overlaps between spoken verbal text and other texts that work auditorially (such as music). Multimodal texts, accordingly, mix constitutive features of the modes on which they build. Table 1 provides an overview of the constitutive features of the verbal mode and the visual mode.

---

Insert Table 1 about here

---

*Semiotic features*. A first substantial difference between the two modes lies in the type of signifiers (Peirce, 1906) each mode provides (e.g., Messaris, 1997). Whereas the basis for understanding verbal signifiers (e.g., words and sentences) are cultural conventions (symbols), visual signifiers can be either iconic (e.g., photo of a house), indexical (e.g., footprints in the snow), or symbolic (e.g., a national flag). This difference in signifiers implies that visual text is more suitable than verbal text for mimicking direct sensory experience (with some exceptions, such as onomotopoetic words).

Second, the two modes structure information differently. Whereas the visual mode presents the world, and thereby suggests meaning, in a primarily spatial manner, the verbal mode's basic structuring principle is sequential and linear (e.g., Jewitt & Oyama, 2001; Kress & van Leeuwen, 2006). Also, verbal text is additive, with each element (e.g., a word) adding a clearly defined contribution to the overall text (e.g. through grammatical rules or word order). Visual text signifies more holistically than verbal text and can present its content in multidimensional form, including spatial depth and perspective, to convey information about potentially complex relations among many elements. Whereas visual text is better suited to capture spatial information, verbal text excels at conveying temporality. The use of tenses makes it possible to mark verbal text with temporal characteristics. Visual text does not provide equally sophisticated tools for temporality. Most impressions within a single visual text are perceived as occurring simultaneously (with the exception of some stylistic ways of evoking the past, like picture quality, art styles, or nostalgic filters, or by using timelines or arrows in figures to represent temporality).

Third, the modes differ in the way in which they convey perspective and attitude. Verbal text uses particular grammatical forms (pronouns) to indicate perspectives. Pronouns offer readers a variety of differentiated attitudes by providing different roles and positions within a single text. Visual text, in contrast, communicates perspective and attitude in an embodied form, meaning that it positions the viewer spatially and corporeally with regard to the depicted scene and/or objects (e.g., Kress & van Leeuwen, 2006). This embodied nature, representing a specific perspective or 'gaze' (e.g., Styhre, 2010), enhances the experience of direct interaction between the observer and the observed.

Overall, the semiotic features of the two modes are most clearly differentiated by the complex and clearly defined grammar and syntax of the verbal mode. Even if the visual mode is characterized by its own 'grammar' (e.g., Kress & van Leeuwen, 2006), such rules of

expression are much less explicit and regulated, which makes for a more open-ended interpretation of visual text.

*Cognitive features*. Semiotic theory commonly claims that visual text has a more immediate effect on perception and comprehension than verbal text. Barthes (1991: 108), for instance, noted that the visual mode gives the impression of imposing "meaning at one stroke, without analyzing or diluting it" (see also Mitchell, 1984; Rowley-Jolivet, 2004). Such claims can be further substantiated with insights from cognitive psychology. Verbal written text is perceived as a sequence of imprints, whereas images are perceived as a single imprint (Liversedge *et al.*, 2004). Even if, in the case of written verbal text, we use the same sensory repository for verbal and visual text, one important difference remains: In absorbing verbal written text, we move our eyes along the line of the text, which requires us to process multiple visual imprints sequentially. In contrast, no differential emphasis, or sequencing, is required to decode visual text, where all perceptual information is assigned equal importance (Coltheart, 1980) and material is stored in our visual short-term memory in the form of coherent, holistic objects (Luck & Vogel, 1997). These differences make the perception of visual text much more immediate than that of verbal text.

Additionally, scholars have emphasized the polysemy of visual as opposed to verbal meaning (Eco 1995; Barthes, 1977). As Messaris (1997) notes, the lack of a clear visual 'syntax' makes visual meaning fluid and indeterminate, and strongly dependent on the viewers' interpretational predispositions. Cognitive psychology explains this interpretive openness through the ability of the iconic repository (i.e., the holder of our initial visual impressions) to retain an impressive, if not unlimited, amount of detail (Coltheart, 1980). However, repositories have rapidly fading contents (Sperling, 1960), which means that only a small fraction of visual information can be transferred to short-term memory for processing (Atkinson & Shiffrin, 1968). Since meaning-making consists of associating new information

with elements retrieved from our long-term memory (e.g., Coltheart, 1980), we tend to more readily perceive those elements with which we are already familiar. This implicit selection entails that meaning construction only has access to a small, highly selective number of items conveyed in an image. Since individuals make their own implicit selection as to which few elements should be transfered from the visual repository to short-term memory, visual text opens itself to multiple interpretations and potentially quite divergent meaning construction. Visual text, accordingly, has an enhanced potential for polysemy based on the excess of information it contains, only part of which can be processed and therefore used for constructing meaning.

*Cultural features.* The character of different modes is also influenced by their status and social regulation within a specific community (Kress, 2010). Societies can differ considerably in the ways in which, and the degrees to which, they utilize and legitimize different modes. In Western societies, the verbal mode is subject to far stronger regulation than the visual mode (e.g., Kress & van Leeuwen, 2006). Elaborate rules exist on the appropriateness of verbal utterances and arguments, and boundaries of what can be said and written are in place and are sanctioned. In contrast, rules are far less elaborate on how and when to use visual texts, as well as on the limits of depiction, with notable exceptions primarily related to religion and the realms of death and eroticism. Examples include the controversial cartoon depictions of the Prophet Mohammed in Danish and French newsmedia in 2005-2006, or the worldwide depiction in 2015 of a drowned Syrian refugee child on the Turkish seashore. The lack of explicit rules reduces the accountability of visual text producers (and diffusers). Like the difference in regulation, evaluation differs between the visual mode and the verbal mode. In Western societies, the word is regarded as more precise than the image – hence it is considered to be the dominant form of communication and construction of meaning (e.g., Berger & Luckmann, 1967). Visual text is often relegated to ‘less serious’ areas of social life, such as art, advertising, and entertainment.

Sign-making is key to processes of institutionalization, which are characterized by both cognitive and cultural features (Berger & Luckmann, 1967). The semiotic, cognitive, and cultural features of each mode, i.e., their communicative infrastructure, resonate with the requirements of institutionalization. In the following, we distill, for each of the two modes, a set of affordances that each rely on one or more of these constitutive features. We then develop propositions about how, and when, these affordances impact on processes of institutionalization.

## Affordances of Verbal and Visual Text

Social semiotics borrows the concept of *affordances* from Gibson (1986) in order to denote a mode's potential in representation and communication. Affordances, which are derived from a mode's constitutive features, are relational rather than inherent, and they manifest themselves only in specific interactions between audiences, particular texts, and contexts (McDonnell, 2010). In other words, semiotics locate affordances *within* visual and verbal texts, but the effect of these affordances will depend on *relations* between text producer, audiences, texts, and contexts. The potential of affordances for meaning-making can be both enabling and constraining. Their relational character and cultural quality imply that the properties of specific texts invite, but never determine, possibilities for interpretation and action in relation to these texts (e.g., Hutchby, 2001). Inspired by Kress (2010: 96), we suggest that identifying the affordances of a particular mode is equivalent to answering the questions 'what can be done with this mode?' and 'what are the best means for achieving my rhetorical aims in my designs of communication?'.

We highlight in the following those affordances that are specific for one of the two modes and that we deem to be particularly relevant for the institutionalization of novel ideas. In so doing, we refrain from mentioning affordances that offer no clear advantage of one mode over the other in relation to institutionalization processes, i.e., where the relevance for

our research question is unclear. In formulating the meaning-making potential of each mode as distinct affordances, we draw on and synthesize a substantial body of existing literature. We label affordances as verbs. This underscores that they – when realized by audiences in a social situation – suggest particular processes of interpretation and possibly action. Figure 1 summarizes the core constitutive features and affordances of the visual mode and the verbal mode that we identified as being particularly relevant for the process of institutionalization. We start by briefly summarizing verbal affordances and subsequently contrast them with visual affordances. The most relevant affordances of verbal text for processes of institutionalization are *argue, specify, narrate,* and *abstract*.

---

Insert Figure 1 about here

---

*Argue.* The strong social regulation of verbal expression, combined with its higher determinacy, afford the construction of systematic formal arguments. Aristotelian rhetoric has strongly influenced Western understandings of what constitutes the 'logic' of argumentation. Contemporary approaches (e.g., Toulmin, 2003) contend that culturally legitimated support and backed-up warrants are required for linking data to a qualified conclusion. Purely visual text cannot entirely reproduce such formal structures (e.g., Birdsell & Groarke, 1996; Blair, 1996; Messaris, 1997). For instance, institutionalist literature on rhetorical approaches stresses the relevance of arguments and causal explanations for advancing institutionalization (e.g., Green et al., 2009; Suddaby & Greenwood, 2005). Work on theorization is similarly inclined toward arguments and explanations (e.g., Greenwood, Suddaby, & Hinings, 2002; Strang & Meyer, 1993).

*Specify*. The additive and temporal signification of verbal text enables a high degree of specificity in communication. This affordance operates through clear syntax and grammar and lower polysemy. For instance, the verbal mode provides highly differentiated resources for expressing transitivity, modulation, and time (e.g., Halliday & Hasan, 1989). In comparison, visual text operates in a much less sophisticated and precise way in this respect (e.g., Kress & van Leeuwen, 2006). Probabilities (e.g., whether an event is certainly, likely, or unlikely to happen) or modal verbs and conjunctive clauses (e.g., whether something can, could, shall, should, or needs to happen), and especially the differences between them are difficult to express visually. Additionally, verbal language can specify objects, attributes, and ideas that cannot be seen with the naked eye. Such specificity allows for the construction of elaborate vocabularies that support, for instance, theorization (e.g., Ocasio & Joseph, 2005).

*Narrate.* The sequential structure of verbal text, combined with its strength in conveying temporality, makes it an excellent resource for the creation of elaborate narratives. The ability to provide multiple perspectives for readers further supports this affordance, which consists in aligning events in specific plots and clear temporal structures (e.g., Vaara, Sonenshein, & Boje, 2016), thereby assembling a variety of elements into coherent wholes. Narratives are essential tools for sensemaking (e.g., Cornelissen, 2012; Vaara & Tienari, 2011), translation (e.g., Wedlin & Sahlin, 2017), and, more generally, institutional work (e.g., Zilber, 2009). Consequently, narrativization is a verbal affordance that plays a central role in the institutionalization of novel ideas.

*Abstract.* Since the relationship between signifier and signified in verbal text is conventional (e.g., Peirce, 1906), verbal text is able to convey ideas and concepts that have no direct correspondence in the world of sensory perception. Through its ability to abstract, verbal text facilitates typification (e.g., Berger & Luckmann, 1967; Tolbert & Zucker, 1996) and categorization (e.g., Durand & Paolella, 2013; Kennedy & Fiss, 2013). As prerequisites

for the emergence of complex vocabulary structures (e.g., Loewenstein, Ocasio, & Jones, 2012), typification and categorization are central elements in the institutionalization of novel ideas (e.g., Phillips *et al.*, 2004; Munir & Phillips, 2005), and part of the 'toolkit' for institutional work (e.g., Weber, 2005).

Visual text, on the other hand, provides its own specific affordances that have not yet been recognized as central aspects of institutionalization processes. We argue that they are – in their own way – equally important to the affordances of verbal text in institutionalization. The affordances of visual text encompass *infiltrate, spatialize, captivate,* and *materialize*.

*Infiltrate.* Visual text is subjected to less scrutiny than verbal text due to the weaker social regulation that applies to visual text in Western cultures (e.g., Kress & van Leeuwen, 2006) and its enhanced potential for polysemy. On the one hand, these features enable experimentation and the transgression of norms, suggesting that visual text can transport messages that cannot be legitimately verbalized (e.g., McQuarrie & Phillips, 2005). On the other hand, these features make visual text more amenable to communicating tacit, aesthetic, or embodied knowledge (e.g., Toraldo, Islam, & Mangia, 2016) that is difficult if not impossible to articulate explicitly in verbal form. Note, however, that the ability of visual text to 'fly under the radar' (e.g., Meyer *et al.*, 2013) does not mean that it remains hidden or attracts less attention (see 'captivate' as affordance). Rather, the indeterminacy of visual meanings combined with the lower social regulation of visual text, enable deviation from norms and the communication of intangibles, even if they stimulate a high degree of attention.

*Spatialize*. Its inherently spatial and holistic signification makes visual text particularly useful for foregrounding the importance of space(s) (e.g., Shortt, 2015). Information about spatial depth and configuration, in particular, is very difficult to express verbally. This difficulty makes visual arrangements especially well-suited for communicating complex and multidimensional relationships. Additionally, the spatial setup of visual text also allows for

the bridging of individual elements through composition and positioning (e.g., Höllerer *et al.*, 2013). In this way, visual text not only materializes ideas, but also locates them in specific environments, thereby allowing the unfamiliar to be 'toned' with familiar elements. In our context, the ability to spatialize facilitates, on the one hand, the literal 'placement' of a particular idea in a broader system of ideas and, on the other hand, the intuitive establishment of a variety of relationships without explicit argument.

*Captivate.* Embodied subjectivities, immediacy of perception, and simultaneity of signification endow visual text with eminent potential to captivate audiences. More specifically, visual text is perceived rapidly (e.g., Edell & Staelin, 1981; McQuarrie & Mick, 1992), attracts attention quickly (e.g., Bloch, 1995), and expresses and elicits attitudes and emotions powerfully and in a way that often precedes active awareness (e.g., Blair, 1996; Hill, 2004; Raab, 2008). Moreover, visual text creates situated perspectives, which suggest a sentiment of involvement and personal relevance (e.g., Kress & van Leeuwen, 2006). As a result, this affordance enables visual texts to generate an immediate and powerful impact that surpasses a purely cognitive processing of its content, whereby audiences become affectively, aesthetically, and corporeally engaged.

*Materialize*. In depicting certain ideas iconically, visual text is able to achieve verisimilitude (i.e., truthlikeness), which suggests 'objective' representation (e.g., Mitchell, 1984; Raab, 2008) and 'facticity' (e.g., Graves, Flesher, & Jordan, 1996). First, this affordance firmly anchors novel ideas in the 'here and now', even if they refer to future visions as well as past and/or geographically distributed events (e.g., Höllerer, Jancsary, & Grafström, 2017). It does so by giving them tangibility and shape in the concrete materiality of people, objects, and events in a way that verbal text is unable to do. Second, visual text is unable to remain completely abstract. Photographs, for instance, cannot show a 'pure' type without also illustrating it. As Metz (1971; cited in Machin 2004: 320) notes, "the image of a

house can never mean ‘house’ but only ever ‘Here is a house’”. Inescapably, an image also provides additional information, like architectural style or size. Finally, although the interpretation of visual text is culturally and temporally specific, materialization helps overcome language barriers and establishes visual text as a kind of ‘global visual language’ (Machin, 2004).

Summing up, we have identified a number of constitutive features that differentiate the visual from the verbal mode. We have derived from these constitutive features specific affordances that each mode can offer for the process of institutionalization. In the following section, we briefly revisit the stage model of institutionalization (Berger & Luckmann, 1967; Tolbert & Zucker, 1996) and then outline a set of characteristics of the communicative situation that influence which communicative acts are more likely to propel a novel idea to the next stage of institutionalization. Subsequently, we develop a set of propositions that outlines which affordances of verbal and visual text, respectively, are most likely to further institutionalization in specific communicative situations at each stage of institutionalization.

## THE ROLE OF MULTIPLE SEMIOTIC MODES IN INSTITUTIONALIZATION

Institutionalization is a complex process that involves “a typical pattern of events and relationships among them” (Lawrence, Winn, & Jennings, 2001: 626). In more detail, all established models of institutionalization entail three ‘ideal-typical’ *stages* that describe the path, or trajectory, for novel ideas to eventually become institutions: pre-, semi-, and full institutionalization (Berger & Luckmann, 1967; Tolbert & Zucker, 1996; see also Barley & Tolbert, 1997; Greenwood *et al.*, 2002; Li, 2016; Meyer, 2008). Three components are of particular relevance to our model: the *idea* as the object of institutionalization, stage specific *requirements* for progressing in the process of institutionalization, and the relevant *characteristics of the communicative situation*.

According to Berger and Luckmann (1967), institutionalization describes the reciprocal typification of actor and action, that is, activities become typified into practices, which are then performed by typified actors as a part of social roles. At the same time, subjective meaning becomes social meaning and, eventually, sediments in social knowledge. In the process of institutionalization, typified activities, categories of actors and meanings continuously become more closely connected to one another, forming a 'package' of ideational and behavioral, material and symbolic features of an emergent institution. For purposes of clarity, we use the term *'idea'* to denote this 'package' throughout the process of institutionalization. Communication is central throughout institutionalization; in fact, it is so central that Luckmann in later years relabeled the 'social construction of reality' into the 'communicative construction of reality' (Luckmann, 2006). Although the terminology varies and no systematic overview exists, literature suggests that the role of text varies according to the stage of institutionalization (e.g., Green *et al.*, 2009; Sillince & Barker, 2012).

The conceptualization of stages is important for our theory development in two respects. First, each stage can be characterized by a number of challenges that need to be overcome in order to progress to the next stage. We model these challenges as stage-specific *requirements* that need to be collectively, and discursively, accomplished. For example, a novel idea must be successfully placed within existing knowledge in order to become comprehensible in a specific cultural setting. Second, the accomplishment of each stage constitutes the point of departure for the following stage. For instance, in a semi-institutionalization stage, we can presume that the idea has had substantial exposure and that it is regarded as relevant and meaningful within the field.

The relevance of verbal and visual texts for accomplishing these requirements is related to *characteristics of the communicative situation*. A communicative situation denotes a situation characterized by certain features, and not a specific natural setting (a distinct 'here and now'). Our theory development, accordingly, is in this sense also meant to apply across

natural settings. First, characteristics of the text producer(s) ("who is speaking?") refer primarily to the field position of speakers and are particularly important prior to full institutionalization. Actors with high prominence have a higher likelihood of being perceived and imitated than those at the margins of the field do (e.g. Haveman, 1993; Phillips *et al.,* 2004). Research also suggests that theorizations of novel ideas are more likely to be accepted if they stem from actors with a high standing in the field, such as experts and other legitimated theorists (e.g., Strang & Meyer, 1993). Secondly, characteristics of the novel idea ("what is spoken about?") substantially influence the relevance of text, especially before an idea becomes fully institutionalized. To make it appealing, a novel idea needs to be communicated differently depending on whether it is associated with a clear and predictable outcome or with a vision that carries aesthetic and/or affective appeal (e.g., Eisenman, 2013). In addition, its generalization depends on whether its characteristics are primarily anchored in sensory perception or in relation to other ideas (Gentner & Kurtz, 2005). Finally, the match between the novel idea and the field ("what is the audience?") plays an important role across all stages of institutionalization. Since fields work as discursive 'resonance chambers', the match of the novel idea and existing meanings (e.g., Hargadon & Douglas, 2001), the alignment of novel ideas with shared understandings in the field (e.g., Meyer & Höllerer, 2010), the potential for contestation (e.g., Maguire & Hardy, 2009), and the match of vocabularies between contexts and audiences as a prerequisite for broad translation and diffusion (e.g., Wedlin & Sahlin, 2017) influence which affordances are most useful for furthering institutionalization.

Table 2 schematically summarizes which characteristics of the communicative situation are relevant for the accomplishment of the requirements associated with each of the three stages of institutionalization. In order for an idea to move through pre-institutionalization, it must be aligned with field level meanings (*placement),* become cognitively available to the target audiences (*exposure*), and be appreciated and deemed

relevant by audiences (*appeal/mobilization*). To advance through semi-institutionalization, the idea must acquire a generalized meaning (*typification*) and be perceived as a pertinent solution to a general problem (*explanation*) that carries legitimacy (*justification*). For full institutionalization to set in, finally, the idea must be maintained as a 'background program' (*taken-for-grantedness*) and be comprehensible across audiences and contexts (*diffusion/translation*).

---

Insert Table 2 about here

---

In the remainder of this section, we structure our theorizing according to the three stages of institutionalization and further divide each stage into its central requirements. For each requirement, we first outline which characteristic of the communicative situation primarily impacts the role of text. We then discuss how particular affordances of verbal and visual text, respectively, help meet the requirements of specific communicative situations. We summarize our suggestions in propositions that outline *when* (i.e., which characteristic of the communicative situation) and *how* (i.e., based on which affordance) verbal and visual text, respectively, has eminent potential to support the accomplishment of a particular requirement, hence advancing the process of institutionalization. In the next section, we bring together all components of our model. More precisely, each proposition contains 1) the situational characteristic that conditions the validity of the proposition, 2) the affordance(s) of the respective mode, and 3) the communicative requirement that needs to be accomplished at a given stage in order for institutionalization to move to the next stage.

## Pre-institutionalization

New ideas gain social existence through externalization, a concept which refers to the process of projecting ideas that are imbued with subjective meaning into the intersubjective and social realm (e.g., Berger & Luckmann, 1967). The production, dissemination, and reception of text is particularly salient for externalization. More precisely, text is vital for accomplishing three requirements related to externalization: *Initial placement* embeds the novel idea in existing knowledge (e.g., Munir & Phillips, 2005; Hargadon & Douglas, 2001). Through *exposure*, potential adopters become aware of the novel idea and existing understandings become destabilized (e.g., Greenwood *et al.*, 2002). In order for an idea to gain acceptance (e.g., Gondo & Amis, 2013), such exposure needs to be complemented with *appeal* and *mobilize consensus,* both of which make it possible to rally allies, garner attention, and stress the relevance of the novel idea.

*Initial placement*

A basic requirement of any text in the pre-institutionalization stage is an alignment of the novel idea with field level meanings, which makes it comprehensible to audiences. This alignment is accomplished by references to well-established and shared understandings (e.g., Cornelissen, Holt, & Zundel, 2011; Phillips *et al.*, 2004). Initial placement may be facilitated, for instance, by analogy and metaphor (e.g., Boxenbaum & Rouleau, 2011; Etzion & Ferraro, 2010), by historicizing the novel idea in elaborate narratives (e.g., Foster, Suddaby, Minkus, & Wiebe, 2011), or by mimicking features of existing institutions (e.g., Hargadon & Douglas, 2001).

Verbal text *narrates* the novel idea as part of overall knowledge. Narratives create continuity elaborately and systematically (e.g., Garud, Gehman, & Giuliani, 2014), make sense of the unfamiliar (e.g., Vaara & Tienari, 2011), and emphasize common qualities (Lounsbury & Glynn, 2001). However, they require an adequate vocabulary to capture the novel idea, and a general match of the idea with existing understandings in order to

meaningfully weave it into more comprehensive story lines. Additionally, verbal text is also very well suited to *specify* novel ideas, for instance, by describing features in detail that are invisible and therefore impossible to visualize. However, ideas that are radically novel may be hard to force into narratives without losing this very novelty. For instance, a novel idea that supports equality of women in the workplace may be easy to align narratively with existing ideas and practices in contexts where diversity and equality is already firmly established, like in Scandinavia. In contrast, specificity may trigger resistance if meanings are contested within the field, such as when novel concepts stressing and specifying the primacy of shareholder interests are introduced into stakeholder-oriented contexts (see, e.g., Meyer & Höllerer, 2010). Accordingly, the affordances of verbal text are particularly useful for initial placement when the existing stock of knowledge provides a fitting vocabulary, and when the suggested meaning of the novel idea resonates well with the audience. In such situations, verbal text enables refined and elaborate story lines.

P1a: The higher the match between the novel idea and existing meanings within a field, the more supportive narration and specification and, hence, verbal text are for initial placement.

Visual text, in contrast, bestows upon a novel idea a particular shape or form by *materializing* it through concrete people, objects, actions, and settings, thereby making it more tangible and projecting it into the 'here and now'. In visual text, the old and the new are linked through *spatializing* and juxtaposition, which de-problematizes novelty by embedding it in familiar settings and making it part of 'everyday life' (e.g., Hargadon & Douglas, 2001; Höllerer *et al*, 2017). This embedding is aided by the ability of visual text to *infiltrate* discourse by flying 'under the radar'. By refraining from explicit argument and specification, visual text achieves initial placement without challenging entrenched understandings or wrestling with known lines of contestation. Early proponents of CSR in a corporatist context, for instance, used

highly polyvocal imagery that combined the visual depiction of company signifiers with nature, religion, and local settings, instead of placing the concept clearly in relation to the existing stakeholder orientation (Höllerer *et al.,* 2013). Consequently, visual text is particularly useful for initial placement when the novel idea is difficult to align with established meanings – for instance, when an idea clearly deviates from an established consensus, or when the field is so fragmented that alignment is difficult (e.g., Seo & Creed, 2002; Purdy & Grey, 2009).

P1b: The lower the match between the novel idea and existing meanings within a field, the more supportive materialization, spatialization, and infiltration and, hence, visual text are for initial placement.

*Exposure*

The exposure of audiences to new ways of thinking is a basic requirement for overcoming inertia and rigidities, for initiating sensemaking, and for accomplishing change (Ocasio, 2011). In fact, radically new ideas may require a disruption of field stability (Greenwood *et al.*, 2002) in the sense of unfreezing shared stocks of knowledge, disturbing field-level consensus, and/or breaking locked-in patterns of behavior and thought (Seo & Creed, 2002). Making the novel idea available for target audiences is the main challenge at this early stage of institutionalization. The likelihood of an idea being perceived and remembered is strongly related to the position of the text producer in the respective field. Prominent, that is highly visible, actors are more likely to attract attention (e.g., Neidhardt, 1994), and organizations that are considered to be at the forefront of the industry are more systematically observed (e.g., Haveman, 1993). For instance, 'management gurus' are a substantial part of the management-fashion-setting community due to their high visibility in terms of bestselling books and appearances at seminars and colloquia (Kieser, 1997). Since it takes longer to decode, verbal text does not have the same potential to attract immediate attention as does

visual text (e.g. Liversedge et al., 2004). However, entrepreneurs have been known to create elaborate projective *narratives* (e.g., Garud, Schildt, & Lant, 2014; Vaara *et al.*, 2016) in order to expose their ideas. We therefore suggest that verbal text is particularly effective when text producers enjoy a high degree of visibility.

P2a The higher the prominence of a text producer in the respective field, the more supportive narration and, hence, verbal text is for the exposure of a novel idea.

In contrast, visual text gains attention more immediately since it imprints itself more rapidly and is less filtered than verbal text. These characteristics enhance the exposure of novel ideas through unexpected and/or otherwise disrupting stimuli that *captivate* audiences. For instance, the first images of humans landing and walking on the moon have captivated audiences and stimulated imagination in a more immediate, and far more powerful, way than verbal text has ever achieved. Additionally, since visual text is less socially controlled than verbal text (e.g., Kress & van Leeuwen, 2006), its use grants more leeway for experimentation through *infiltrating* the established discourse in a field. The lesser social control of visual text makes it possible for text producers to distance themselves from interpretations that their visual texts evoke (e.g., McQuarrie & Phillips, 2005; Messaris, 1997). This affordance is particularly vital for more marginally positioned actors who are vulnerable to legitimacy discounts.

P2b The lower the prominence of a text producer in the respective field, the more supportive captivation and infiltration and, hence, visual text are for the exposure of a novel idea.

*Appeal and mobilization*

The exposure of a novel idea is a necessary but insufficient criterion for 'kick starting' its process of institutionalization. Ideas that are generally perceived as irrelevant for broader

concerns, unappealing, or unable to captivate audiences emotionally are unlikely to move beyond idiosyncratic and sporadic use (e.g., Green, 2004). The key to achieving appeal, accordingly, is to convey that the novel idea provides a desirable outcome (e.g., Munir & Phillips, 2005) and to mobilize consensus regarding its relevance.

Similar to initial placement, verbal text is particularly useful when relevance and appeal of a novel idea can be *specified* with existing vocabulary. This specification is particularly effective when the idea's outcomes can be presented as clear and predictable (e.g., reduced use of resources, enhanced speed of procedures, etc.). Additionally, the ability of verbal text to specify becomes salient when the appeal of novel ideas is based on features that *can* be described in detail, but that cannot (e.g., changes on the molecular level in technological innovations) or should not be seen (e.g., mimicking existing visual styles to avoid resistance; see Hargadon & Douglas, 2001). Verbal text, in summary, is particularly relevant at the early stage of institutionalization when appeal and mobilization are based on deliberation, rather than emotion.

P3a The more a novel idea suggests a clear and predictable outcome, the more supportive specification and, hence, verbal text is for appeal and mobilization.

Turning to the potential of visual text, aesthetic appeal that triggers emotions and mobilizes consensus around a shared problem has been found to be particularly important as long as no dominant model or theorization exists (e.g., Eisenman, 2013). Visual text is well suited to *captivate* audiences and create engagement on an affective, rather than a cognitive level (e.g., Schill, 2012). Especially first-person perspectives help the audience 'feel' the relevance of a novel idea by drawing the viewer into embodied perspectives and making her part of the scene. This embodied engagement may not only put "fire in the belly and iron in the soul" (Gamson, 1992: 32) regarding the problems of one's own community, but also engage the

audience directly and physically in the problems of others, sensitizing them more strongly to such grievances than similar appeals expressed in verbal text would. For instance, Dogra (2007: 165) mentions a fundraising campaign by an international NGO that "shows needy looking dark children staring at the camera though it has been 'cropped' innovatively with only one eye of each child shown". This is an excellent example of visual text creating engagement and urgency, thereby supporting the mobilization of consensus. The use of visual text, accordingly, plays a crucial role at the early stage of institutionalization when the outcome of ideas is more ephemeral, visionary, and/or its appeal is based more on 'feeling' than 'deliberation'.

P3b The less a novel idea suggests a clear and tangible outcome, the more supportive captivation and, hence, visual text is for appeal and mobilization.

## Semi-institutionalization

After successful pre-institutionalization, ideas become detached from their origin and novices begin to experience them, during semi-institutionalization, "as existing over and beyond the individuals that 'happen to' embody them at the moment" (Berger & Luckmann, 1967: 58). The predominant task becomes the theorization of the novel idea, which is a prerequisite for its further diffusion (e.g., Greenwood *et al.*, 2002; Strang & Meyer, 1993; Tolbert & Zucker, 1996) and which encompasses the following requirements: *typification* of abstract categories, *explanation* including clarification of causal relationships between categories, and *justification* of the idea.

### *Typification*

Through typification, singular empirical occurrences become understood in terms of more abstract social categories. Typification of ideas extends their meaning beyond the idiosyncratic, temporally and spatially bounded case. Furthermore, in order to imbue categories with social meaning and make them applicable, examples link abstract categories to perceptual parts of social reality (e.g., Loewenstein *et al.*, 2012; Nigam & Ocasio, 2010). The central challenge for typification is the reduction of an idea to its generalized meanings. Texts achieve this reduction by stressing the most salient and central characteristics of an idea and removing its contextualized 'flavor' (e.g., Wedlin & Sahlin, 2017; Meyer, 2014). We suggest that typification is likely to draw on different textual means depending on whether an idea is centered around either its perceptual attributes or its relational embeddedness in a system of social categories (e.g., Loewenstein, 2014; Gentner & Kurtz, 2005).

Verbal text is excellently suited to establish particular labels as signifiers for novel ideas. Labels distinguish novel ideas from each other, contribute to perceived homogeneity within a category, and bring the "brute fact [of the referent] into the web of signifiers that acquire signified meanings only in relation to one another" (Li, 2016: 25). A management idea called 'CSR', for instance, does not have a particular material form, so it cannot be depicted *per se*. Ideas that derive their meaning from their embeddedness in systems of social categories are typified as 'relational' categories (e.g., Gentner & Kurtz, 2005). Since words are generally independent from the material manifestation of a particular idea, verbal text can create and play with signifiers more freely through its *abstracting* affordance. This affordance permits the communication of ideas that are not grounded in a perceivable reality at all. Additionally, the ability to *specify* allows verbal text to provide precise and comprehensive definitions of relational features. Such avenues towards typification are only viable, however, when the novel idea is not centrally constituted by its material, spatial, or aesthetic properties – or when already established vocabulary exists to describe these properties.

P4a The less a novel idea is anchored in sensory perception, the more supportive abstraction and specification and, hence verbal text are for its typification.

Visual text communicates meaning in a distinctly different way. Whereas verbal text enables pure abstraction, visual text communicates typification through *materialization*. This is achieved by combining abstraction with an illustration of the attributes that anchor the novel idea in a specific materiality. For instance, the category of 'physician' may be depicted with typical props of the trade (e.g., a stethoscope) and in a typical setting, such as a hospital room. Variation in the spatial composition (e.g., by showing the person either in front of an operating table or behind an office desk) suggests different aspects of the category as most salient. In contrast to verbal text, visual text will therefore be more effective in conveying ideas as categories that center around particular material, spatial, or aesthetic features – 'entity categories' (Gentner & Kurtz, 2005) – rather than their relation to other categories. In the words of Nigam and Ocasio (2010), such ideas have a greater need for representation, i.e., making use of exemplars and illustration of specific properties of a category as resources that are provided for sensemaking.

P4b The more a novel idea is anchored in sensory perception, the more supportive materialization and, hence, visual text is for its typification.

*Explanation*

In addition to typification, the theorization of a novel idea also involves the creation of particular relationships, such as concrete chains of cause and effect (Strang & Meyer, 1993). Whereas initial placement in pre-institutionalization articulates a vision in the form of a desired outcome, explanation seeks to cast plausibly and compellingly the novel idea as a solution to a common problem (e.g., Greenwood *et al.*, 2002; Meyer, 2008). For example, the

theorization of Shareholder Value not only addresses the concerns of particular types of companies (publicly listed corporations), but also explains how and why (focusing on Discounted Free Cash Flows, introducing Stock Option Plans, etc.) this is accomplished (Meyer & Höller, 2010). Accordingly, the challenge for text producers is to create text that adequately conveys the complexity of the problem-solution link, i.e., that communicates the causalities implied by the novel idea in a way that is comprehensible for targeted audiences.

Verbal text excels in the construction of explications and formal arguments, an affordance that has already been recognized in rhetorical approaches to institutionalization (e.g., Green *et al.*, 2009; Harmon *et al.*, 2015). The verbal affordance to *argue* the connection between problems and solutions according to established rules (e.g., Toulmin, 2003) cannot be easily mimicked by other semiotic modes (but see Birdsell & Groarke, 1996; Blair, 1996 on the idea of visual arguments). In particular, institutional theory has highlighted the role of explanatory accounts as reasoning devices that help develop and stabilize a novel idea over the course of institutionalization. Such accounts are building blocks of a potential argument and become quasi 'ready-made' when ideas are fully institutionalized (Creed, Scully, and Austin 2002; Meyer, 2014; Meyer & Rowan, 1977). As Scott and Lyman (1970: 107) emphasize, an account is likely to be comprehensible and accepted "(t)o the extent that everyone in the audience to whom an account is given shares a common universe of discourse and a common basis of beliefs". The limitation of verbal explanation is that every meaning system deems relevant only particular premises and forms of backing claims (e.g., Sillince, 1999; Toulmin, 2003). This means, for instance, that 'higher efficiency' may be a valid backing in an economic context but not in matters of love and family. Verbal arguments are, accordingly, more salient when the field already provides the resources for plausible arguments, and when audiences are likely to understand and share the suggested causal links.

P5a The higher the match between idea and shared understandings in the field, the more supportive argument and, hence, verbal text is for explanation.

In contrast, visual images are able to retain the multidimensionality of complex relationships and reduce complexity through their *spatial* character. This key affordance of visual text is commonly utilized in academia, where elaborate graphs and diagrams are employed in order to explain complex relationships in articles and public talks (e.g., Rowley-Jolivet, 2004; Swedberg, 2016). Since visual text is based on juxtaposition and holistic integration rather than formal logic, it can establish plausibility without recourse to formal rules. Juxtaposition can reduce complex technical arguments through visual associations; for instance, warning signs often visually juxtapose dangerous behavior and its consequences without providing the respective technical arguments. Although the rhetorical power of such simplified explanations is limited, they are particularly useful when the more fine-grained causalities have not yet been established in a field's stock of knowledge and/or when audiences cannot be expected to be knowledgeable enough to understand the technical arguments.

P5b The lower the match between idea and shared understandings in the field, the more supportive spatialization and, hence, visual text is for explanation.

*Justification*

In a similar way to initial placement, but more elaborately and reflectively, justification relies on the nesting and alignment of new ideas within prevailing normative prescriptions (Greenwood *et al.*, 2002). The challenge is to create legitimacy. Prestige and standing of the particular text producer in the field (e.g., Neidhardt, 1994) is key to determine whether visual or verbal text will be more effective for legitimizing a novel idea.

Many studies furthermore show the ability of verbal text to achieve justification. Rhetorical strategies for legitimation (e.g., Suddaby & Greenwood, 2005; Vaara & Tiennari, 2008; van Leeuwen & Wodak, 1999) draw on the affordances of verbal text related to *arguing*

and *narrating*. However, such justifications are more likely to be accepted and to support institutionalization when they are created and diffused by legitimated theorists (Strang & Meyer, 1993), such as acknowledged experts (e.g., Phillips *et al.*, 2004; Zald & Lounsbury, 2010). Lefsrud and Meyer (2012), in their study on the framings of professional experts, for instance, have shown how science is used to back claims regarding climate change. They also show that delegitimation is often required to demolish the opponent's status as a scientist. We therefore suggest that the potential of verbal text to justify unfolds primarily when used by prestigious text producers.

P6a The higher the standing of text producers, the more supportive argument and narration and, hence, verbal text are for the justification of a novel idea.

The visual mode, on the other hand, provides two distinct affordances that help alleviate a lesser prestige of text producers in the field. First, visual text is able to *materialize* certain pragmatic outcomes of the novel idea as a fact, and thereby detach justification from the author (e.g., Graves, Flesher, & Jordan, 1996; Jones, Meyer, Jancsary, & Höllerer, 2017). Second, through *spatializing*, visual text implicitly establishes relationships through composition. Positioning in close proximity to highly legitimate persons or objects enables spillovers of their legitimacy (e.g., Haack, Pfarrer, & Scherer, 2014). For instance, visual 'testimonials' by highly reputed experts, or the linking of ideas to widely shared values can compensate a lower standing of the text producer. In their study on CSR in Austria, Höllerer *et al.* (2013) found that visual advocacy and testimonials in favour of the concept were created, for instance, by visualizing adherence to the UN Global Compact with an image of the then UN Secretary-General Kofi Annan.

P6b The lower the standing of text producers, the more supportive materialization and spatialization and, hence, visual text are for the justification of a novel idea.

## Full institutionalization

Finally, if the process of institutionalization is successful, the idea becomes fully institutionalized. During full institutionalization, ideas integrate cognitive, structural, and behavioral elements that are sufficiently generic and legitimate to remain enacted within a field. The main requirements in this stage are the deepening of the taken-for-grantedness of an idea, and its further diffusion and translation. When sedimentation into the social stock of knowledge is accomplished, scrutiny declines, and ideas become *taken-for-granted* and appear to be given objectively. In order to persist, institutionalized ideas must be transmitted to new generations (e.g., Berger & Luckmann, 1967) as well as broadly *diffused* or *translated* to new contexts.

*Taken-for-grantedness*

Once institutionalized, ideas are perceived as 'exterior' to any subjective reasoning and are able to endure because they are taken for granted. However, taken-for-grantedness is a matter of degree rather than a discrete state, and social knowledge is unequally distributed within any field. Consequently, literature suggests that the continued existence of an idea is not guaranteed, but depends on institutional maintenance, even if it is supported by powerful actors (e.g., Lawrence, Suddaby, & Leca, 2009; Zilber, 2009). The lower the degree of taken-for-grantedness of an idea within a field, or the more it is unequally distributed across sub-communities, the greater the potential is for contestation to flare up even at this late stage of institutionalization.

The role of verbal text with regard to taken-for-grantedness is by the same token inherently precarious. When deeply ingrained, ideas manifest themselves in practices without discussion, which results in an absence of debates when ideas are fully institutionalized (e.g., Green, 2004; Harmon et al., 2015). Discursive engagement with a fully institutionalized idea opens a window of opportunity for contention, one that may trigger a process of de-institutionalization (Berger & Kellner, 1984). However, as long as the idea remains shared and uncontested within the field, its vocabulary is repeated in an unquestioned manner. In such situations, *abstraction* helps build fundamental legitimating 'explanations' directly into the vocabulary (Berger & Luckmann, 1967) and can act as a way of normalization (e.g., Lawrence et al., 2001).

P7a The lower the potential for contestation of an idea among audiences, the more supportive abstraction and, hence, verbal text is for maintaining the taken-for-grantedness of fully institutionalized ideas.

The use of visual text, in contrast, enables the application of knowledge and information that is either impossible (for instance, because it is pre-reflective) or forbidden (for instance, due to existing taboos) to verbalize. This affordance is one of *infiltrating*. In addition, visual text can illustrate the enactment (i.e., reproduction) of institutions without it being rendered explicit. It thereby supports the maintenance of institutions as taken-for-granted 'background programs' (Berger & Kellner, 1984), especially when verbalization would bear the risk of triggering debate. To illustrate, the simple depiction of a leadership team engaged in a strategy meeting, for instance, may reify ideas about gender, ethnicity, and hierarchy without explicitly pointing to them.

P7b The higher the potential for contestation of an idea among audiences, the more supportive infiltration and, hence, visual text is for maintaining the taken-for-grantedness of fully institutionalized ideas.

*Diffusion and translation*

A second core characteristic of full institutionalization is sedimentation, which occurs when theorized ideas become part of the social stock of knowledge. It ensures that ideas and knowledge are passed down through generations as objective social facts. Such transmission is accompanied by legitimation, which means the use of more or less elaborate communication that reifies the idea. The most basic level of legitimation – what Berger and Luckmann (1967: 112) call "incipient legitimation" – concerns the simple assertion that 'this is how things are'. Scholars emphasize that it is not practices that diffuse, but actually the abstracted and theorized models that underpin them (e.g., Strang & Meyer, 1993; Strang and Soule, 1998; Wedlin & Sahlin, 2017). When institutionalized ideas diffuse across field boundaries, their meanings and theorizations need to be 'translated' for the ideas to gain legitimacy locally (see e.g., Boxenbaum, 2006; Czarniawska & Joerges 1996; Sahlin-Andersson 1996; Meyer and Höllerer 2010). The central challenge for diffusion and translation, accordingly, is to make institutionalized ideas in their abstracted form understandable, acceptable and desirable to multiple new audiences. This process hinges most importantly on the ability of texts to connect with new audiences, i.e., to present an idea through signs that are comprehensible for them.

Without doubt, verbal text is more precise than visual text in transmitting institutionalized ideas to new audiences across time and space. Through *abstracting* ideas from material and sensory reality, verbal text creates labels that can achieve object-like status, hence facilitating the ability of ideas to travel (e.g., Czarniawska & Joerges, 1996). For example, labels like 'corporation', 'internet', or 'Shareholder Value' have achieved such a status. The establishment of a shared label for 'brute facts' is a prerequisite for the diffusion of novel ideas (Li, 2016). The ability of verbal text to *narrate* contributes to legitimation by embedding ideas in 'meta-narratives' (e.g. Zilber, 2009), 'editing' success stories (Sahlin-

Andersson, 1996), (re-)contextualizing particular rational myths (e.g., Zilber, 2006), and creating broader story lines that support discourse coalitions (Meyer & Höllerer, 2010). However, the comprehensibility of verbal narratives requires a certain match in vocabularies (i.e., the availability of the same specialized verbal language). If vocabularies differ widely across audiences, the resonance of narratives accordingly suffers.

P8a The greater the match of vocabularies between contexts and audiences, the more supportive abstraction and narration and, hence, verbal text are for the transmission and diffusion of fully institutionalized ideas.

Turning to visual text, one of its distinct communicative advantages is its ability to naturalize social reality and to reify ideas as material facts (*materialize*). By continuously materializing an idea, visual text facilitates transmission and reinforces the inevitability of institutions. Literature claims that visual text has the potential to emerge as a kind of global language, because it is easily conflated with objective reality (Machin, 2004). This ability to directly connect to the concrete materiality of the local life-world makes visual text uniquely suited to help ideas travel across time and space, continually reaching new audiences. For instance, on the basis of an analysis of the Getty image bank, Machin (2004: 334) exemplifies how rather abstract ideas such as 'professionally successful women' diffuse in the global corporate world through imageries in which "women in white-light diffused abstracted rooms speak on telephones or laugh with other women while pointing at computer screens". At the same time, visual text facilitates the (re-)contextualization of institutionalized ideas by imbuing such images with a distinct local flavor (e.g., Höllerer *et al.*, 2013).

P8b The lower the match of vocabularies between contexts and audiences, the more supportive materialization and, hence, visual text is for the transmission and diffusion of fully institutionalized ideas.

Figure 2 summarizes our theory development by schematically depicting the role of verbal and visual affordances in the accomplishment of specific requirements for each stage of institutionalization in distinct communicative situations.

---

Insert Figure 2 about here

---

## DISCUSSION AND CONCLUSION

The central aim of this article has been to develop a more systematic account of the respective impact of the verbal and visual mode in the process of institutionalization. In more detail, we addressed the question of how, and under which conditions, the use of verbal text, respectively visual text, has the highest potential to facilitate the institutionalization of novel ideas. Research on the performative role of communication has, for some time, acknowledged that texts rely on a variety of semiotic modes of communication. However, their specific generativity (e.g., Kress, 2010; Kress & van Leeuwen, 2001) is, so far, rather understudied. Instead, it has more or less implicitly been assumed that different semiotic modes play quite similar roles in institutional processes. Our work challenges such an assumption and extends previous insights into institutionalization as a communicative accomplishment. Firstly, we disentangle the central parameters: the communicative requirements at different stages of institutionalization, the characteristics of the communicative situation, and different semiotic modes. Secondly, we contribute to a better understanding of institutions as inherently multimodal achievements by systematizing the specific features and affordances of the verbal mode and the visual mode, and their respective potential for advancing the process of institutionalization. Finally, our theory development also has implications for research on institutional emergence and innovation in as much as it opens a new fruitful path to examine processes of meaning making.

## Institutionalization as communicative accomplishment

Previous research has stressed that institutions are communicative (e.g., Cornelissen *et al*., 2015; Luckmann, 2006), discursive (e.g., Phillips *et al.,* 2004; Hardy & Maguire, 2010), and rhetorical (e.g., Green & Li, 2011; Sillince and Barker, 2012) achievements, and that the production and dissemination of texts is central to the process of institutionalization. A first major contribution of our work is to substantially extend such insights by specifying how different semiotic modes of communication impact on processes and stages of institutionalization.

The model of discursive institutionalization proposed by Phillips *et al.* (2004) is a milestone of discursive institutionalism and serves as an excellent point of departure to explicate how we extend existing knowledge. Figure 3 visually represents our contribution. Elements that have already been discussed extensively are de-emphasized through grey color whereas our extensions and specifications are highlighted in black. The model suggests a mutually constitutive relationship between text, discourse, and institution. We focus on and further specify particular components of the overall model and also introduce new components that have so far been neglected. First and foremost, we add the insight that texts may draw from different semiotic modes, each of which provides distinct affordances for the institutionalization of new ideas. Second, we integrate the stage model of institutionalization and suggest that texts have to meet different communicative requirements at different stages of institutionalization in order to facilitate the process of institutionalizing new ideas. Third, and finally, we outline how specific characteristics of the communicative situation call for certain affordances, and hence shape which semiotic mode is most relevant for meeting the requirements at a particular stage of institutionalization.

---

Insert Figure 3 about here

Although our extensions are aligned with the theorectical insights provided by existing literature, they make a fundamentally novel contribution to this literature in the form of adding an entirely new modality for communication. Not only do we differentiate between semiotic modes, but we also differentiate discursive aspects of institutionalization according to the particular stages of institutionalization. Scholars have already begun to pursue such an agenda. Previous work suggests, for instance, that the role and the form of argument, or rhetorical strategies more generally (e.g., Green *et al.*, 2009), change during the course of institutionalization, that the process of institutionalization consists of linked signification processes (Li, 2016), and that different rhetorical tropes are prevalent in specific stages of institutionalization (e.g., Sillince & Barker, 2012). We extend these emerging insights and formulate a theoretical account of how distinct semiotic modes of communication impact differently on the process of institutionalization. Most importantly, our theorizing provides a more fine-grained treatment of this process than the simple specification of which textual characteristics matter at progressive stages of institutionalization. This contribution consists essentially of proposing that texts impact on institutionalization processes by means of interactions among (a) the semiotic mode on which a text draws, (b) the particular stage of institutionalization, (c) the requirement that needs to be accomplished at each stage, and (d) the specific communicative situation in which text production (and reception) takes place.

These extensions enable more systematic insights into the role and effects of texts, but also prompt future research to delve deeper into specific characteristics of the communicative situation. A novel idea may, for instance, be borrowed from another field, or be developed top-down or in a bottom-up process. It can have individual or collective 'authorship', diffuse through structural or cultural bases (Strang & Soule, 1998), and have very different areas of application. These additional characteristics of novel ideas may also influence the respective relevance of different semiotic modes. Future research may therefore consider how different

cultural and institutional contexts impact on the use and effects of verbal and visual text respectively. Semiotic modes are bound to particular historical and cultural conditions (Kress & van Leeuwen, 2006), and we have defined their cultural features according to the Western context. In addition, vocabularies (Loewensten *et al*., 2012) and 'visual codes' (Jones *et al.*, 2017) are distinct to specific institutional domains or value spheres. We encourage research that empirically compares the distribution, use, and interaction of semiotic modes across various cultural and institutional contexts.

Somewhat relatedly, different communication arenas need to be studied in more depth, e.g., arenas where visual text is common (like the news media; e.g., Gamson, 1992) as compared to arenas in which verbal text is predominant and visual text is rather frowned upon or downright prohibited (like in courtrooms; e.g., Suddaby & Greenwood, 2005). In line with further specifying communicative situations, we also suggest a more systematic engagement with audiences. Our performative framework based on affordances assumes that audiences respond somewhat similarly to the affordances in a given text. A fruitful line for future research would be to further 'unpack' audience responses in line with more recent developments in communicative institutionalism (e.g., Cornelissen et al., 2015). Such engagement would bring to the fore more contextualized elements of visual literacy, cultural features, uncertainty, and other aspects of sensemaking, all of which may influence how audiences respond to specific affordances.

Although we focus on how both verbal and visual text can positively impact the successful progression between stages of institutionalization, we also lay the groundwork for future research to explore how verbal and visual affordances may hinder institutionalization, or contribute to deinstitutionalization (e.g., Maguire & Hardy, 2009; Brown, Ainsworth, & Grant, 2012).

## Institutionalization and multimodality

A second contribution stresses that not only are institutions communicative, discursive, and rhetorical achievements, but also, in most cases, multimodal accomplishments, meaning that they are constructed, enacted, reproduced and altered, using a multiplicity of semiotic modes simultaneously. According to social semiotics, the relevance of multimodality for institutionalization lies in the fact that each mode offers a distinct potential for constructing and representing social reality (e.g., Kress, 2010; Kress & van Leeuwen, 2001). Research on institutions has started to explore the differentiated impact of semiotic modes on institutions and their dynamics (e.g., Höllerer, Daudigeos, & Jancsary, forthcoming). We contribute to this field of research by systematizing the role of different modes in processes of institutionalization. Drawing from social semiotics, we develop a set of constitutive features for the verbal mode and the visual mode, which allows for a systematic comparison between them. We further outline how these constitutive features endow texts with particular affordances that impact institutionalization in distinct ways.

Our conceptualization of institutions as multimodal achievements also points at 'blind spots' in current institutional theory. Existing studies largely neglect that social meaning is constructed, maintained, and transformed on the basis of multiple sign systems, and that these sign systems interact in various ways. For instance, while the role of verbal text in early stages of institutionalization is well documented, our theory development shows the relevance of visual text. Visual text not only facilitates experimentation and meaning alignment, but also allows marginal voices to join the conversation. During full institutionalization, explicit verbal discussions may cease (e.g., Green, 2004) while visual traces of institutionalized ideas remain and may constitute mechanisms of inconspicuous institutional maintenance.

Our arguments in this article are – necessarily – foundational and abstract; they have, however, considerable implications for a wide range of core concepts in institutional theory. Core concepts that draw on multiple semiotic modes, and which can therefore gain from a better articulation of a multimodal perspective, include identity-work (e.g., Glynn, 2017),

legitimation strategies (e.g., Lefsrud & Meyer, 2012; Vaara & Tienari, 2011), framing (e.g., Cornelissen & Werner, 2014; Meyer & Höllerer, 2010), and struggles for 'truth' (e.g. Lefsrud, Graves, & Phillips, 2013). Many other theoretical constructs can also benefit from our conceptual work in as much as institutional orders are not only characterized by specific verbal vocabularies, but also by registers provided by other semiotic modes (e.g., Jancsary, Meyer, Höllerer, & Boxenbaum, forthcoming). Conversely, a lack of attention to non-verbal sign systems, or attention to only one semiotic mode, increases the risk that we reconstruct social meaning incompletely or even inadequately, and that we miss important mechanisms and processes of institutionalization.

Our framework furthermore facilitates a more thorough understanding of multimodality in several ways. First, we link affordances of each mode to characteristics of the communicative situation, which is a prerequisite for analysing interactions between modes. In concrete situations, such characteristics may be combined in complex ways. For instance, a text producer with little visibility may put forward a radically novel idea that cannot build on any shared knowledge within the field, but that suggests a very clear and measurable outcome (for instance, a major technological advancement). Two of these situational characteristics suggest the use of visual communication to meet requirements at the pre-institutionalization stage, whereas one situational characteristic calls for the use of verbal text. Although it seems at first sight that a combination of the two semiotic modes could easily meet the different requirements, there is actually, at present, no systematic research on the dynamic interaction between semiotic modes. It remains to be studied whether (or when) verbal and visual affordances are additive and able to support each other, as opposed to being neutral or even contradictory to each other (see e.g., Höllerer *et al.*, 2017). There is also merit in examining when the mix of modes in multimodal communication gives rise to new affordances, and how the communicative situation impacts on this process. Such questions require detailed theoretical and empirical work in the coming years.

Second, our characterization of modes by their constitutive features and the associated affordances that they inscribe in texts may enable future research to explore the role of other semiotic modes than the verbal or visual in institutionalization. Pioneering research has started to engage with modes such as sound (e.g., Pinch & Bijsterveld, 2012) or scent (e.g., Islam, Endrissat, & Noppeney, 2016). Our systematic treatment of semiotic modes according to their core features and affordances can facilitate the integration of a greater variety of communicative and semiotic resources in empirical research, as well as in theory development.

Finally, our multimodal understanding of institutions also has substantial methodological implications. A large part of methodological tools available for institutional inquiry – although broad in scope – are tailored to the analysis of verbal text. The specificities of visual sign systems – and those of other modes – require equally sophisticated methods of analysis. Literature in the field of discourse analysis provides researchers with valuable insights on how to study different modes (e.g., Jancsary, Höllerer, & Meyer, 2016; Machin & Mayr, 2012). Additionally, literature in social semiotics (e.g., Kress & van Leeuwen, 2006) offers more in-depth insights into the specific 'grammar' of the visual, insights that are conducive to institutional analysis. Ultimately, multimodal institutional analysis constitutes an emerging field of research with a substantial potential – and necessity – for systematic methodological development.

## Innovation and institutional emergence

Our theory development extends, as mentioned, primarily insights on processes of institutionalization by articulating the respective affordances of the verbal and visual modes of communication at different stages of institutionalization. However, our work also has implications for institutional emergence more broadly.

*Optimal distinctiveness*. We suggest that a differentiation of modes is of significant interest for research on the optimal distinctiveness of novel ideas (e.g., Alvarez, Mazza, Strandgaard Pedersen, & Svejenova, 2005). Navis and Glynn (2010) find that once a new category has achieved legitimacy, organizations within that category shift their focus to achieving individual distinctiveness, rather than conformity to the category. Studies on management fashion (e.g., Heusinkveld, Benders, & Hillebrand, 2013) show that ideas are often rebranded by using new names for old ideas, or broadened by including additional ideas under the same label. Research on innovation stresses the substantial role of metaphor and bricolage in the exercise of creativity, suggesting that the 'layering' of multiple metaphors into complex theoretical constructs supports the inception of new theory (e.g., Boxenbaum & Rouleau, 2011). The acknowledgement of different semiotic modes offers an interesting new lens for studying the delicate balance between uniqueness and isomorphic adaptation when introducing novel ideas. Our insights are strongly sympathetic to studies on technological innovation that suggest aesthetic innovation as a possible avenue for achieving distinction when technical innovation is unfeasible (e.g., Eisenman, 2013), or aesthetic similarity as a way to enhance legitimacy when technical innovation breaks with highly prevalent norms (e.g., Hargadon & Douglas, 2001). In more detail, we suggest that the affordances of different semiotic modes provide specialized resources for creating similarity or uniqueness, and often simultaneously. For instance, the spatial arrangement of information in the visual mode provides resources for bricolage that are quite different from the linear and sequential organization of the verbal mode. A similar pattern pertains to the holistic presentation of elements that the visual mode affords. Undoubtedly, further specifications of the multimodal communication of optimal distinctiveness would be fruitful as an avenue for future research.

*Sources and spaces of innovation.* An important question in more recent institutional theory is how innovations come to be, and where they emerge (e.g., Padgett & Powell, 2012). Empirical research has shown that novel ideas triggering institutional change sometimes stem

from actors at the center of their field, such as powerful corporations (e.g., Munir & Phillips, 2005), and at other times from more peripheral sources (e.g., Khaire & Wadhwani, 2010). Our theory development contributes to reconciling these seemingly inconsistent findings by suggesting that the structural position of innovators may be linked to the use of specific semiotic modes of communication. For instance, we have argued that visual text may help text producers in less favorable field positions achieve both exposure for, and justification of, their novel ideas; in contrast, verbal text is potentially more useful for highly prominent and prestigious text producers in that it allows them to communicate their ideas with more precision and sophistication.

The insights we have produced are also relevant for more agentic approaches to institutions. Since meaning is always negotiated and co-constituted in interaction with audiences, the characteristics of different semiotic modes can shed new light on how actors contribute to relevant processes, such as institutional work (e.g., Lawrence *et al.*, 2011), institutional entrepreneurship (e.g., Battilana, Leca, & Boxenbaum, 2009; Greenwood & Suddaby, 2006), cultural entrepreneurship (Lounsbury & Glynn, 2001), or institutional defense (Lefsrud & Meyer, 2012). Future research may also integrate the analysis of multiple semiotic modes with the recently emerging literature on emotions in institutional theory (e.g., Voronov & Vince, 2012; Creed, Hudson, Okhuysen, & Smith-Crowe, 2014). As in this instance, attentiveness to semiotic modes and their textual affordances may illuminate not only institutional dynamics but also a much broader range of organizational phenomena.

## Conclusion

Institutions are not only communicative achievements, they are also multimodal in nature. The ever broadening accessibility of visual and digital forms of expression render the construction, maintenance, and transformation of institutions increasingly diverse in terms of the semiotic resources at play. Institutional research needs to integrate this multimodal

diversity into conceptual frameworks and methodological designs. Without multimodal 'literacy' and the corresponding analytical competency on the part of researchers, institutional inquiry is prone to missing substantial aspects of how institutions emerge, solidify, and eventually become sedimented. In this article, we have provided a foundation for further engagement with multimodal forms of communication by elaborating how verbal and visual forms of expression play different – and potentially complementary – roles in the institutionalization of novel ideas. In doing so, we not only enhance institutional inquiry but also join others in developing a genuinely multimodal research agenda for organization theory, the prospects of which are very promising indeed.

## REFERENCES


Alvarez, J. L., Mazza, C., Strandgaard Pedersen, J., & Svejenova, S. 2005. Shielding idiosyncrasy from isomorphic pressures: Towards optimal distinctiveness in European filmmaking. ***Organization***, 12(6): 863-888.

Ashcraft, K. L., Kuhn, T. R., & Cooren, F. 2009. Constitutional amendments: "Materializing" organizational communication. ***The Academy of Management Annals***, 3: 1–64.

Atkinson, R. C., & Shiffrin, R. M. 1968. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), ***The Psychology of Learning and Motivation (vol. 2)***: 89–195. New York: Academic Press.

Barley, S. R., & Tolbert, P. S. 1997. Institutionalization and structuration: Studying the links between action and institution. ***Organization Studies***, 18(1): 93–117.

Barthes, R. 1991. ***Mythologies.*** New York: The Noonday Press.

Barthes, R. 1977. ***Image, music, text***. London, UK: Fontana Press.

Battilana, J., Leca, B., & Boxenbaum, E. 2009. How actors change institutions: Towards a theory of institutional entrepreneurship. ***Academy of Management Annals***, 3: 65–107.

Bell, E., Warren, S., & Schroeder, J. E. (Eds.). 2014. ***The Routledge companion to visual organization.*** Abingdon and New York: Routledge.

Berger, P. L., & Kellner, H. 1984. ***Für eine neue Soziologie***. Frankfurt am Main, Germany: Fischer.

Berger, P. L., & Luckmann, T. 1967. ***The social construction of reality: A treatise in the sociology of knowledge.*** New York: Anchor Books.

Blair, J. A. 1996. The possibility and actuality of visual arguments. ***Argumentation and Advocacy***, 33(1): 23-39.

Birdsell, D. S., & Groarke, L. 1996. Toward a theory of visual argument. ***Argumentation and Advocacy***, 33(1): 1-10.

Bloch, P. H. 1995. Seeking the ideal form: Product design and consumer response. ***Journal of Marketing***, 59: 16-29.

Boxenbaum, E. 2006. Lost in translation? The making of Danish diversity management. ***American Behavioral Scientist***, 49(7): 939-948.

Boxenbaum, E., & Rouleau, L. 2011. New knowledge products as bricolage: Metaphors and scripts in organizational theory. ***Academy of Management Review***, 36(2): 272-296.

Brown, A. D. Ainsworth. S., & Grant, D. 2012. The rhetoric of institutional change. ***Organization Studies***, 33(3): 297-321.

Coltheart, M. 1980. Iconic memory and visible persistence. ***Perception & Psychophysics***, 27 (3): 183–228.

Cooren, F., Kuhn, T., Cornelissen, J. P., & Clark, T. 2011. Communication, organizing and organization: An overview and introduction to the special issue. ***Organization Studies***, 32: 1149–1170.

Cornelissen, J. P. 2012. Sensemaking under pressure: The influence of professional roles and social accountability on the creation of sense. ***Organization Science,*** 23(1): 118-137.

Cornelissen, J. P., Durand, R., Fiss, P. C., Lammers, J. C., & Vaara, E. 2015. Putting communication front and center in institutional theory and analysis. ***Academy of Management Review***, 40(1): 10-27.

Cornelissen, J. P., Holt, R., & Zundel, M. 2011. The role of analogy and metaphor in the framing and legitimization of strategic change. ***Organization Studies***, 32(12): 1701-1716.

Cornelissen, J. P., & Werner, M. D. 2014. Putting Framing in Perspective: A Review of Framing and Frame Analysis across the Management and Organizational Literature. ***The Academy of Management Annals,*** 8(1): 181-235.

Cowan, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. ***Behavioral and Brain Sciences***, 24: 87–185.

Creed, W. E. D., Hudson, B. A., Okhuysen, G. A., & Smith-Crowe, K. 2014. Swimming in a sea of shame: incorporating emotion into explanations of institutional reproduction and change. ***Academy of Management Review***, 39(3): 275-301.

Creed, W. E. D., Scully, M. A., & Austin, J. R. 2002. Clothes make the person? The tailoring of legitimating accounts and the social construction of identity. ***Organization Science,*** 13(5): 475-496.

Czarniawska, B., & Joerges, B. 1996. Travel of ideas. In B. Czarniawska & G. Sevón (Eds.), ***Translating organizational change***: 13-48. Berlin, Germany, New York, NY: de Gruyter.

Dogra, N. 2007. 'Reading NGOs visually'—Implications of visual images for NGO management. ***Journal of International Development***, 19(2): 161-171.

Drori, G.S. 2006. Governed by governance: The institutionalization of governance as a prism for organizational change. In G. S. Drori, J. W. Meyer, H. Hwang (Eds.), ***Globalization and***

***organization: World society and organizational change***: 91-118. Oxford: Oxford University Press.

Durand, R., & Paolella, L. 2013. Category stretching: Reorienting research on categories in strategy, entrepreneurship, and organization theory. ***Journal of Management Studies***, 50(6): 1100-1123.

Eco, U. 1995. ***The search for the perfect language.*** Oxford: Oxford University Press.

Edell, J. A., & Staelin, R. 1983. The information processing of pictures in print advertisements. ***Journal of Consumer Research***, 10(1): 45–61.

Eisenman, M. 2013. Understanding aesthetic innovation in the context of technological evolution. ***Academy of Management Review***, 38(3): 332-351.

Etzion, D., & Ferraro, F. 2010. The role of analogy in the institutionalization of sustainability reporting. ***Organization Science***, 21(5): 1092-1107.

Foster, W. M., Suddaby, R., Minkus, A., & Wiebe, E. 2011. History as social memory assets: The example of Tim Hortons. ***Management & Organizational History,*** 6(1): 101-120.

Gamson, W. A. 1992. ***Talking politics***. Cambridge, UK, New York, NY, Melbourne, Australia: Cambridge University Press.

Garud, R., Gehman, J. & Giuliani, A. P. 2014. Contextualizing entrepreneurial innovation: A narrative perspective. ***Research Policy***, 43(7): 1177-1188.

Garud, R., Schildt, H. A., & Lant, T. K. 2014. Entrepreneurial storytelling, future expectations, and the paradox of legitimacy. ***Organization Science***, 25(5): 1479-1492.

Gentner, D., & Kurtz, K. 2005. Relational Categories. In B. C. Goldstone, A. B. Markman & P. W. Wolff (Eds.), ***Categorization Inside and Outside the Lab***: 151-175. Washington DC: APA.

Gibson, J. J. 1986. ***The ecological approach to visual perception.*** New York and Hove: Psychology Press.

Glynn, M. A. 2017. Theorizing the Identity–Institution Relationship: Considering Identity as Antecedent to, Consequence of, and Mechanism for, Processes of Institutional Change In R. Greenwood, C. Oliver, R. E. Meyer, & T. Lawrence (Eds.), ***The SAGE Handbook of Organizational Institutionalism*** (2nd ed.). Los Angeles et al.: Sage Publications.

Gondo, M. B., & Amis, J. M. 2013. Variations in practice adoption: The roles of conscious reflection and discourse. ***Academy of Management Review***, 38(2): 229-247.

Graves, O. F., Flesher, D. L., & Jordan, R. E. 1996. Pictures and the bottom line: The television epistemology of U.S. annual reports. ***Accounting, Organizations and Society***, 21(1): 57–88.

Gray, B., Purdy, J., & Ansari, S. 2015. From interactions to institutions: Microprocesses of framing and mechanisms for the structuring of institutional fields. ***Academy of Management Review***, 40(1): 115-143.

Green, S. E. 2004. A rhetorical theory of diffusion. ***Academy of Management Review***, 29(4): 653-669.

Green, S. E., & Li, Y. 2011. Rhetorical institutionalism: Language, agency, and structure in institutional theory since Alvesson 1993. ***Journal of Management Studies****, 48*(7): 1662-1697.

Green, S. E., Li, Y., & Nohria, N. 2009. Suspended in self-spun webs of significance: A rhetorical model of institutionalization and institutionally embedded agency. ***Academy of Management Journal***, 52(1): 11-36.

Greenwood, R., & Suddaby, R. 2006. Institutional entrepreneurship in mature fields: The Big Five accounting firms. ***Academy of Management Journal***, 49(1): 27-48.

Greenwood, R., Suddaby, R., & Hinings, C. 2002. Theorizing change: The role of professional associations in the transformation of institutionalized fields. ***Academy of Management Journal,*** 45(1): 58-80.

Haack, P., Pfarrer, M. D., & Scherer, A. G. 2014. Legitimacy-as-feeling: How affect leads to vertical legitimacy spillovers in transnational governance. ***Journal of Management Studies***, 51(4): 634-666.

Hardy, C., & Maguire, S. 2010. Discourse, field-configuring events, and change in organizations and institutional fields: Narratives of DDT and the Stockholm Convention. ***Academy of Management Journal***, 53(6): 1365-1392.

Halliday, M. A. & Hasan, R. 1989. ***Language, context, and text: Aspects of language in a social-semiotic perspective***. Oxford, NY: Oxford University Press.

Hargadon, A. B., & Douglas, Y. 2001.When innovations meet institutions: Edison and the design of the electric light. ***Administrative Science Quarterly***, 46(3): 476-501.

Harmon, D. J., Green, S. E., Jr., & Goodnight, G. T. 2015. A model of rhetorical legitimation: The structure of communication and cognition underlying institutional maintenance and change. ***Academy of Management Review***, 40: 76–95.

Haveman, Heather A. 1993. Follow the leader: Mimetic isomorphism and entry into new markets. ***Administrative Science Quarterly***, 38(4): 593-627.

Heusinkveld, S., Benders, J. & Hillebrand, B. 2013. Stretching concepts: The role of competing pressures and decoupling in the evolution of organization concepts. ***Organization Studies***, 43(1): 7-32.

Hill, C. A. 2004. The psychology of rhetorical images. In C. A. Hill & M. Helmers (Eds.), ***Defining visual rhetorics:*** 25–40. New York: Routledge.

Höllerer, M. A., Daudigeos, T., & Jancsary, D. (Eds.). forthcoming. ***Multimodality, Meaning, and Institutions*** (Research in the Sociology of Organizations). Bingley: Emerald.

Höllerer, M. A., Jancsary, D., & Grafström M. 2017. 'A picture is worth a thousand words': Multimodal sensemaking of the Global Financial Crisis. ***Working paper, WU Vienna University of Economics and Business***.

Höllerer, M. A., Jancsary, D., Meyer, R. E., & Vettori, O. 2013. Imageries of corporate social responsibility: Visual re-contextualization and field-level meaning. ***Research in the Sociology of Organizations***, 39(B): 139-174.

Hutchby, I. 2001. Technologies, texts and affordances. ***Sociology,*** 35(2): 441-456.

Islam, G., Endrissat, N., & Noppeney, C. 2016. Beyond 'the Eye'of the Beholder: Scent innovation through analogical reconfiguration. ***Organization Studies,*** 37(6): 769-795.

Jancsary, D., Höllerer, M. A., & Meyer, R. E. 2016. Critical analysis of visual and multimodal texts. In R. Wodak & M. Meyer (Eds.), ***Methods of Critical Discourse Studies*** (3rd ed., pp. 180-204). Los Angeles: Sage Publications.

Jancsary, D., Meyer, R. E., Höllerer, M. A., & Boxenbaum, E. forthcoming. Institutions as multimodal accomplishments: Towards the analysis of visual registers. ***Research in the Sociology of Organizations.***

Jewitt, C., & Oyama, R. 2001. Visual meaning: A social semiotic approach. In T. van Leeuwen, & C. Jewitt (Eds.), ***The handbook of visual analysis:*** 134-156. London *et al.*: SAGE Publications.

Jones, C., Meyer, R. E., Jancsary, D., & Höllerer, M. A. 2017. The material and visual basis of institutions. In R. Greenwood, C. Oliver, R. E. Meyer, & T. Lawrence (Eds.), ***The SAGE Handbook of Organizational Institutionalism*** (2nd ed.). Los Angeles et al.: Sage Publications.

Kaplan, S. 2011. Strategy and PowerPoint: An inquiry into the epistemic culture and machinery of strategy making. ***Organization Science***, 22(2): 320-346.

Kennedy, M. T., & Fiss, P. C. 2013. An ontological turn in categories research: From standards of legitimacy to evidence of actuality. ***Journal of Management Studies***, 50(6): 1138-1154.

Khaire, M., & Wadhwani, R. D. 2010. Changing landscapes: The construction of meaning and value in a new market category—Modern Indian Art. ***Academy of Management Journal***, 53(6): 1281-1304.

Kieser, A. 1997. Rhetoric and myth in management fashion. ***Organization,*** 4(1): 49-74.

Kress, G. 2010. ***Multimodality: A social semiotic approach to contemporary communication.*** Abingdon *et al.*: Routledge.

Kress, G., & Van Leeuwen, T. 2001. ***Multimodal discourse: the modes and media of contemporary communication.*** London: Hodder Education.

Kress, G., & van Leeuwen, T. 2006. ***Reading images: The grammar of visual design*** (2nd ed.). London *et al.*: Routledge.

Lawrence, T. B., Suddaby, R., & Leca, B. (Eds.) 2009. ***Institutional work: Actors and agency in institutional studies of organizations***. Cambridge, UK: Cambridge University Press.

Lawrence, T. B., Winn, M. I., & Jennings, P. D. 2001. The temporal dynamics of institutionalization. ***Academy of Management Review,*** 26(4): 624-644.

Lawrence, T. B., Suddaby, R., & Leca, B. 2011. Institutional work: Refocusing institutional studies of organization. ***Journal of Management Inquiry***, 20(1): 52-58.

Lefsrud, L., & Meyer, R. E. 2012. Science or science fiction? Professionals' discursive construction of climate change. ***Organization Studies,*** 33(11): 1477-1506.

Lefsrud, L., Graves, H., Phillips, N. 2013. Dirty oil, ethical oil: Categorical illegitimacy and the struggle over the Alberta oil sands. ***Academy of Management Proceedings***, 2013: 13283.

Li, Y. 2016. A semiotic theory of institutionalization. ***Academy of Management Review***, doi: 10.5465/amr.2014.0274

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. 2004. Eye movements when reading disappearing text: Is there a gap effect in reading? ***Vision Research***, 44(10): 1013–1024.

Loewenstein, J. 2014. Take my word for it: How professional vocabularies foster organizing. ***Journal of Professions and Organization*****,** 1(1): 65-83.

Loewenstein, J., Ocasio, W., & Jones, C. 2012. Vocabularies and vocabulary structure: A new approach linking categories, practices, and institutions. ***Academy of Management Annals***, 6: 41-86.

Lounsbury, M., & Glynn, M. A. 2001. Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. ***Strategic Management Journal***, 22(6/7): 545-564.

Luck, S. J., & Vogel, E. K. 1997. The capacity of visual working memory for features and conjunctions. ***Nature***, 390: 279–281.

Luckmann, T. 2006. Die kommunikative Konstruktion der Wirklichkeit. In D. Tänzler, H. Knoblauch, H-G. Soeffner (Eds.), ***Neue Perspektiven der Wissenssoziologie***: 15-26. Konstanz: UVK Verlagsgesellschaft.

Machin, D. 2004. Building the world's visual language: The increasing global importance of image banks in corporate media. ***Visual Communication***, 3(3): 316-336.

Machin, D., & Mayr, A. 2012. ***How to do critical discourse analysis: A multimodal introduction.*** Los Angeles et al.: Sage Publications.

Maguire, S., & Hardy, C. 2009. Discourse and deinstitutionalization: The decline of DDT. ***Academy of Management Journal,*** 52(1): 148-178.

McDonnell, T. E. 2010. Cultural objects as objects: Materiality, urban space, and the interpretation of AIDS campaigns in Accra, Ghana. ***American Journal of Sociology***, 115(6): 1800-1852.

McQuarrie, E. F., & Mick, D. G. 1992. On resonance: A critical pluralistic inquiry into advertising rhetoric. ***Journal of Consumer Research***, 19(2): 180-197.

McQuarrie, E. F., & Phillips, B. J. 2005. Indirect persuasion in advertising: How consumers process metaphors presented in pictures and words. ***Journal of Advertising***, 34(2): 7-20.

Messaris, P. 1997. ***Visual persuasion: The role of images in advertising.*** Thousand Oaks, CA: SAGE Publications.

Meyer, R. E. 2014. 'Re-localization' as micro-mobilization of consent and legitimacy. In G. S. Drori, M. A. Höllerer, & P. Walgenbach (Eds.), ***Global themes and local variations in organization and management: Perspectives on glocalization***: 79-89. New York, NY, Abingdon, UK: Routledge.

Meyer, R. E. 2008. New sociology of knowledge: Historical legacy and contributions to current debates in institutional research. In R. Greenwood, C. Oliver, K. Sahlin, & R. Suddaby (Eds.), ***The SAGE handbook of organizational institutionalism***: 519-538. London, UK, Thousand Oaks, CA, New Delhi, India, Singapore, Singapore: Sage Publications.

Meyer, R. E., & Höllerer, M. A. 2010. Meaning structures in a contested issue field: A topographic map of shareholder value in Austria. ***Academy of Management Journal***, 53(6): 1241-1262.

Meyer, R. E., Höllerer, M. A., Jancsary, D., & van Leeuwen, T. 2013. The visual dimension in organizing, organization, and organization research: Core ideas, current developments, and promising avenues. ***Academy of Management Annals***, 7: 489-555.

Meyer, J. W., & Rowan, B. 1977. Institutionalized organizations: Formal structure as myth and ceremony. ***American Journal of Sociology,*** 83(2): 340-363.

Mitchell, W. J. T. 1984. What is an image? ***New Literary History***, 15(3): 503-537.

Munir, K. A., & Phillips, N. 2005. The birth of the 'Kodak Moment': Institutional entrepreneurship and the adoption of new technologies. ***Organization Studies***, 26(11): 1665-1687.

Navis, C., & Glynn, M. A. 2010. How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990-2005. ***Administrative Science Quarterly***, 55(3): 439-471.

Neidhardt, Friedhelm 1994. Öffentlichkeit, öffentliche Meinung, soziale Bewegungen. In F. Neidhardt (Ed.), ***Öffentlichkeit, öffentliche Meinung, soziale Bewegungen, Kölner Zeitschrift für Soziologie und Sozialpsychologie***, Sonderheft 34: 7-41. Opladen: Westdeutscher Verlag.

Nigam, A., & Ocasio, W. 2010. Event attention, environmental sensemaking, and change in institutional logics: An inductive analysis of the effects of public attention to Clinton's Health Care Reform Initiative. ***Organization Science,*** 21(4): 823-841.

Ocasio, W. 2011. Attention to attention. ***Organization Science***, 22(5): 1286-1296.

Ocasio, W., & Joseph, J. 2005. Cultural adaptation and institutional change: The evolution of vocabularies of corporate governance, 1972–2003. ***Poetics,*** *33*(3-4): 163-178.

Ocasio, W., Loewenstein, J., & Nigam, A. 2015. How streams of communication reproduce and change institutional logics: The role of categories. ***Academy of Management Review***, 40(1): 28-48.

Padgett, J. F., & Powell, W. W. 2012. ***The emergence of organizations and markets.*** Princeton & Oxford: Princeton University Press.

Peirce, C. S. S. 1906. Prolegomena to an apology for pragmaticism. ***The Monist***, 16(4): 492-546.

Phillips, N., Lawrence, T. B., & Hardy, C. 2004. Discourse and institutions. ***Academy of Management Review***, 29(4): 635-652.

Pinch, T. & Bijsterveld. K. (Eds.). 2012. ***The Oxford Handbook of sound studies***. Oxford: Oxford University Press.

Purdy, J. M., & Gray, B. 2009. Conflicting logics, mechanisms of diffusion, and multilevel dynamics in emerging institutional fields. ***Academy of Management Journal***, 52(2): 355-380.

Puyou, F. R., Quattrone, P., McLean, C., & Thrift, N. (Eds.). ***Imagining organizations: Performative imagery in business and beyond:*** 1-18. New York: Routledge.

Raab, J. 2008. ***Visuelle Wissenssoziologie: Theoretische Konzeption und materiale Analysen.*** Konstanz: UVK.

Rowley-Jolivet, E. 2004. Different visions, different visuals: A social semiotic analysis of field-specific visual composition in scientific conference presentations. ***Visual Communication***, 3(2): 145-175.

Sahlin-Andersson, K. 1996. Imitating by editing success: The construction of organizational fields. In B. Czarniawska & G. Sevón (Eds.), ***Translating organizational change***: 69-92. Berlin, Germany, New York, NY: de Gruyter.

Schill, D. 2012. The visual image and the political image: A review of visual communication research in the field of political communication. ***Review of Communication***, 12(2): 118-142.

Scott, M. B., & Lyman, S. M. 1970. ***Accounts, deviance and social order***. New York, NY: Basic Books.

Seo, M.-G., & Creed, W. E. D. 2002. Institutional contradictions, praxis, and institutional change: A dialectical perspective. ***The Academy of Management Review***, 27(2): 222-247.

Shorrt, H. 2015. Liminality, space and the importance of 'transitory dwelling places' at work. ***Human Relations***, 68(4): 633-658.

Sillince, J. A. A. 1999. The organizational setting, use and institutionalization of argumentation repertoires. ***Journal of Management Studies***, 36(6): 795-830.

Sillince, J. A., & Barker, J. R. 2012. A tropological theory of institutionalization. ***Organization Studies***, 33(1): 7-38.

Sperling, G. 1960. The information available in brief visual presentations. ***Psychological Monographs: General and Applied,*** 74(11): 1–29.

Strang. D., & Meyer, J. W. 1993. Institutional conditions for diffusion. ***Theory and Society***, 22(4): 487-511.

Strang, D., & Soule, S. A. 1998. Diffusion in organizations and social movements: From hybrid corn to poison pills. ***Annual Review of Sociology***, 24(1): 265-290.

Styhre, A. 2010. Knowledge work and practices of seeing: Epistemologies of the eye, gaze, and professional vision. ***Culture and Organization***, 16(4): 361-376.

Suddaby, R., & Greenwood, R. 2005. Rhetorical strategies of legitimacy. ***Administrative Science Quarterly***, 50(1): 35–67.

Swedberg, R. 2016. Can You Visualize Theory? On the Use of Visual Thinking in Theory Pictures, Theorizing Diagrams, and Visual Sketches. ***Sociological Theory***, 34(3): 250-275.

Tolbert, P. S., & Zucker, L. G. 1996. The institutionalization of institutional theory. In: Clegg, S. R., Hardy, C., & Nord, W. R. (Eds.), ***Handbook of organization studies***: 175-190. London *et al.:* SAGE Publications.

Toraldo, M. L., Islam, G., & Mangia, G. 2016. Modes of knowing: Video research and the problem of elusive knowledges. ***Organizational Research Methods***, doi: 10.1177/1094428116657394.

Toulmin, S. E. 2003 [1958]. ***The uses of argument*** (updated edition). Cambridge: Cambridge University Press.

Vaara, E., Sonenshein, S., & Boje, D. 2016. Narratives as sources of stability and change in organizations: Approaches and directions for future research. ***The Academy of Management Annals***, 10(1): 495-560.

Vaara, E. & Tienari, J. 2011. On the narrative construction of multinational corporations: An antenarrative analysis of legitimation and resistance in a cross-border merger. ***Organization Science***, 22(2): 370-390.

van Leeuwen, T. 2007. Legitimation in discourse and communication. ***Discourse & Communication***, 1(1): 91-112.

van Leeuwen, T., & Wodak, R. 1999. Legitimizing immigration control: A discourse-historical analysis. ***Discourse Studies,*** 1(1): 83-118.

Voronov, M., & Vince, R. 2012. Integrating emotions into the analysis of institutional work. ***Academy of Management Review***, 37(1): 58-81.

Weber, K. 2005. A toolkit for analyzing corporate cultural toolkits. ***Poetics,*** 33: 227-252.

Wedlin, L., & Sahlin, K. 2017. The Imitation and Translation of Management Ideas. In R. Greenwood, C. Oliver, R. E. Meyer, & T. Lawrence (Eds.), ***The SAGE Handbook of Organizational Institutionalism*** (2nd ed.). Los Angeles et al.: Sage Publications.

Wodak, R. & Meyer, M. (Eds.) 2016. ***Critical discourse studies: History, agendy, theory and methodology***. London et al.: SAGE Publications.

Zald, M. N., & Lounsbury, M. 2010. The wizards of Oz: Towards an institutional approach to elites, expertise and command posts. ***Organization Studies***, 31(7): 963-996.

Zilber, T. 2006. The work of the symbolic in institutional processes: Translations of rational myths in Israeli high tech. ***Academy of Management Journal***, 49(2): 281-303.

Zilber, T. B. 2009. Institutional maintenance as narrative acts. In T. B. Lawrence, R. Suddaby, & B. Leca (Eds.), ***Institutional Work: Actors and Agency in Institutional Studies of Organizations***: 205-235. Cambridge, UK: Cambridge University Press.

TABLE 1

Constitutive features of the visual and the verbal mode

| | **Dimensions** | **Verbal** | **Visual** |
|---|---|---|---|
| **Semiotic features** | *Basis of signification* | symbol | icon, index, symbol |
| | *Structure of information* | linear/additive/temporal | spatial/holistic/simultaneous |
| | *Perspective* | descriptive (pronouns) | embodied (gaze) |
| **Cognitive features** | *Perception and processing* | sequential | immediate |
| | *Determinacy of cues* | lower polysemy | higher polysemy |
| **Cultural features** | *Social regulation/ accountability* | strong | weak |

TABLE 2

Stages, requirements, and characteristics of the communicative situation

| **Stage** | **Requirement** | **Characteristics of the communicative situation** |
|---|---|---|
| **Pre-institutionalization**: projecting ideas imbued with subjective meaning into the intersubjective and social realm | **Placement**: Aligning the novel idea with field-level meanings | **Idea-field relationship**: Match between idea and existing meanings |
| | **Exposure**: Making the novel idea available to the target audiences | **Text producer**: Prominence of text producer |
| | **Appeal/Mobilization**: Creating liking and relevance for the novel idea | **Idea**: Clarity and predictability of outcome(s) |
| **Semi-institutionalization**: detaching ideas from their origin and establishing them as objective parts of social reality | **Typification**: Generalizing the meaning and applicability of an idea | **Idea**: Anchoring in sensory perception |
| | **Explanation**: Creating a plausible link between the idea and a relevant problem | **Idea-field relationship**: Match between idea and shared understandings in the field |
| | **Justification**: Creating legitimacy for the novel idea | **Text producer**: Standing of text producer |
| **Full institutionalization:** integrating cognitive, structural, and behavioral elements, which are sufficiently generic and legitimate to remain adopted | **Taken-for-grantedness:** Establishing and maintaining the idea as 'background program' | **Idea-field relationship**: Potential for contestation of idea |
| | **Diffusion/Translation**: Making the novel idea comprehensible across audiences/contexts | **Idea-field relationship**: Match of vocabularies between contexts and audiences |

FIGURE 1

The affordances of verbal and visual text

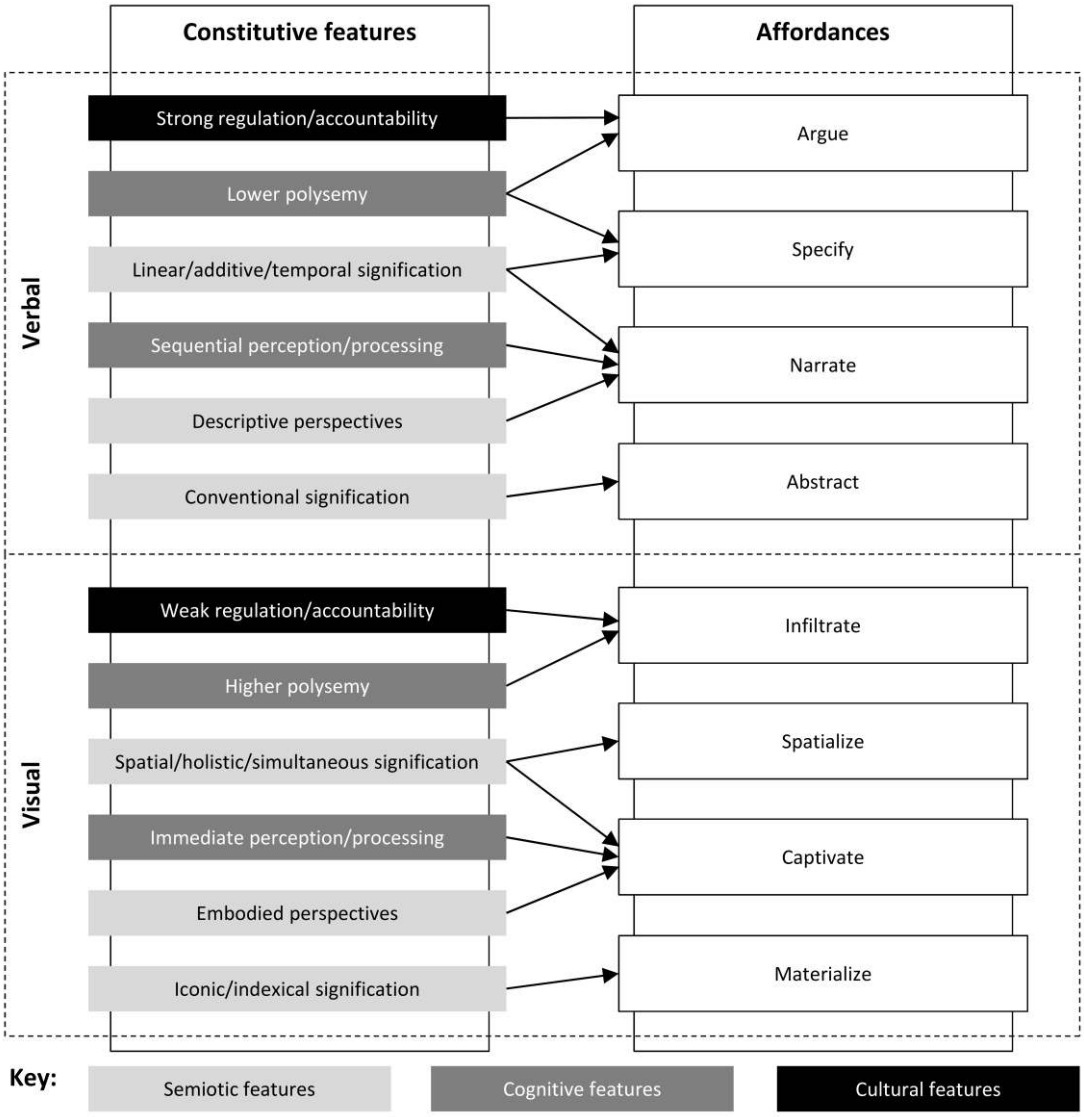


FIGURE 2

The contributions of verbal and visual text to the institutionalization of novel ideas

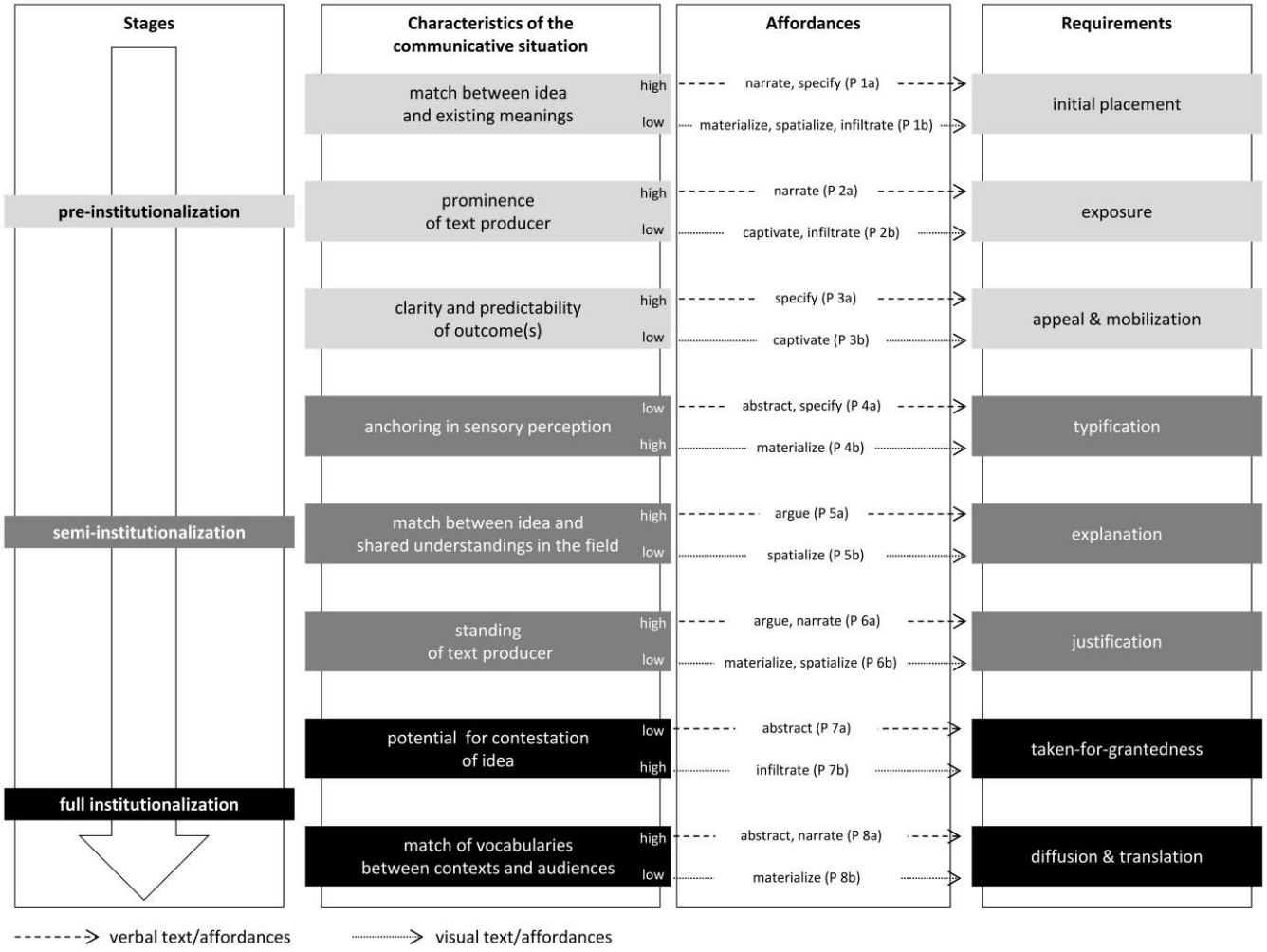
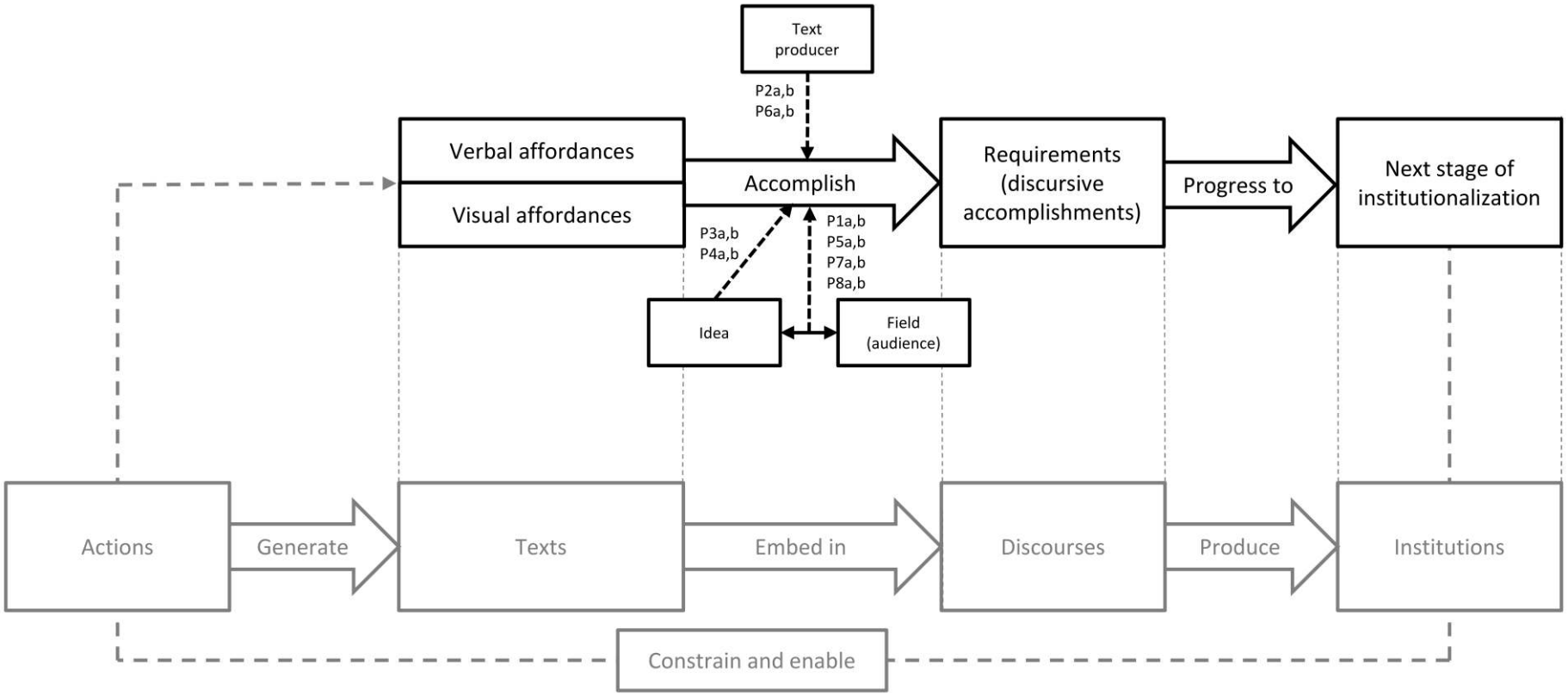


FIGURE 3

Adapted model of discursive institutionalization (based on Phillips *et al.,* 2004)



**Renate Meyer** (renate.meyer@wu.ac.at) is Professor for Organization Studies at WU Vienna and Permanent Visiting Professor at CBS. She focuses on meaning structures and has recently studied structural forms of institutional pluralism, institutions as multimodal accomplishments, novel organizational forms and patterns of management ideas mostly in areas of urban governance challenges.

**Dennis Jancsary** (dennis.jancsary@wu.ac.at) is Assistant Professor at the Institute for Organization Studies at WU Vienna University of Economics and Business. His research focuses on the communicative dimension of institutions and organizations, specifically the role of verbal, visual, and multimodal forms of rhetoric and symbolism in the construction of management knowledge.

**Markus A. Höllerer** (markus.höllerer@wu.ac.at) is Professor of Public Management and Governance at WU Vienna University of Economics and Business and holds a position as Senior Scholar in Organization Theory at UNSW Sydney Business School. His scholarly work is focused on the study of institutions, meaning, and novel forms of organization and governance.

**Eva Boxenbaum** (eva.boxenbaum@mines-paristech.fr) is Professor of Management at MINES ParisTech and affiliated also with Copenhagen Business School. She conducts research on the socio-cognitive, visual and material dimensions of institutional innovation processes, notably how organizational actors shape the emergence and spread of new management practices.