Frames within Themselves: Treating Visual Imagery as a Variable in IR

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts
in the Graduate School of The Ohio State University

By
Andrew Glenn Dombrowski, A.B.
Graduate Program in Political Science

The Ohio State University
2015

Thesis Committee:
Alexander E. Wendt, Advisor
Jennifer Mitzen
Abstract

What implications, if any, has the growing ubiquity of visual images had for state security practices and behaviors? To answer this question, this paper considers how social scientists should treat visual imagery as a causal variable in its own right. Of the two literatures that have seriously considered the implications that the visual has for IR, one treats it as an intervening variable and leaves it both under-conceptualized and deeply entangled with the discourse variable (the “CNN-effect”), while the other argues that visual images have no effects apart from the discursive practices that render them intelligible (the “discursive approach”). To date, no study has treated visual imagery as the independent variable. It is argued below that the independent effect that visual imagery enjoys finds its source in the sense experience that it produces in the human subject. Certain sense experiences will elicit universally shared responses from human beings. How an event or object appears is argued to constrain what it can mean.
Dedication

Dedicated to Kavya
Vita

May 2007………………………………Freedom High School
June 2011………………………………University of Chicago

Field of Study

Major: Political Science
# Table of Contents

Abstract ........................................................................................................................................ ii  
Dedication ..................................................................................................................................... iii  
Vita................................................................................................................................................ iv  
Table of Contents ....................................................................................................................... v  
Introduction ................................................................................................................................... 1  
Conceptualizing Visual Images .................................................................................................... 5  
Visual Imagery in IR ..................................................................................................................... 11  
Material Properties of Visual Images ........................................................................................ 25  
The Causal Chain ......................................................................................................................... 31  
Applications for IR ...................................................................................................................... 35  
Conclusion ................................................................................................................................... 46  
References ................................................................................................................................... 47
Introduction

Over the past century, visual imagery has increasingly come to challenge print-text’s hegemony as the world’s dominant mode of communication. The atom bomb burst into the American consciousness not through the written word, but a photo. \(^1\) Television was an integral part of what made the terrorist attacks on the World Trade Center (WTC) the event that was “9/11”. \(^2\) And the immediateness experienced when viewing images of famine, torture, combat loss, civil war, and genocide seems to affect public emotions in ways that print-text is unable to do. What implications, if any, has the growing ubiquity of visual images had for state security practices and behaviors?

To answer this question, this paper considers how social scientists should treat visual imagery as a causal variable in its own right. How, we might ask, should we conceptualize “visual imagery”? What makes it different from other forms of communication? And what effects, if any, should we expect to observe at the international level?

---

\(^1\) "The Bomb". *Time*. 20 Aug 1945  
International relations theorists have largely ignored the potential effects that visual imagery might have for the conduct of foreign policy and state security practices. And of those who have, no work has been done to disentangle the predicted causal effects of visual imagery from the effects of other variables like discourse. Of the two literatures that have seriously considered the implications that the visual has for IR, one treats it as an intervening variable and leaves it both under-conceptualized and deeply entangled with the discourse variable (the “CNN-effect”), while the other argues that visual images have no effects apart from the discursive practices that render them intelligible (the “discursive approach”). Here “visual practices”, not “visual images”, are the variable of interest. To date, no study has treated visual imagery as the independent variable.

But conceptualizing what it would mean for this to be the case presents theorists with a puzzle. To the extent that the effects of visual imagery lie in its ontology as a communicative medium, these effects are necessarily dependent upon the social discourses that imbue them with meaning. Understood as a form of communication, discourse, not visual imagery, becomes the variable of interest. Thus, if visual imagery is to be treated as an independent variable, its causal effects must lie outside of its semantic properties. The independent effects of visual imagery, if it has any, must be non-discursive.

Does visual imagery have non-discursive effects, and if so, what are they? It is argued below that the independent effect that visual imagery enjoys finds its source in the sense
experience that it produces in the human subject. Certain sense experiences will elicit universally shared responses from human beings, namely a heightened state of emotional arousal. This paper considers four such “sense experiences” which will be referred to as “properties” of the visual image. These properties are: scale, speed, force, and novelty. All four, by heightening a subject’s emotional arousal, naturally directs the subject’s attention to some visual imagery rather than others, meaning that the “privileging” of certain objects and events over others is not entirely reducible to the social discourses that render such events and objects intelligible. How an event or object appears is found to constrain what it can mean.

The remainder of the paper will be devoted to developing a research agenda around visual imagery as an independent variable. I argue that a better appreciation of the visual might yield surprising results for the securitization, risk analysis, emotions, and communicative action literatures.

This paper is divided into four sections. The first conceptualizes what visual imagery is and the second critically engages the research that has already been done on the visual and IR. The goal here is to highlight how little has been said by way of theory or empirics about visual imagery as an independent variable and what effects such a variable is hypothesized to produce. Third, I spell out the four properties of visual imagery that produce the hypothesized emotional effect and then consider what implications this

---

3 Though I discuss these four, there very well might be others that I have yet to consider.
causal relationship would have for social outcomes. The fourth closes with a brief discussion on the implications that this argument has for IR.
Conceptualizing Visual Images

Visual images are understood by this paper to be both mental and material representations of the external world. They re-present a portion of the world from a particular perspective for a particular amount of time. They possess both a sensible and a discursive dimension. Sensibly speaking, they produce a particular visual sense-experience in the human subject. Discursively, visual images are understood and treated as possessing inter-subjectively constructed meanings, which makes them a form of communication. In visual theory’s accepted parlance, the sensible dimension described here refers to the “image as icon” and the discursive refers to the “image as symbol”.

Two initial clarifications are necessary. First, “visual imagery” refers to representations of an event that has not been directly experienced by the individual. Social actors treat them as proxies for the actual event. Second, to split images into the mental and the material is simply to say that we can either internally imagine or externally observe a visual image. A mental representation of an event is one where the individual visually

---

4 In Pierce’s terminology the distinction is being made between the iconic and the symbolic. For Hume, the former refers to our external sense whereas the latter our inner sense.
6 Indeed, the orthodox view in epistemology is that all sense experiences are representations of the world “out there”. Though endorsing this broader view, this paper focuses on material representations of reality (e.g. photography and moving images) that afford subjects sensible knowledge of events that were not experienced first-hand.
imagines what the event looked like. For the purposes of this paper, a material representation is either a photograph or a motion picture. Both materially mimic what the individual would have observed had the individual been present during the event.

Because mental representations are dependent upon sense experience, this paper will treat them as secondary to material representations because observing a material representation of an event is more sensibly immediate than imagining it. Oftentimes only a very small proportion of a nation-state’s population experiences an international event in person. Almost all knowledge of such events is communicated to individuals through various media. Because of this fact, all mental representations of an event will be assumed to have been first observed as a material image (either a photograph or a motion picture). Thus, the variable of interest will be visual images that are represented materially (either photographic or televisual). Social scientists can study our reactions to the photographs and video images taken of an event because these visual images serve as a proxy for the appearance of the event itself.

This position might be met with some resistance, especially by social constructivists. Indeed, it has been persuasively argued that how a visual image is “framed” is largely the result of social practices. To say, as I do, that visual images can serve as a proxy for the event itself is to suggest that there is a natural frame for an event and that perception is

---

7 Understood here simply as a continuous movement of connected images through time.
8 Though they are obviously framed from a particular point of view
not inherently perspectival. There are two possible ways to respond to this charge. The first is to concede this particular point and clarify my own position by saying that (a) because an extremely high majority of publics and decision-makers in nation-states learn about “events” via visual images as opposed to direct experience, we are in fact interested in how they respond to these images and not the event “as it actually happened”; and (b) insofar as these visual images possess one of the four properties discussed below, universal (i.e. cross-cultural) responses will be produced by them. A second tactic is to not concede the point and take a stronger line by arguing that insofar as the events themselves possess one of the four properties, people will have an innate need to frame the visual images that capture these events in a way that accentuates and draws attention to at least one of these four properties. Ultimately this is an empirical question, albeit one that is difficult to settle. For the purposes of this paper, the first line is taken largely because we are interested in the possibly universal effects that visual images might produce, though I consider the implications that the second later on in this essay.

Relatedly, the term “representation” is controversial in some social constructivist circles. It brings to mind the so-called positivist pretension to know the social world “as it is”, uncontaminated by the human observer (i.e. the scientist).\textsuperscript{10} It purportedly assumes a possible symmetry between the signifier and signified. These criticisms exist both in the visual theory and aesthetic IR literatures, and, at least from a visual theory standpoint, the above formulation is in the minority.

All that is meant here by “representation” is that when the human subject observes or imagines an object, similar neural responses are elicited regardless of whether or not the object itself is directly experienced or viewed as a photograph, motion picture, or imagined thought. Because the visual image of an object is not the object itself, it is taken to represent the object to the observer. But it also represents the object in the sense that visual imagery is always a simplification of the object itself. This formulation should not be construed as a wider epistemological stance that is hostile to more hermeneutic or interpretive approaches to the social sciences. It does not deny that there is a hazy line between subject and object when meaning is constructed. For instance, in Jan Mieszkowski’s study of visuality and war, he argues that publics are as “inside” of a war as anyone else. What a war is necessitates the collective creative construction on the part of the public as much or possibly even more so than by those who actually fought in the war. My theory attempts to see if certain forms of spectacle produce systematically observable outcomes in human behavior. Furthermore, it does not reject the constructivist position that images are necessarily perspectival in presentation and meaning. What it does suggest, however, is that there is an irreducible distinction between the sense experience that an object produces and the social meaning that the object possesses. This distinction has implications for the discussion at hand.

12 For instance, neither photographs nor motion pictures possess a third dimension.
13 Though, I would certainly label my work as positivist in the same sense that Richard Ashley does his own. See “The Poverty of Neo-Realism” f.n. 14
14 Mieszkowski, Jan *Watching War*
15 Though perhaps only to a certain extent, as will be hypothesized below.
In what has been said about visual imagery in the social sciences, almost all attention has been directed toward the possible implications that the “image as symbol” hold for social outcomes. To date, very little has been said about the possible causal effects that the “image as icon” may have for human behavior. This is probably because both rationalists and constructivists treat subjective experience as epiphenomenal to social outcomes.\(^\text{16}\)

But if understood only symbolically, the appeal to treat visual imagery as a variable is greatly diminished. This is because, as a symbol, images are either dependent upon discourse (from a constructivist standpoint), or simply a form of information (from a rationalist standpoint). Thus, for visual imagery to be treated as an independent variable, its causal power, if it is to have any, must lie in its sensible effects. And as I argue later in section III, the four sense experiences of scale, speed, force, and the novelty provoke innate behavioral responses in human subjects by heightening their emotional arousal.

That the “image as icon” affects human behavior has, to a certain extent, been documented by psychological findings (though the literature on this topic is comparatively small). One study finds that subjects who are instructed to visualize while they read written statements experience heightened emotional states compared to those in the control who do not.\(^\text{17}\) Studies in both social psychology and advertising have found that visual communication is more emotionally stimulating than written

\(^{16}\) This is why recent attempts to “get at” emotions have proven to be so difficult.

Considering that all of these studies control for semantic content, the differing emotional states that visual imagery produces compared to text is striking. If the effects of visual imagery were purely semantic, then they would be indistinguishable from the effects of written texts that enjoy the same semantic content.

These findings are taken as the starting point of the two literatures that have addressed the potential impact that visual imagery might have on the conduct of foreign policy and state security practices. Both the “CNN effect” and the discursive literatures hold that the sensible immediateness of visual imagery produces emotional effects in subjects that are not explainable by appeals to discourse alone. Neither, however, differentiates between the “image as icon” and the “image as symbol”. This has led to both theoretical and methodological shortcomings in the two literatures, which I will now focus on in turn.

---

Visual Imagery in IR

To date, two general approaches to the study of visual imagery have emerged within the IR literature. The first assesses claims regarding the so-called “CNN effect” which argues that the growth of 24-hour cable news has had an observable impact on the conduct of American foreign policy. At play is a causal story: does the way that the mass media covers events shape the public’s attitude toward the need for humanitarian and military interventions?

The other approach has stressed the discursive dimensions of visual imagery, namely the fact that visual images cannot be understood apart from the social discourses that constitute their meanings. This literature has largely downplayed visual imagery’s causal elements in favor of its constitutive and hermeneutic dimensions. The question here is: how have the discourses surrounding visual imagery changed security practices?

Neither literature addresses what I argue to be the more important question of whether or not visual images on their own affect human behavior and cognition independent of cultural context and, if so, how. Without answering this question, we are prone to make type II errors (as is the case with CNN effect research) or commit theoretical over-reach (as is the case with more discursive, “ideas all the way down” approach). The remainder
of this section will explicate why failing to consider this central question produces theoretical issues for both the causal and discursive approaches.

The “CNN effect”

The notion of the CNN effect—the claim that the media has been a necessary cause of some foreign policy outcomes—originated in Realist critiques of the media after the first Gulf War and Somalia. Foreign policy scholars such as George F. Keenan argued that the new 24-hour news media had deleterious effects on the conduct of American foreign policy because the need to sensationalize news coverage elicited heightened public emotions which constrained the ability of foreign policy elites to conduct Realpolitik.\textsuperscript{19}

The fickle heart of the democratic public was overruling the cool, calculative reason of the policymaker’s head—and only bad outcomes would result.

Of course, the media prided itself on its purportedly newfound influence and reported it as fact. Only within the past fifteen years or so have social scientists evaluated the validity of such claims. Evidence for the CNN effect has been mixed largely because of methodological issues, but also because communication theorists and political scientists have had a difficult time theorizing what in fact the CNN effect is supposed to be explaining.\textsuperscript{20} Determining what qualifies as (dis)confirming evidence has been a contentious issue. Though, as one scholar explains, “[i]n recent years…researchers have

\textsuperscript{19} Keenan, George F. “Somalia, Through a Glass Darkly.” \textit{New York Times} 30 Sep 1993

predominantly associated global real-time news coverage with forcing policy on leaders and accelerating the pace of international communication”.

Nevertheless, evidence is still murky.

This formulation, however, is useful to consider because it isolates what is taken to be the effect’s two primary features. The first is salience: the contemporary mass media’s reliance on visual imagery in particular is thought to heighten the emotional response that the audience has to news events. The second is speed: the growing pace of communication may create a “time pressure” on decision-making, forcing decision-makers to choose policies less carefully. The causal chain is understood as follows:

Media (IV) → [Visual Imagery + Speed → Public Opinion/Elite Perception of Public Opinion] → Intervention (DV)

Though the issue of speed is important in its own right, the focus here will be on the contention that news events enjoy a heightened emotional salience in the 24-hour news era specifically because of its heavy reliance on televisual imagery. Indeed, case studies in this literature largely point to the role that the growing use of visual images has played when arguing for the validity of the CNN effect. Speed on its own is not thought to explain the phenomenon outright.

---

22 Though the causal mechanisms at play are still disputed.
23 Some have argued that the media does not influence outcomes, but it does affect process. For the purposes of this discussion, the variables of process and outcomes are interchangeable as preferring one over the other has no bearing on the discussion at hand.
24 Robinson, 2002

13
Considered this way, scholars need to disentangle the potential confounding effects that discourse may have when studying visual imagery’s effects. This is because as a form of communication visual images are only rendered intelligible when mediated through language. And images are taken by this literature to exert communicative effects. Thus, in order for images to mean anything, they are necessarily dependent upon discourse. To show that visual imagery has its own effects requires the social scientist to show how such imagery is able to operate independently of discourse. If not, then images have no role to play in our theory because we can instead conduct a discourse analysis. Here we are interested in what is said about the image and not the image in itself. For instance, though news telecasts and reporting have been shown to run more emotionally gripping news stories than does print journalism, this may owe to the differing discursive practices of the two institutions as opposed to a more fundamental ontological difference between the two media.\(^{25}\)

We must show how this mechanism:

\[
\text{Media (IV)} \rightarrow \text{[Visual Imagery + Discourse + Speed} \rightarrow \text{Public Opinion/Elite Perception of Public Opinion]} \rightarrow \text{Intervention (DV)}
\]

Is not reducible to this:

\[
\text{Media (IV)} \rightarrow \text{[Discourse + Speed} \rightarrow \text{Public Opinion/Elite Perception of Public Opinion]} \rightarrow \text{Intervention (DV)}
\]

\(^{25}\) i.e. it may simply be the case that television news feels more free to frame their stories in emotional terms than print journalism does. This is borne out by studies conducting content analyses of the two media.
This is not to suggest that the validity of the CNN effect hinges solely on whether or not it can be shown that visual imagery produces such an independent causal effect. The CNN effect is not a theory of visual images *per se*. It may be the case that the speed through which discourses are transmitted globally explain the phenomenon. In this respect discounting the role of the visual would not disprove the CNN effect, but it would raise the question as to why, when making their case, so many empirical studies rely upon the purported importance of the visual aspects of news events. And it would also undermine the core intuition held by CNN effect proponents that the graphic nature of some visual images adds something to the equation in a way that narration and print text do not. So in this respect disentangling the effects of images from the effects of discourse would at least prove to be a welcome clarifying move.  

But investigating the role of visual imagery may prove to be more than a clarifying move. For one, if it turns out that visual imagery does not possess the causal effects that some assume, then our predictions may need to be adjusted. Though social scientists treat intervention as the dependent variable in part because of methodological reasons, it is also assumed that the power of visual imagery, if channeled correctly, could produce this outcome. If it turns out that this is not the case, then we risk committing a type II error. The CNN effect may be of a lesser magnitude than is presently theorized, but this on its

---

26 Such a move would also apply to CNN effect detractors because, though they argue against the supposed autonomy of the mass media, they are in agreement with its proponents in that both assume that the growing ubiquity of visual imagery has mattered for the conduct of foreign policy. Detractors only argue that the mass media is not a constraint on the government, but merely an instrument that foreign policy elites use to whip up popular support for certain state actions.
own does not suggest that such effects do not exist at all. Other more circumscribed dependent variables would need to be found.\textsuperscript{27}

Second, and more importantly from the perspective of this paper, is the possibility that the CNN effect frame has actually narrowed our understanding of the possible ways that visual imagery can affect foreign policy behavior and thinking. By treating visual imagery as an intervening variable that either the media or governments wield, we neglect the possibility that visual imagery might exert its own effects on these two actors. But we will be unable to get at these effects insofar as the image is bound up in surrounding social discourses. Gaining a better understanding of how the two interact will help us better understand the independent effects of both and it could broaden the possible applications that the study of visual imagery could have for the study of foreign policy and security practices. The question that needs to be asked, then, is how the two, in fact, interact.

\textit{Discursive Approaches}

While scholarship on the CNN effect has been the only sustained effort to study (at least indirectly) the causal effects produced by visual images, it does not exhaust what has been said about images in IR. Discursive approaches too have begun to study the impact of visual communication on IR practices.

\textsuperscript{27} Which is one of the reasons why scholars like Strobell have emphasized process over outcomes; \textit{See} Strobel, Warren. \textit{Late-breaking foreign policy: The news media's influence on peace operations.} Washington D.C.: United States Institute of Peace Press, 1997.
Beginning with Williams’ call to theorize what the rise of visual images as a communicative medium has meant for security practices and communicative action more generally, discursive IR theorists have focused on what implications this development has had for how actors construct shared meanings. This approach correctly emphasizes the need to place visual images within a larger discursive context as a precondition for their intelligibility. Indeed, as mentioned above, an image’s meaning is necessarily mediated through language and is therefore dependent upon the background of prevailing social discourses. Thus, the focus has been on constitutive, as opposed to causal, analysis.

But their focus on how the meanings of individual images are constructed through discourse is only one part of their research agenda. More important for our purposes is what they have to say about how the “practices of seeing” constitute the social ontology of visual imagery more generally. It is held that the material effects that are produced by visual imagery are highly mediated by (and vary according to) the discursive practices that instantiate this social ontology. That is, the effects that the “image as sign” will have will vary depending upon our shared understanding of what visual images are and the “practices of seeing” that we perform. Though discursive theorists recognize the distinction between the “image as sign” and “image as symbol”, they want to problematize the idea raised by this paper that the “image as sign” can, independent of a discursive context, produce social effects.

Hansen, to date, has provided the most definitive explication of such a position, melding together securitization theory with the work that post-structuralists like David Campbell have done on visuality. Departing from the linguistic constructivism of the Copenhagen School, she seeks to distinguish the ontology of the visual from that of the textual. The sensible immediacy of visual imagery is said to produce emotive effects in subjects that language alone is incapable of producing. This is the same claim that motivates both this paper and work on the CNN effect. But Hansen’s explanation for why this is rests upon an ontological distinction that differs from the one that this paper makes. This is because Hansen writes that the visual needs to be treated as “an ontological-political condition rather than a variable”; the differing effects that images produce relative to texts stem from their differing social, as opposed to natural, ontologies. What is of interest here, then, are the differing discursive practices that constitute these distinct ontologies. Indeed, this brings to mind Bourdieu’s oft-quoted line: “[t]he ‘eye’ is a product of history reproduced by education”. That the sensible immediacy of the visual matters for social outcomes is itself a result of practice, particularly the practice of visual authenticity.

As I understand their argument, discursive theorists like Hansen hold that though visual imagery, as a material object, produces heightened sense experiences (known here as

---

30 Hansen, 52
immediacy)\textsuperscript{32} in the subject, whether or not this sense experience translates into a performative action is necessarily a result of the discursive context that produces the social ontology of visual imagery more generally. Visible imagery, with respect to its sensible dimension, cannot cause action on its own and is therefore unable to yield social effects.

Defined in this manner, I do not disagree with the conclusion. But given the rather technical nature of the terminology employed above, taking a step back to unpack this line of reasoning is necessary. Doing so will help clarify not only the thrust of this particular discussion, but the overall claim of this paper as well. Following Adler and Pouliot, we need to distinguish among behavior, action, and practice.\textsuperscript{33} Behavior is defined by them as the “material dimension of doing.” It is “a deed performed in and on the world.” Action is a specific form of behavior, a doing that adds “an ideational layer, emphasizing the meaningfulness of the deed at both the subjective and intersubjective levels.” Practices are the “patterned nature of deeds in a socially organized context.” Whereas behaviors are merely material doings, actions are purposive material doings. And practices are patterned actions.

Understood this way, Hansen’s logic begins to make sense: since we are studying the social, as opposed to the natural, world, we are interested in how social actors act when placed in certain material and discursive circumstances. Social actors only act to bring

\begin{flushright}
32 Hansen, 55-7  
\end{flushright}
about or react to material effects when these material effects are understood within a
discursive context. Thus, visual imagery, insofar as it affects social *actions*—and thereby
produces social effects—necessarily does so in virtue of its social, and not its natural,
ontology. Therefore, the social effects produced by the “image as sign” are dependent
upon the discursive practices that constitute the social ontology of visual imagery.
Hansen writes at one point, “‘immediacy’ in the form of emotion thus enters discourse
not as a simple fact but as something that is constituted in discourse.”

This view does not deny that these claims are subject to empirical investigation, only that
“images” proper cannot be treated as a variable because their effects lie in their social
constitution, something that is inevitably in flux. In effect, then, discursive theorists like
Hansen are setting very stringent scope conditions on the study of visual images. The
empirical outcomes of social life are taken to be so contingent upon discursive production
that very little general theorizing can result. And an investigation into the discursive
production itself is not itself causal and scientific, but hermeneutic—meaning that much
of our analysis lies in excavating an object’s social ontology.

Specifically, the visual practice that is often cited as constituting the image’s social
ontology is the visual practice of authenticity. Following visual cultural theorists, critical
IR scholars like Hansen, Campbell, and der Derian\(^\text{35}\) argue that the privileged epistemic
status that visual imagery enjoys vis-à-vis text owes to the modern visual practice of

\(^{34}\) Hansen, 57

treated visual imagery as authentic. That the visual image is more sensibly immediate is taken to grant it its supposed authority over other forms of communication, in that visual imagery is taken to show what ‘really happened’. Thus, the heightened emotional effects that are evoked by the visual are not produced by the image *per se*, but by this discursive practice of authenticity. The only reason we feel the emotions we do when gazing upon certain images is because we are educated to believe that visual imagery reproduces the truth. For instance, der Derian takes the standard visual culture line when he writes, “How long before photographic immanence loses its power of authentication and stimulation, we stop believing what we see, and the significance of the image is itself called into question? How many times can the truth take a beating before the public just stops believing *anything* it hears, reads, and sees? Not soon enough?”

If this current visual practice were deconstructed, so the thinking goes, the social effects of the “image as sign” could very well be different.

But is this actually true? Are our acts of privileging and valuing certain social and natural features of ourselves solely produced by discursive practices? The mind is not a blank slate, so there are inevitable cognitive hard limits found in humans. My fear is that the discursive approach is committing theoretical over-reach by ascribing effects to discursive practices that are in reality being caused by the visual image’s materiality. For their argument to be correct, the visual image “variable” must be “ideas all the way down”; if not, then some causal role must be granted to the image *qua* material object. Now, discursive theorists will reply with what they take to be their “ace-in-the-hole”: that

36 der Derian, 35
materiality in itself cannot explain social outcomes because such outcomes require a discursive context. And this is correct up to a point. But this still leaves open the possibility that the materiality of an object can constrain possible social outcomes. Critical IR theorists are not idealists in any metaphysical sense; no one argues that discourse is what prevents humans from flying or jumping one hundred feet in the air (a naïve idealism indeed). So we would be remiss to neglect the possible role that materiality plays in constraining social outcomes.

More specifically for our purposes, the claim that the visual practice of authenticity is what allows images to elicit heightened emotional states in subjects is a causal claim that needs to be substantiated by the empirics. In this respect discursive theorists have merely noted a possible correlation between certain practices and certain social outcomes. More rigorous theorizing and testing are needed.

Perhaps the “image as sign” is, like any other material object, constraining the possible social actions that can be performed. I want to suggest that, contra the discursive approach, the independence of the visual image lies in its ability to act as a material constraint on social actors.

But how does the visual image constrain? Our typical understanding of “material constraint” privileges brute physicality: tanks, regardless of the social context, can pulverize human bodies; walls are very good at keeping people either in or out; roads and
then train tracks moved artillery faster. Visual imagery certainly lacks physicality in this respect, which is probably why the materiality of the visual image has been neglected by the literature. Though, sheer brute presence does not exhaust the potential ways that objects can materially constrain social actors. Another way—one that is rarely ever considered by the social constructivist literature—is behavioral. That is, visual imagery elicits certain innate behavioral responses in human beings that constrain the future actions that social actors can perform.

This possibility, though rarely considered, is allowable within the frameworks presented by Adler and Poulion as well as Wendt and Wight. For Adler and Poulion, innate behaviors would fall under non-purposive behavior. Though I do not believe Wendt ever addresses innate behaviors specifically, they are an instance of his larger “rump materialism” category. And my argument would be considered an elaboration upon Wight’s notion of Agency1.

The causal chain, which is elaborated upon in the forthcoming section, flows as follows: visual imagery produces certain sense experiences in subjects. Certain sense experiences, which I label scale, speed, force, and the novelty, elicit an innate behavioral response: emotional arousal (i.e. excitability) is heightened. Conscious attention is thereby directed toward the stimulus. If correct, this suggests that the objects and events that human

---

beings privilege with their attention is not solely a product of discourse. Indeed, we might possess our own natural framing devices.

If social constructivism is about excavating the logics and rhetorics invented by human beings to justify what \textit{matters} (i.e. what is ethically privileged by society), then it is worth considering if this phenomenon of privileging is solely a product of discourse or discursive practices, or if human beings possess predispositions to attend to and epistemologically privilege certain events that assume particular visual forms.
Material Properties of Visual Images

The task so far has been to highlight a lacuna in the literature: that though we have begun to recognize the causal and constitutive importance of visual imagery for social processes and outcomes, as of yet no one has conceptualized what makes visual imagery distinct as a variable. And though the discursive approach is apprehensive about doing just this, having not done so leaves others to ask what exactly it is that they are arguing against. What would it mean for visual imagery to be a variable in the first place?

I have suggested that visual imagery’s causal independence lies in its distinct material effects. Visual images produce certain sense experiences that are hypothesized to elicit a shared, cross-cultural behavior in human beings. These four visual properties are thought to innately privilege certain objects and events over others because they direct human attention to them. I will now focus on each one in turn:
Scale

The first feature to consider is *scale*. Scale results from the juxtaposing of the immensity or vastness of the visually captured object with the perceived insignificance of the human subject when placed beside it (either in reality or the subject’s imagination). Historically, scale has referred mostly to natural objects like mountains, oceans, and vistas, though man-made objects—especially since the advent of modernity—too are able to produce this effect. Indeed, we see throughout history that scale has been used by some individuals and groups to exercise their power over others. From the pyramids of Egypt to Nazi architecture, the scale of these structures effects certain emotional responses in those who gaze upon them in virtue of the dynamic of juxtaposition described above.

Speed

*Speed* refers to how quickly the visually captured object produces its effects. This feature is best captured by video imaging, though some photographs also convey this phenomenon.

The immediacy of this effect is only further highlighted by witnessing it on film. It should be noted, however, that speed is, to a certain extent, a relative concept. Society understands the speed of an object in relation to the speeds of other objects that have already been experienced. We should be cognizant of this fact when assessing the speed conveyed by images taken in past “eras”. For instance, though turn of the century steam
engines were quite slow by today’s standards, their speed provoked a level of excitation in subjects that had to date been unparalleled.³⁹

Paul Virilio has famously argued that the logic of modernity lies in the continuous and never-ending drive to compress time. Technology, in both its analogue and digital forms, operate according to a logic of ever increasing speed. Some of this speed can be observed, or at least, directly experienced. Here one recalls the German blitzkrieg or supersonic jets. But sometimes this time compression is less than visible: computational technology comes to mind. Indeed, much of contemporary social life is experienced as the effects of techno-material causes that operate beyond the radar of everyday human perception.

Virilio here advances his own causal claim: that the human body is necessarily experiencing particular behavioral and cognitive effects because of this time compression in their daily activities and the techno-material forces that produce and mediate their everyday lives. It is this very conjecture that I wish to test.

³⁹ An interesting question to consider here is where human psychology sets hard limits on the perception and apprehension of speed. For instance, the speed of computer networks today seems to be beyond the level of human comprehension. This raises a further issue of the extent to which the material forces of the future will be “unobservable” to human sense experience and what effects that might have on social life.
Force

*Force* refers to the level of destruction that an image conveys. Images are best able to capture *force* when it is concentrated in both time and space. For instance, though car crashes kill anywhere between thirty to forty thousand motorists each year, these images are spread out over both time and space. The force of such a statistic is difficult to collapse into an image for these reasons. Indeed, there are no thirty thousand car pile-ups, which, if there were, would almost certainly provoke a response from the public and decision-makers *even assuming that the total number of deaths each year remained constant*. What counts is the visual perception of potency.

Novelty

Novelty is a relative concept. An image can only be “novel” in the sense that the individual or group of individuals observing it possesses a low level of prior sensible experience of that particular object or event. Novelty is contingent upon what has been experienced in the past.

Therefore, unlike the properties of scale, speed, and force (though they themselves are too—to a certain extent—are relative), the researcher cannot determine whether or not a visual image is “novel” by the appearance of the image itself. Instead, we must unearth the discourses that surround the images in order to assess whether or not they were
experienced as such. Were they treated as or thought to be out of the ordinary? How was the uncanny expressed?

For instance, in his discourse analysis of the early atomic tests, Joseph Masco finds that atomic scientists, policy-makers, and journalists alike all treated the visual imagery of the bomb as a “cognitive rupture”, a tactile experience that exceeded the current discursive categories and thus required new forms of social consciousness to neutralize the threatening nature of this “rupture”.

Quite simply, the image of the bomb was taken to be visually novel: nothing like it had been experienced before.

Similarly, after the 9/11 attacks, Donald Rumsfeld—perhaps unwittingly—expressed this logic in a more home-spun manner: “What I do know is the standard words jangle in my head when I hear them, and then I put them onto the subject they’re relating to…and I think to myself, ‘Gee, that isn’t really as good a word as we ought to be able to find.’

Of the four properties discussed, novelty is the most studied by constructivists. The “practice turn” has given it a privileged position in its own thinking, oftentimes treating novelty as the engine behind agency. The purpose here is not to engage in the larger debate over what novelty might mean for agency and the conceptualization offered is not all that different from what has already been said. The particular focus on novel sense-


experiences may be a slight departure from the more discourse-centric formulations, but this amounts to no more than an adjusted angle of focus. Nevertheless, the differing effects of visual novelty have yet to be theorized more fully and it is here where the research agenda being proposed hopes to exercise its comparative advantage.
The Causal Chain

Separately (or together) these four visual properties—when observed—produce innate behavioral responses in human subjects. The hypothesized causal chain flows as follows:

Visual Imagery (IV) → [Sense experience → Emotional Arousal] → Attention (DV), where “sense experience” refers to at least one of the visual properties discussed above.

If correct, this would suggest that the privileged standing granted to certain objects and events in social discourses is not solely dependent upon discursive practices. Events like 9/11 or objects like the atom bomb cannot be fully deconstructed away because the attention that is given to them is pre-cognitive and pre-linguistic.

To a certain extent, this seems obvious and verges on the banal. But the hypothesis offered here illustrates one of the mechanisms behind the *a priori* material constraints on discourse, which is both tougher to show (both empirically and theoretically) and something that is rarely ever done. Indeed, though it is old hat to say that ideas interact with the material world thereby producing observable material effects, more works needs to be done that investigates and unearths how these interactions actually take place. And though we want to say that the meaning of social objects is socially constructed, we
hesitate to say that our interactions with them are purely ideational. Once this uneasiness sets in, the door has been opened to theorizing the mechanisms that underlie the interactions between materiality and ideational variables. The solution offered here is to say that there are innate behavioral constraints that inherently limit our fields of attention and which privilege the objects and events that are later rendered intelligible by social discourses.

Whether or not the mechanism outlined above accurately captures the dynamics of human social interactions is at the moment undecided. Sufficient empirics are simply lacking. There is, however, suggestive evidence that has come out of the fields of social and cognitive psychology that point towards this paper’s thesis. Nevertheless, even here the output on this subject has been relatively minor. Much more experimental research needs to be done.42 Some of this evidence should still be considered in turn.

To date, much of the psychological work done on human behavior and visual imagery focuses on our shared reactions to what I have called “scale”. It has been found that the enlargement of visual images in various forms (e.g. screen, photograph) produces heightened emotional arousal in subjects.43 Now, the parallels between the size of the

---

42 What is being called “neuro-aesthetics” is a research program that has been in development over the past ten years or so and, hopefully, the evidence will begin to bear out some of the claims made in this paper. See Chatterjee, Anjan. "Neuro-aesthetics: A Coming of Age Story." Journal of Cognitive Neuroscience. 23.1 (2010): 53-63; and Jacobsen, Thomas. "Beauty and the Brain: culture, history, and individual differences in aesthetic appreciation." Journal of Anatomy. 216. (2010): 184-91.

actual image and the relative size of the objects represented in the image could be limited. No psychological study, as far as I know, has addressed the latter question. And until that is the case, the jury will be out on whether or not the concept of scale elaborated here has any sort of theoretical purchase.

There is one paper, produced by Keltner and Haidt, that shares a similar understanding of scale and speed as this paper—albeit couched in different terminology. They draw on early Enlightenment aesthetic theory and operationalize the concept of the “sublime” to explain the emotional effects that intense visual imagery produces in human subjects. Because of the dense philosophical baggage that this concept carries, I am wary of importing it into the social sciences. But at a very basic level they seek to capture the same dynamics as this paper does. For them, both scale and speed should produce heightened emotional states in subjects, though because they do not disaggregate visual imagery from discourse, they argue that particular visual images can elicit particular emotional states—a claim from which this argument wants to distance itself.

Unfortunately, their paper amounts to little more than a sketch for a more ambitious research agenda—one that is often cited, but never has been put into practice.

---


45 This has not stopped others in IR from importing this aesthetic category into IR. Termed the “aesthetic turn” by its adherents, a special 2006 issue of Millennium was devoted in turn to the sublime in IR. Though it is not useful to go into a full-blown criticism of this literature here, two things bear mentioning: (1) to date, the Kantian sublime is a validation of the postulate that human beings are rational creatures, and (2) it is unclear what the empirical implications of the Kantian sublime would be. See Steele, Brent. Defacing Power: The Aesthetics of Insecurity in Global Politics. Ann Arbor: University of Michigan Press, 2010 for a similar criticism.
With respect to the actual mechanism at work, however, there is both more research and a rather strong consensus throughout the psychological community that visual imagery can produce emotional responses either prior to or by bypassing entirely our cognitive faculties. And it is fairly well-established that heightened emotional arousal correlates with greater attentiveness and heightened focus on the stimulating image(s). If it can be substantiated that the properties of images have the effect argued here, then the rest of the mechanism could, in practice, follow.

Applications for IR

A closer study of visual imagery holds possible implications for a variety of IR literatures. This section addresses four possible conversations that might be held: between the study of visual imagery and (1) securitization theory, (2) work on emotions and IR, (3) the dialogue between risk and security studies, and (4) communicative action approaches.

(a) Discursive Approaches

The argument presented here in no way diminishes the importance that discursive practices and the discourses that instantiate them have for the study of social life. But as more and more of the world’s communication is mediated through visual imagery, discursive approaches are going to need to be more and more supplemented by visual approaches.

One counter-factual might illustrate this intuition nicely: would 9/11 have provoked the response it did had it not been broadcast over live television in real-time? Would al-Qaeda have even attempted the attack had televisual imagery not been so culturally ubiquitous? Of course, we can never know for sure. But we would be more puzzled to discover that it had no added effect whatsoever.
The problem encountered when trying to get at this, however, is the same problem that plagues the emotions in IR research (indeed, this argument would fit within this growing literature): how do we apply the findings from controlled experiments to the uncontrolled world that is international politics? If this innate behavior is observed cross-culturally, what has this psychological fact really told us about the workings of international politics?

One way to solve this problem is to follow what some communication theorists have done and map the communication pathways that undergird international knowledge exchange. Of course, this will prove to be difficult and unwieldy, but simplified models may still yield interesting insights. The goal then is to see how the cross-cultural psychological tendencies that I argue exist are channeled throughout the social sphere and what macro-level outcomes might result from them. The approach then would be the reverse of the one advocated by the CNN effect literature: here visual imagery is treated as the independent variable and the media (a proxy for information pathways) is treated as one of the intervening variables.

Currently I have two rough “models” in mind. The first being what I will term the “modern” communicative system and the other being the “post-modern”. What characterizes the modern is that most of the knowledge that it produces originates and is disseminated from a few dominant nodal “clusters” in the communicative network. Knowledge networks of the 50s-80s were much more “top-down” owing to the

---

hegemony of the mass media as well as a strong centralized government. The “post-modern” system is characterized by its more decentralized nature, and the shift from the modern to the post-modern tracks largely with the gradual and growing predominance of the Internet. As communication becomes more dependent upon visual mediums (and mixed mediums), these communicative networks are not only channeling knowledge, but heightened emotions as well. Emotional contagion is the phenomenon that theorists of visual imagery should be looking to map.

One hypothesis that follows from the argument is that as visual imagery becomes more ubiquitous, the prevalence of heightened emotional states and attention may be on the rise, which may translate into different foreign policy decisions than may otherwise have been performed. Does visual imagery qualitatively change issue cycles? Are we in an age when the passions have become unleashed? Is the international arena less “detached” and calculating than it was purported to be before? How would we go about showing this?

One potential entry-point may be to look at how risk is understood and perceived. The past decade has seen a growing conversation between the literatures in risk and security studies. One of the puzzles that has resulted from this engagement is why foreign policy-makers treat some risks disproportionately to others. According to the risk literature, the consequences of 9/11 and the risk of terrorism more generally, for instance, did not justify the level of resources that were dedicated to dealing with the threat.49

Over-securitizing the terror risk sapped state resources that could have otherwise been spent on more imminent and consequential domestic and foreign policy risks. Why do some extremely low-risk international events provoke over-securitizing behavior from states?

A substantial literature has grown around the study of risk and disaster over-response. One strand of this literature has considered the role that emotions play in distorting risk calculations. High emotionally-laden outcomes have been shown to dramatically alter the decision-calculus of actors when compared to less emotionally-laden outcomes of equal consequence and probability. Little attention, however, has been paid to the origins of these emotions. Does the fear of such outcomes stem from a common source? Or is there great variance owing to variables like culture? Answering this question holds significant implications for how states deal with and understand the nature of security risks.

Given the logic and mechanisms discussed and illustrated above, I hypothesize that equally risky and consequential outcomes are able to provoke either overreactions or under-reactions from actors, in part, because of their visual appearance. A greater appreciation of the (possibly) universal effects of visual imagery on risk perception is


Mueller and Stewart, 103-7


needed if we are to understand why foreign policymakers and publics alike over-
securitize after certain international events.

The overreaction to low risk events is a common social and political occurrence that
confounds expected utility models. Some have suggested that bounded rationality is to
blame, arguing that overreactions result from the miscalculation of probabilities. But
psychological studies have shown that even when presented with objective probabilities,
subjects continue to overestimate the costs of certain rare events. Furthermore, studies
have also shown that “people are insensitive to variations in low probabilities”. And
Sunstein provides evidence suggesting that in certain highly emotional circumstances,
people neglect probabilities entirely. That actors are, to a certain extent, ‘statistics
resistant’ sheds doubt on the imperfect rationalizer claim.

Prospect theory does predict that actors are willing to incur costs upfront in order to
insure against a small risk of catastrophe. But as Sunstein notes, prospect theory does not
consider the effects that emotions potentially have on risk perception. Indeed, “prospect
theory predicts the same reaction to risks that produce strong emotional reactions and to
statistically equivalent risks that do not produce such reactions”. There is ample
evidence to suggest that such reactions vary depending upon whether or not the potential
outcome is emotionally laden or not. And given the oftentimes emotional character of

---

53 Kahneman, Daniel and Amos Tversky, “Prospect Theory: An Analysis of Decision Under Risk”
54 Sunstein, 123; Cf. Kunreuther, Howard, Nathan Novemsky and Daniel Kahneman “Making Low
Probabilities Useful” *Journal of Risk and Uncertainty* 23 (2001): 103-120
55 Sunstein, 123-6
56 id. 123
international events and crises, this variance is important to consider. For instance, risk perception has been shown to be heightened when the subjects are primed with affective language.\textsuperscript{57} And when two outcomes are framed as equally costly and likely the emotionally laden outcome has been shown to elicit significantly higher perceptions of risk and willingness to act.\textsuperscript{58} One question that arises from these results is why some low risk events are emotionally salient whereas others are emotionally neutral. What mediates these responses?

One possible explanation—given by the “risk as feeling” literature—points to a person’s ability to visualize potential outcomes. The vividness of an imagined outcome correlates with heightened emotional responses in subjects.\textsuperscript{59} And it follows that higher emotional responses and greater image recall result from the more vivid and immediate an imagined or described outcome is. Examples of such behavior usually come from the insurance literature. For instance, flood insurance is always undervalued in relation to earthquake insurance, which is always over-valued. One explanation as to why this is the case is because visualizing two feet of water in one’s house, though a costly inconvenience, is nowhere near as fear inducing as a tornado or earthquake. Experiments have shown that subjects view the possibility of a nuclear plant meltdown to be more risky than radon poisoning, even when the probabilities and costs of each are equal. Again, an explanation

given for this finding in the literature is that nuclear meltdowns (and explosions) are far easier to visualize and come far more readily to mind than images of silent radon poisonings.

In the IR context, as John Mueller has noted, an American citizen has a greater chance of dying from freak accidents like falling into a toilet than she does from a terrorist attack. Nevertheless, we in the United States fear the latter outcome far more.\(^{60}\) This is also the case for risks of fatality that are also far higher and—from the perspective of the individual(s) killed—equally gruesome: traffic accidents.\(^{61}\) But Americans are not willing to spend billions of dollars on traffic reform.\(^{62}\) The visual vividness of the 9/11 attacks comes far more readily to mind when we calculate the risk of both.\(^{63}\) The question that arises here is this: what makes some images more vivid than others? What is it about events like 9/11 that differentiate it from more likely outcomes of fatality?\(^{64}\) The four properties of visual images I propose above offers such an explanation.

In the end I think that the explanation for this behavior is evolutionary: on the savanna and in the forests, events that possessed these visual characteristics triggered instinctual

---

\(^{60}\) Though both statistics are extremely negligible

\(^{61}\) For a discussion on this see David Campbell

\(^{62}\) As far as I know, no such term even exists

\(^{63}\) Prospect theory offers the availability heuristic as an explanation for this dynamic. But here we are interested in why some images become more available than others. The availability heuristic also does not address the emotional aspects of decision-making.

responses. The advantage seems to lay in the fact that, back then, the only time human beings would witness such events is when they were close at hand. This visual adaptation circumvented our reflective and cognitive faculties because in such situations there was little time for reflective thought.

We humans possess the same hardware as our hunter-gatherer forefathers, though we obviously live in radically different environments. One such environment is that of mass media: it broadcasts to us, from great distances, events that, at times, trigger these biological adaptations. And because we crave emotionally heightening experiences, there is a market imperative to frame stories and broadcast images that fulfill such demand.

For many students of foreign policy, this overvaluing of risk describes the United States’ behavior after the 9/11 attacks. Terrorism seems specifically designed to exploit this relationship between human nature and the changed technological (specifically media) landscape. John Mueller, for one, has forcefully argued that the risks that terrorism poses, though very much real, are dramatically overvalued given the probabilities and costs that are actually associated with terrorist acts themselves. Indeed, he argues that our reactions to terrorism have been far more costly and damaging to human lives and the US economy than any terrorist act has been.

Some have argued that the image of the Muslim “Other” is to be blame and this may be true to a certain extent. But at issue is why the public and its representatives feel
compelled to over-exaggerate the Other’s threats level and capabilities? Surely had al-Qaeda conducted an attack that was imagistically less-fantastical, our reaction would have been muted. I want to suggest that there was something about the way that the event visually appeared to the American public that triggered heightened emotions after the attack, namely fear and anxiety. That is, terrorist attacks like 9/11 produce ready-made images that are designed to capitalize on our biologically innate tendencies and reactions that we have toward visions of scale, speed, force, and novelty.

**Testing these claims**

Without finding confirming or disconfirming empirics, the discussion above will never amount to more than theoretical conjecture. This paper proposes two empirical tests to adjudicate the claims made by the theory proposed above.

First, though there have been many psychological studies that provide suggestive evidence for the claims offered in this paper, none so far have explicitly tested some of the proposed links. For instance, studies on scale have only increased the actual size of the image being shown to subjects (i.e. projecting it on a larger screen), rather than relying on subject’s perspectival intuitions. Indeed, much of what is seen today is mediated by small screens. What is of question is not the actual imagistic size, but whether or not the scale between an object and a human being has any sort of psychological effect on how humans process certain visual information.
One experiment would see if an object’s scale has this hypothesized effect controlling for all other variables (risk primarily). Does the physical scale of an event lead to differing assessments of risk? Other experiments would test whether or not these same dynamics are found with respect to an event’s speed and force (as described above).

Developing experiments are, of course, not enough. Though yielding fascinating insights into human cognitive dynamics, psychological experiments nevertheless experience an inevitable hurdle: translating the findings produced from “artificial” environments to the far-less “controllable” real world. Though experimental findings are indeed suggestive, they are oftentimes only merely that. To find, for instance, that the strength of a subject’s emotional response differs between visual and textual stimuli does not prove that the visuality of a terrorist attack directly produces over-securitizing behavior on the part of states. Instead, we would want to find evidence from past terrorist attacks that demonstrate a link between the two.

One dependent variable worth considering is the change in military funding between years before and after a terrorist attack. Though time consuming, it is possible to gather data on the independent variables proposed above: each attack produces a certain volume of deaths in a measurable amount of time. Each target has a particular size. Each type of an attack is considered by newspapers afterward to follow a pattern or to be wholly novel. Controlling for a host of control variables (country, attack type, a state’s economic output, the ethnicity of the attacker, and whether or not it was international or domestic terrorism), we can observe if any of these variables have any noticeable effect on whether
or not states spend too much of their resources to combat a threat after a particular type of attack.
Conclusion

This paper has aimed to lay out a possible research agenda for the study of visual imagery in the social sciences, more specifically, international relations. In this respect no empirical facts have been validated, nor have any theoretical claims been substantiated. This essay has primarily served to clarify the issue. How should social scientists treat visual imagery as an autonomous variable? The above discussion has attempted to provide this clarification.

Many will likely not find my formulation to be satisfactory. Discursive theorists will likely point to my glossing over their epistemological claims (which is really where they take refuge against more “positivist” critiques). Nonetheless, it is clear, theoretically, at least, that a distinction is required between the sensible effects that visual imagery has on human behavior and the hermeneutic effects their meanings produce. And insofar as this distinction holds, the former is not reducible to the latter and warrants the attention of serious social scientific scholarship.
References

"The Bomb". *Time*. 20 Aug 1945


Liefeld JP. The Relative Importance of the Size, Content and Pictures on Cigarette Package Warning Messages. Department of Consumer Studies, University of Guelph, Prepared for Health Canada, 1999


Ruiter RAC, Kok G. Saying is not (always) doing: cigarette warning labels are useless. *Eur J Public Health* 2005;15:329


