

Discourse Processes in Online Behavioral Advertising

by

Lisa Dawn Harper

December 2013

Presented to

Science, Information Arts and Technologies

University of Baltimore

In Partial Fulfillment

of the requirements for the degree of

Doctor of Science, Information and Interaction Design

Approved by:

Advisor: Kathryn Summers

Committee Member: Deborah Kohl

Committee Member: Eric Roberts

Copyright © December 2013 Lisa D. Harper: *Discourse Processes in Online Behavioral Advertising*. All rights reserved.

ACKNOWLEDGEMENTS

Acknowledgements here.

PREFACE

This dissertation sits at the nexus of linguistics, psychology, and interaction design. Ostensibly, the topic centers on user confusion as it relates to certain practices in online behavioral advertising. While linguists have studied press, television, and radio advertising, behavioral and interactive advertising stand on new ground yet unexplored. Advertisers are no longer restricted from a broadcast-style communication with consumers in ads, but are silently watching behavior and driving interactions in close to real time. From this perspective, advertisers are influencing thought through an interactive medium. If this is true, there is more to Internet language than David Crystal surmises in his 2006 book on the topic. And interaction designers should become as attuned to the effects a user interface may have on comprehension as they are to other sorts of cognitive effects.

ABSTRACT

With every generation of new media, advertisers find new ways to persuade and manipulate consumers to buy. Online behavioral advertising is concerned with fine-grained targeting of consumers by examining and tracking their online behavior over time. The ecosystem of behavioral advertising contains both interactive ads as well as mechanisms designed to enhance and protect privacy. Despite efforts by the Federal Trade Commission (FTC) and Advertising Industry to make privacy controls more transparent and easy to use, there is much confusion about how these work.

The aim of this thesis is to prove that certain problems and confusions that users face interacting with technology associated with online behavioral advertising can best be understood as problems in discourse understanding. Through the use of quantitative methods, I demonstrate that users are subject to the same context-dependent aspects of meaning that arise with the use of language.

With the rise of new interactions fueled by behavioral advertising, I suggest that advertisers may look toward exploiting new opportunities for manipulating meaning and understanding. Because the same cognitive architecture that supports language understanding also supports and facilitates understanding of user interfaces, discourse theory helps illuminate effects that small changes in context may have on user understanding during the course of interaction.

CONTENTS

1	INTRODUCTION	2
1.1	Mis-Communication in Interaction Design	2
1.2	Discourse Understanding	4
1.3	Confusion in Online Behavioral Advertising	4
1.4	Empirical Study	5
1.5	Thesis Aim	6
2	ONLINE BEHAVIORAL ADVERTISING	8
2.1	Intimate Capital	8
2.1.1	Science in Advertising	9
2.1.2	The Power of Data	11
2.1.3	The Rise of Behavioral Advertising	12
2.1.4	Tracking Users	14
2.1.5	The Nature and Use of Behavioral Data in Advertising	15
2.1.6	New Markets	20
2.2	Privacy and Policy Issues	24
2.2.1	The Internet is Free	24
2.2.2	Privacy and Personally Identifiable Infor- mation (PII)	25
2.2.3	Growing Public Awareness	27
2.2.4	Third-Party Data Collection	28
2.2.5	Cookies and Privacy Law	31
2.2.6	Policy Legislation - Do Not Track	32
2.2.7	Building Privacy Policies into Browser Pro- tocols	34
2.2.8	Choice and Self-Regulation	38
2.2.9	Opt-In Versus Opt-Out	39
2.3	User Confusion	40
2.3.1	Confusion About Third Party Tracking	40
2.3.2	Confusion about Defenses	43
2.3.3	Are Privacy Concerns Affecting User Be- havior?	45
2.4	The Future of Interactive Advertising	47
2.5	Summary	49
3	THEORETICAL FRAMEWORK	51
3.1	Pragmatics	52
3.2	Language in Interaction	55
3.3	Discourse	59
3.3.1	Discourse Continuum	60

3.3.2	Discourse Models	61
3.4	Applications	63
3.4.1	Graphics Communication	63
3.4.2	Advertising and Language	66
3.5	Conclusion	71
4	METHOD	73
4.1	Research Program	73
4.2	Design of Experiments	74
4.3	Mechanical Turk	77
4.3.1	The AMT Marketplace	78
4.3.2	General Concerns with the Use of AMT for Experimental Research	81
4.4	General Procedure	84
4.5	Data Storage and Privacy	88
4.6	Population Sample	89
4.7	Validity	90
4.7.1	Internal Validity	90
4.7.2	External Validity	91
4.7.3	Construct Validity	91
4.8	Summary	93
5	STUDY ONE: MODAL DIALOG BOXES	94
5.1	Motivation: Privacy as User Choice	94
5.2	Choice Design Problem	95
5.3	Forced Choice	96
5.4	Related Research	96
5.4.1	Heuristics and Biases	96
5.4.2	Pragmatic Reasoning	98
5.4.3	Deontic Reasoning and Knowledge	99
5.4.4	Content Affects Reasoning	99
5.4.5	Graphical Implicature	100
5.5	Pre-Pilot	101
5.5.1	Method and Procedure	101
5.5.2	Results and Discussion	102
5.6	Aims	103
5.7	Experimental Design	103
5.8	Method	104
5.8.1	Experiment 1A Method	104
5.8.2	Experiment 1B Method	107
5.9	Data Collected	111
5.10	Results	111
5.10.1	Experiment 1A Results	111
5.10.2	Experiment 1B Results	114
6	STUDY TWO: HYPERDEIXIS IN ADVERTISEMENTS	117

6.1	Motivation	117
6.1.1	Advertiser Self-Regulation and AdChoices	118
6.1.2	Effectiveness of the AdChoices Icon	119
6.1.3	AdChoices and Confusion	120
6.1.4	What Makes a Good Ad Design?	120
6.2	Related Research	122
6.2.1	Salience	122
6.2.2	Deixis	124
6.2.3	Indexicality and Knowledge	125
6.2.4	Demonstrative Reference	125
6.2.5	Hyperdeixis	126
6.2.6	Usability of Hyperlinked Icons	129
6.3	Pre-Pilot	129
6.3.1	Method and Procedure	129
6.3.2	Results and Discussion	130
6.4	Aims	136
6.5	Experimental Design	136
6.6	Method	138
6.6.1	Experiment 2A Method	138
6.6.2	Experiment 2B Method	141
6.7	Data Collected	142
6.8	Results	143
6.8.1	Experiment 2A Results	143
6.8.2	Experiment 2B Results	144
7	STUDY THREE: MULTIPARTY DISCOURSE IN THE BROWSER	146
7.1	OBA is “Creepy”	146
7.2	Related Privacy Research	147
7.3	Related Linguistic Resesearch	149
7.3.1	The Stateful Web	149
7.3.2	A Model of MultiParty Interaction	150
7.3.3	Modeling the Hearer	152
7.4	Aims	154
7.5	Experimental Design	154
7.6	Method	155
7.6.1	Settings and Participants	155
7.6.2	Procedure and Materials	155
7.6.3	Instrumentation	157
7.7	Data Collected	158
7.8	Results	158
8	DISCUSSION	162
8.1	Summary of Findings	162
8.2	Discussion of Results	164
8.2.1	Experiment 1: Modal Dialog Boxes	165

8.2.2	Experiment 2: Embedded Icons in Graphical Ads	167
8.2.3	Experiment 3: Third-Party Bystanders Versus Eavesdroppers	168
8.2.4	Use of AMT	169
8.2.5	Potential Population Bias	170
8.3	Cognitive Architecture	170
8.3.1	Situation Models	170
8.3.2	Interactive Alignment and Priming	171
8.4	Themes for Future Areas of Study	173
8.4.1	Micro-Interactional Design Patterns	173
8.4.2	Interactive Ads	179
8.4.3	Display Adaptation	179
8.4.4	Social Microinteractions	180
8.4.5	Socially Aware Browsers	182
8.5	Conclusion	184
A	PARTICIPANT CONSENT FORM	186
B	AMT TERMS OF SERVICE	187
C	SAMPLE AMT ASSIGNMENT DEFINITION	188
D	SURVEY RESPONSE	189
E	EXPERIMENT 1 CONDITIONS	193
F	EXPERIMENT 1 RAW RESULTS	196
G	PARTICIPANT SURVEY	198
	BIBLIOGRAPHY	199

INTRODUCTION

All forms of communication entail design, as the intent of communication is to be understood by others or by one's self at another time. Communication design, then, is inherently social, because to be understood by another or by self at another time entails fashioning communications to fit the presumed mental states of others or of one's self at another time. (B. Tversky, 2010, p. 6)

1.1 MIS-COMMUNICATION IN INTERACTION DESIGN

Good interaction design takes into account user cognitive processes and limitations. Though designers spend considerable effort toward producing user interfaces and interactions that communicate effectively, the tools that they use are flexible can easily bend to cause users to make incorrect inferences about the situation at hand. Because graphical users interfaces (GUIs) share properties with other forms of communication such as language, gesture, diagrams and action — and are inherently interactional — *I argue that they are also subject to the same context-dependent aspects of meaning that arise with language use.* The same cognitive architecture that supports understanding of language also supports and facilitates understanding of user interfaces. This observation forms the basis for claims made in this dissertation.

In traditional GUI-based interaction, user interactions are designed. When a user interacts with a website, that interaction is structured according to that site designer's plan. Ideally, information has been organized with user expectation in mind. The designer has considered the sorts of tasks users intend to perform. Sometimes, navigation is guided. Sometimes not, but designers leave behind contextual cues and guideposts to aid users in their task.

In fact, it is easy to influence users to act in predictable ways. For example, A. Tversky and Kahneman (1981) showed that predictable, dramatic shifts in preference could be generated simply by changing the ways options are framed: in

early framing experiments, subjects showed a preference toward risk aversion when lotteries were framed as gains, and risk seeking when lotteries were framed as losses.

As noted by Thaler, Sunstein, and Balz (2010), designers have the potential for great control over user decision-making by manipulating behavior using psychological tools described by psychologists and behavioral economists. However, though Kahneman and Tversky (1984) and others have noted that small changes in context affects decision-making, there is considerably less research on how *small changes in context affect understanding in the context of decision-making*.

Language operates using the same principles and processes as other cognitive functions. It is not surprising to learn that language can be used to manipulate hearer understanding. For example,

- (1) The Sheriff caught Robin red-handed. He is now serving time in prison.

Sentence 1 above appears to convey the following facts that weren't explicitly stated:

- (1a) It is Robin who is serving time in prison.
- (1b) Robin is serving time in prison as a result of being caught red-handed by the Sheriff.

These un-stated propositions (1a) and (1b) are known as *implicatures*, a well-known pragmatic phenomenon in language understanding. In fact, depending on situational context, either of these implicatures may not be true.¹ Implicatures operate under principles such that we may presume their truth even if un-stated. But if the truth value of either implicature (1a) or (1b) turns out to be false, this does not mean that (1) is false.

Because what we know is so often derived by inference, it is easy to see how language can be used to manipulate understanding. But is it possible for GUIs to communicate in the same sort of way? I believe so. This dissertation will show that GUI-based mis-communications may be brought about by faulty inferences.

¹ Imagine you are watching a humorous movie and (1) is a caption. The Sheriff caught Robin Hood returning illegal taxes back to the citizenry. Later, the Sheriff was jailed for fraud. Though we might ordinarily expect the second sentence in this example to have a causal relation to the first, the visual scene serves as a contextual backdrop cancelling the implicature.

1.2 DISCOURSE UNDERSTANDING

Discourse is generally said to encompass language use beyond the bounds of a sentence or proposition. It is often associated with the study of pragmatics: “Syntax studies sentence. Semantics studies propositions. Pragmatics is the study of linguistic acts and the contexts in which they are performed” ([Stalnaker, 1999](#), p. 34). Pragmatics attempt to account for meaning of messages where actual meaning is underdetermined from what is said.

Pragmatics is fundamentally concerned with reasoning processes that go beyond conventional meaning to interpretation in social, situational, and belief contexts. As such, it is founded on the notion of language as action with communicative goals ([H. H. Clark, 1996](#); [Grice, 1975a](#); [Levinson, 2000](#)). In this view, “hearers” infer a “speaker’s” meaning on the basis of evidence provided.

Discourse is also studied by sociologists, anthropologists, and sociolinguists: it is bound to socio-cultural knowledge and governed by social norms ([Gumperz, 1982](#); [Hymes, 1974](#)). Even so, discourse communication is seen as structured: both in terms of conversational events and learned, ritualistic schemas ([Goffman, 1981](#), e.g., making a reservation). In this tradition, conversational inferences are also context-bound, but conceived more generally as preferences, maxims, or tendencies.

1.3 CONFUSION IN ONLINE BEHAVIORAL ADVERTISING

Of concern to nearly everyone on the Internet today is privacy. Polls conducted over the last decade indicate that the majority of Internet users report that they do not wish to be tracked across sites ([TRUSTe, 2012](#)). At the same time, Internet advertising has entered a boom for mass data collection and predictive analytics. While it may be in the best interest of consumers to restrict online data collection and tracking, in practice, the legal system has been unable to keep pace. Even so, businesses that engage in online behavioral advertising (OBA) commonly employ techniques that influence user behavior by manipulation. The goal of advertisers, after all, is to influence consumers and sell products.

Pertaining to OBA, the following sorts of questions are addressed in this dissertation:

- Why might users say privacy is important but not choose to block cookies when given a choice?
- Why is the Interactive Advertising Bureau (IAB) AdChoices opt-out icon confusing to so many users?
- Why might users, who acknowledge that the Internet is not private, behave as if it is? Such questions relate to user belief and decision-making in the context of website interactions on the Internet.

Though advertising language has been studied in the *content* of ads (Cook, 2001; Geis, 1982; Harris, 1983; Leech, 1966; Tanaka, 1999; Vestergaard & Schröder, 1985a), the study of how advertisers might influence users during the course interaction has not yet been addressed.

1.4 EMPIRICAL STUDY

In this dissertation, I use randomized controlled experimentation to observe user behavior under specific conditions. Each experiment asks: is there mis-understanding and is such confusion pragmatic in nature?

The first experiment hypothesizes the interpretation of **implicature** in “do not track” dialog boxes. Such dialog boxes are often presented as modal dialogs, thus, as an unforced “yes-no” choice design. In experiment 1A, I examine whether graphical representations of choice in this context evoke implicature in the same manner as textual representations. Following this, experiment 1B considers what people believe is the consequence of “no choice” in task-based interaction.

The second experiment focuses on the role of **deixis** in hyper-linked ad images. The Interactive Advertising Bureau (IAB) AdChoices icon associated with behavioral advertisements is designed to give users the means to control behavioral tracking preferences. However, relatively few people notice such icons,² let alone click them (Evidon, 2011; Turow, 2012). Blase Ur (2012); Leon, Cranshaw, et al. (2012) describe communicative flaws with AdChoices icon. While they considered the effect of iconic versus textual communication, they only speculate on causes for confusion. Experiment 2A hypothesizes that knowledge plays a role in whether an icon

² This is based on a national wireline and cell telephone survey of 1,503 adult Americans during April and May, 2012 (Turow, 2012).

is taken to be indexically linked to a website. Experiment 2B considers whether familiar iconic representations affect how users interact with advertisements in context.

Finally, Experiment 3 focuses on the pervasive phenomenon of third-party tracking in the browser itself. In this study, user interaction on the Internet is analyzed as a form of multi-party discourse leading to issues in **conversational inference** where the user does not infer the presence of unrati-fied participants monitoring web interactions. Manipulating the user's perceptual awareness of third party participants is hypothesized to have an effect on user behavior. More broadly, this experiment attempts to show that user behavior, while answering sensitive questions, is affected in the presence of visible observers. I hypothesize that their propensity for disclosure of sensitive information is reduced under such conditions.

In each experiment described above, I show how advertisers may either benefit from poor design decisions or deliberately mis-lead users through the manipulation of linguistic understanding. Though the context for experimentation is situated in the domain of online behavioral advertising, the method for analysis is general in nature.

1.5 THESIS AIM

The aim of this thesis is to prove that certain problems and confusions that users face interacting with technology associated with online behavioral advertising can best be understood as problems in discourse understanding. Using quantitative methods, I demonstrate that users are subject to the same context-dependent aspects of meaning that arise with the use of language. With the rise of new interactions fueled by behavioral advertising, I argue that advertisers may look toward exploiting new opportunities for manipulating meaning and understanding.

This thesis does not intend to catalog the full range of discourse phenomena extant in user interaction — nor does it attempt to solve the problems described. It also doesn't intend to cover the entire range of phenomena within the context of OBA. The intent is to reveal how a linguistic analysis may account for user confusion not fully explained otherwise. Such confusions emphasize the need for designers to consider the importance of discourse inference in GUI-based interaction.

Chapter 2 introduces background literature on the topic of online behavioral advertising. This chapter lays the foundation for analyzing observed phenomena under an inferential model of communication. Chapter 3 describes the discourse-theoretic framework necessary for testing hypotheses about pragmatic processes. Chapter 4 details a general method for study. Chapters 5, 6, and 7 each describes an experiment addressing a particular problem. Chapter 5 is concerned with whether implicature can be observed in choice processes via a non-forced choice modal dialog interaction. Chapter 6 considers knowledge and the role of deixis in hyperlinked ads. And Chapter 7 highlights the role of conversational inference in an experiment designed to ascertain whether user behavior changes in the presence of observers. Finally, Chapter 8 presents a discussion of results and offers direction for future research. It is my belief that this thesis lends deeper insight into communicative challenges for user interaction design and bridges these to sound theoretical tools for addressing a specific class of problems previously unidentified.

2

ONLINE BEHAVIORAL ADVERTISING

This chapter provides an introduction to the concept of online behavioral advertising (OBA). This domain is the backdrop against which experimental hypotheses in this dissertation are tested. Since I began writing this dissertation, OBA has grown from relative obscurity to notoriety. With the rising prominence of mobile browsing, cookie tracking is predicted to disappear within a few years. The surge of public awareness of OBA has recently led to new default settings in some browsers where third-party cookies are blocked by default.¹ Though particular issues discussed in this chapter will quickly stale, research in this dissertation should not as it relies upon theoretical tools and methods generalizing to phenomena not yet observed.

This chapter is organized in the following way. First I discuss the phenomenon of OBA and circumstances that led to its meteoric rise. Second, I introduce privacy issues in terms of how data is collected and used by third party advertisers. This has played an important role in shaping policy, leading to the design of interactions examined in this dissertation. Therefore, I follow by outlining the scope of user confusion as observed by privacy researchers. Finally, I conclude by summarizing the sorts of interactions behavioral advertising has engendered and speculate on what we may see in the future. Advertising is becoming increasingly becoming less of a broadcast art and more a personalized one. For this reason, we need to take a closer look at how advertisers may manipulate our thinking and beliefs during the course of interaction.

2.1 INTIMATE CAPITAL

A new market-driven ecosystem of targeted advertising has emerged, spanning the divide of Internet and brick-and-mortar

¹ Mozilla and Microsoft have moved to this in the past year, while Apple Safari has blocked cookies by default for some time. Meanwhile, Microsoft, Google, and Apple are all exploring new technologies amenable to tracking across platforms to include mobile, gaming, and video services (D'Orazio, 2013).

business. This ecosystem is fueled by the unprecedented availability of data, algorithms, storage, and computing power. The forbidden fruit, and yet golden chalice, is the ability for the advertiser to know the target intimately. Of serious concern outside the advertising industry is that behavioral data may be combined with data owned by other content publishers or advertisers to actually identify a particular user. The more an advertiser knows the user — and not just as a demographic profile but as an individual — the more fine-grained the targeting. So it would not be surprising to learn that publishers and advertisers treat information about consumers as a form of capital — an “intimate capital... likely to be worth something to others” (Locke, 2010, p. 127). This section focuses on the sorts of data advertisers track and how this data is used in behavioral advertising. The availability of such data, coupled with advances in computing technology have given rise to new market economies struggling to maximize a return-on-investment (ROI) while the very measures defining ROI and ad effectiveness are in flux.

2.1.1 *Science in Advertising*

The quote below from a 1901 article in Publicity seems almost prescient by today’s standards:

The time is not far away when the advertising writer will find out the inestimable benefits of a knowledge of psychology. The preparation of copy has usually followed the instincts rather than the analytical functions. An advertisement has been written to describe the articles which it was wished to place before the reader; a bit of cleverness, an attractive cut, or some other catchy device has been used, with the hope that the hit or miss ratio could be made as favorable as possible. But the future must needs be full of better methods than these to make advertising advance with the same rapidity as it has during the latter part of the last century. And this will come through a closer knowledge of the psychological composition of the mind. The so-called ‘students of human nature’ will then be called successful psychologists, and the successful advertisers will be likewise termed psychological advertisers. The mere mention of psycholog-

ical terms, habit, self, conception, discrimination, association, memory, imagination and perception, reason, emotion, instinct and will, should create a flood of new thought that should appeal to every advanced consumer of advertising space (as cited in [Scott, 1904](#)).

That these concepts should transfer so easily to the 21st century Internet is startling: what held true in advertising more than 100 years ago still rings true today.

In 1957, when Vance Packard wrote the runaway best seller *The Hidden Persuaders*, he exposed advertising not as a huckster bag of tricks, but as a subtle and calculated science with deep roots in psychoanalysis, sociology, and ethnographic anthropology. He honed in on a group of psychologists known as “the depth boys” who believed that to understand the consumer, you needed to find out what they really wanted at an unconscious level. In previous decades, advertisers had found little success interviewing and asking people what they wanted in a product. From *Advertising Age*, “In very few instance do people really know what they want, even when they say the do” (as cited in [Packard, 1957](#), p. 37). According to Packard, what marketers and psychologists had been learning is that what people tell interviewers has little bearing on how they would actually behave. “What you are more likely to get, they decided, are answers that will protect the informants in their steadfast endeavor to appear to the world as really sensible, intelligent, rational beings” ([Packard, 1957](#), p. 35). What advertisers ultimately learned from motivational research was how to find psychological and emotional levers that would generate an affinity to a particular product over a myriad of close alternatives — and how to use those levers to trigger an action to buy.

Recently, author [Duhigg \(2012b\)](#), *The Power of Habit: Why We Do What We Do in Life and Business*, described advances in the brain and behavioral sciences relating specifically to habit and learning. Habits, he says, are essentially a mechanism by which the brain encodes basic behaviors (in behavior chunks) in order to improve the efficiency of our brains. Marketers and advertisers, he says, leverage a three-step habit loop in order to reach “in market” customers and make sales. In this process, there is a *cue*, or trigger, for a particular pattern or behavior chunk. As [Duhigg \(2012b\)](#) relates, the basal ganglia plays a central role in recalling patterns and acting on them

as stored habits. A stored habit, or *routine*, which can be physical, mental or emotional, is linked to both the cue and subsequent *reward*. The reward itself plays a vital role in remembering such that a habit becoming “intertwined until a powerful sense of anticipation and craving emerges” (Duhigg, 2012b, p. 19). To no surprise, habit plays a central role in market strategy and advertising success. But the tools with which advertisers discover habits and influence habit loops, have become quite sophisticated.

2.1.2 *The Power of Data*

Discovering and tracking customer behavior is not new. Brick and mortar institutions have been doing this for the last century using purchase history to develop and refine models as well as coupons and mail to target specific customers. What has changed is technology. When Netscape engineer Lou Montulli invented the cookie and wrote the original specification, he was solving a state management problem in web applications (Kristol, 2001). Using cookies, websites could remember users and pages they had visited. Moreover, cookies were simply part of a browser protocol hidden from the user. Between 1994 and 2000, working groups were established in the Internet Engineering Task Force (IETF) centered on standards for browser state management: cookies had become the central mechanism by which all browsers managed state. By this time, America Online was a huge presence on the Internet and many businesses scurried to establish an online presence.

In 2000, Google began selling advertisements associated with search keywords (Google, 2010). Google also claimed the world’s largest search index — exceeding one billion pages. By 2008, it had reached a trillion pages: it had solved a myriad of technical problems to reach this scale (Google, 2010). Chiefly, through innovations in computer hardware and processing, Google designed a network infrastructure capable of supporting the sharing of trillions of pages of web content, as well the means to aggregate and index that content. One such innovation was the means to process very large data sets using a model of distributed computing on large clusters of computers. Concurrently, advances were made in algorithm development, particularly in the area of data mining and machine learning.

Even as technical advancements have lead to innovation at the corporate level, so they have became more accessible to small companies and individuals. "I can get a hundred machines and I can sequence my own DNA for \$150, or analyze the rise of Justin Bieber links across the Web over the last two years really quickly," says Hilary Mason, chief scientist at bitly (as cited in Woods, 2012). "If we didn't have that kind of commodity access to computer power and commodity access to analytics tools, we wouldn't be able to do the things we're able to do, and we certainly wouldn't be able to do them at startups with small budgets" (Woods, 2012).

Thus, in the past few years, there has been a surge of industry interest in the burgeoning field of "data science." According to Mason, data science is a blend of analytics ("counting things") and statistical machine learning algorithms. And a really good data scientist is a master at asking the right questions (H. Mason, 2012). In the world of advertising, its not surprising that the "right" questions are deeply concerned with understanding people.

2.1.3 *The Rise of Behavioral Advertising*

In the early years of the Internet, web advertisements were characterized by large banner ads and ad spaces reminiscent of print news and magazine media. A publisher sold space (inventory) on its site to advertisers, who filled that space with banner (or pop-up) ads. Cookies quickly changed the landscape of web advertising by making banners clickable and trackable, and this basic form of online advertising remains prevalent today. However, cookies also made it possible for advertisers to target ads to particular users, user agents, and devices. This has stimulated third-party (non-publisher) tracking and has led to changes in the economy and format of advertising, as well.

Though traditional advertising revenue via newspapers, television, radio and cable have been steadily dropping for several years, Internet advertising revenues continue to grow (IAB, 2012). In 2011, the Interactive Advertising Bureau (IAB) reported Internet advertising revenues of 31.74 billion, up 21.9% from 2010, and increasing at a linear rate from 2002 (IAB, 2012).

In terms of ad formats, search revenue in 2011 accounted for 46.5% of the total, with display advertising at 34.8%. Mo-

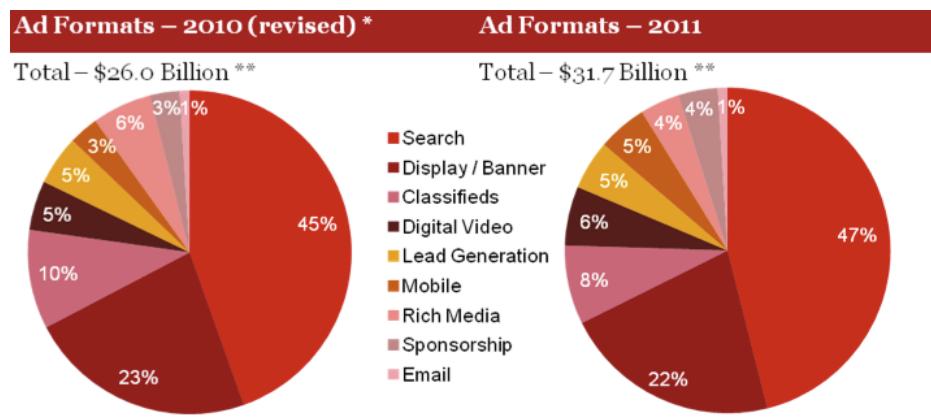


Figure 1: Revenue According to Ad Format (Image credit: [IAB, 2012](#))

bile has emerged as a relevant category at 5%, while classifieds, lead generation, and email account for the remaining amount ([IAB, 2012](#)).

Because advertisers are most concerned with efficiently reaching high value consumers (those most likely to buy), targeting is important across all forms of advertising. On the surface, targeting is about breaking the market down into groups, or segments, which share certain characteristics. Typical market segments includes geographic (e.g., region, climate, urban / suburban / rural), demographic (e.g., age, gender, ethnicity, education), contextual (e.g., web page content) and behavioral (e.g., browsing history, content accessed over time). Ideally, advertisers target an advertisement to users most likely to be influenced by it. To accomplish this, advertisers attempt to collect and correlate information about users in order to best segment them.

Intrinsically, targeting seems linked with profiling and the potential for the accumulation, aggregation, and storage of personal data. Advertising networks claim that such information is anonymous, while consumers and policy makers struggle with the question of whether advertising networks should be entitled to assemble profiles and whether that should entail gaining user permission first. In this balance of this section of Chapter 2, we will put aside privacy concerns and instead focus on:

1. What it means to track users and why this is important;
2. The nature and use of behavioral data in advertising; and,

3. New market opportunities and potential effects.

2.1.4 *Tracking Users*

Defining “tracking” is a contentious exercise. The World Wide Web Consortium (W3C) Tracking Protection Working Group has been struggling with a definition for over a year (W3C, 2012). But, in its broadest sense, tracking is the collection, correlation, or transfer of data about Internet activities of a particular users, user agent, or device. Tracking may or may not be consensual and the users may or may not be aware of who is tracking, how tracking occurs, or what data is tracked. Tracking in the browser is most commonly associated by the use of HTTP cookies, but other means include IP address tracking, flash cookies (using the Adobe Flash plug-in), and web bugs or beacons (e.g., images retrieved from a third party website). Of potential future concern may be browser fingerprinting (Eckersley, 2010), though it does not seem that advertisers are using this technique now.

Publishers and advertisers alike track user data. Publishers, or first-party, tracking may include very fine-grained information about a user such email address, name, navigation behavior, etc. Personally identifiable information (PII) may be collected with user consent. And companies with both a brick-and-mortar and significant online presence likely combine information about what they know from shopping habits in the store with other data obtained online. More recently, even purely online businesses are able to tap offline data. For example, FaceBook has recently partnered with Datalogix to link online and offline consumer data (Reitman, 2012).

First party tracking may also be combined with data collected and aggregated by third party trackers. Third party trackers do not communicate directly with a user but monitor a user’s actions in other ways. Generally, third party trackers are explicitly given consent to track by a website publisher. But they may also acquire user data via data exchanges and from Internet Service Providers (ISPs; e.g., via deep packet inspection such as that described in Sesto & Frankel, 2008) or other vendor sources. It is also possible for trackers to acquire data by taking advantage of security vulnerabilities and via information leakage (as described in Krishnamurthy & Wills, 2009, for example).

2.1.5 *The Nature and Use of Behavioral Data in Advertising*

In fact, how behavioral tracking is accomplished is anything but transparent. The image below in [Figure 2](#) depicts the many HTTP requests called from the New York Times (NYT) homepage in November of 2012.

A large number of the domains included are third party trackers engaged in analytic or personalization services and advertising. But what makes behavioral tracking feel so insidious is that user activity on one site *can directly influence content displayed in completely un-related sites*.

For example, on October 5, 2012, after visiting the NYT, Washington Post, and Mozilla home pages, I generated a search query for “waterproof boots” on Google search. On October 6, 2012, I refreshed the NYT homepage and was presented an ad for Marc Jacob waterproof boots prominently displayed in the top banner ad space in [Figure 3](#) below. This may or may not be co-incidence. There is no way to know.

A graph diagram generated from the Firefox Collusion extension in [Figure 4](#) offers few clues about how my Google query might have caused this ad to be displayed on the NYT homepage.

Possibly, Google’s advertising subsidiary DoubleClick may have been involved. When seen in the context of this collusion diagram, DoubleClick appears to know of visits to the Washington Post and NYT, but not what search query was made. Yet this is still possible. Business relationships outside of the communicative context also account for data transactions and exchanges.

Behavioral tracking concerns technologies and methods centered on capturing user data when users interact with web content. For example, information captured may include: user visits to a web site, specific page content, visit recency / frequency of visits, links clicked, searches, form data, and other interactive content. This data, plus metadata such as IP address may be combined to create a ‘profile’ linked to that user, browser user agent, or device. The goal of behavioral targeting is to show ads only to users of high value (those who are likely to buy the product) combined with a large number of opportunities to show such ads to that user.

An informal survey of agencies by the IAB suggests that behavioral advertising is widespread: up to 80% or more of advertising campaigns conducted in 2009 involved some form

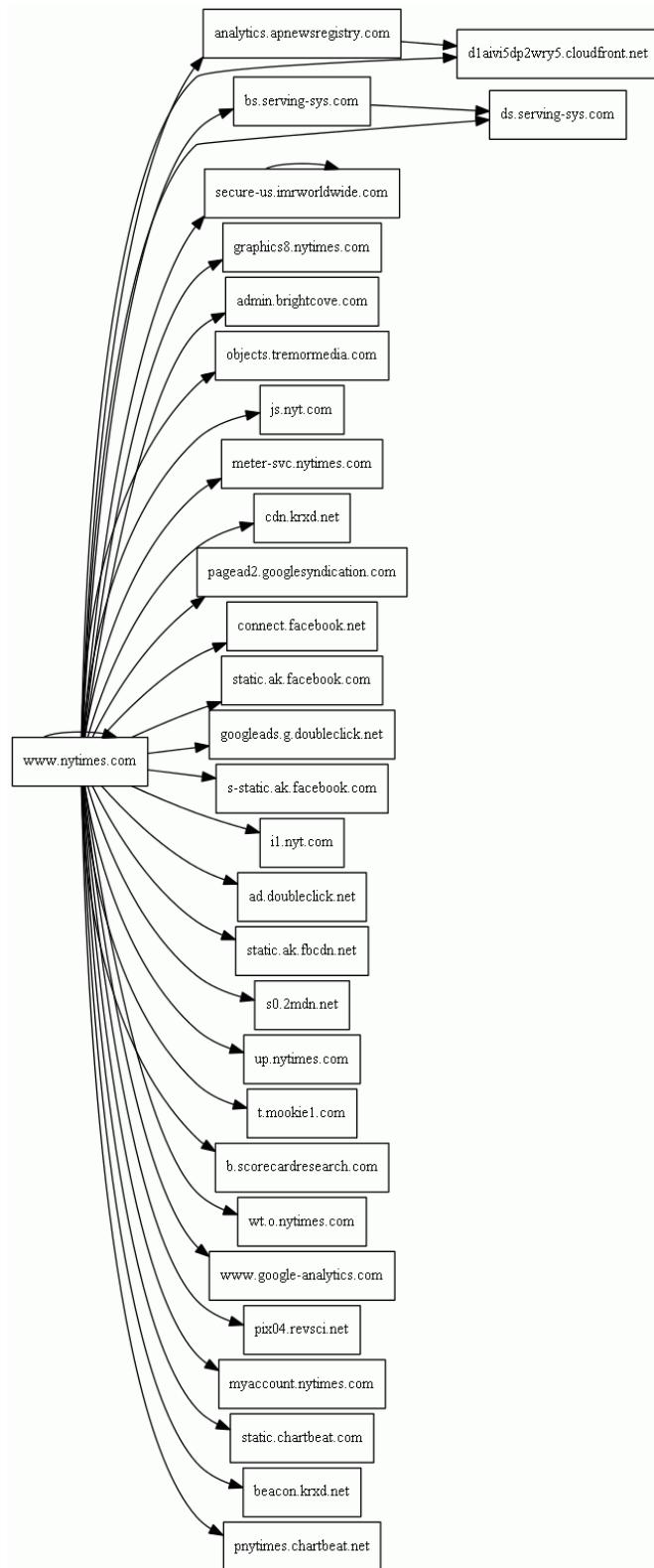


Figure 2: HTTP Requests from the New York Times Homepage



Figure 3: New York Times Homepage with Advertisement for Marc Jacob Boots

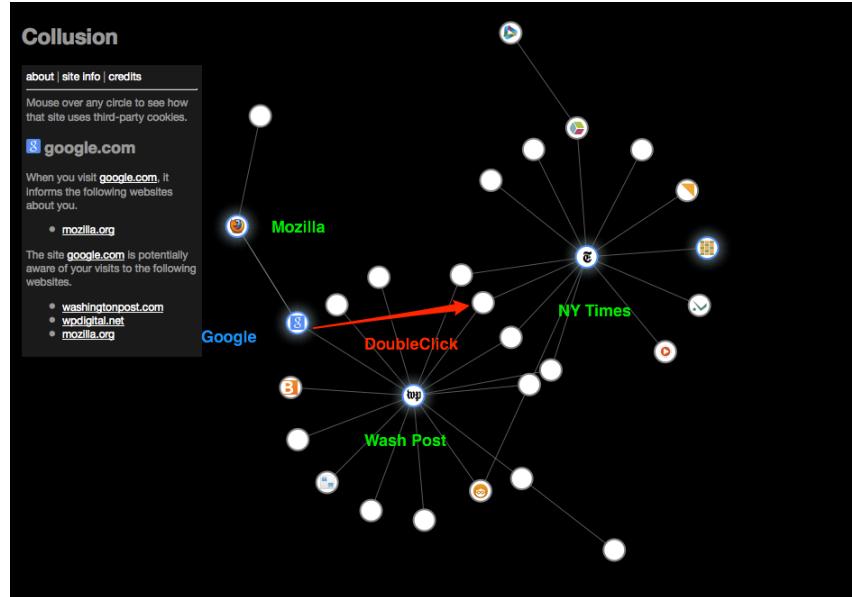


Figure 4: Firefox Collusion Diagram

of tracking (IAB, 2010). But until recently, we've had little real insight into how businesses, to include advertisers, used behavioral tracking.

In Feb 2012, The New York Times published an article written by Duhigg (2012a) telling the story of how Andrew Pole, a data scientist in Target's marketing analytics department, developed a model that could successfully predict whether a female customer was in the second trimester of pregnancy.

According to Duhigg (2012a), there are periods in life when routines and buying patterns change. Advertisers are quick to target customers during major life changes such as the purchase of a house or vehicle, or birth of a child. Retailers are particularly interested in acquiring and retaining new customers during such life changes when habits are more malleable. In the case of a birth event, parents will buy all sorts of items such as maternity clothing and prenatal vitamins — transitioning to baby care products soon after. Pole began analyzing consumer spending to identify 25 products, in aggregate, predictive of pregnancy and within a narrow window of time. As Duhigg (2012a) notes, "it's difficult to attribute how much behavior modeling contributed to Target's revenue, but between 2002 — when Pole was hired — and 2010, Target's revenues grew from \$44 billion to \$67 billion" (Duhigg, 2012a). Beyond anecdotal accounts such as pregnancy prediction by Target, there is relatively little information available about how businesses and advertisers profile users through online tracking.

Not all businesses have such fine-grained information as customer data to enrich predictive models. While Target may have access to name, address, demographics, geography, contact history and many other pieces of information (Pole, 2010), search engines typically have access to different sorts of data. In Yan et al. (2009), a team from China reportedly published the first academic, and empirical study, to address the question of the whether online behavioral targeting (OBT) can help in online advertising. In this paper, a basic OBT assumption was addressed. The basic assumption is this: "users who have a similar search or browsing behavior will have similar interests and thus have higher probability to click the same ad than the users who have different online behaviors" (Yan et al., 2009).

The team used seven days' ads click-through log data from a commercial search engine recording user search click be-

havior to include both web page clicks and ad clicks. They represented user behavior by page views and created a behavioral profile by considering all terms appearing in a user's query as previous behaviors. Both users and queries were represented as numerical vectors so that similarity between users could be easily calculated. Using simple clustering techniques, users were segmented according to behavior. Finally, user within-ad similarity was compared with user between-ads similarity. The result was that the within ads similarity, represented by user search behavior, was around ninety times larger than between ads similarity. Thus [Yan et al. \(2009\)](#) concluded that users with similar search behavior would indeed be distinguishable from other users and this could be used to predict the clickability of an ad.

[Chen, Pavlov, and Canny \(2009\)](#) described "massive improvements" in the ability to do offline training of OBT models. They noted, "behavioral data is intrinsically in large scale (e.g., Yahoo! logged 9 terabytes of display ad data with about 500 billion entries on August, 2008)" with a sparse click through rate (CTR) of about 0.05% ([Chen et al., 2009](#)). To build the entire 450 OBA category models from Yahoo! on fine-grained (ad clicks and search queries) at this scale would take about a week before the innovations they describe were implemented. By using a MapReduce learning algorithm, improved feature vector generation algorithm, better in-memory caching, and more efficient data structures and models they were able to reduce offline model building to about one day.

Scarcely a year later, [Pandey et al. \(2011\)](#) demonstrated the ability to include even more behavioral context to achieve, what they describe, as better results on live data. They created a general purpose model which allowed for optimizing three strategies of behavioral targeting: property (document context), segment (user demographics), and behavior (user past behavior). Each strategy was estimated to successively encode deeper context, thereby potentially improving the richness of the model. User events were modeled as an event stream of three types of events: pages visited, search queries, and graphics ads. By modeling in this manner, the research team examined the relative performance of specific event types (using extracted features), within specific temporal windows with respect to a variety of advertising campaigns of various types and sizes. In a live experiment spanning three months, they generated user models spanning eight weeks of user history.

For each of four ad campaigns, old models (like those described in the previous two paper described above) and new models received at least 1M ad impressions on a monthly basis. The conversion rate for new models was calculated to be considerably higher than for the older models, ranging from 57% to 264% higher.

Pandey et al. (2011) note, the problem of predicting clicks and predicting conversion have been split along clear lines of information available to publishers and information available to advertisers. There has been growing pressure from within the ad industry to change this. Moving to payment by conversion is an attractive option for advertisers because: 1) it helps to prevent click fraud; 2) can be used to analyze the effectiveness of the advertisement (was there an actual sales conversion and not just a click); 3) can prevent the user from being inundated by too many of the same ad; and, 4) and is compatible with re-targeting across sites (Pandey et al., 2011). Accordingly, advertisers have been recently more willing to share individual responses to ads to publishers since it facilitates the use of conversion-optimized models.

2.1.6 *New Markets*

The market of online advertising today is, in fact, much more complex than simple supply (publisher) and demand (advertiser) economics. This model, known as direct buy, was dominant in the early days of online advertising (Mayer & Mitchell, 2012). By the late 1990's advertising networks emerged allowing advertisers to place ads with many publishers — and publishers to work with many advertisers — through a common network. In such a model, it is easier for advertisers to target users along multiple dimensions (e.g., page context, demography, geography, behavior) for ad slots.

The advertising community as a whole sees targeting as an opportunity funnel (see Figure 5). Relating this back to psychology, advertisers are aware that to target effectively, then need to consider consumer:

1. Daily activities;
2. Online activities and habits; and
3. Research time, whether purchase is online or offline (Jupiter Research, 2010).

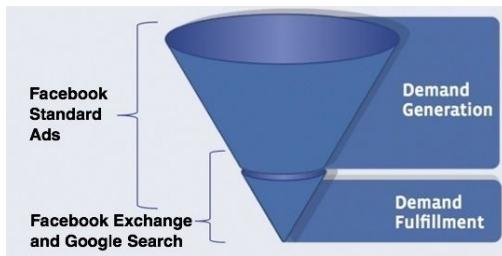


Figure 5: Facebook Re-targeting (Image credit: [Constine, 2012](#))

There is a clear “in market” time of opportunity when users are researching purchase decisions. During this period of time, their behavior changes in predictable ways. For example, initially, the user may broadly review a general product space. Then move to a more comparative or winnowing down of choices. And finally, that consumer will look for a store and make a purchase.

Cookie re-targeting takes advantage of this purchase funnel. Recently, Facebook announced a new Facebook Exchange program, a real-time bidding ad system, where visitors with exchange party cookies can then be shown ads related to their web browsing when they return from those sites to Facebook ([Constine, 2012](#)). The basic idea is to re-target customers that visited a commercial site but did not purchase at that visit. This effectively positions Facebook from working in the broad part of the funnel (demand generation) to the more narrow part of the funnel (demand fulfillment; [Constine, 2012](#)).

To do this, Facebook doesn’t share biographical information with advertisers, but takes cookies and combine them with Facebook data ([Constine, 2012](#)). The emergence of an *exchange market* — where buyers and sellers converge — is attributable to Google’s DoubleClick Ad Exchange in 2009 ([Duggal, 2012](#)). The perceived value of an exchange over an advertising network is that the exchange operates on the behalf of any number of buyers, sellers, and middlemen alike. [KruX \(2012\)](#) estimates that real-time bidding exchanges now account for 40% online data collection. This is up from 0% three years ago.

Another sort of emergent market is a data market. *Data management platforms* serve as “a unified technology platform that intakes disparate first-, second-, and third-party data sets, provides normalization and segmentation on that data, and allows a user to push the resulting segmentation into live interactive channel environments” ([O’Connell, 2011](#)). KruX

(2012) found that more than 300 companies collected data on its users, up from 167 companies in 2010.

Many data collectors also piggyback on each other. By piggy-backing, one tracker invites other trackers to the trough. Because of this, publishers and users may not be aware of how much data is tracked on a given site (Angwin, 2012).

Not surprisingly, it is difficult to measure the effect of behavioral advertising. Referencing a study from Kazienski and Adamski (2007), Farahat and Bailey (2012) observe that click-through-rates (CTR) have declined from 3% to less than 1% (since the 1990s). CTR have traditionally been a measure of ad performance, though ultimately, advertisers prefer a more concrete measure such as sales conversions. In the Yan et al. (2009) experiment described earlier, ad CTR improvements through behavioral tracking were measured as high as 670% on the basis of their simple user segmentation strategy from search queries and clicked pages.

However, Farahat and Bailey (2012) question these ad CTR results stating that any study that naively looks at response lifts between targeted and un-targeted groups will greatly overestimate the effects of advertising since there is an inherent selection bias: the targeted users' behavior is likely to be highly correlated with the measured response ("click"). Farahat and Bailey (2012) ask, how do you know that the targeted segment isn't also the most clicky? Since what advertisers care more about are sales conversions than clicks, this is an important question.

In a field experiment on the Yahoo! front page, Farahat and Bailey (2012) try to weed out selection effects in terms of clickiness of the targeted population, brand interest, and category interest. They find strong evidence that user clickiness and brand interest are determinants of variation in CTR. This has bearing on the projected ROI of behavioral advertising for a given campaign. Furthermore, they question the ad CTR improvements of Yan et al. (2009), suggesting that the targeted lift of brand-related searches may be over-estimated by almost 1000%.

Farahat and Bailey (2012) conclude that advertisers would benefit from comparing how targeted and non-targeted populations respond to advertising in any particular campaign in order to gauge the cost effectiveness of their targeting strategy. "Given an industry average of a three times price premium for targeting, we might conclude targeting is more cost ef-

fective, but of course this depends greatly on the targeting product" (Farahat & Bailey, 2012). To date, relatively little is published about proprietary behavioral tracking model development nor the effectiveness of such models in actual ad campaigns. Nonetheless, the value of behavioral data is vigorously defended by marketers in public media as "the fuel that drives the Internet".²

What we really have little understanding is how to deal with the socio-technological effects of the connectedness of many different sorts and bits of data across vast networks of inter-connected participants. Currently, the heavy emphasis is on *who* to target. But there is mounting research to suggest that *when* (in the purchase funnel), *where* (on what site and how the ad should be placed), *what time of day*, *at what geographic location*, *how often*, and *how much* to pay (via auction strategy) are also important to advertisers. Furthermore, with the market moving away from CTR (click through rate) and CTM (cost per mile / "impression") towards a sales conversion rate (CVR), there will pressure for even greater information flow between producers and advertisers. Impediments to behavioral tracking become no longer technological, but sociological. And of course, these barriers will fall if both producers and advertisers believe they can profit.

In purely statistically-driven systems, the amount of data, not the algorithm, is king. However, there are many examples of where social network topology provides information that dramatically improves performance over pure volume. When Google created the PageRank algorithm it used page outbound links as implicitly encoded information about what page creators thought important. So what might the effect be when social networks such as Facebook harness vast networks of inter-connected consumers to the problem of behavioral advertising? How likely are we to behave like our friends? Or to what extent can they be used to influence us?³ What if producers and advertisers find secure and effective means of sharing private information about users and their purchasing and other behavior? If the time has not yet come that trans-

² This is an oft repeated phrase in the media and not attributable to a particular person or source.

³ A NYT article (published Oct. 13, 2012) reports on how the current presidential political campaign is heavily leveraging social and behavioral information to both reach "in market" voters as well as pressure them to vote. <http://www.nytimes.com/2012/10/14/us/politics/campaigns-mine-personal-lives-to-get-out-vote.html?pagewanted=1&hp&r=0>

actional, behavioral, social, geographic location, and other offline third party data sources are not combined and used by business to more closely target individual consumers — that time will come shortly.

2.2 PRIVACY AND POLICY ISSUES

Every day, major newspapers across the globe report on some concern over digital data and privacy. Included in this global discussion are talk of law, policy, and technology — to include threats to privacy as well as proposed solutions. According to [Gomez, Pinnick, and Soltani \(2009\)](#), consumer reports and polls repeatedly show “overwhelming concern by users about the collection of personal information and behavioral profiling” ([Gomez et al., 2009](#), p. 17). While website publishers typically offer a privacy policy to inform consumers what sorts of data may be collected and how that data may be used or shared, there is nonetheless much confusion and debate regarding what data is considered private and how it is handled. Though how personal data is collected and stored by first-party websites is in question, this section focuses more on the problem of data collected and used by third party trackers as well as public attitude toward this collection. However, in current privacy debates, there has been intrinsically little difference in the stance and attitude between large publishers and advertisers.

2.2.1 *The Internet is Free*

For some years, the Internet has been a free and open place. Access to services and information has been largely free. While many were initially skeptical about online commerce, convenience and ease won us over. With the advent of mobile technology and GPS, our applications became time and location aware. We can scan barcodes in stores and have applications tell us whether cheaper products are found close by. We embrace the wonder of connecting our virtual lives with our physical lives. But we feel uneasy, as well. The idea of being personally *surveilled* as we move about in the physical world is deeply unsettling — whether at home or in public. But to be monitored in a web browser leaves open the possibility that strangers will learn, in an intimate sense, *who we are*. They will do this by knowing what we read, what activ-

ties we pursue, where we live and travel, where we bank, where we work, what sorts of medical issues we research — and, naturally, what we buy. This information can be aggregated and stored by strangers — and then shared with other strangers: without consent, without review.

The delicate balance between public and private life may, in fact, be a relatively recent phenomenon. [Locke \(2010\)](#) examines closely the inherent tendencies of biological organisms to eavesdrop on others. Privacy is likened to a “regulatory process that serves to selectively control access of external stimulation to one’s self or the flow of information to others” ([Klopfer & Rubenstein, 1977](#), p. 53). One of the key functions of privacy is the withholding of information that might provide an opponent with a competitive advantage ([Klopfer & Rubenstein, 1977](#)).

In small, egalitarian societies, the ability to constantly monitor and eavesdrop “rendered trust unnecessary” ([Locke, 2010](#), p. 74). Everyone shares access to the same resources. According to Locke, the pervasiveness of eavesdropping serves as a form of social control with an inhibitory effect on behavior. In more complex societies, where walls and buildings function as a sort of social technology, it has become less possible to overhear others and the notion of privacy has became mired with secrecy. At the same time, since one does not know everyone in the community, eavesdropping has become a means for knowing and understanding the private lives of others. Ultimately, “walls and population increases made it necessary to take in information about others when they were physically absent, relying on the representations — the gossip — of intermediaries” ([Locke, 2010](#), p. 192).

2.2.2 *Privacy and Personally Identifiable Information (PII)*

Dan Solove, a George Washington University Law professor and author of *Understanding Privacy* ([2008](#)), frames privacy in terms of legal history and case law. The Fourth Amendment provides for the fundamental right to a “reasonable expectation of privacy” to the United States Constitution (as cited in [Solove, 2008](#), p. 2). In *Katz v. United States*, 1967, the Fourth Amendment protection was extended to include Fourth Amendment protection to public places, such as private conversations in public phone booths (as cited in [Solove, 2008](#), p. 22). However, the Supreme Court observed, “What

a person knowingly exposes to the public, even in his own home or office, is not a subject of Fourth Amendment protection" (as cited in Solove, 2008, p. 22).

In fact, it is the inherent difficulty in resolving what "privacy" is, that drove Solove to write his book. He notes "there is no overarching conception of privacy — it must be mapped like terrain, by painstakingly studying the landscape" (Solove, 2008, p. ix). This is particularly relevant when considering user browser-based interaction with websites over the Internet.

As Solove (2008) notes, there are literally hundreds of laws at state and federal levels to protect privacy:

Congress has passed several dozen statutes to protect the privacy of government records, student records, financial information, electronic communications, video rental data, and drivers' records, among other things. Furthermore, privacy is recognized as a fundamental human right. According to the United Nations Declaration of Human Rights of 1948, "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation." (Solove, 2008, p. 3).

At particular risk or bits of information known as "personally identifiable information" (PIIs) typically associated with harms such as identify theft and mis-use. In May 2007, the Deputy Director of the Office of Management and Budget (OMB) sent a memorandum on the subject of "safeguarding against and responding to the breach of personally identifiable information" noting,

The term "personally identifiable information" refers to information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information. (OMB, 2007)

The purpose of this memorandum was to reinforce requirements of the Privacy Act of 1974 to both safeguard and establish rules of conduct for the handling of personally identifiable information. The particular challenge was in response to

the ease to which technology has made it possible to personally identify particular individuals by combining small pieces of information. It is not that this is a new challenge: for example, brick-and-mortar businesses have long asked customers for a zip code when a customer pays with a credit card. These two pieces of information can be easily combined to find a home address. What's changed is that technology and the Internet has made it far easier to identify and, potentially harm, users at vast scale. And the idea of PII now seems to extend to seemingly innocuous bits of information such as Internet browsing behavior. What was once viewed as categorically deterministic is now probabilistic in terms of re-identifiability and risk.

2.2.3 *Growing Public Awareness*

Steward Cheifert (Interviewer): How concerned should I be as a user about what I put up on the Net?

John Markoff (West coast correspondent, New York Times): I think it depends on what you do. There's better security coming for economic business transactions, but I'd be careful about putting my password on the net, I'd pick a password that's a safe password, and I wouldn't put my credit card up until there is security software up that will protect my credit card. (*The Internet, 1995*)

In the early 1990's, the Internet exploded into public awareness. 1993 signaled the first graphical web browser and the world was mesmerized. There was a growing excitement for interaction on the World Wide Web, and public talk was dominated by the promise of possibility. The large concern was whether it was safe to shop online — though the focus was on the encryption of credit card transactions, rather than the seemingly benign micro-exchanges of a user clicking through web pages.

In November of 1999, DoubleClick, an advertising network of over 11,000 publishers, purchased Abacus Direct a company maintaining a database of detailed consumer profiles on approximately 90% of American households ("In re: DoubleClick Inc.", 2001). DoubleClick announced that it planned to merge "anonymous" online data (100 million user profiles) with Abacus profiles (including names, addresses, phone

numbers, etc.). The media roiled and public sentiment was overwhelmingly negative. Ultimately, DoubleClick put these plans on hold: its stock dropped dramatically after the merger was announced since neither the public nor Wall Street had reacted favorably to this plan (Chief Marketer, 1999). It was about this time that the public became aware of how their behavior on the Internet might be used by commercial entities. A Federal Trade Commission (FTC) investigation in 2001 ended its investigation with no finding that DoubleClick had engaged in unfair or deceptive trade practices. It concluded:

Based on this investigation, it appears to staff that DoubleClick never used or disclosed consumers' PII (personal identifiable information) for purposes other than those disclosed in its privacy policy. Specifically, it appears that DoubleClick did not combine PII from Abacus Direct with clickstream collected on client Web sites. In addition, it appears that DoubleClick has not used sensitive data for any online preference marketing product, in contravention of its stated online policy ("In re: DoubleClick Inc.", 2001).

This case, and others like it, now had the public's eye. By the year 2000, 54% of Americans used the Internet at home (Simms, 2000). In a March 2000 poll, 82% of web users polled indicated that they were uncomfortable if their browsing habits and shopping patterns were linked to their identities (Business Week / Harris Poll, 2000). 74% indicated that they were very uncomfortable if their information were sold to other organizations. Though only 40% of users had heard of browser "cookies", many were aware that data was being collected silently and that data could be both sold and merged with identifiable information. Public negative sentiment toward tracking has changed little over the years. Over 70% in 2012 say they find little value in online ads (Hoofnagle, Urban, & Li, 2012). And Microsoft reports that 75% of its US and European users would like to opt-out of online advertising (Hill, 2012).

2.2.4 *Third-Party Data Collection*

Browser tracking is not the only means by which websites identify and collection user information. But it's, perhaps, the

sort of data collection people are intuitively least comfortable with. Third-party trackers are trackers on websites that collect information about users for a third party. They are not officially part of the website visited, but are generally associated with advertisers or web analytics. Typically, they take the form of cookies. Cookies can be set by embedded JavaScript and allow a site to essentially “remember” arbitrary bits of information from a previous session. Advertisers can utilize cookies in a cross-domain manner by creating a unique ID that is set by one website and then picked up by that same advertiser on another website. In this way, the advertiser knows that you’ve visited both sites.

The problem with cookies is that it is a browser-mediated information exchange *hidden from view*. Some cookies serve useful functions for sites with which we have trusted relations. But many are set and tracked by organizations that we may not recognize and for purposes that are unclear. Of great concern is that these entities may be buying, selling, and trading browser profiles which can be linked to personally identifiable information (PII). [Narayanan \(2011\)](#) proposed a five-fold taxonomy in which browsing history might become pseudoanonymously⁴ identifiable to a particular device or user.

1. **The third party is sometimes a first-party.** A third party tracker such as Facebook may be a first-party in other contexts. If you visit Facebook and login, it sets a cookie that can be used to track you as you browse the web. [Roesner, Kohno, and Wetherall \(2012\)](#) describe the particular situation where sites have buttons such as Facebook, Twitter, or Google embedded on their web pages. These buttons are trackers and used to track you even if you never click on them. If you visit a website, and are also logged into Facebook in the same browser, Facebook will know that you’ve visited that site.
2. **A first party leaks data to third-parties.** [Krishnamurthy, Naryshkin, and Wills \(2011\)](#) found that 56% of 100 popular, non-social network sites leaked private information. This number grows to 75% if a site userid is included as a PII.

⁴ “Psuedoanonymous” unique IDs may be used to identify a particular device or user agent. These may potentially be linked to other, personally identifying information later.

3. **A first party sells the user's identity.** The previously mentioned case of Abacus and DoubleClick serves as a good example. Also, [Narayanan \(2011\)](#) notes that survey sites (e.g., "win a free iPod!") can also act as an identity provider to sites on which they have a third-party presence.
4. **Hacks and exploits.** A third-party might exploit a cross-site vulnerability on a first-party website to learn the user's identity ([Narayanan, 2011](#)).
5. **Deanonymization (re-identification).** A third party could match browsing histories against datasets to re-identify particular persons, as [Narayanan and Shmatikov \(2010\)](#) did with the Netflix Prize dataset. In this experiment, they demonstrated that it's possible to identify a particular subscriber with only a little bit of information.⁵

Though Netscape never intended cookies to be privacy-invasive, enterprise-minded businesses quickly realized that cookies could be used for identifying browsers. In fact, the ecology of tracker technology in the wild is surprising complex and, potentially, adaptable.

[Roesner et al. \(2012\)](#) performed a study in which they observed tracker behavior from the point-of-view of the client browser. They classified this behavior into five behavioral patterns in which a given tracker site might exhibit variable behavior depending on contexts. They note complex behaviors such as that of an aggregate tracker referring data to other trackers. A few trackers stored data in multiple locations simultaneously (e.g., cookies, flash storage, HTML5 local storage) or exhibited spawning behavior when cookies were cleared. When they asked the question of how much data can any one tracker collect, using 2006 AOL search query data, they found that, on average, DoubleClick could track a user across 39% of pages visited, and a maximum of 66% of pages visited.

Though cookies are the most prevalent trackers, "web bugs" serve a similar function. They are traditionally embedded as small (often 1 pixel) GIF or PNG images and can also be used in HTML mail, informing advertisers whether an email

⁵ In fact, this is only one example of many. Re-identification is another huge challenge for privacy law. For discussion in greater depth, see [Ohm \(2010\)](#).

has been opened and when. In March 2010, Ghostery identified 117 unique web bugs on nearly 400,000 unique domains (Gomez, Pinnick, & Soltani, 2009). Though privacy policies may state that information will not be shared with third parties, many of these sites nonetheless allow third-party tracking through web bugs (Gomez et al., 2009).

Given the complex behavior of trackers, coupled with complex business practices in which first-parties and third-parties may mingle or exchange data, it is not surprising that users may be confused about who has collected, stored, traded, or used their browsing profile.

2.2.5 Cookies and Privacy Law

Recent litigation over OBT invokes Wiretapping acts both federal and state. At the federal level, the Electronic Communications Privacy Act ("ECPA") Wiretap Act, Stored Communications Act ("SCA"), and the Computer Fraud and Abuse Act ("CFAA") are frequently invoked.

The federal Wiretap Act (as amended by the ECPA) protects the privacy of wire, oral, and electronic communication. The latter is defined as "any transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature transmitted in whole or in part by a wire, radio, electromagnetic, photo-electronic or photooptical system" (Title 18, Part 1, Chapter 119 §2510(12)). Under §2511(3), "a person or entity providing an electronic communication service to the public shall not intentionally divulge the contents of any communications... while in transmission on that service to any person or entity other than an addressee or intended recipient..."

The ECPA has been criticized lately by the media:

Many Internet companies and consumer advocates say the main law governing communication privacy — enacted in 1986, before cellphone and e-mail use was widespread, and before social networking was even conceived — is outdated, affording more protection to letters in a file cabinet than e-mail on a server." (Helft & Miller, 2011)

Such criticism had been directed at Government data requests, yet numerous cases solely involving commercial enti-

ties abound.⁶ Of issue for OBT is an exception known as the *consent exception*: so long as one of the parties of electronic communication has given prior consent to the interception, and the interception was not intended for criminal or “tortious act”, then it is exempt from the Wiretap Act (Center for Democracy and Technology, 2008).

One notable case which invoked ECPA, SCA, and CFAA was the case of “In re Doubleclick, Inc.”. In 2000, a class action suit was brought against Doubleclick Inc., which used cookies for behavioral advertising. With respect to ECPA Wiretap law, Doubleclick successfully argued that its affiliated websites were “users” of the Internet and that communications accessed by Doubleclick’s cookies were “of or intended for” these websites (“In re: DoubleClick Inc.”, 2001). The Stored Communications Act was found not to apply to Doubleclick, as well. The SCA was intended to address temporary and transactional records of internet service providers. Cookies stored on personal hard drives and were out of scope for the intent of SCA (“In re: DoubleClick Inc.”, 2001). Finally, with respect to the Computer Fraud and Abuse Act, the plaintiff failed to provide evidence that they had incurred harm.

As it stands today, federal laws do not address concerns raised by online behavioral tracking on the Internet. Furthermore, given the single-party consent exemption of Federal Wiretap law, communications between Internet users and websites is not private if the website owner has given its consent for monitoring.

2.2.6 Policy Legislation - Do Not Track

Though the United States does not have specific federal law that applies to behavioral tracking, the Federal Trade Commission (FTC) can enforce the terms of privacy policies under Section 5 of the FTC Act prohibiting “unfair or deceptive” marketing practices (*United States Code Title 15 §45*). This applies to FTC handling of behavioral tracking. In 2011, the FTC found three parties in violation for the FTC act relating to third-party tracking (Mayer & Mitchell, 2012).⁷

⁶ For example, "In re Doubleclick, Inc." and others referenced at http://www.internetlibrary.com/topics/electronic_cpa.cfm.

⁷ A violation is determined if a company states in a website notice or privacy policy that it does not engage in a particular practice, when in fact it does.

In 2010, the Federal Trade Commission proposed “do not track” technology as “likely a persistent setting on consumers’ browsers — so consumers can choose whether to allow the collection of data regarding their online searching and browsing activities” (Federal Trade Commission, 2010). This represents a form of self-regulation insofar that trackers are expected to respond appropriately to browser requests not to track. In its March 2012 report Federal Trade Commission (2012), the FTC outlined a broader privacy framework to address a myriad of concerns surrounding data brokering, mobile, internet service providers, and enforceable self-regulatory codes. The scope of this framework applies to:

[...] all commercial entities that collect or use consumer data that can be reasonably linked to a specific consumer, computer, or other device, unless the entity collects only non-sensitive data from fewer than 5,000 consumers per year and does not share the data with third parties. (Federal Trade Commission, 2012)

The FTC notes that this framework applies to offline as well as online data and applies to data that is reasonably linkable to a specific consumer, computer, or device.

As proposed by FTC (Federal Trade Commission, 2012), Do Not Track should include five key principles:

1. It should be implemented universally to cover all parties that would track consumers.
2. The choice mechanism should be easy to find, easy to understand, and easy to use.
3. Any choices offered should be persistent and should not be overridden if, for example, consumers clear their cookies or update their browsers.
4. It should be comprehensive, effective, and enforceable: it should opt consumers out of behavioral tracking through any means and not permit technical loopholes.
5. It should go beyond simply opting consumers out of receiving targeted advertisements; it should opt them out of collection of behavioral data for all purposes other than those that would be consistent with the context of

the interaction (e.g., preventing click-fraud or collecting de-identified data for analytics purposes; [Federal Trade Commission, 2012](#)).

Since the FTC released its preliminary findings, the W3C Internet standards body convened a working group toward the adoption of an industry-wide Do Not Track standard. This group includes representatives from the advertising industry, online businesses, academia, privacy advocates — both nationally and internationally. The challenge for this group is to devise a standard that meets FTC guidelines, while satisfying a diverse array of agendas.

2.2.7 Building Privacy Policies into Browser Protocols

In the late 1990s, there were a couple of attempts to address the problem of privacy invasion introduced by third-party cookies via technical means. Introduced in 1997, RFC 2109 was a proposal aimed at putting strict controls on cookie usage. Commercial entities participating in standards discussion objected to portions of this proposal and it was, ultimately, not supported by browser developers. “Where the browser manufacturers had been asked to redesign their software to reject cookies automatically, Netscape and Microsoft Internet Explorer instead included options for users to reject cookies if they so chose” ([Eichelberger, n.d.](#)). As cookie inventor Lou Montulli stated, “browser default would still be set to accept cookies, since it was felt by the designers that if we were to unilaterally disable this feature, existing content on the Web would no longer work” ([Bruner, 1997](#)).

From 1998 - 2003, the W3C organized a standards committee to propose a Platform for Privacy Preferences (P3P) standard ([A. Schwartz, 2009](#)). The standard aimed to create a machine-readable policy that captured the complexity of human-readable privacy policies and boiled them down to a multiple-choice set of options. What resulted from long and heated discussion was highly controversial and touted as overly complex — and unlikely to be adopted as a self-regulatory option with no degree of enforcement.

About the same time, advertisers began to take a more active role in self-regulation. The National Advertising Initiative (NAI) set up a website to allow users to download opt-out cookies for participating networks. Their principles have been

to notify users and allow them a choice. Opt-out does not mean that you will not be tracked or that ads will not be served. It means that ads will not be tailored to your browsing behavior. Opt-out choices are stored in browser cookies, so they recommend to visit their opt-out page periodically for updates.

Other advertising networks and companies have since also offered similar opt-out cookie services.⁸ Large advertising industry groups formed a new parallel self-regulatory program for behavioral advertising, the Digital Advertising Alliance (DAA). The DAA and the Interactive Advertising Bureau (IAB), a self-regulatory for advertising networks, advocate the use of an advertising option icon to be displayed on web pages with tracking scripts. A user who clicks on the icon sees a disclosure statement and can click through to an opt-out page. However, according to a 2011 study at the Center for Internet and Society at Stanford University, only 11.3% of members to date offer cookie-based opt-out. (Mayer, 2011).

In 2007, the Center for Democracy and Technology (CDT) sent a note to the Federal Trade Commission (FTC) requesting that “the online tracking and targeting of consumers — both in its current form and as it may develop in the future — needs to be limited so that consumers can exercise meaningful, granular preferences based on timely and contextual disclosures that are understandable on whichever devices consumers choose to use” (Center for Democracy and Technology, 2007). They called for a national Do Not Track list similar to the national Do Not Call list. Their proposal called for a machine readable list of the domain names which used cookies and other means to track users. Browsers and third party software could then use this list in order to limit tracking (Center for Democracy and Technology, 2007).

Thus, the W3C chartered the Tracking Protection Working Group (TPWG) “to improve user privacy and user control by defining mechanisms for expressing user preferences around Web tracking and for blocking or allowing Web tracking elements” (W3C TPWG, n.d.).

In effect, discussion in this group mirrors, and perhaps fuels, public debate in media and society.⁹ Chairs for the group

⁸ For example, see <http://www.worldprivacyforum.org/cookieoptout.html>.

⁹ The public-tracking list is archived at: <http://lists.w3.org/Archives/Public/public-tracking/>.



Figure 6: Tweet, 7 September 2012

serve to organize and moderate debate. As questions become more defined, formal “issues” are raised and “actions” assigned to particular members. For example, ISSUE-5, “What is the definition of tracking?” was opened on September 21, 2011. As of October 2013, there are 446 emails pertaining to this issue alone, and it is still open more than two years later.

Because each member has his or her own platform or stance on any particular matter, debate sometimes bleeds into other venues. For example, as of this writing, the current DNT draft standard says the following (regarding determining a person’s DNT preference): “A user agent must have a default tracking preference of ‘unset’ (not enabled) unless specific tracking preference is implied by the decision to use that agent” ([W3C TPWG, 2011](#)). When TPWG member Roy T. Fielding, who works for Adobe but who also contributes to the open source Apache webserver codebase, committed a patch to this codebase which disabled DNT if the browser requesting data was Internet Explorer 10 (which at the time set DNT to 1 by default), he caused a ripple tide of effects. His patch, labeled “Apache does not tolerate deliberate abuse of open standards” ([Fielding, 2012a](#)), clearly reflects his own position on the implementation of DNT. This hit C/NET news and Twitter almost simultaneously. On Twitter, Jonathan Mayer ([Mayer, 2012b](#)), fellow TPWG member responded with [Figure 6](#).

About the same time, C/NET reported the event, noting Roy T. Fielding’s position as an author of the DNT standard and principal scientist at Adobe Systems ([Shankland, 2012](#)). Of course, there was an immediate back-lash on the public-tracking list between Mayer and Fielding, in which Mayer ([Mayer, 2012a](#)) partially retracted his comment on Twitter [Figure 7](#).

From Fielding on the public-tracking list:

Our charter forbids us from specifying UI requirements. That does not mean any of the following excerpts are ambiguous:



Figure 7: Tweet, 7 September 2012

A user is an individual human. When user-agent software accesses online resources, whether or not the user understands or has specific knowledge of a particular request, that request is made "by" the user.

The goal of this protocol is to allow a user to express their personal preference regarding tracking to each server and web application that they communicate with via HTTP ...

Key to that notion of expression is that it MUST reflect the user's preference, not the choice of some vendor, institution, or network-imposed mechanism outside the user's control. The basic principle is that a tracking preference expression is only transmitted when it reflects a deliberate choice by the user. In the absence of user choice, there is no tracking preference expressed. (Fielding, 2012b)

On the basis of this personal conviction, Fielding argued that if DNT was switched on by default, it should be ignored, regardless of other considerations. Since Microsoft had announced that it would turn on DNT by default, Fielding retaliated using the technological means at hand — direct modification of code. Naturally, this led to heated debate within the Apache http server developer community. As of this writing, there were 364 comments on this commit. Fielding's changes are reflected in the current configuration of the main branch of code, but are commented out.

Lorrie Cranor, one of the original developers of P3P is skeptical that DNT will succeed. To her it seems a simpler, but not better, replay of the P3P proposal (Federal Trade Commission, 2012; Fulton, 2012).

As we walk about in the physical world, we raise and lower our voice and we raise and lower our window shades and we turn our faces, and we are all constantly adjusting to regulate our exposure and our privacy, Dr. Cranor tells RWW. "And

it comes naturally; we don't spend a lot of time thinking about it. We just sort of naturally do it. But when we go online, it's no longer natural, because we don't have these readily apparent, physical things where you can just easily close that shade, and it's obvious what you're doing. So we have to rely on software tools to help us with this privacy regulation process. (Fulton, 2012)

2.2.8 *Choice and Self-Regulation*

While policy debate rages, the advertising industry continues to advocate self-regulation. Central to self-regulation is the notion of choice. The DAA and NAI naturally view behavioral advertising in a very positive sense such that better targeted ads are preferable to ads that do not match a consumer's interests. But if a user does not wish to participate, then they should have the choice to opt-out. However, privacy advocates cite two large concerns with respect to industry self-regulation.

The first argument against the effectiveness of self-regulation is non-compliance. A Carnegie Mellon University (CMU) report by Komanduri, Shay, Norcie, and Ur (2012) finds that, despite attempts at self-regulation by bodies such as the DAA and NAI, there are still many examples of non-compliance. Stanford University website <http://donottrack.us> believes that further steps are required: namely, regulation enforcing compliance with user choice.

A second, perhaps more compelling argument is that the advertising industry relies on technologies that seem to ensure that users remain unaware of them. When consumers attempt to block tracking, other more resistant methods are developed which circumvent these defenses. While DNT would make it very easy for users to decide not to be tracked, market choice seems to make it easier for them to be tracked. As Hoofnagle, Soltani, Good, Wambach, and Ayenson (2012) note, "those who argue that consumers can negotiate the nuances of privacy and tracking online assume that the online world is similar to the offline world." In the offline world the consumer can leave without leaving behind a data trail. But in the online world, invisible attributes leave marks that are easy to follow. They further note that surveys of top websites in 2009 and 2011, revealed new tracking mechanisms resistant

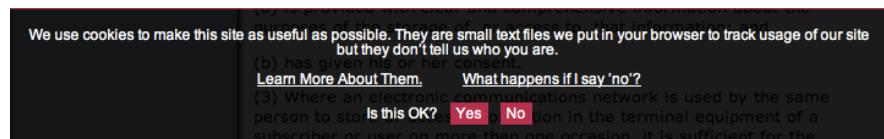


Figure 8: Example cookie opt-in notification from: <http://www.jonwallacedesign.com/privacypolicy.html>

to the strongest of privacy settings. Or, perhaps, advertisers simply know that once online shopping is habit, consumers will adapt.

2.2.9 *Opt-In Versus Opt-Out*

On the 26th of May 2012 in the U.K., a new European Cookie Law, E-Privacy Directive 2009/136/EC, came into effect.¹⁰ This law requires the consent of users before cookies are set (see Figure 8). In effect, this is an unequivocal “opt-in” system of notification and consent (with the exception of cookies “strictly necessary” for the operation of a website ([Official Journal of the EU, 2009](#))).

Unlike current trends in US policy and law, this puts the burden of consent on the website publisher — not the user. Though this issue permeates the entire European Union (EU), discussion of the UK “Cookie Law” requirement for consent has finally reached US audiences on the Internet.

Perhaps, this EU Directive gives a sense of what may be required if US lawmakers decide to place the burden of consent on website publishers and not users. Such a shift would also acknowledge that Internet transactions between users and websites are currently not “private” as Federal Wiretap law has held. The EU Directive seems more in line with user expectations, but given current criticism on the difficulty with compliance, falls short with respect to implementation. The hope is that the W3C Tracking Protection Working Group (TPWG) will provide strong enough guidance to satisfy this Directive. However, US members (and especially those representing business interests) may decide to act in an insular fashion and have, thus far, shown little motivation to join a more global debate ([W3C TPWG, 2012](#)).

¹⁰ There is a high-level summation here: <http://gigaom.com/europe/cookie-law-explainer/>.

2.3 USER CONFUSION

There has long been an asymmetry in communication between web publishers and users. Interactive page elements such as “contact us” via web forms, instant chat, web forums, and email have enabled users to connect directly to persons behind the scenes. More recently, publishers utilize personal messaging through other channels such as Twitter, Facebook, or SMS for allowing users to offer direct and personal feedback. These are direct and intentional communications between parties. But the act of visiting a webpage in itself also has communicative value.

In this section, I highlight user confusion regarding indirect communications where advertisers and other commercial actors silently monitor web interaction via the use of cookies. Not only is the phenomenon of behavioral tracking difficult to understand and see, but user defenses against ubiquitous monitoring are confusing, as well. Of particular interest are user opt-out mechanisms provided for by emerging DNT standards and also the advertising industry efforts toward self-regulation of online behavioral advertising. The vehicle for this is “consumer choice” and the Advertising Option Icon (viz. AdChoices icon; see <http://www.aboutads.info> for general information).

2.3.1 *Confusion About Third Party Tracking*

In a 2009 study of Internet user understanding of behavioral advertising, [McDonald, Reeder, Kelley, and Cranor \(2009\)](#) found that only 51% believed that the ability of an ad company to determine which ads to show them based on the history of prior websites they visited was something that “happens a lot right now” ([McDonald et al., 2009](#)). 46% of the same study participants found the idea of behavioral advertising “creepy”, and another 40% agreed or strongly agreed they would be more careful online if they knew advertisers were collecting data.

In the virtual online world where billions of people have access to the web through the view of what may often be a personal device, it’s not surprising that the experience feels, to some degree, private. As privacy advocate Christopher Soghoian remarks,

Consumers treat the search engine box like their psychiatrist, their rabbi, their priest, their doctor. People type the most intimate things into search engines and other websites primarily because they think they're anonymous. They type in things on WebMD that sometimes they wouldn't even ask their own doctors... And in fact, we are not anonymous, these sites are tracking us." (as cited in Kessler, 2010)

The UC Berkeley School of Information KnowPrivacy Project (Gomez et al., 2009) looked specifically at user concerns and knowledge. They examined 50 of the most visited websites and their privacy policies. They also considered specific practices such as third-party tracking and sharing with affiliates. Not surprisingly, they found that most users are concerned about data collection on themselves and their control over that collection and use of information. Furthermore, they found overwhelming evidence (from prior surveys) that users lack knowledge and understanding about data collection practices and policies.

One source of obvious confusion has to do with terminology. What is tracking? Does that include collection? Most users think not (McDonald, 2011; McDonald & Peha, 2011). What is the difference between "third party" and "affiliate"? Privacy law treats these categories differently: third-party information sharing is subject to more restriction. According to Gomez et al. (2009), websites make distinctions between sharing with affiliates and third parties. 29 of 50 companies examined state that they do not share data with unrelated third parties, 45 state that they share data with affiliates, and 36 state that they allow third-party tracking. This last category falls outside of privacy policy coverage. As Gomez et al. (2009) note, it's very unclear what it means to not share data with unrelated third parties yet permit third party tracking.

Despite assurances that affiliates and third parties are treated differently, website providers often leak personal information to the less trusted, embedded third party trackers such as username, login time, and other information. Just as troubling is that users have no practical way of knowing who affiliates are and what sorts of information are passed to them. In the Know Privacy Report (2009), many of the websites examined are owned by companies with hundreds of subsidiaries: NewsCorp has over 1500 subsidiaries while Bank of America

has over 2300.¹¹ The pervasive presence of third party trackers on the Internet means companies have ample opportunity to construct long term profiles of things we do online.

Privacy policies often go unread, or may be difficult to understand (Milne & Culnan, 2004; Sherman, 2008). Fernback and Papacharissi (2007) conducted an in-depth discourse analysis of privacy statements of three large companies. They also noted discrepancies regarding the use of the term "third-party". From their study of the Real Network privacy policy: "We will never sell, rent or disclose to third parties our customers' personally identifiable information ... gathered on a RealNetworks Website unless we are required to do so by law or receive your advance informed consent" (Fernback & Papacharissi, 2007). Yet, in what appears a blatant contradiction, later in the same privacy policy:

Your personally identifiable information may be transferred in connection with a sale, merger, transfer, exchange or other disposition (whether of assets, stock or otherwise) of all or a portion of a business of RealNetworks, Inc. and/or its subsidiaries. (Fernback & Papacharissi, 2007)

In general, Fernback and Papacharissi (2007) note that privacy policies are seen to protect the company and typically avoid mention of specifics such as precisely what information might be collected and for what purpose. Given that the FTC has a history of prosecuting companies that violate their terms of agreement,¹² this is not terribly surprising.

McDonald, Reeder, Kelley, and Cranor (2009), examined the readability of online privacy statements. They examined standardized text formats and compared these to free text policies. They found that experiment participants could more readily read and comprehend standardized formats, but at the expense of accuracy. And users, in general, did not like either standardized or free text formats. But when it comes to purchasing privacy-sensitive items, Egelman, Tsai, Cranor, and Acquisti (2009) found that users actually paid attention to graphical privacy indicators.

¹¹ These numbers reportedly do not include subsidiaries of subsidiaries.

¹² See <http://www.ftc.gov/os/caselist/index.shtm> for example. Google and Facebook have both earned penalties.

2.3.2 *Confusion about Defenses*

When it comes to defending against third party tracking and privacy, users have a number of choices. [Roesner et al. \(2012\)](#) identified the following options:

- **Pop-up blockers:** Pop-up blockers can stop trackers from forced pop-ups, but there are similar methods such as site re-directs that may not be stopped;
- **Browser third-party cookie blocking:** Many third party cookies can be blocked if users select third-party cookie blocking in their browsers. However, this will not stop the sort of tracking when a Facebook tracker is accepted as a first-party tracker and its role changes later to a third-party tracker;
- **Private browsing:** Private browsing mode was designed to protect users from having their browse state examined by physical access. But it does not keep browsing state from being examined online;
- **Opt-out cookies (and AdChoice icon):** The Digital Advertising Alliance (DAA) is an industry funded policy group that hosts an opt-out web page. From this page, users can click a button to set opt-out cookies;¹³
- **Clearing browser state:** Clearing cookies when closing the browser is a simple means to reduce the effects of tracking. However, this may also remove opt-out cookies and is also not effective against re-identification by trackers such as the Facebook like icon when logged in;
- **Do not track (DNT):** The proposed FTC "Do Not Track" DNT policy is designed to give users a way to opt out of web tracking. This is accomplished via an http request header with a DNT=1 header to inform the remote server that the client wishes to opt out. DNT is not mandatory and requires no compliance. The DAA has committed only to stop content personalization, if it receives a DNT signal. (<http://www.aboutads.info/choices/>); and,

¹³ There are other similar sites that offer opt-out cookies. For example, Evidon Global, and others have web pages with such links.

- **Blocking plug-ins:** There are a number of browser plug-ins which are designed to block trackers.

Lorri Craner directs the CyLab Usability Privacy and Security (CUPS) Laboratory at Carnegie Mellon University. She and her students have conducted a series of studies centered on user understanding of online behavioral advertising and usability of blocking tools (Blase Ur, 2012; Cranor, 2012; Leon, Ur, et al., 2012; McDonald, 2010). Leon, Ur, et al. (2012) studied the usability of Ghostery, Adblock Plus, and the Internet Explorer Tracking Protection List blocking tools. They found that self-help blocker tools have significant issues in terms of user understanding:

- Users don't recognize the names of the majority of companies that they can opt-out;
- Some of the tools use terms that were meaningless to participants: for example, "web tracker, web bug, flash cookie, silverlight cookie, tracking cookie, script, iframe, and targeted ad network.;"
- Participants testing opt-out tools did not understand what the tools would opt them out of, mistakenly believing that they were protected against tracking;
- Opt-out tool users thought deleting cookies would protect their privacy even more, not realizing that deleting their cookies would also delete their opt-out cookies and undo their opt-out;
- Users were left unaware whether or not most tools were working, and oblivious to what was happening behind the scenes;
- None of the opt-out tools tested notify users while they are browsing that their preferences are being respected; and,
- Participants who tested the browser cookie settings also had no mechanism for understanding what exactly was happening behind the scenes unless websites didn't work (Leon, Ur, et al., 2012).

In a related study conducted by Leon, Cranshaw, et al. (2012), 45% of participants who saw "AdChoices" believed

that it was intended to sell advertising space, while only 27% believed that it was intended to stop tailored ads. A significant part of the problem is that 1) users don't really understand the mechanisms behind tracking and so misunderstand the nature of the problem; and 2) users don't have a clear idea of what options are available to help alleviate the problem (Blase Ur, 2012; McDonald, 2010). Moreover, users have no practical way of knowing how effective blocking is when it works. Mayer and Mitchell (2012) examined the effectiveness of 11 blocking tools and found significant variability in performance. Most of these self-help tools work about the same way. They consist of a black list that is modifiable by the user. The lists vary widely and account for much difference seen in performance. Since it's not obvious that these tools essentially work in the same way, users may be tempted to install more than one. However, doing so may be risky: Mayer and Mitchell (2012) report that TRUSTe's tool actually over-rode blocking lists by other tools allowing tracking by several large third party trackers.

2.3.3 *Are Privacy Concerns Affecting User Behavior?*

When questioned in polls and studies, users overwhelmingly share a negative attitude toward online tracking. In a 2010 CUPS study:

64% found the idea invasive, and we see signs of a possible chilling effect with 40% self-reporting they would change their online behavior if advertisers were collecting data. We found a gap between people's willingness to pay to protect their privacy and their willingness to accept discounts in exchange for private information. 69% believe privacy is a right and 61% think it is extortion to pay to keep their data private. Only 11% say they would pay to avoid ads. We found participants are comfortable with the idea that advertising supports free online content, but they do not believe their data are part of that exchange. (McDonald & Cranor, 2010)

Though, a large proportion of users claim that they would change their online behavior if they believed advertisers were collecting data, there appears to be no general usage statistics to indicate that so large a population employs existing opt-out

technology beyond what the browser may natively provide. In 2011, Fowler (2011) reported that only 5.6 million Firefox desktop users had turned on DNT. In terms of more aggressive blocking, when I started this dissertation in late 2012, 14 million Firefox users had installed the most popular blocker extension (Ad Blocker). But this was only approximately 3% of the 450 million Firefox users which represented only 20–24% of desktop browsers in use (Mozilla, n.d.; Statcounter, 2012). However, trends do indicate that blocking trackers is growing in popularity (Acohido, 2011). In late 2013, DNT adoption is approximately 17% (Fowler, 2013b). And the new default setting for Firefox (as of Feb 2013) is not to allow all third party domains, but to allow cookies only from visited domains (Fowler, 2013a).

In some contexts, users may be willing to make very clear choices when confronted with a privacy trade-off. Egelman et al. (2009) found that laboratory study participants of an online shopping task were willing to pay more for sensitive purchases when confronted with a choice between a site with less privacy but cheaper prices, and a site with more privacy but more expensive prices. Hoofnagle, King, Li, and Turow (2010) likewise found that more than half of online poll participants reported changing their minds about the purchase of a product online because of privacy concerns.

As users awareness grows, advertisers find new strategies to evade blocking (Hoofnagle, Urban, & Li, 2012; Leon, Ur, et al., 2012; Soltani, Carty, Mayo, Thomas, & Hoofnagle, 2009). What users want is not in accordance with what publishers and advertisers are currently willing to do. In a survey from McDonald and Peha (2011), 72% expected that regulatory “do not track” efforts limit data collection, while 34% of respondents expected that “do not track” would prevent data from being collected by websites and advertisers. Since industry proponents currently interpret “do not track” as not affecting data collection but simply use of data for presenting advertisements, there exists an impasse between consumer advocates and industry proponents in policy-oriented regulatory efforts by the W3C.

Central to this debate is the notion of “consumer notice and choice”. Advertising self-regulation is firmly based on this notion. As we will see in chapters to come, there are strong reasons why this is so.



Figure 9: Interactive Print Advertisement

2.4 THE FUTURE OF INTERACTIVE ADVERTISING

Though the goal of advertising is ultimately to sell products, advertisers also have other goals such as gaining wider audiences and positive brand engagement. Interactive advertising plays a role in the space by providing consumers opportunities to interact with advertisements directly.

I live in the country. Recently, I was finally able to have high-speed Internet installed at my house. For the first time, I could stream broadcast content over my AppleTV. While I was poking around, I found I could access a PBS channel for free. In order to activate it, I was shown a code and asked to visit the PBS website to register. After registering, content was made available over my AppleTV. Now PBS knows who I am, what I watch, and when I watch it. Though, I am not engaging with ads over PBS yet, Apple has transformed the broadcast experience by opening a window between PBS and me, whereby we can affect each other's actions.

In the past year, we've seen other, more unusual, examples of interactive advertising. Lexus has created an ad which "comes to life" when an iPad is placed behind the printed magazine ad.¹⁴

Car manufacturers had already engaged a potential audience by creating game-like experiences where consumers could design their own vehicle and examine it in 3D. Cross-media experiences such as this, transform print into something new.

¹⁴ <http://www.lexus.com/stunning>



Figure 10: Print-Embedded Smartphone

On October 5, 2012, *Entertainment Weekly* placed a smartphone inside its print magazine. The phone contained a digital ad running video and live tweets (Ulanoff, 2012).

It has become easier to imagine a world where print seems fluid and magic — as envisioned by J.K. Rowling portrays the Daily Prophet in the world of Harry Potter.

Not surprisingly, there is a strong interest by advertisers to engage over mobile platforms. Utilizing location-based coupons, the iButterfly app engages customers by gamifying coupons.¹⁵ Users flick their phone to hunt virtual 3D butterflies.

While interactive advertising may appear tangential to behavioral advertising, there is a common theme underlying both: *interactivity*. Behavioral advertising relies on the ability to observe users engaging in everyday activities while interactive ads provide a means for advertisers to invite consumers to engage with brands and products.

Why should this parallel concern us? It does not seem far-fetched to imagine the content of interactive ads become more tailored to the interests and expectations of individuals. Advertisers already argue that OBA benefits the consumer by delivering a more tailored (and presumably less annoying) experience. The argument posed by advertisers is fairly simple: wouldn't you rather see advertisements about products that interest you rather than advertisements about things that do not?

¹⁵ <http://www.cherrypicks.com/products/ibutterfly>

Suppose that advertisers collect images of your friends and family and create facsimiles which appear in advertisements.¹⁶ In June 2013, Alessandro Acquisti spoke of the following research in progress.

Imagine that an organization has access to your list of Facebook friends, and through some kind of algorithm they can detect the two friends that you like the most. And then they create, in real time, a facial composite of these two friends. Now studies prior to ours have shown that people don't recognize any longer even themselves in facial composites, but they react to those composites in a positive manner. So next time you are looking for a certain product, and there is an ad suggesting you to buy it, it will not be just a standard spokesperson. It will be one of your friends, and you will not even know that this is happening. (Acquisti, 2013)

The problem is that, even with policy efforts toward transparency of collection and use, it is remarkably easy to manipulate people to behave in predictable ways. As advertisements become increasingly interactive — and as methods for collecting and using online behavioral data become more sophisticated — we will need to become much more attuned to ways in which this data may be used to drive behavior.

2.5 SUMMARY

This chapter introduced the phenomenon of online behavioral advertising to include issues leading to the formation of law and policy as well as issues concerning user notice and choice — principles espoused by policy makers and advertiser self-regulatory bodies alike.

Two mechanisms for user control were introduced: "do not track" and the AdChoices advertising option icon. Both present a means for users to opt-out of behavioral advertising. Both present user interactive mechanisms of choice. And both present opportunities for manipulating behavior in subtle ways.

¹⁶ FaceBook may already use your posts and personal data for advertising (Goel, 2013). Google also includes Google user names, faces, and content in ads (Kelly, 2013).

The remainder of this dissertation is largely concerned with the study the cause of user confusion in specific contexts. The next chapter first introduces a theoretical framework for meaning and understanding in interaction. In addition, I briefly cover previous work in the areas of graphical communication and language and advertising.

3

THEORETICAL FRAMEWORK

Language has long played a role in successful advertising. Not surprisingly, as new media have emerged, advertisers have skillfully adapted to form.

Vestergaard and Schröder (1985b) studied the language of press advertisements from 1976–1977. Though noting the contextual role of visual elements, the primary focus of their work was textual.

There is no reason to believe that TV and press advertising differ in their persuasive methods in a basic way, although an analysis of TV adverts, owing to the processual character of the TV commercial and its use of both sound and picture, requires an additional body of analytic procedures.
(Vestergaard & Schröder, 1985b, p. 10)

They considered advertising as, fundamentally, a message-based, one-way communication. Though they acknowledged psychological effects of messages, textual content was seen as meaning "transfer".

About the same time, Geis (1982) studied linguistic communication in television advertisements. He proposed that advertising discourse was better analyzed under a cooperative model of communication. Commercials both directly address viewers and also communicate indirectly by promoting products through dialogue between characters. He focused strongly on the persuasive use of language, in particular, highlighting the subtle ways in which language can be used to mis-lead consumers.

This chapter presents the notion that some user interactions associated with online behavioral advertising evoke discourse processes using mechanisms similar to those described by Geis (1982). Furthermore, such phenomena are representative of a new sort of advertising discourse. First, I introduce the notion of pragmatics and inferential communication. Then, I describe properties of discourse under a cooperative model. Following this, I discuss the notion of discourse structure and summarize factors contributing to successful communication.

Finally, I review two applications of discourse understanding upon which this dissertation builds: graphics communication and the use of language in advertising.

3.1 PRAGMATICS

Pragmatics is the systematic study of meaning by virtue of, or dependent on, the use of language.
(Huang, 2007, p. 2)

To understand pragmatics, it's helpful to understand a bit of history in linguistic theory. The study of meaning has long been the domain of semantic theory. Moreover, semantics arose as an extension of philosophical and mathematical logic. In 1905, Bertrand Russell published a short work, "On Denoting", in the journal *Mind* which held influence over semantics through much of the 20th century (Russell, 1905). Russell argued that meaning was propositional and that most statements have truth-value. Such thinking led philosophers to devise a logic-based approach to meaning that forms the basis of formal semantics today.

In a somewhat different vein, J.L. Austin also studied language use in Philosophy. Contrary to Russell, Austin (1975), in "How to Do Things With Words", argued that the meaning of sentences cannot be characterized solely by truth value. In a theory of speech acts, Austin regarded utterances as performing additional functions. For example, an interrogative cannot have a value of truth or false, but is used to perform an act for the purpose of eliciting some other action (e.g., to obtain information).

Like Austin, Grice and White (1961) were also concerned with understanding "non-natural" meaning. They distinguished between two sorts of meaning: the kind of meaning associated with individual words of that utterance (conventional implicature) and what a speaker means by an utterance (conversational implicature). This seems a subtle distinction in the abstract, but is actually an important distinction which had great impact on the discipline now known as pragmatics.

Conventional implicature relates to the meaning inherent in the words of an utterance. For example, Grice demonstrates using the word "but" in the utterance "she was poor but honest" (Grice & White, 1961, p. 127). What is meant literally,

is “she was poor” and “she was honest”. However, “but” *implicates* a contrast between “poor” and “honest”, suggesting that honesty is a trait not generally associated with poverty. He notes further that if one were to change “but” to “and”, the literal meaning of the utterance would not change — but the implication would be gone (Grice & White, 1961, p. 129).

Grice (1975b) also identified a class of implicatures known as **conversational implicatures**. This sort of implicature is exemplified by the following example:

Speaker A: Can you close the windows?

Speaker B: Sure.

In this example, speaker A is understood by speaker B to be making request. According to Grice, speaker B reasons that speaker A has some reason to make this utterance: it must be somehow relevant to the situation. Conversation implicatures are understood by reference to a *conversational situation* and are inferred via a cooperative principle and general maxims of conversation (Grice, 1975b, p. 46; covered in more depth later in this chapter). To this end, conversational implicatures also depend on conversants understanding each other’s goals and intentions.

In contrast to a formal study of semantics, the discipline of pragmatics holds that the study of language must systematically study meaning *in the context of use* (Huang, 2007). Under this view, pragmatics is considered a component of an integrated theory of language much as phonetics, phonology, morphology, syntax, and semantics. Phenomena typically considered pragmatic include **implicature**, **presupposition**, **deixis**, and **speech acts**.

Presuppositions share some properties of implicature. They are also derived by inference. The famous sentence “the king of France is bald” presupposes that there is a king of France, though this is not explicitly stated. Unlike implicature, the presupposition does not change under negation. “the king of France is not bald” still presupposes a king of France.

Deictic expressions are those that make reference to some aspect of context. This context may be expressed symbolically as in “tomorrow”, “here”, “you”, or “over there”, accompanied by gesture. Deixis is a universal phenomena expressed in all languages. Deixis exists because “a language without deictics cannot serve the communicative needs of its users

as effectively and efficiently as a language which does have them" (Huang, 2007, p. 132).

Speech acts derive from Austin's theory of performatives. Under a more general theory of speech acts, speech acts follow certain conventions. Searle (1969) defined five types of (illocutionary) speech acts:

1. **Assertives** - speech acts which express a proposition with a truth value and play the role of asserting, claiming, reporting, etc.
2. **Directives** - speech acts which represent an attempt for the speaker to get the addressee to do something as in a question, request, command, etc.
3. **Commissives** - speech acts which commit the speaker to do something in the future as in a promise, offer, threat, etc.
4. **Expressives** - speech acts which expresses a psychological state or attitude such as praising, thanking, apologizing, etc.
5. **Declarations** - speech acts which intend to change bring about change as in a declaration, nomination, etc.

Functionally, speech acts express strategies for communication. While speech acts have long been considered linguistic acts (Levinson, 1983; Searle, 1969, and many others), they also have a social function.¹ In standard speech act theory, illocutionary acts are directed toward addressees. H. H. Clark and Carlson (1982) argued, however, that such acts may also be directed toward other participants that are not addressees. Furthermore, they argued that there must be an act called an *informative act* which is directed toward participants who are not addressees. Consider this excerpt from the 1998 movie, *The Truman Show*, where the lead character is unaware that he is starring in a live television drama:

T: Why do you want to have a baby with me? You can't stand me!

M: That's not true! ... Why don't you let me fix you some of this Mococoa drink, all natural cocoa beans from the

¹ In fact, Geis (1995) argued that speech acts are not linguistic acts at all — but social acts.

upper slopes of Mount Nicaragua, no artificial sweeteners!

T: What the hell are you talking about?!? Who are you talking to?!?

M: I've tasted other cocoas, this is the best!

Truman becomes confused because he recognizes that his wife, Meryl, is directing information towards an audience while speaking with him. It's easy to see that participant roles have bearing on how we produce language in interaction.

One of the great challenges of studying pragmatics is the use of language in social contexts. Of the three experiments in this dissertation, the first concerns the hypothesized incidence of implicature, the second deixis, and the third concerns participant roles in interaction. In Chapter 8, I suggest that these several case studies are potentially indicative that pragmatic reasoning occurs during user interaction online. As we will see, such reasoning may also lead to mis-understanding.

3.2 LANGUAGE IN INTERACTION

While philosophy was the well-spring of pragmatics, sociology and anthropology were more concerned with interactional aspects of language. John Gumperz is widely regarded as the "father" of interactional sociolinguistics. He and Dell Hymes studied how situations and cultures affect meaning in social interaction (Gumperz, 1982; Hymes, 1974). In addition to how words and linguistic expressions affect meaning, they also studied how "contextualization cues" in prosody and register (i.e., language used in particular social settings) affect discourse. Gumperz widely influenced the perspective that inferential processes in interaction could be studied systematically in natural communicative contexts. For example, he gives examples of how microinteractions, such as the timing of prosodic and nonverbal cues, give evidence of breakdowns in conversational coordination (Gumperz, 1982).

Both Gumperz and Hymes strongly referenced pioneering work of Sacks, Schegloff, and Jefferson (Jefferson, 1972; Sacks, Schegloff, & Jefferson, 1974; Schegloff, Jefferson, & Sacks, 1977; Schegloff & Sacks, 1973). This body of work itself was inspired by the work of sociologist Harold Garfinkel, who was concerned with how activities and routines in daily life

inform social interaction. Garfinkel (1967) describes research conducted in the 1960s at the Los Angeles Suicide Prevention Center (SPC) and UCLA Outpatient Clinic. He asked the question, “by what criteria are its applicants selected for treatment?” (Garfinkel, 1967, p. 18). Garfinkel’s work directly inspired Sacks, Schegloff, and Jefferson to engage in a detailed analysis of conversation which revealed basic organizational features of conversation (Sacks et al., 1974).

Their basic insight was that speakers organize contributions in discrete units analyzed as turns. Turns themselves are organized into sequences. According to Sacks, each turn adds to the context affecting what will be done or said in the next turn. Speaker’s orient toward each other’s turns revealing systematic processes regulating both speech acts contributions as well as mechanisms for error detection and repair (Levinson, 1983). Pioneering work in conversation analysis in the tradition of ethnomethodology led to deeper understanding of how people manage conversation in interaction and how this system for management helps us to recognize error and repair it when detected.

Sacks, Schegloff, and Jefferson (1974) suggested that “rules” (what we now think of patterns) governing the organization of dialogue form the basis of how people manage conversation. For example, short silences or overlap in speech affect how the next speaker in a conversation is selected. Also, prosodic and gestural cues serve as signals during the coordination of exchanges. In addition, structural features such as turn transitions play a role in the detection of error and constrains repairs in dialogue. And repair sequences themselves respect the turn-taking system (Schegloff et al., 1977).

Herb Clark, noted cognitive psychologist from Stanford University, studies cognitive and social processes in language use. Building off sociolinguistic theory (H. H. Clark & Schaefer, 1989; Jefferson, 1972; Sacks et al., 1974; Schegloff et al., 1977; Schegloff & Sacks, 1973), he developed a theory of conversation as joint action² (H. H. Clark, 2001). H. H. Clark and

²

Clark on Joint Acts: Conversations reflect the joint activities they coordinate. Every joint activity has participants — the people actually taking part, who are distinct from nonparticipants (bystanders, onlookers, overhearers) — and so do the conversations that emerge from them. The participants take particular roles, such as doctor and patient, teacher and student, or friend calling and friend called, and the roles con-

Schaefer (1989) observed that basic turn construction units patterned into larger chunks of discourse they termed "contributions". These units are composed of a distinct presentation phase followed by an acceptance phase. Contributions form a hierarchy where every signal presented (verbal or non-verbal) belongs to the presentation phase of a contribution.

In this example from Clark, the presentation of "how much does Norman get off – " ends with "oh".

PRESENTATION PHASE

- A. is it, how much does Norman get off –

ACCEPTANCE PHASE

- B. pardon
- A. how much does Norman get off
- B. oh

Clark noted that the acceptance phase, generally initiated by B, offers A positive evidence that he understands by what A meant. This could be a verbal response or demonstration of some sort. When B has difficulty understanding, the acceptance phase may become prolonged.

H. H. Clark and Schaefer (1989) posited abstract levels of communication based on patterns of contributions. While a contribution is closed with acceptance, evidence of understanding has the properties of "upward completion" and "downward evidence". For any utterance, B may believe he is in one of four successive states of understanding.

STATE 0. B didn't notice that A uttered something.

STATE 1. B noticed that A uttered something (but wasn't in state 2).

strain what the participants do and say. Every joint activity has public goals — mutually agreed-upon purposes for carrying them out. The overall goal may be to exchange gossip, plan an outing, or negotiate a contract, and these have sub-goals. Although some of these goals are set from the start, most get established as the participants go along. The participants also have private goal — to be polite, not to lose face, or to finish quickly, for example — and these, too, constrain what the participants do and say. Finally, people often engage in two or more joint activities at a time — such as gossiping and eating dinner together — so their conversation switches back and forth between them. (2001, p. 2744)

STATE 2. B correctly heard the utterance (but wasn't in state 3).

STATE 3. B understood what A meant by the utterance.

Understanding in each successive state presupposes understanding in the prior state (unless evidence from a lower state suggests otherwise).³

The contribution model serves as a means for adding to discourse. However, the amount of evidence required for demonstrating understanding may vary depending on the task. Task-oriented dialogues often require stronger evidence of understanding than other sorts of dialogues (H. H. Clark & Wilkes-Gibbs, 1986a). By the principle of least collaborative effort, people ground with as little combined effort as needed for the situation at hand (Brennan & Clark, 1996).

When communicating through a different medium (e.g., phone), H. H. Clark and Brennan (2005) predicted that people ground using whatever techniques are available to ground with least collaborative effort. They noted that while a hearer backchannel response such as "okay" may be little effort face-to-face or on the phone, it may take considerable effort when teleconferencing or in a chat. The cost of acknowledgement is higher since it may interrupt the speaker.

H. H. Clark (1996) observed that dialogue gives insight to factors contributing to successful communication. This includes:

1. How speakers and listeners cooperate to exchange information;
2. How speakers design utterances to help listeners understand; and,
3. How listeners provide speakers with evidence of understanding.

For conversations to succeed, participants **ground** what they say: they establish the mutual belief that addressees have understood the speaker well enough for current purposes (H. H. Clark & Schaefer, 1989). This **Grounding Hypothesis** suggests information acquired by participants accumulates in a principled way: "each party is responsible for keeping track

³ (H. H. Clark, 1996) proposed one final state: state 4: B considers taking up A's proposed joint project.

of what is being said, and for enabling everyone else to keep track of what is being said" (H. H. Clark, 2001).

With each contribution to the conversation, the current speaker presupposes the common ground already established; and all the parties, the speaker included, add what is new in that contribution to their common ground. (H. H. Clark & Schaefer, 1992, 208)

Thus, in its broadest interpretation, this view of discourse corresponds to *meaning exchange* using linguistic messages in social and cultural contexts. Importantly, speakers design their utterances taking into account their potential listeners — listeners themselves are agents in particular roles (H. H. Clark & Schaefer, 1992).

3.3 DISCOURSE

Discourse is a term that is broadly concerned with the use of language. In its broadest, most functional sense, discourse has been viewed as a kind of social practice (Brown & Yule, 1983; Goffman, 1981; Gumperz & Hymes, 1972; Hymes, 1974; Leech, 1983). In interactional sociolinguistics, discourse has been studied as "forms of talk" (Goffman, 1981), characterized as ritualized, or institutionalized talk across social and cultural dimensions. Such talk draws on socio-cultural knowledge governed by social norms which are shared, culturally specific aspects of interpretation (Gumperz, 1982; Gumperz & Hymes, 1964; Hymes, 1974). Discourse in this perspective is deeply interactive and contextual, acting as a mechanism or process for social exchange.

In conversation analysis, discourse is seen as structured; for example, as interactive frames or schemata such as routinized exchanges in Goffman (1981) "replies and responses" or Sacks, Schegloff, and Jefferson (1974) "adjacency pairs". The recognition of such schemata are learned and form the basis for "co-occurrence expectations" (Gumperz, 1982, p. 162). Such conversational inferences are context-bound and conceived as preferences, maxims, or tendencies.

Discourse is also the focus of study in text analysis. "In its most general significance, a text is a sociological event, a semiotic encounter through which the meanings that constitute the social system are exchanged" (Halliday, 1977, p. 199).



Figure 11: Discourse Continuum (adapted from H. H. Clark, 1996)

Such study is concerned with phenomena relating meaning and structure beyond the bounds of a simple turn or sentence — “stretches of language received to be meaningful, unified, and purposive” (Cook, 1989, p. 156). This includes textual structures playing a role in cohesion, coherence, narrative, and rhetorical structure; cognitive structures such as theme-rheme, topic-focus, foreground-background, given-new, contrasts; and also, syntactic and semantic structures participating in discourse-level anaphora, presupposition, and tense.

3.3.1 Discourse Continuum

Discourse analysis widely includes practices listed above — interactional sociolinguistics, conversation analysis, text analysis, and pragmatics. From a cognitive perspective, H. H. Clark (1996) considers language as a form of **joint action** embodying individual and social processes. The use of language in joint activities lies on what he calls a *discourse continuum* (11 H. H. Clark, 1996, p. 50).

To the far right, all activity is accomplished using language. Moving to the left, language is augmented by diagrams, gesture, video etc. In the middle, language is balanced by action, while further to the left, language has less of a role in the content and coordination of activity. But, according to Clark, “**if we include any signal — any communicative act — then language use is present across the entire discourse continuum**” (H. H. Clark, 1996, p. 51).

In this dissertation, *I claim that user interface components, to include interactive graphics, can be considered as units of discourse* falling somewhat on the right side of this scale. Undoubtedly, such components contain both graphical and textual elements, but they are also interactive, and as such, trigger patterns of thought and behavior associated with language use. Though such interaction is not strictly conversation —

since applications do not have beliefs and intentions operating in coordination with the user — application designers model user tasks relying on the user's basic linguistic competence to interpret cues during interaction. This idea will be discussed further Chapter 8.

3.3.2 Discourse Models

Ultimately, underlying language use are speaker intentions. Such intentions account for why someone can express the same concept in different ways. For example, [Donnellan \(1966\)](#) gave the famous example, "Smith's murderer was insane" as an attributive description versus referential description (e.g., "Jones"). In his view, to say "Smith's murderer" is to say something about this person rather than to pick out or identify him.⁴

But this sort of indefinite reference raises problems for a semantic analysis of reference. To account for a range of discourse phenomena such as this, [Kamp and Reyle \(1993\)](#) proposed a theory of discourse representation which includes a representation of a **discourse model**. This structure contains a set of discourse referents (entities under discussion) and conditions which represent information that has been given about referents.

[Kamp and Reyle \(1993\)](#) and, independently [Heim \(1982\)](#), made it possible to analyze chunks of discourse using a semantic analysis. But the idea of discourse models was not entirely new. "Discourse models make explicit the structure of not of sentences but of situations as we perceive and imagine them" ([Johnson-Laird, 1983](#), p. 419). Such models help to account for how people comprehend language. One piece of evidence given for discourse models are **bridging inferences** — a sort of implicature by which people flesh out missing details in discourse ([H. H. Clark, 2002](#)). [Stenning and Van Lambalgen \(2008\)](#) and [Johnson-Laird \(1989\)](#) cited work from [Bransford, Barclay, and Franks \(1972\)](#) giving experimental evidence that recalled sentences are inferences from explicitly presented material. An example used was,

- (a) Three turtles rested on a floating log and a fish swam beneath them.

⁴ [Kripke \(1979\)](#) distinguished between these two descriptions as contrasting a semantic referent with a speaker referent.

Subject later confused this sentence with:

- (b) Three turtles rested on a floating log and a fish swam beneath it.

[Bransford et al. \(1972\)](#) argued that memory for language draws on the use of representations organized around referents (entities) in scenes and events. Models are updated incrementally and inferences about participants are made when needed. In this way, its possible to remember things about a discourse referent while not remembering how the information was conveyed.

The notion of direct and indirect representations made by [Stenning and Oberlander \(1995\)](#) is important. In their view, while indirect representations such as those used in language have a syntax, direct representations do not. It is the aim of a discourse model to provide a direct representation. As discourse proceeds, the discourse model is revised to accommodate new beliefs, but is also subject to revision when conflicting information arises. The notion of a discourse model has been crucial to theory in linguistics that accommodates a wide range of semantic-pragmatic phenomena to include anaphora, deixis, tense, and presupposition. Discourse models have also been applied to graphical reasoning where such representations are posited as a means to achieve efficiency in cognitive processing given constraints in working memory ([Stenning & Oberlander, 1995](#)).

Discourse models have been formalized in semantics ([Kamp & Reyle, 1993](#)), computational linguistics and artificial intelligence ([Grosz, Pollack, & Sidner, 1989; Hobbs, 1985](#)), cognitive science ([Johnson-Laird, 1989; Stenning, 2002; van Dijk & Kintsch, 1983](#)) and cognitive psychology ([H. H. Clark, 1996](#)). In these models, discourse is structured to account for informational links as well as the role of context in meaning ([Grosz & Sidner, 1986](#)). Discourse is also considered fundamentally cooperative, thus models take into account belief and attentional states of addressees ([H. H. Clark, 1996](#)). Finally, such models account for the role of lexical, grammatical, and structural properties of language in pragmatic inference.

Thus, *discourse processes concern how information is comprehended and produced*. This involves creating, accessing, and using discourse representations (or models) ([Graesser, Gernsbacher, & Goldman, 2012](#)). As [Graesser, Gernsbacher, and Goldman \(2012\)](#) observed, discourse processing research is

multi-disciplinary encompassing linguistics, psychology, computer science, and neuroscience. Techniques used in study include corpus analysis, mathematical and statistical modeling, “on-line” experimentation (e.g., during comprehension), and brain imaging.

While discourse researchers have busily been studying new forms of communication — some broadcast-oriented, such as lectures, radio /podcast, print advertising — others more interactive, such as those found in computer-mediated discourse (e.g., discussion forums, chat, microblogging, and blogging), little attention has been yet afforded to direct interaction with media such as multimedia advertisements. To do so would necessitate analyzing the coordinated use of both graphics and language. The next section of this chapter specifically looks at work associated with graphics and diagram understanding and the role of discourse models as a unified representation for textual and graphical elements.

3.4 APPLICATIONS

In this section, I consider two applications of discourse understanding relevant to claims made in this dissertation: graphics communication and advertising language. The first summarizes work claiming that implicatures may be understood in graphical diagrams. The second describes how advertisers make use of language to persuade and manipulate. Both are relevant to a discussion of how people may be manipulated in user interaction in the context of online behavioral advertising.

3.4.1 *Graphics Communication*

There is reason to think that messages produced by graphical user interfaces may be interpreted under the same inferential model of communication as language. Stenning, Lascarides, and Calder (2006) define graphics as “planar displays of information that use the distribution of shapes, patterns, and annotations and the relation between them to convey information” (Stenning et al., 2006, p. 476). In simpler terms, graphics use spatial relations, in addition to text labels, to convey meaning. Good examples are maps, tables, and charts.

Stenning et al. (2006) note that their definition also fits textual language. While text on a page is essentially organized

linearly, there is an abstract syntactic structure which encodes meaning relations between parts – thus mapping to space indirectly. Graphics, on the other hand, encode meaning in a direct spatial representation.

Diagrams are limited in the number of dimensions they can use to map relations. This might include icon shape, color, order, etc. Each dimension is directly mapped to a semantic representation. Phrases, words, morphemes, and phonemes, on the other hand, pack many relations across multiple abstract levels, each syntactic relation having a distinct semantic representation.

B. Tversky (2004, 2010) summarized a body of psychological experimentation demonstrating parallels between graphics, gesture, and language. Shared cognitive principles allow for mapping between visual and verbal modes in a number of domains. She described work from Denis (1997) in which route directions and maps could both be decomposed into a series of segments. Each segment was described to have four elements: a start point, orientation, progression, and end point. For example, “exit the Central Square station, turn left, go down Mass Avenue until you come to Cafe Centro” (B. Tversky, 2004, p. 146). The correspondence between both descriptive textual elements and graphical depictive elements suggest a common, underlying conceptual structure.

Tversky argued that such meanings of schematic visual forms (glyphs and other devices) both convey and constrain meaning B. Tversky (2010). Glyphs combine into complex diagrams using domain-specific syntax rules in a manner similar to linguistic structures (B. Tversky & Lee, 1999). Similarly, B. Tversky (2010) noted that glyphs may be used to abstractly represent a variety of concepts such as objects, collections, relations, and processes. Such representations participate in cognitive processes such as “inference, analogy, generalization, transfer, and insight” (B. Tversky, 2010, p. 2).

Though graphics and language differ in expressiveness and information packaging, they can both be studied using pragmatic theory (Stenning, 2002; Stenning & Oberlander, 1995). Johnson-Laird (1989) described the machinery by which linguistic representations (logical syllogisms expressed graphically) can be used to construct and update a mental model⁵.

⁵ Mental models are considered equivalent to discourse models (Johnson-Laird & Garnham, 1980). However, the term "discourse model" is pre-

A syllogism consists of two premises followed by a conclusion as in:

1. All A are B
2. All B are C
3. Therefore, all A are C

Syllogisms are defined from a narrow set of “moods”: All A are B, No A are B, Some A are not B (Khemlani & Johnson-Laird, 2012). It is also possible to make constructions that look like syllogisms such as:

1. In some cases when I go out, I am not in company.
2. Every time I am very happy I am in company.
3. Therefore, in some cases when I go out, I am not very happy. (Khemlani & Johnson-Laird, 2012, p. 1)

What makes example above different is the application of *quantifiers* (e.g., more than) and *determiners* (all, every, some, etc.) and over more complex objects such as sets and events. Contrast this with the limited logic of “all”, “some”, and “no” in a syllogism. The construction seems very similar — but its meaning potentially much more complex.

Such constructions are interesting to study because they are informative of how people reason and also amenable to online (realtime) experimentation. Interestingly, not only does content affect reasoning (e.g., people are more likely to accept a believable conclusion than unbelievable) but when encountering such constructions, people are also highly susceptible to pragmatic inference (Khemlani & Johnson-Laird, 2012).

Another interesting property is that they can be expressed graphically. In fact, such graphics are effective methods for teaching deductive logic (Stenning & Oberlander, 1995). Euler diagrams, such as Figure 12 helps aid comprehension by providing a more direct representation of the syllogism at the top of this section.

Johnson-Laird (2007) argued that, in fact, people do not learn as children a “mental logic”. They acquire understanding of lexical structures (including connectives, quantifiers,

ferred here as it is has been more widely adopted by researchers in the academic traditions discussed.

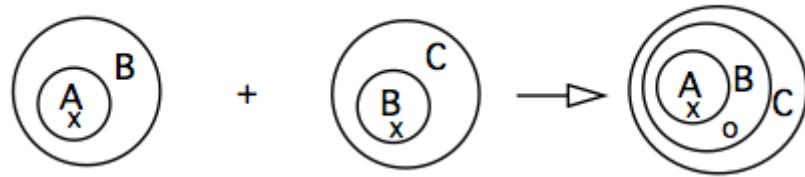


Figure 12: Graphical Syllogism (Stenning & Oberlander, 1995)

relation terms, etc.) and use these to construct models of premises. Inferential schemas and the ability to calculate semantic validity are what's necessary for logical reasoning.

While Oberlander (1995) argued for graphical implicature in the use of graphical notations for computer-assisted design in electronics, I argue that implicature can also be found in user interfaces. This makes them particularly amenable for deliberate manipulation. The next section focuses on how advertisers use language to persuade and manipulate. The use of graphics and language for advertising have been a rich area of study, though not yet in the context of interaction.

3.4.2 Advertising and Language

Cook (2001) argued that advertising is, in itself, a form of discourse. Though previous linguistic analysis tackled such discourse by focusing primarily on textual or verbal elements, he argues that other contextual elements should not be divorced from analysis. This is in contrast to earlier attempts to study language use in advertising where linguistic analysis had focused primarily on textual content and style (Leech, 1966; Vestergaard & Schröder, 1985b).

As Geis (1982) observed in his study of the language of television advertising,⁶ the goal of commercial advertising is to get consumers to buy. Geis (1982) focuses especially on the arguments that advertisers use in order to persuade users to buy products. By persuasion, Geis meant the process by which a message is taken up such that the recipient is convinced of its truth or is moved to act. He distinguished this from manipulation in that *persuasion involves conscious evaluation by the recipient on the source message and manipulation does not.*

⁶ Geis's (1982) study is based on the analysis of 800 commercials made between 1978 and 1981.

Not surprisingly, language can and does play an important role in the process of persuasive advertising. Advertisers use language to convey messages, but also for more subtle purposes. As Geis noted, language performs a role in getting viewer attention, supporting arguments used to persuade, and playing a role in facilitating consumer memory on the desirability of the product or service (Geis, 1982, p. 23).

Central to how advertisers use language for persuasion is the notion of implication. Advertisers are bound legally not to use deception in order to sell. Brandt and Preston (2013) observed that the FTC took a strong position on "misleading advertising" during the 1970s. During this time, it recognized the role of cognitive research in demonstrating how consumer world knowledge might interact with ads to form false beliefs. Thus the FTC began to rely on behavioral research to determine whether an ad was misleading. In fact, there are many ways ads can be misleading without direct assertion.

Probably a fundamental reason has been the more subtle forms of deception that the FTC has become desirous of controlling. Was it deceptive for Wonder Bread (ITT, 1973) to claim it is high in nutrition when the claim is literally true yet might imply falsely that the quality is unique to itself? Did the name "Safety Champion" (Firestone, 1972; Firestone, 1973) imply that a tire is perfectly safe, or that it is safer than other tires? (Brandt & Preston, 2013, p. 5)

Persuasion Through Implicature

Geis adopted a Gricean position on advertising precisely because advertisers often persuade through the use of implicatures. He gives the following example,

- (a) Wartsoff contains Vivaline and you know that Vivaline removes warts instantly. (Geis, 1982, p. 26)

The implication is that Wartsoff removes warts instantly. An unsuspecting consumer might reason that since the advertiser used the verb "know", there must be solid evidence that Vivaline removes warts instantly.

Geis focused on three sorts of implicatures: 1) conventional implicatures; 2) theoretical implicatures (those which depend

on listener belief for validity); and, 3) conversational implicatures.

The Wartsoff statement is an example of a conventional (embedded) implicature by the use of "know". Because the word "know" has been used, it is implicated that there is good evidence supporting that Vivaline removes warts instantly.

A theoretical implicature, on the other hand, is given by the example:

- (b) Choosy mothers choose Jif. (Geis, 1982, p. 48)

The implication is that if a mother doesn't choose Jif, she isn't a good mother.

Finally, the following is an example of a conversation implicatures:

- (c) We're building a reputation, not resting on one. (Geis, 1982, p. 50) citing an ad in Newsweek, 2/6/1978

Here the implication is that Ramada Inn has a leading competitor that rests on its reputation. But, of course, this is not explicitly said. However, because resting on a reputation was mentioned — it must in some way be relevant to the listener.

To back off just a bit, Grice (1975a) proposed that language is governed by the Cooperative Principle. According to Grice, the cooperative use of language in conversation appears to follow certain maxims, or rules. In particular,

Quantity: Say no more and no less than is necessary.

Quality: Say what you believe to be true; do not say what lacks adequate evidence.

Relation: Say what is appropriate at the appropriate time; be relevant.

Manner: Be clear; be brief; be perspicuous; avoid obscure expressions.

Grice hypothesized that such maxims generate implicatures. In particular, he focused on examples of when such maxims were "flouted" such that the listener is expected to understand the message, despite violation of conversational rules.

In fact, we flout these maxims on a regular basis. Examples of obvious flouting are in the use of irony and metaphor.

- (d) Brilliant! (Exclaimed after tripping)

- (e) Great weather we're having. (After sudden clap of thunder)
- (f) Those two are peas in a pod.

Tanaka (1999) noted the use of metaphors and puns in advertisement.

- (g) What makes a shy girl get Intimate (Intimate, Revlon) (Tanaka, 1999, p.104)

She suggested that the reader is expected to get and reject:

- (h) What makes a shy girl become intimate.

in favor of:

- (i) What makes a shy girl buy Intimate perfume.

Likely, such use of implicature is intended to attract attention and intrigue and not to mislead.

Geis (1982) argued that while advertisers may be held accountable for misleading the public through conventional and theoretical implicature, this may not be the case for conversational implicature.

- (j) Wet feet? LOOK OUT FOR A COLD — Gargle with LISTERINE QUICK (Geis, 1982, p. 49)

In the example above, attributed to B. Clark (1944), Listerine is conversationally implicated by the Maxim of Relevance. In fact, Geis mentions that after 40 years of advertising in this manner, the advertisers were finally compelled by the FTC to add the disclaimer:

While Listerine will not help prevent colds or sore throats or lessen their severity, Listerine's strong formula keeps your breath clean for hours — it kills the germs that can cause bad breath.

In this disclaimer, the advertiser still exploits a conventional implication by stating that Listerine does not help prevent colds. As Geis notes, conventional implicatures are perceptually less salient than assertions though, nonetheless, implicate a relation.

Overt and Covert Communication

Sperber and Wilson (1986) argued that there is reason to recognize a distinction between an *informative intention* and *communicative intention*. The first makes “manifest or more manifest to the audience a set of assumptions.” (Sperber & Wilson, 1986, p. 58) The second makes it “mutually manifest to audience and communicator that the communicator has this informative intention.” (Sperber & Wilson, 1986, p. 61)

In **ostensive-inferential communication**, as described by Sperber and Wilson (1986), an ostensive act by the communicator is intended to be recognized. Thus, a pointing gesture intends to draw attention to some aspect of the physical context that the communicator thinks worthy of attention.

In Sperber and Wilson (1986) view, **relevance** is the key to human cognition. Humans pay attention to things most relevant. A communicator that *manipulates context* has an affect over what is perceived more or less relevant. This means that the communicator needs to be aware of what contextual information is accessible to the addressee in order affect the process of comprehension.

From this point-of-view, ostensive communication is successful if the hearer is able to “recover the speaker’s informative intention”. **Covert communication**, on the other hand, is described as:

A case of communication where the intention of the speaker is to alter the cognitive environment of the hearer, i.e., to make a set of assumptions more manifest to her, without making this intuition mutually manifest. (Bencherif & Tanaka, 1987, as quoted in Tanaka 1994, p. 41)

Tanaka (1999) argued that covert communication is used in advertising for two main purposes. First, to make the addressee forget that he is trying to sell her something. Second, to “avoid taking responsibility for the social consequences of certain implications arising from advertisements” (Tanaka, 1999, p. 44). Tanaka follows with a number of examples centered on the use of sexual imagery and captioning to promote products which are not obviously sexual. For example, she shows an image from a watch ad where there is a background image of a man and woman in swimsuits embracing.

A subtle caption indicates: “designed to perform.” Though the ad overtly communicates that these watches are designed to perform well in aquatic sports, there are strong suggestive overtones in the image clearly implicating sexual performance.

As Sperber and Wilson (1986) noted, a large difference between Grice’s approach (to include neo-Gricean theories proposed by Horn, 2006 and Levinson, 2000) and theirs (i.e., **Relevance Theory**) is that Grice’s principle and maxims are norms that all people are expected to know and recognize. Implicatures are recognized and understood as violation of these known norms. By contrast, Relevance Theory is general to the point that people don’t need to recognize or know anything. All that is necessary is that each ostensibly communicated act have a presumption of relevance.

As advertising has adapted to new forms of media, so has the means by which advertisers harness language. Though researchers have made study of press and television advertising, little attention has been paid to behavioral and interactive advertising (with the exception of privacy notices). One can already imagine ads where people who look like one’s friends recommend products used by those friends. More subtle are ad interactions, hidden in plain sight, that alter one’s beliefs in fleeting moments in time. It would not be surprising to find that advertisers exploit cognitive tendencies in inferential communication activated during the course interaction.

3.5 CONCLUSION

On the basis of reviewed literature, I propose that there are gaps in knowledge that have not yet been addressed: first, that interaction with graphical user interfaces involve discourse processes affecting comprehension; and second, that online behavioral advertising participates in a new form of advertising discourse where user’s beliefs are affected during the course of interaction. I show that this view gives insight into communication errors difficult to explain otherwise.

The next chapter describes my research program and method for data collection and analysis. To follow in Chapters 5 through 7, I focus on specific sorts of pragmatic misunderstandings in the domain of online behavioral advertising. Each time a user opens a web browser and interacts with a website, that user is challenged to make decisions about

what to communicate — and to whom — in an environment where commercial actors actively encourage disclosure. Because discourse comprehension relies much on inference, manipulation may subtly take advantage of automatic processes such that the user is unaware of the potential for misunderstanding. Though Kahneman and Tversky (1984) and others have observed that small changes in context affect judgement and decision-making, there is considerably less research on how small changes in context affect understanding during the course of interaction in decision-making tasks. This is the topic under discussion in the remainder of this dissertation.

METHOD

In this chapter, I outline a general research program and associated substantive and experimental hypotheses. Following, I describe the use of Amazon Mechanical Turk for the collection of behavioral data, general procedure followed, and summary of the population sample observed.

4.1 RESEARCH PROGRAM

There is little question that interaction with computers change what we know and believe. Usability has long been concerned with the fit between user “mental model” and the design of user interfaces. Typically, such studies include the use of qualitative research techniques such as cognitive walk-throughs, eye-tracking, and card sorting. Such studies are often task-oriented and situated to address the use of specific products.

Dan Saffer, in his book *Microinteractions* (2013), touches upon user interaction with user interface components, or features, that perform only one small task. For instance, a login screen embodies a self-contained interaction that may be quite small, but extremely important for user experience. This microinteraction has a clear start and end. As Saffer notes, some triggers to enter a microinteraction are user-initiated, and some may be system-initiated. Regardless, I believe there is a clear parallel to conversational discourse discussed in the previous chapter: users exchange information interactively.

This dissertation examines several microinteraction in the domain of online behavioral advertising. Of particular interest are phenomena relating to user confusion in the context of online interaction.

The substantive hypothesis for questions asked in this dissertation is that *some user confusion stemming from online behavioral advertising is caused by error in discourse understanding*. I contend that by regarding user interaction with graphical user interfaces as a form of discourse, we can account for some sorts of user confusion that may be difficult to see or may otherwise go un-noticed.

The next section summarizes specific research hypotheses for this dissertation. The research approach used to examine these hypotheses is *randomized experimental*. The purpose for adopting such an approach is to determine causality for some sorts of user confusion. Each research hypothesis is paired to some aspect of discourse theory. Therefore, Chapters 5, 6, and 7 more fully motivate the connection between theory and practice prior to a presentation of results.

4.2 DESIGN OF EXPERIMENTS

The substantive hypothesis (user confusion stems from errors in discourse understanding) implies a set of research hypotheses, which in turn imply individual study hypotheses.

RESEARCH HYPOTHESIS 1:

Graphical user interfaces have properties comparable to other forms of linguistic communication.

RESEARCH HYPOTHESIS 2:

Graphical user interfaces can cause users to make faulty pragmatic inferences.

RESEARCH HYPOTHESIS 3:

A pragmatic inference derived while interacting with a graphical user interface may be affected when given immediate and direct feedback.

Each of the three sets of experiments described in Chapters 5-7 support one or more of the research hypotheses above. Specific research manipulations are described below. The number in the left column refers to the experiment number.

Research manipulation (by experimental condition):

- 1A: Suppose that pragmatic inference only occurs in linguistic (text or speech) interaction. A graphical user interface is used in the experimental condition whereas text is used in the control condition.
- 1B: Suppose that system feedback adds information to a situation model. Feedback is used in the experimental condition while none in the control condition.

- 2A: Suppose images and icons are objects that are both indexical. Images and icons are combined such that contrastive relations exist in the experimental condition while none exists in the control condition.
- 2B: Suppose advertising images are implicitly indexical. Explicit indexical element appears in the experimental condition while none exists in the control condition.
- 3: Suppose that a user designs utterances taking into account potential listeners. Listener roles are explicit in the experimental condition while implied in the control condition.

Finally, the table below summarizes experimental designs. Each should be considered a pilot study. That is, experiments described are new and there are no other results against which to compare. The goal is to apply difference statistics to determine whether the variables described are independent (null hypothesis) or dependent. A relation between variables indicates some degree of predictiveness and causality. We can use information about causality to guide, or motivate, a strategy for reducing the possibility of miscommunication.

Table 1: Experiment Design

1A	2x2x2 between group posttest-only design where the control group is presented with a set of textual expressions and asked to answer questions about their meaning. Treatment groups are presented with either textual expressions or a dialog box expressing the same set of choices and asked the same questions. Independent Variables: Modality, Deontic force, Attitude toward privacy. Dependent Variable: Pragmatic implicature
1B	Between group posttest-only design where two groups are presented with a cookie banner and later asked about whether or not they believe the website placed "cookies" in their browser. The treatment group is presented with feedback about the consequence of their action / non-action following presentation of the banner. Independent Variable: Feedback. Dependent Variable: Pragmatic implicature.

-
- 2A Between group posttest-only design where five groups are presented an advertisement in the context of a webpage and asked to identify hyperlinks. Treatment groups are presented an advertisement with embedded image icons at four levels (known icon + different company, unknown icon, known call-to-action (CTA), “DAA opt-out icon”) while the control group is presented with an embedded image from the same company as the advertiser. Independent Variable: Icon type. Dependent Variable: Indexicality of icon.
-
- 2B Between groups posttest-only design where three groups are presented an advertisement in the context of a webpage and asked to identify hyperlinks. Treatment groups are presented an advertisement at two levels (iconic CTA, textual CTA) while the control group is presented with an advertisement with no CTA. Independent Variable: Modality of CTA. Dependent Variable: Click target.
-
- 3 Between group repeated measures posttest-only design where a control group is asked to respond to a survey containing questions relating to activities (using questions drawn from [Acquisti, John, and Loewenstein \(2012\)](#); embarrassing, socially / ethically questionable, illegal). Participants are notified in a privacy statement that there may be ad trackers on the site. The treatment group receives exactly the same notification and survey. However, tracker presence is indicated in a visual display throughout the participant’s session. Independent Variable: Visual presence. Dependent Variable: Propensity to respond in the affirmative to engaging in specific behaviors.
-

Of course, not every usability problem needs be tested using experimental methods. But there are two good reasons to do so here:

1. No causality has yet been determined for presumed cases of mis-communication studied here. In fact, it may not even be apparent that there is a problem at all.

2. Though linguistic theory offers explanation for the sort of miscommunication described in experiments here, there are no earlier studies with which to compare. This dissertation intends to make such a relation clear and provide a path for future exploration.

Each of the research hypotheses above has bearing on current practice in online behavioral advertising as represented in technical specifications of existing or proposed standards. Though self-regulatory bodies for OBA suggest that specific principles are adhered, effectiveness is in question. Marketing advocates affirm that self-regulatory practices are effective ([Tribal Fusion, 2012](#)). However, research studies by policy advocates (e.g., [Komanduri et al., 2012](#); [Mayer & Mitchell, 2012](#)) have raised issues of industry non-compliance. Even if compliance were universal, usability studies have uncovered inadequacies in terms of communicative efficacy ([Blase Ur, 2012](#); [Hastak & Culnan, 2010](#); [Leon, Ur, et al., 2012](#)). The benefit of the approach in this dissertation is the ability to test technical specifications, and ascertain effectiveness in context.

One of the particular challenges of the randomized experimental approach is the need for a larger number of participants than one might expect from a traditional usability experiment. Increasingly, psychologists and linguists have come to rely on the large pool of volunteers available via the Amazon Mechanical Turk (AMT) crowdsourcing platform. The next section addresses both advantages and disadvantages of AMT for experimentation. Characteristics of the user sample studied have particular bearing on the research validity of experiments presented in this thesis.

4.3 MECHANICAL TURK AS A PLATFORM FOR HUMAN INTELLIGENCE TASKS

The original Mechanical Turk constructed in the 18th century was a hoax. In all appearance, he was a mechanical man dressed in an turban and robe, seated behind a large cabinet. Despite being mechanical, he played chess like no other machine could. Of course, unknown to most at the time, inside the cabinet was a tiny space where a real human played the game by controlling the automaton's arms on the chessboard.

By contrast, Amazon's Mechanical Turk (AMT) is no hoax. From a simple command line interface, requester's can create

“human intelligence tasks” and farm them out to a sea of workers. Within seconds, a task may be accepted, completed, and remunerated without the requester ever knowing or communicating with the invisible worker. AMT shares a simple design concept with the original Turk: a human is in the machine. The idea is a “job posting board” where human intelligence is needed. Any of a variety of “Human Intelligence Tasks” (HITs) can be posted along with small monetary rewards ranging from cents to dollars. AMT is a prototypical example of crowdsourcing — outsourcing to a large, undefined community of people (i.e., crowd).

4.3.1 *The AMT Marketplace*

In 2006, Wired magazine published the influential *The Rise of Crowdsourcing* (Howe, 2006), which tells the story of the rise of the amateur on the web. Highlighted, is the example of a stock photo site iStockphoto which created a marketplace for amateur photographers. Professional grade camera technology at affordable prices combined with the ability to upload content to the Internet, search and categorization, and micro-payments changed the market for photographers and photo purchasers. Over the last decade, the power of many users over distributed networks has changed the world. Today, FaceBook and eBay could not exist without the contributions of their users.

The irony is, of course, that by engaging in the creation of user-generated content on the web, people have become part of the machine. This has led to the well-known technology humanist Jaron Lanier to rail against the de-humanization of the web. His position is — to design applications that treat humans as machines devalues both people and content “making ourselves into idiots” (Lanier, 2006). Lanier particularly dislikes Wikipedia which creates a false sense of authority behind information by removing any connection between the real author of the information as well as a subjective context for the interpretation of content (Lanier, 2010).

Nevertheless, crowdsourcing has become a viable economic model on the web. More than a million workers login to crowdsourcing platforms to complete short tasks for pay-per-task compensation (Munro & Tily, 2011). And who is the invisible worker? As reported on The Mechanical Turk Blog in August of 2012, “turkers” comprise a workforce approxi-

mately 500,000 persons (at any one time) across 190 countries. In an analysis of the AMT marketplace over multiple studies, [W. Mason and Suri \(2011\)](#) report that the majority of turkers reside in the United States or India and tend more likely to be female (55%) with a median average age of 30 and, on average, earning 30K per year. According to Ipeirotis ([2010a](#); [2010b](#)), workers in the US are more likely to use AMT as a secondary source of income while those in India use it as a primary source of income. Workers are self-selected on the marketplace, working for cash payments in US dollars and Indian rupees. All studies agree that workers are represented across a wide range of ages, ethnicities, countries, languages and income.

The AMT marketplace has been active since 2005. Among the most well-known documented instances of use, turkers were asked to help search through sections of satellite imagery to flag images with anomalies. The hope was that turkers might be able to help pinpoint the 2007 crash site of aviator Steve Fossett. This event so permeated the media at the time, I signed up to become a turker in order to participate. In the end, Fossett was not found in the region examined by turkers. In fact, there was some criticism toward the use of turkers for annotating satellite imagery. With little understanding of what constitutes anomaly in satellite imagery, the usefulness of generated leads were considered by some to be questionable.

Since then, AMT has been used for a variety of tasks, some of which call for specific skill or expertise, such as knowledge of a foreign language. More common are tasks that involve transcription, rating, editing, and image annotation. The AMT website provides a basic web interface for loading HITs, though it is also possible for requesters to direct workers to an external website or directly embed that website as an iframe on the AMT site.

Workers themselves comprise a community that interact on public forums, such as *Turker Nation*, and through networks such as Twitter. Once on the AMT website, workers track their earnings (managed by Amazon), status (HIT success), qualifications, and the ever revolving database of available tasks. [Figure 13](#) is a screengrab taken from the AMT developer sandbox, a place for requesters to test tasks before deploying on the AMT website.

4.3 MECHANICAL TURK

The screenshot shows the Amazon Mechanical Turk (AMT) Requester interface. At the top, there's a navigation bar with links for 'go to MTurk.com', 'Mechanical Turk Developer Sandbox' (which is noted as being for test and development only), 'Learn more', 'Lisa Harper | Account Settings | Sign Out | Help', and the 'amazonmechanical turk' logo.

The main area has tabs for 'Home', 'Create', 'Manage', 'Developer', and 'Help'. Below these, there are two buttons: 'New Project' and 'New Batch with an Existing Project'. A link 'Create HITs individually' is also present.

A sidebar on the left titled 'Start a New Project' lists various 'Categorization' options: Data Collection, Moderation of an Image, Sentiment, Survey, Survey Link, Tagging of an Image, Transcription from A/V, Transcription from an Image, Writing, and Other. The 'Other' option is currently selected.

The central panel displays an example of categorization. It shows a photograph of a bedroom with a bed, a desk, and a chair. The task asks to 'Choose the best category for this image'. The categories listed are: kitchen, living, bath, bed, and outside. Below the image, a note says 'You must ACCEPT the HIT before you can submit the results.' A large orange 'Create Project +' button is at the bottom right of this panel.

At the bottom of the page, there are links for 'Help | Contact Us | Policies | Press Inquiries | Blog | Careers', 'MTurk.com | Requesters | Workers | Developers', and social media links for 'Follow Us on Twitter' and 'Become a Fan on Facebook'.

Figure 13: AMT Requester Website

The screenshot shows the Amazon Mechanical Turk (AMT) Worker HITs interface. At the top, it says 'You are using the Mechanical Turk Developer Sandbox. This site is for test and development only. Learn more'.

The main area displays a list of HITs. The first item is 'Write the words shown in an image' with requester 'Will Leong', expiration date 'Oct 17, 2013 (23 hours 58 minutes)', reward '\$0.02', and 2 HITs available. There are buttons to 'Request Qualification (Why?)' and 'View a HIT in this group'.

Below this are several other HIT descriptions, each with requester name, expiration date, reward, and number of HITs available. Examples include:

- '(beta) Mark nose tip in 72 faces' by 'Turk Experiment' with reward \$0.15 and 2 HITs available.
- '(beta) Mimic 10 facial expression' by 'Turk Experiment' with reward \$0.65 and 93 HITs available.
- '(beta) Label 3D orientation of 25 heads' by 'Turk Experiment' with reward \$0.25 and 50 HITs available.
- 'Phrase Pleasantness' by 'Lorenzo Gatti' with reward \$0.01 and 1 HIT available.
- 'Translation from English to Arabic (default)' by 'None' with reward \$0.75 and 30 HITs available.
- 'Describe an image - 2013-10-16 20:58:36' by 'IvanBoyke' with reward \$0.00 and 10 HITs available.
- 'What is the date on this receipt? (B1) (qa)' by '411Richmond' with reward \$0.02 and 2 HITs available.
- 'Transcribe the Store, Date, Total and Address from the receipt (B2) (qa)' by '411Richmond' with reward \$0.03 and 7 HITs available.

At the bottom, there are links for 'FAQ | Contact Us | Careers at Amazon | Developers | Press | Policies | Blog' and a note '©2005-2013 Amazon.com, Inc. or its Affiliates'. It also says 'An amazon.com company'.

Figure 14: AMT Worker HITs

Requesters define the task, number of tasks available, payment, qualification, and expiration date. Requesters must provide a validated US bank account and address in order to initiate tasks. Some basic worker qualification types are provided by Amazon out-of-the-box: worker HIT success and demographic region are commonly used by requesters. However, requesters also use qualifications as a mechanism for training workers. By requiring workers to "train" and go through testing, workers with special skills may be utilized in future tasks.

Because the worker success metric is often used as a filter for ensuring higher quality work — and as a means for obtaining higher paid HITS — workers take care to protect their rating. This means that requesters are also held accountable in their dealings with workers, via informal ratings over social media channels. Though some HITs may be paid automatically, it is fairly common for requesters to do quality control before accepting HITs. HIT rejection has negative consequences and is not taken lightly by the turkers.

There are a number of obvious advantages to using AMT. The workforce is roughly half a million strong and available around the clock. They represent a wide demographic range, many quite well educated. Amazon manages and automates payments and ensures the pool of workers abides by its terms of use. [W. Mason and Suri \(2011\)](#) observed that, for experimental research, a key advantage is faster iteration between hypothesis formulation and testing. However, there exists a number of real concerns and issues with crowdsourced experimentation. These are discussed below.

4.3.2 General Concerns with the Use of AMT for Experimental Research

First, and foremost, AMT is an Internet-based platform. [Reips \(2006\)](#) noted advantages as well as disadvantages to experimentation over the Internet (relative to the first four bullet points below). However, there are also number of other issues particular to Mechanical Turk.

- 1. User Fraud.** There is the possibility of worker "gaming" through multiple submissions. The Turk Requester's Blog

¹ gives an example from the Romanian Mturk Forum, in which on every page (on a total of more than 300), are discussions of how to break Amazon's terms of service. As a result, Amazon adjusted their policy to an invitation-only registration for International workers (Chiarella, 2013).

2. **Self-Selection.** Users self-select by reviewing titles, payment, time expected, and instructions. Unless controlled, users are generally able to preview external websites (or embedded iframes) to decide whether they wish to accept a given HIT.
3. **Drop-outs.** AMT workers have the ability to return or abandon HITs without affecting their reputation. However, requester's don't automatically see which workers accept — though don't complete — assignments, so it's difficult to gauge dropouts.
4. **Mis-communication.** Reduced communication between requesters and workers may mean that directions are not well understood or followed. This can be difficult to detect and debug.
5. **Screening.** Amazon provides no general mechanism for screening aside from a general mechanism for filtering via qualifications. Of note, is the need to screen for demographic criteria and prior participation. This study has the particular need for the latter: excluding workers who have participated in one experiment from participating in another. The mechanism that AMT provides is exclusion by assignment: once a worker has accepted an assignment, that particular assignment is no longer available. For example, Experiment 2A of this dissertation required at least 300 subjects. Once a worker accepted this HIT the other 299 assignments were no longer available. But, without cross-experimental controls, it would be possible for some worker to accept a HIT from a closely related experiment, potentially biasing results.
6. **Compensation.** The very fact that AMT is fee-based places obligation on both the requester and worker.

¹ <http://turkrequesters.blogspot.com/2013/01/the-reasons-why-amazon-mechanical-turk.html>

W. Mason and Watts (2009) investigated the relationship between compensation and performance in two experiments on AMT. They observed no interaction between difficulty of task and compensation on performance. They also found that increasing compensation alone did not improve accuracy. But how workers were paid (pay-per-word, pay-per-puzzle) did have an impact on output and accuracy.

While it is common for workers to accept tasks for pennies each, some workers find by working many low paying tasks in quick succession, they are able to earn close to minimum wage.² More typically they earn much less (Paolacci, Chandler, & Ipeirotis, 2010). Though most workers appear attracted by the earnings potential, others participate for reasons such as boredom, fun, curiosity, and even education (Behrend, Sharek, Meade, & Wiebe, 2011).

The studies in this dissertation limit participation to workers with a high HIT acceptance rate. From discussion on turker forums and blogs (e.g., *Turker Nation*, *mTurk Forum*, and *Turkkit-Reddit*), its clear that these workers expect to be paid more than other workers. And many take this work quite seriously.

7. **Quality and Reliability.** Beyond recruiting from the most reliable workers available, how much should quality be of concern? From prior crowdsourcing experiments on Amazon Mechanical Turk (AMT), quality is a valid concern (Callison-Burch & Dredze, 2010; Gormley, Gerber, Harper, & Dredze, 2010). In a study of translation tasks, researchers consistently receive poor and noisy translations including blank annotations, misspellings, copy-pasting of machine translations, and downright cheating (Ambati & Vogel, 2010). Having workflows for the filtering of noisy data has proven vital in this environment. However, it is feasible to filter noisy judgements of non-experts depending on the task: in an image annotation task Nowak and Rüger (2010) found that, while agreement between experts and non-experts varies depending on the measure used, its influence on image ranking as a whole, is minimal. Other sorts of studies

² See discussion at <http://mturkforum.com/showthread.php?2744-How-much-do-you-earn-per-hour>, for example.

echo this finding including labeling text with emotion (Snow, O'Connor, Jurafsky, & Ng, 2008), search relevance judgements (Alonso & Mizzaro, 2009), and more.

8. **Presentation Consistency.** All Internet experiments conducted in the wild suffer the need for exceptional attention to cross-browser effects and robust services. If the experiment relies on everyone being able to see stimuli at the same resolution and in the same manner, then both scripts and screens need to be tested carefully, both across platforms and browsers. Furthermore, hosted services require reliable access by hundreds simultaneously — and with little risk of downtime.
9. **Anonymity.** Finally, though we would like to believe that the human in the machine is anonymous, reality differs. It is estimated that 50% of workers have been linked to public Amazon user profiles (Lease, Hullman, & Bigham, 2013). This means that Amazon worker IDs are essentially PII. And, in fact, many workers are not blind to this.
10. **The "Superturker".** Legal scholar Dan Kahan warns particularly of a side-effect of prior, repeated exposure to cognitive studies on AMT (Kahan, 2013). Chandler, Mueller, and Paolacci (2013) note that while the probability that any one worker has seen some manipulation, there is a population of "superturkers" (prolific workers) who are significantly more likely to end up in studies. Pooling 16,408 HITs in 132 unique studies, they found that HITs were completed by 7498 unique workers. The top 10% of prolific workers completed 41% of total HITs. On a positive note, Chandler et al. (2013) note that these turkers are less likely to be multi-tasking and more likely to be available for follow-up studies such as required for longitudinal research. On the negative side, these workers are much more likely to have participated in cognitive tasks potentially biasing them for future cognitive tasks (Chandler et al., 2013).

4.4 GENERAL PROCEDURE

Each experiment in this dissertation was delivered over the Internet in a web browser. Subjects (AMT workers) were ran-

domly assigned to groups per study design. Experiments were implemented within a Qualtrics (<http://qualtrics.com>) survey using supplementary, custom JavaScript/HTML/CSS code that I wrote, as needed. Surveys were accessible via an HTTPS encrypted iframe on the AMT website (<https://www.mturk.com>).

In order to participate, AMT workers selected my HIT from a worker queue (Figure 14). Three base qualifications were stipulated:

- workers must be 18 years or older (provided by the Amazon Mechanical Turk Terms of Service in [Appendix B](#));
- workers must be physically located in the United States; and,
- Workers must have a HIT success rate of 90% or better.

The AMT API provided for qualification criteria such as geographic region and HIT success rate. By specifying these as requirements, Amazon automatically matched eligible workers to my “surveys”. A sample assignment (Experiment 1A) is provided in [Appendix C](#).

Using AMT command line tools, I tested each experiment in the AMT sandbox before deploying to AMT. Requesters can monitor assignment progress, approve or reject workers, download results, and assign bonuses from their dashboard. I tested assignments in several browsers both on a Mac and Windows PC to ensure functionality and appearance were consistent. Following successful tests, I pushed assignments to AMT.

Workers who selected one of my HITS were told they would be participating in either a language study, user interface study, image annotation study, or ethics survey, depending on the experiment. Because these were all studies of language use, workers were expected to have basic *linguistic competence* in English. It was assumed that if they were on the Internet capable of responding to a request to participate in an online experiment, and residing in the United States, they had adequate competence in English. However, I also provided a demographic survey question targeting fluency level, if the subject was not a native speaker of English.

Each experimental design was deployed as a single Qualtrics survey. Qualtrics provides for sophisticated functionality such as survey flow, block randomization, display and question logic, and a JavaScript application programming interface (API). Several of my surveys required hand-coded instrumentation not provided by Qualtrics. I discuss instrumentation in the methods section of each experiment described.

Because I conducted multiple, related experiments simultaneously, exposure to one experiment potentially disqualified a worker from other experiments. For example, Experiment 1A has a potential priming effect on the topic of Internet privacy that could potentially bias that same worker in Experiment 1B. In such cases, I included a message on the preview that said something like, “please don’t accept this HIT if you have previously participated in an experiment with a blue rectangle.” As reinforcement, I checked each worker on acceptance of consent against a database of prior workers. If a worker had already participated in a closely related experiment, a message was presented to that worker to please return the HIT. Then they were prohibited from further access. Even so, it was still possible for a single worker to participate in more than one experiment. For example, if a worker signed up for Experiment 1B, they were allowed to also sign up for 2B.

In each experiment, workers previewed basic instructions ([Figure 15](#)) before selecting the HIT. I inserted custom code to prevent participants from advancing beyond this screen until the HIT had been accepted. In order to proceed, the worker must have clicked “Accept HIT”.

After clicking “next”, a consent form was displayed (see [Appendix A](#)). Again, in order to continue, the participant must have selected “yes” to continue.

General experiment flow is depicted below ([Figure 16](#)).

At the completion of each survey, I configured assignments to automatically approve HITS so that Amazon would pay the worker immediately. Experiments were designed to take approximately three minutes to complete. I offered 15 cents for most. While it was also possible to manually review results before approving HITs, I decided this was not necessary for any of the studies included here since workers were recruited from those with known high performance. I did, however, manually assign random dollar bonuses after assignment completion.

4.4 GENERAL PROCEDURE

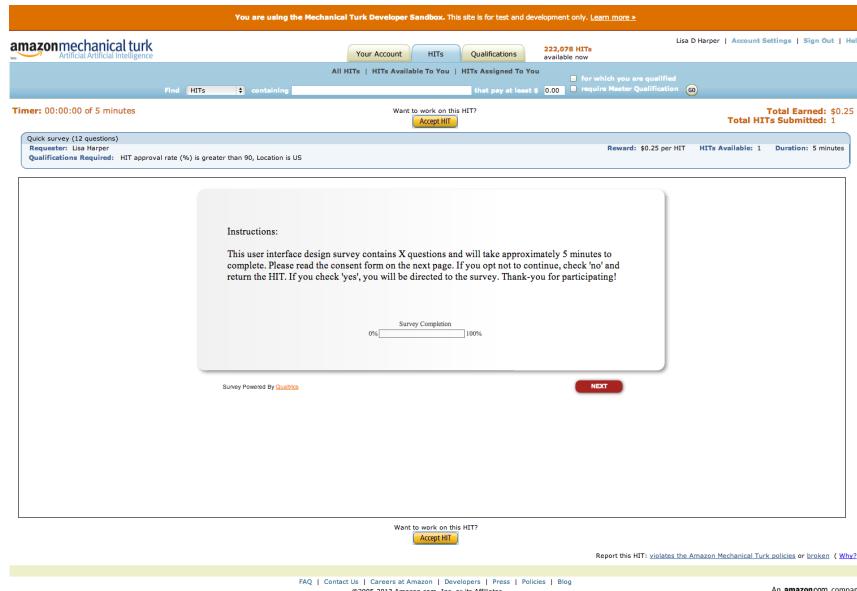


Figure 15: Participant Instructions (Experiment Three)

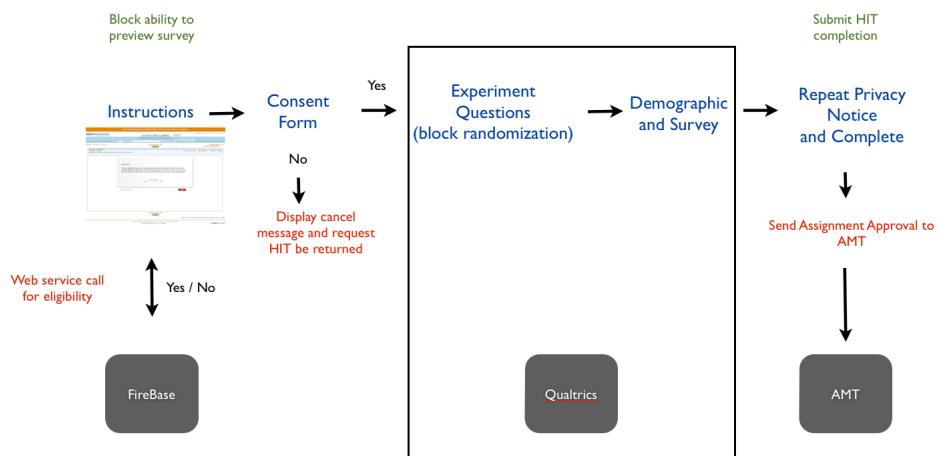


Figure 16: Experiment Flow

Table 2: Data Captured

Stored	Type of Data
Amazon Server	WorkerId, AssignmentId, HitId, Accept / Reject, Fees paid
Qualtrics Server	Random number identifier, Experiment data
FireBase Server	Hash of workerIds for each experiment

4.5 DATA STORAGE AND PRIVACY

Data was stored in three places (Table 2): AMT, Qualtrics, and FireBase. AMT stored a list of workerIds for each experiment, Qualtrics stored survey data (but no workerIds), and FireBase was used to store a list of worker hashes (but no data or workerIds).

There was no connection between AMT and Qualtrics except for URL-encoded data passed to and from the Qualtrics server:

- On acceptance of a HIT, Amazon passed Qualtrics an AssignmentId, HitId, and WorkerId.
- On completion of the HIT, Qualtrics sent back to Amazon the AssignmentId.

Survey data was not stored on Amazon’s servers nor was participant identity accessible to Qualtrics via the AMT API. Because AMT does not offer “verified” user profiles, users can lie about their demographic group to qualify for a study. However, given the relative anonymity of the system, most appear to be honest (Ipeirotis, 2010a).

The Firebase service listed above was required only to disqualify workers from participating in certain experiment combinations. To safeguard identity, my code on Qualtrics converted the AMT WorkerId to a 32-bit hash and passed this hash to the Firebase service. New prospective workerIds were hashed and compared with the hash list on the Firebase service. Hashing is uni-directional, thus it is not possible to recover the original WorkerId from a hash list alone.

On Qualtrics, the WorkerId was discarded and replaced with a random number identifier. WorkerId was treated as

Personally Identifiable Information (PII) and de-linked from the data such that it would not be possible to re-link workerIDs to experiment results. Furthermore, though Qualtrics retains IP address data by default, I turned this, and other browser identifiers, off for these experiments. Because no identifying questions were asked, nor was IP address collected, all experiment data collected is fully anonymized.

The next section addresses characteristics of the population sampled.

4.6 POPULATION SAMPLE

The first and third columns of the table in [Appendix D](#) summarize questionnaire responses across the five studies described in Chapters 5 - 7. Of the 1158 subjects who took the demographic survey, 84% were unique, with 16% duplication due to participation in more than one experiment. This was possible, depending on the order in which workers accepted tasks. For example, if a worker accepted a HIT in Experiment 3, that same worker was later still eligible for any of the other experiments. However, once participating in Experiment 1B, he or she would have been no longer eligible to participate further.

Demographics are comparable to those published in other studies ([Ipeirotis, 2010b](#); [W. Mason & Suri, 2011](#); [Ross, Irani, Silberman, Zaldivar, & Tomlinson, 2010](#)), though the sample I collected was skewed more toward males. Of 1158 subjects, 73% were under 35, 53% male, 73% caucasian, 97% English speaking, 86% having at least some college, and 53% making under 30,000 per year. The profile drawn here is from a set of workers with a 90% or better HIT acceptance rate and from the United States.

In September of 2011, [McDonald and Peha \(2011\)](#) conducted a user survey on AMT to study the differences between what users expect “Do Not Track” to mean versus DNT definitions under debate. At the time, DNT was very new. They studied 304 participants limiting participation to the United States. 81% had never heard of DNT.

Though I did not use those exact questions in my survey, in October of 2013, 91% reportedly knew what a tracking cookie is. 73% admitted to using browser plugins for privacy protection and more than half to configuring their browsers to “opt-out”. However, despite that AMT workers are much

more knowledgeable than two years ago, most still do not realize that DNT, as specified by advertisers, is limited to not showing targeted advertisements: information is still collected, stored, and used.

In an older study (not on AMT), Acquisti and Grossklags (2005) found a correlation between concern for privacy and income. Those with higher incomes were generally more concerned with privacy. However, this may have correlated more strongly with knowledge than income. They noted that their sample particularly lacked knowledge about technological or legal forms of privacy protection.

4.7 VALIDITY

Discussed below are three aspects of validity for studies using Mechanical Turk: internal, external, and construct validity. Questions of statistical validity are addressed in the results sections of experiments.

4.7.1 Internal Validity

Internal validity concerns the equivalence between groups and the control of extraneous variables. Other behavioral experiments utilizing AMT (Crump, McDonnell, & Gureckis, 2013) have suggested that the pool of available workers is sufficiently representative and large such that random assignments consistently produce results equivalent to those produced in carefully controlled laboratory settings. Furthermore, it is possible to conduct experiments where workers may not aware that they are in an experiment (one source of experimental bias). Also, Paolacci et al. (Paolacci et al., 2010) found that retention was particularly high on AMT, compared to other pools compared (university student population and Internet boards). However, they noted that, unlike student populations, turker membership is organic and workers may be potentially available for years. This particularly highlights the need to track responses across experiments.

More insidious to internal validity may be the effects of payment itself. Two effects of *volunteer bias* are of particular concern are: 1) effect of informed consent; and 2) effects of obligation incurred by accepting a HIT and, as a result, payment for services. Rush, Phillips, and Panek (1978) found difference between unpaid volunteers and paid subjects in

a selective attention task. Unpaid volunteers were found to commit fewer omission errors than paid subjects.

Another potential threat to internal validity is the control of the subject environment. Mentioned previously, was the problem of ensuring all conditions are presented consistently across browser types and monitor sizes. In addition, there is no way to control what sorts of environmental factors may be present: workers may be participating in multiple tasks, watching television, or be distracted in a myriad of ways.

To address these concerns, I took the following precautions:

1. using custom code, I limited participation across conditions where prior exposure to privacy questions was potentially biasing; and,
2. to the extent possible, I sandbox tested — not just code consistency, but visual consistency across platforms and browsers.

4.7.2 External Validity

One important aspect of external validity is whether AMT workers are representative of the desired population as a whole. The experiments described in this dissertation rely on basic linguistic competency. For this purpose, native language competency is of concern. Mechanical Turk appears representative of the U.S. population as a subject pool from the perspective of gender, race, and education. Previous studies ([Behrend et al., 2011](#); [Buhrmester, Kwang, & Gosling, 2010](#)) find that turkers are slightly more representative of U.S. population statistics than are standard Internet samples. They are also significantly more diverse than university samples.

The population sample participating in studies described here may, however, differ from the general Internet population in other ways. Turkers polled seemed surprisingly knowledgeable of privacy issues. This indicates that experimental results, in some cases, may not be representative of a more general population.

4.7.3 Construct Validity

To large extent, construct validity is a property of each individual experiment. However, the experiments described in

this dissertation fall into the genre of both linguistic and judgment /decision-making tests. To this end, there exist other studies which are generally comparable.

[Paolacci et al. \(2010\)](#) conducted replication of traditional judgment and decision-making tasks on AMT and compared them with groups from traditional subject pools at a large University and also Internet discussion boards. Tests included the [Kahneman and Tversky \(1984\)](#) Asian disease problem, [Kahneman and Tversky \(1983\)](#) Linda problem, and [Baron and Hershey \(1988\)](#) Physician's problem. Their results confirm that AMT is a reliable source of experimental data in judgment and decision-making. Not only were group results similar across conditions, but the effect size for AMT was the highest. More generally, AMT has been validated as a tool for behavioral cognitive studies such as those in reaction time research (e.g., Stroop, Task-switching, Flanker tasks, etc.; [\(Crump et al., 2013\)](#) and recall of written information ([Tietze, Winterboer, & Moore, 2009](#)).

While researchers in natural language processing were among the first to utilize AMT for the collection and annotation of linguistic resources ([Callison-Burch & Dredze, 2010](#)), its adoption by theoretical linguists has been slow to develop. Traditionally, user judgments supporting theoretical claims have been weakly quantitative relying on a very small number of examples and researchers – often limited to the authors of theoretical studies themselves. [Gibson and Piantadosi \(2011\)](#) demonstrate the utility of AMT for collecting linguistic behavioral judgments for acceptability of sentence / meaning pairs. [Sprouse \(2010\)](#) compares just such a task with two groups of users (AMT versus laboratory) each with 176 participants. Data collected from the two groups was deemed virtually indistinguishable.

Finally, researchers in pragmatics (e.g., in studies of implicature) have begun not only to adopt experimental methods, but have applied these to collections on AMT (for example, [Bergen, Goodman, & Levy, 2012](#); [Degen & Tanenhaus, 2011](#); [Stiller, Goodman, & Frank, 2011](#)). As noted by [Anand and Andrews \(2011\)](#), pragmatic inference depends on a multitude of factors including task structure, social norms, and response elicited. Furthermore, because there are potentially so many parameters, it is difficult to systematically model interactions between linguistic form, context, and pragmatic inference. Crowdsourcing platforms make such study more

tractable, though we have much yet to learn about what specific methods are most amenable. Problematic for pragmatic studies, in particular, is that the subject's knowledge of the experiment itself plays directly into context (Rosnow & Rosenthal, 1976).

4.8 SUMMARY

This chapter outlined a research program with the substantive hypothesis that some user confusion stemming from online behavioral advertising is caused by error in discourse understanding. I described three situations where miscommunication may occur. The first considers the possibility of pragmatic implicature in a "do not track" modal dialog, the second considers the indexicality of an icon embedded in an image-based advertisement, and the third the effect of visual presence on non-ratified participants on a web page.

In this chapter, I also discussed related research utilizing Amazon Mechanical Turk, outlined general procedures and considerations, and described general characteristics of the population sampled.

In each of the next three chapters, after framing experimental hypotheses within the theoretical framework, I detail specifics in terms of instrumentation, collection, analysis and results.

STUDY ONE: MODAL DIALOG BOXES

This study concerns whether people make pragmatic inferences when interacting with user interface components. Elements of language and iconic graphics integrate to communicate what seems a simple message. So simple, users may not even be aware of potential errors in communication.

5.1 MOTIVATION: PRIVACY AS USER CHOICE

The W3C Tracking Protection Working Group (TPWG) has been working toward a “Do Not Track” (DNT) policy intended to allow users to signal their intent with regard to browser-based tracking. DNT is not designed as a general purpose tool for communicating privacy practices. It is intended to simply communicate a user’s preference not to be tracked.

The current Tracking Preference Expression draft specifies three possible states: DNT:1 (do not track), DNT:0 (allow tracking) and un-set. In this third case, tracking preference is not enabled. The TPWG draft posits a number of reasons for why a user agent may not have tracking preference enabled:

1. The browser user agent does not implement DNT;
2. The user has not yet expressed a specific preference; or,
3. The user has not chosen to transmit a preference.

User preference mechanisms specified by the TPWG represent an earnest attempt to place some burden of policy implementation on browser developers rather than publishers: instead of forcing the user to make a choice for every website, the idea is that a user specifies choice in browser preferences and sets exceptions as desired. Nonetheless, some websites offer a tracking preference choice in order to comply with other international regulations, such as the EU cookie law described in Chapter 2.

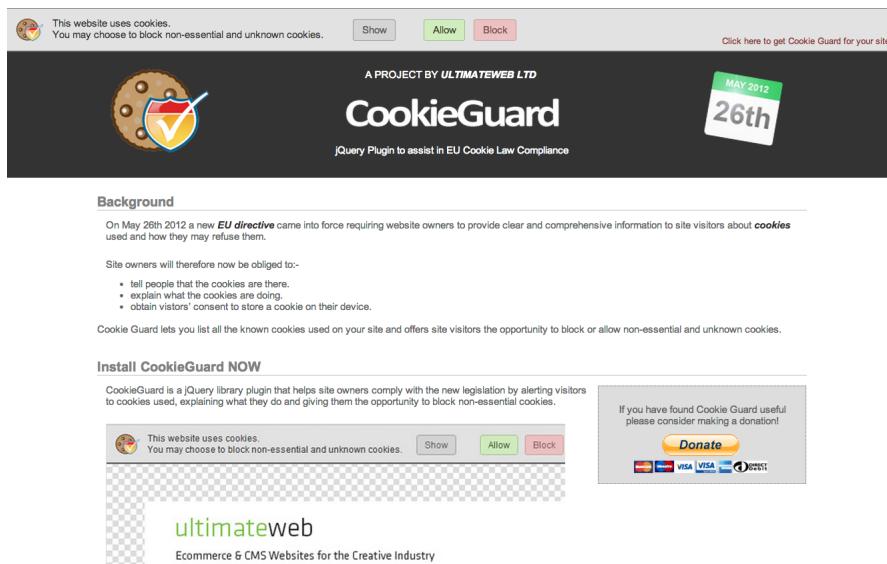


Figure 17: Cookie dialog control from cookiedguard.eu

5.2 CHOICE DESIGN PROBLEM

Though the TPWG intent is to provide a machine readable preference expression mechanism and not a user interface (UI) specification, the three options above map to common UI pattern: **modeless dialog control**.

Given this, if a user is presented with a dialog control presenting a choice between opt-in, opt-out and dismiss, what does the user believe is the consequence of choosing to dismiss?

One way to consider this problem is as a choice design problem ([Figure 19](#)):

1. If I click "allow", I choose "allow" cookies
2. If I click "block", I choose "block" cookies
3. I can do neither ("dismiss")

Fair and unbiased choice design is a tricky problem. Default choices have a dramatic impact on user action ([Johnson, Bellman, & Lohse, 2002](#)). *Heuristics and biases* theories of reasoning account for a number of different situations leading to systematic bias. Biases including loss aversion (e.g., change from status quo), framing, and evaluation of options in relation to reference points (e.g., expectation and social comparison) have been well-described by [A. Tversky and Kahneman \(1981\)](#) and [Kahneman and Tversky \(1984\)](#).

Whether intentional or not, designers – both standards architects and web designers – have tremendous potential to influence choice (Thaler et al., 2010).

5.3 FORCED CHOICE

Given that users may not understand the consequences of not making a DNT choice, one might ask why the TPWG might not specify “forced choice.” That is, why not force the user to make a choice?

As described by Dhar and Simonson (2003), forced choice under preference uncertainty can produce psychological discomfort (referencing work by Festinger, 1964; Janis & Mann, 1977; Lewin, 1951). They note that the no-choice alternative is an attractive way of resolving difficult choices when subjects are forced to choose.

The study described here concerns whether users confronted with a non-forced choice dialog box understand the meaning of their choices in the context of interaction. The next section will cover related research, focusing largely on meaning and understanding.

5.4 RELATED RESEARCH

Prior privacy research in the context of online behavioral advertising has focused on the effectiveness of communicating privacy risks to consumers (McDonald & Cranor, 2009), and confusability in user interface design (Leon, Ur, et al., 2012). Moreover, research on “opt-in” /“opt-out” in the context of privacy-related decision-making has also been studied (Bellman, Johnson, & Lohse, 2001; Lai & Hui, 2006). This body of work is summarized below.

5.4.1 Heuristics and Biases

Bellman et al. (2001) systematically studied the influence of question framing and defaults on consumer privacy preferences. They remarked that the decision to share private information on a site is typically on a case-by-case basis. This is particularly problematic in light of research by (Kahneman & Tversky, 1984), who found that user behavior was susceptible to several important factors, namely the framing of

Figure 1. Checkbox format questions for participation in health surveys.			
No-Action Question	Default	Participation	
<input type="checkbox"/> I. Notify me about more health surveys.		Not Participate	48.2%
<input checked="" type="checkbox"/> 2. Do NOT notify me about more health surveys.		Participate	96.3%

Figure 18: Effect of Defaults on Participation (Image credit: Bellman et al. (2001))

options. In other words, people tend to evaluate options in relation to how a question is asked. What Bellman et al. discovered in the context of privacy preference was that framing the question as opt-out instead of opt-in affected choice.

In Figure 18, from Bellman et al. (2001), 134 participants were given either the first question or the second — which are framed positively and negatively, respectively. By making opt-in the default, 96.3% agreed to participate, while when not the default, only about half agreed to participate.

Equally interesting is that positive (e.g., "notify me about...") versus negative (e.g., "do not notify me about...") framing has an effect on participation rates. In the same task as above, but in a forced choice question format, respondents agreed to participate at a higher rate in a positive frame compared to a negative (Bellman et al., 2001).

As they further note,

Defaults and framing are likely to have even more impact when, as is often the case, the question is set in a miniature font, or answering most questions is optional, or the implications of answering are buried in a large privacy policy document. (Bellman et al., 2001, p. 25)

Lai and Hui (2006) extended this work by investigating the role of “privacy concern” on option frames. They found that the mechanism described by Bellman et al. (2001) had a greater effect in people who were less concerned with privacy than those who were more concerned with privacy. Thus, those with greater privacy sensitivity appeared to be less susceptible to option frames.

To date, “opt-in” versus “opt-out” in the context of dialog interaction has not been studied. However, there has been some relevant research on interaction with modal dialog, in general. In a study comparing techniques for user notification of information updates, Bailey, Konstan, and Carlis (2000) find

dialog windows to be the most most intrusive and distracting of methods compared. Bailey and Konstan (2006) note that: “when peripheral tasks interrupt the execution of primary tasks, users require from 3% to 27% more time to complete the tasks, commit twice the number of errors across tasks, experience from 31% to 106% more annoyance, and experience twice the increase in anxiety than when those same peripheral tasks are presented at the boundary between primary tasks.” This work is particularly relevant from the perspective that people have a choice to even attend a dialog upon visiting a website.

5.4.2 Pragmatic Reasoning

Pragmatics is concerned with reasoning processes that go beyond conventional meaning. It is founded on the notion that communication is essentially cooperative (H. H. Clark, 1996; Grice, 1975a; Levinson, 1983). A common pragmatic phenomenon in linguistic understanding is implicature. An implicature represents a gap between what is expressed and what is communicated. Importantly, whether an implicature is true or not, does not affect the meaning of the message itself. For example,

1. Harry and Sally are married.
2. Tell a friend or colleague.
3. If you mow the lawn, I'll give you five dollars.

In (1) the implication is that Harry and Sally are married to each other. But, if they were not, and both married to someone else, this statement would still true. (2) exemplifies implicature derived by considering “or” as inclusive or exclusive. (3) typifies what is known as *invited inference* (Geis & Zwicky, 1971) or *conditional promise* (Searle, 1971). In (3), hearers understand the conditional relation between getting five dollars and mowing. But they may also infer “not to mow” means they will “not get five dollars.”

Fillenbaum (1975) showed that the obverse of a conditional promise (in the example above, “I won’t give you five dollars, if you don’t mow the lawn”) was an accepted inference for 85% of subjects tested. Invited inference in a conditional statement affects people’s understanding of a situation. Such inferences contain what is known as “deontic force”.

5.4.3 Deontic Reasoning and Knowledge

Interestingly, arbitrary assertions such as:

If a fish is red, it has wings. (Markovits & C, 1990)

elicit far fewer bi-conditional interpretations (i.e., A fish has wings, if and only if it is red) than conditional promises.

What's different? Conditional promises such as the lawn mowing example above incur some sense of social obligation. Lexical items such as "must" and "may" do, as well.

You must finish your homework before you can go out to play.

Obligation is somehow implicit in the word "must" and its meaning is licensed by context — if this is a mother speaking to a child, than that mother is placing constraint on her child.

Likewise, a symbol may reflect social obligation. A stop sign is as explicit social rule and one we obey because of societal consequences. But there may also be contexts where such signs are ignored, for example, if the street is on an abandoned compound.

The meaning of assertions containing concepts of permission, obligation, and prohibition, and release (from obligation) involve what is called *deontic reasoning* (Beller, 2008). This sort of reasoning is concerned with the relation between pragmatic knowledge (i.e., knowledge of the content of arguments) and social obligation.

5.4.4 Content Affects Reasoning

While the *cooperative principle* (Grice, 1975) affects comprehension, knowledge of subject matter does, as well. If both the antecedent and consequent are closely correlated in the real world as in:

If I wear a jacket, then I put on a tie.

Politzer (1981) found more bi-conditional interpretations for this sort of sentence than if they were not closely correlated. Marcus and Rips (1979) found the same for statements where the antecedent and consequent were causally related. *In conditional reasoning tasks, subjects use knowledge of content to reason about the problem.*

In the experiment described in this chapter, two confounding factors are considered: *deontic force* and *privacy bias*. The first is invoked through the use of word traditionally seen in cookie dialog boxes — “allow” /“block”. As control, I substitute the words “on” /“off” which carry not implied deontic force. The second confound is knowledge and attitude towards privacy and tracking cookies. I control this factor by substituting the words “pictures” for cookies.

5.4.5 Graphical Implicature

Visual context is known to influence language processing very early in the processing of language (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In eye-tracking tasks, Tanenhaus et al. (1995) observed comprehension tasks on the millisecond time scale during language comprehension tasks. Initial experiments demonstrated that subjects made saccadic eye movements to objects immediately after hearing particular relevant words in instruction. Such movements are closely time-locked to linguistic objects. Eberhard, Spivey-Knowlton, Sedivy, and Tanenhaus (1995) found that listeners immediately integrate lexical, sub-lexical, and prosodic information in speech with information from visual context to reduce a set of referents to the intended one.

There is relatively little empirical research directed at graphical implicatures and mixed-modal representations, in particular. Stenning et al. (2006) and Stenning and Oberlander (1995) posit that diagrammatic representations may be studied using the same kinds of semantic techniques used in linguistics. Oberlander (1995) discusses the notion of *graphical implicature* in diagrams while Marks and Reiter (1990) discuss methods for avoiding unwanted implicatures in generating text and graphics.

While this experiment aims to show whether pragmatic processes are in play, it makes no particular claim with regard to the processing of implicature; only that once a dialog box has been comprehended, this understanding has bearing on decision-making and belief. Furthermore, an implicature generated during understanding may not be remembered as such — since people generally can't distinguish between assertions and implications in memory (Brewer, 1977) Indeed, it is likely that people are not even aware of any confusion or mis-understanding during or after comprehension.

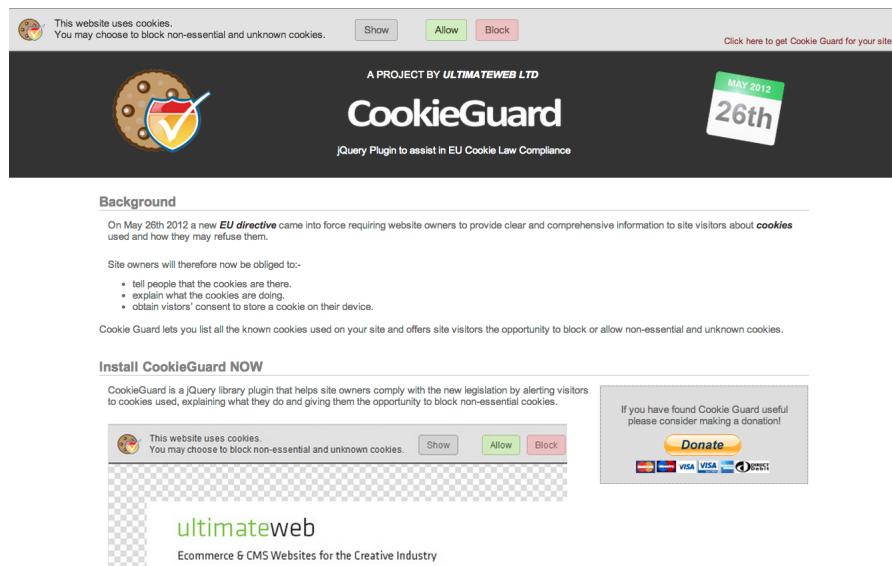


Figure 19: CookieGuard Website

5.5 PRE-PILOT

The purpose of this pre-pilot was to assess the feasibility of conducting a larger, online pilot study. Specific questions addressed were: 1) whether the dialog was placed in such a way that subjects would likely read it; and, 2) whether subjects would consider the dialog as independent of the experiment. I needed to be confident that users would not be influenced by the environment to select a particular choice. To this end, the study was posed as a “survey”.

5.5.1 Method and Procedure

In December 2012, 17 subjects were recruited at the University of Baltimore from university business offices. Employees and students were invited to participate in a “10 minute survey” in the Information Arts and Technologies usability lab in exchange for a \$5 dollar gift card. All were native English speakers who were comfortable using the Internet.

Using a Tobii T60 eye-tracker, subjects were presented with a choice banner (“This website uses cookies”) preceding a short survey on Internet privacy. The choice banner was modified from the CookGuard plugin¹ Figure 19 designed to help website owners comply with European Union directives.

¹ <http://cookieguard.eu>

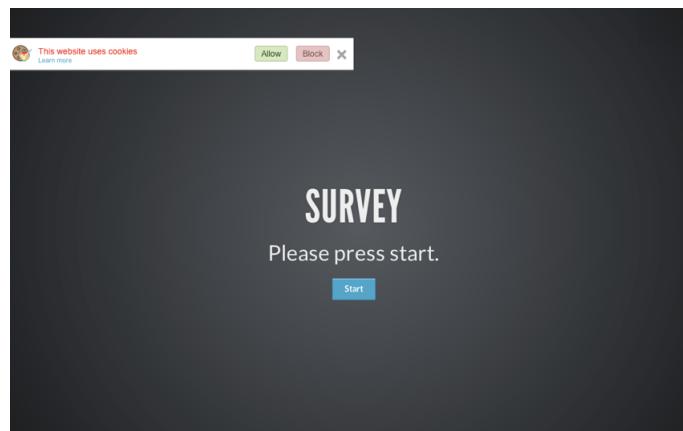


Figure 20: User Display

A modification was made to include an “x” so that a third option – “if I click the x, I dismiss the control” – was explicit.

This banner was placed prominently on the start page of the Internet survey [Figure 20](#). Generally, such banners are placed at the very top of a website, but I was concerned that subjects might not notice it there, so it was positioned it in such a way to make it more visually distinct.

5.5.2 *Results and Discussion*

Of 17 subjects, 5 did not know what browser cookies were, 14 reported that privacy was very important to them, and 15 reported that they would turn off tracking if it were easy. Notably, only 1 subject selected any option other than “dismiss” on the cookie banner.² He selected “block” cookies because “he didn’t like cookies.” No one clicked the provided link “learn more”.

For the first goal of assessing likelihood that a subject would read the banner, I learned that, despite a sparsity of information on the start page, the first 9 subjects did not see the banner. For the remaining, subjects were verbally cued that there would be a cookie banner on the start page and that they could “choose however they wanted.” Indeed, all but one did then see and read the banner.

The second goal was more successful. When presented the banner, subjects reported that they did not suspect that the cookie banner had anything to do with the following survey.

² Interestingly, many subjects who dismissed the cookie banner without making a choice later indicated that privacy was very important to them.

After presentation of the cookie banner and a number of demographic questions, subjects were asked whether they believed ad trackers were present on the site or not. 8 subjects believed that ad trackers were present while 9 did not. Of those that believed ad trackers were present, 1 subject believed that this was the case since he did not “block” cookies. Of the 9 that did not believe ad trackers were present, 6 believed this because they did not “allow” cookies. Accordingly, 7 of 17 (41%) believed there either were or were not ad trackers on the site based on *not* clicking “allow” or “block” cookies, respectively. These results suggest that a pragmatic implicature was in play — information was suggested via the cookie banner though not explicitly stated.

5.6 AIMS

In this pilot, I consider whether a non-forced choice (modal) dialog box has the potential to generate an implicature in user understanding. One way to view the choice problem above is as a discourse reasoning task where more than one conditional is given for interpretation in a single turn. In the cookie dialog choice decision described in this paper, not only must subjects interpret the meaning of each conditional independently, but they must do so in the context of choice between an additional explicit conditional and graphical third choice (“dismiss”). The particular question addressed here is what a user believes his choice to mean when he neither selects “allow” nor “block”. What do users believe is the meaning of the graphical third choice?

5.7 EXPERIMENTAL DESIGN

This study has two experimental hypotheses:

HYPOTHESIS 1A

A GUI dialog box, which communicate through a combination of linguistic elements and graphical elements, is subject to the communication of pragmatic implicature.

HYPOTHESIS 1B

The interpretation of a pragmatic implicature in a GUI dialog influences user belief about the presence or absence of ad tracking cookies relative to the site they are viewing.

Experiment 1A is a $2 \times 2 \times 2$ random control group posttest-only design. A control group is presented with a set of linguistic expressions and asked to answer questions about meaning. Treatment groups are presented with either linguistic expressions or a dialog box expressing the same set of choices and asked the same questions.

The **independent variable** (IV) is “presentation form” (textual or mixed-modal) and the **dependent variable** (DV) is inference of a pragmatic implicature.

This is a factorial design considering not only modality but:

1. knowledge and attitude toward browser cookies (privacy bias); and,
2. deontic force of the words "allow" and "block".

Experiment 1B is a single factor random control group posttest-only design. Both control and treatment groups are presented with a cookie banner and later asked about whether or not they believe the website placed “cookies” in their browser. The treatment group is presented with a textual display that communicates the outcome of their action.

The **independent variable** is “feedback” and the **dependent variable** is inference of a pragmatic implicature.

This design varies from the previous experiment since participants are asked not directly asked about the meaning of the dialog box, but about their beliefs about the presence or absence of cookies in their browser based on their previous behavior during an actual task.

5.8 METHOD

These two experimental designs follow the general procedure outlined in Chapter Four.

5.8.1 *Experiment 1A Method*

Settings and Participants

Using Amazon’s Mechanical Turk, 213 workers were paid \$0.15 to participate in one of eight conditions of the $2 \times 2 \times 2$ design previously described. Participants were assigned randomly via Qualtrics block randomizer. Results were collected

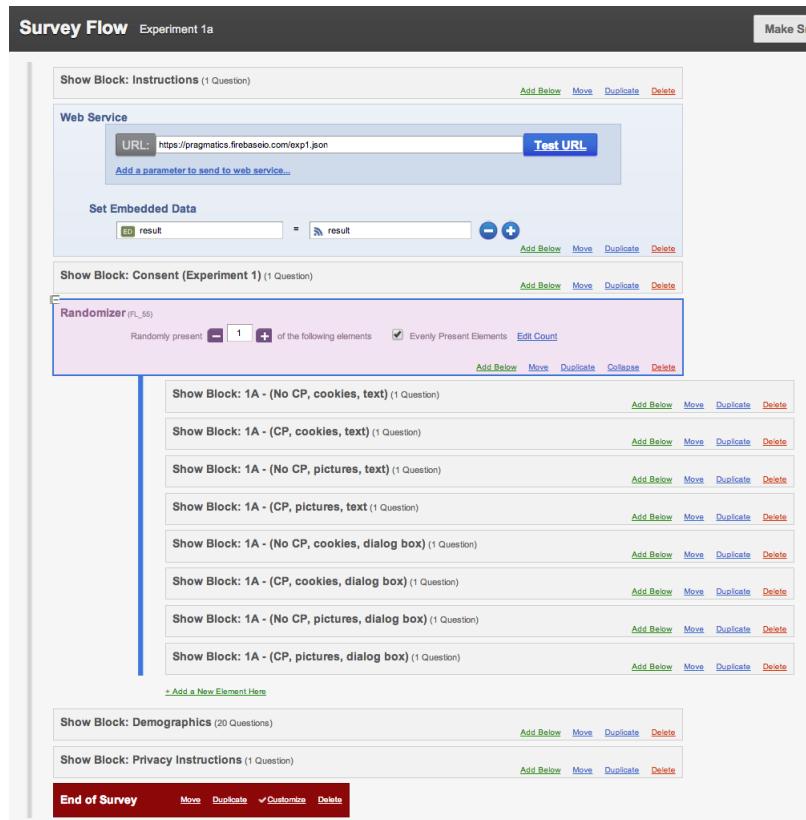


Figure 21: Experiment 1A Flow

during the period of 7–14 October 2014. The task was expected to take approximately three minutes, though workers were allowed up to 10 minutes to complete.

Procedure and Materials

The general procedure follows that described in Chapter Four. [Figure 21](#) is a more detailed graphical depiction of flow.

Following presentation of instructions and consent form, participants were each presented one question, as illustrated by [Figure 22](#) and [Figure 23](#), followed by survey questions. Each of eight variants are presented in [Appendix E](#) of this document. The cookie banner is the same as the one described in discussion of the pre-pilot earlier in this Chapter, though text was altered using PhotoShop to create variant conditions.

Answering either:

- The website will allow cookies because I didn't select block
- The website won't allow cookies because I didn't select allow

You go to a website which presents the following choices:

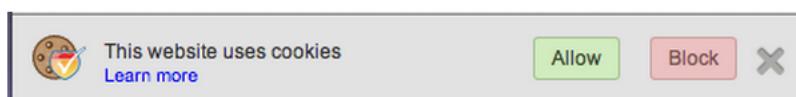
- Allow cookies
- Block cookies
- Cancel

You pick '**cancel**'. Assuming the website respects your wishes, what does this mean?

- The website will allow cookies because I didn't select block.
- The website won't allow cookies because I didn't select allow.
- The website will allow cookies or block them.
- The website will neither allow cookies nor block them.

Figure 22: Textual, Deontic Force, Cookie Condition

You go to a website which presents the following:



You pick '**x**'. Assuming the website respects your wishes, what does this mean?

- The website will allow cookies because I didn't select block.
- The website won't allow cookies because I didn't select allow.
- The website will allow cookies or not.
- The website will neither allow cookies or not.

Figure 23: Mixed-Modal, Deontic Force, Cookie Condition

was considered evidence of pragmatic implicature.

Instrumentation

In addition to features provided by AMT and Qualtrics, my scripts performed the following:

1. Disable preview using a custom CSS class blocking Qualtrics controls
2. Check worker hash against a FireBase web service and request worker to return HIT if the hash is in the exclusion list
3. Add worker hash to FireBase
4. Submit results to AMT on completion of the Qualtrics survey

No other instrumentation was required for this experimental design.

5.8.2 Experiment 1B Method

Settings and Participants

Using Amazon's Mechanical Turk, 200 workers were paid \$0.15 to participate in one of two conditions of the single factor design previously described. Participants were assigned randomly via Qualtrics block randomizer. Results were collected between the days of 7–14 October 2013. The task was expected to take approximately three minutes, though workers were allowed up to 10 minutes to complete.

Procedure and Materials

The general procedure follows that described in Chapter Four. [Figure 53](#) is a more detailed graphical depiction of flow.

Following presentation of instructions and consent form, participants were each presented a start screen with a cookie banner (appearance animated to draw attention), as illustrated in [Figure 25](#).

Immediately following this, workers were presented with the following question:

Do you think there are ad trackers on this site?

Survey Flow Experiment #1b

The screenshot displays the SurveyFlow interface with the following structure:

- Show Block: Instructions (1 Question)**
- Show Block: Consent (Experiment 1) (2 Questions)**
- Web Service** (highlighted):
 - URL: <https://pragmatics.firebaseio.com/exp1.json>
 - Test URL
 - Add a parameter to send to web service...
- Set Embedded Data** (highlighted):
 - cookies = 1
 - Add a New Field
- Randomizer (FL_34)** (highlighted):
 - Randomly present 1 of the following elements
 - Evenly Present Elements (checkbox checked)
 - Edit Count
- Set Embedded Data** (highlighted):
 - persistence = 0
 - Add a New Field
- Set Embedded Data** (highlighted):
 - persistence = 1
 - Add a New Field
- + Add a New Element Here**
- Show Block: 1B banner (1 Question)**
- Show Block: Question (3 Questions)**
- Show Block: Demographics (20 Questions)**
- Show Block: Privacy Instructions (1 Question)**
- End of Survey**

Figure 24: Experiment 1B Flow

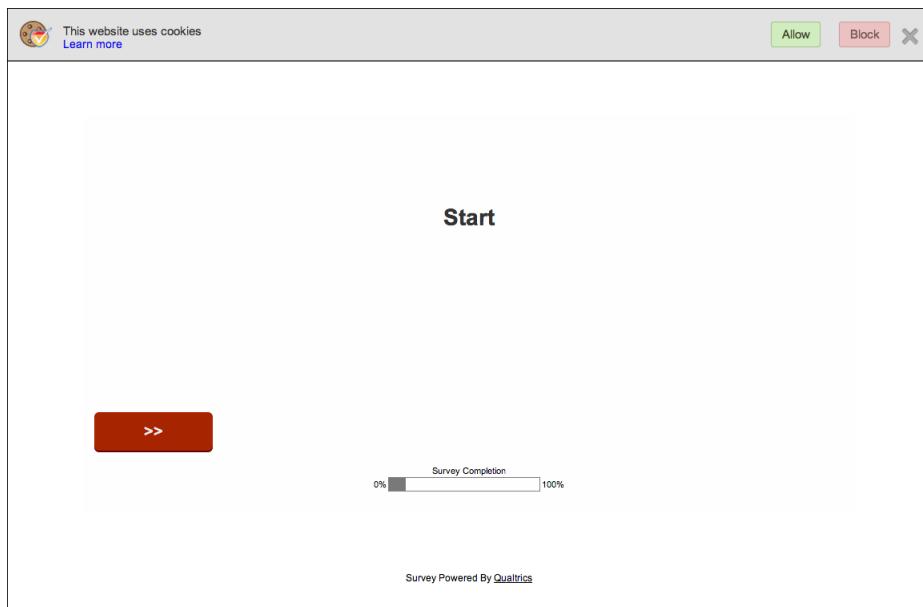


Figure 25: Experiment 1B Stimulus

In the treatment condition ([Figure 26](#)), participants were given feedback from their action on the previous screen.

They were then asked if they believed there were ad trackers on the site ([Figure 27](#)). Finally, participants were asked about their beliefs on the basis of this response ([Figure 28](#)).

Responses were coded as implicature for the control group if the subject did nothing and responded with “I chose allow” /“I didn’t choose allow” (Yes) or “I didn’t choose block” /“I chose block” (No). In the treatment group “I chose allow” /“I didn’t choose block” (Yes) were not coded as implicature since feedback had been provided giving the subject knowledge of the consequences of their action.

Instrumentation

In addition to features provided by AMT and Qualtrics, my scripts performed the following:

1. Disable preview using a custom CSS class blocking Qualtrics controls
2. Check worker hash against a FireBase web service and request worker to return HIT if the hash is in the exclusion list
3. Add worker hash to FireBase

you allowed
cookies

Do you think there are ad trackers on this site?

Yes
 No

>>

Survey Completion
0%  100%

Survey Powered By [Qualtrics](#)

Figure 26: Experiment 1B Question (with feedback)

Why not?

- I didn't choose "allow"
- I chose "block"
- I am using an ad blocker

>>

Figure 27: Experiment 1B Cookie Belief

Why?

- I chose "allow" cookie
- I didn't choose "block"
- There are probably always ad trackers

>>

Figure 28: Experiment 1B Belief Justification

4. Submit results to AMT on completion of the Qualtrics survey
5. Display cookie banner on start page
6. Record button presses on the cookie banner (accept, block, learn more, close)
7. In the treatment condition, display feedback on the basis of cookie banner interaction

5.9 DATA COLLECTED

Using G* Power 3 chi square goodness of fit test ([Berger, 2012](#)), for both experiments I estimated a sample size of approximately 220 was necessary in order to detect a medium effect with power of 0.95. All surveys initiated were completed: there were no known drop-outs.

Once all assignments had been completed, (as indicated on my AMT requester dashboard), I downloaded data as a single CSV (comma separated values) file from the Qualtrics website. Data was organized such that each participant's data was on its own row.

5.10 RESULTS

In this section, I present experimental results for both experimental designs.

5.10.1 *Experiment 1A Results*

The two tables, [Table 3](#) and [Table 4](#), present raw percentages of implicature calculated for the three-factor design with 213 participants. Determination of implicature is dichotomous — either the participant interprets an implicature or not. Thus, the top two answers for each question were rolled into a single dependent variable (implicature) for which a percentage value is displayed. Within [Table 3](#) and [Table 4](#) are two contexts for interpretation: the participants believes that a cookie will be deposited in the browser or not.

For analysis, a **basic difference statistics** is required. Because the IV is categorical (binary), a non-parametric statistics is required. A Pearson's chi-square test for independent samples is appropriate. However, its not possible to apply

Table 3: Implicature in Text Conditions

Text		
Cookies	DF	48%
	No DF	50%
Pictures	DF	67%
	No DF	50%

Table 4: Implicature in Mixed-Modal Conditions

Mixed-Modal		
Cookies	DF	63%
	No DF	65%
Pictures	DF	72%
	No DF	67%

a chi-square test to tables with three or more discrete variables. In such a case, a **log linear model** is more useful. It has both the characteristics of a chi-square test determining the fit between observed and expected frequencies, as well as features of an ANOVA such that it is possible to do simultaneous testing of main effects and interactions within a fully factorial design.

For hypothesis 1A above, of primary interest are group differences associated with the main effect mixed-modal versus text. But we can also examine the 2-way interaction of privacy and 2-way interaction with deontic force in both textual and mixed-modal conditions.

In Table 5, independent and dependent variables are treated equally. It includes main effect plus 2-way and 3-way interactions. Where residuals are close to 0, the model has a better fit. I discarded the 4-way interaction (Graphics:DF:Privacy:Implicature) since it did not improve model fit.

From the table above, *there is a significant difference between text and mixed-modal conditions for implicature where $p < .01$: implicature is higher in conditions where users are presented mixed-modal information*. This validates my hypothesis that implicature may be communicated in mixed-modal conditions, just as it would in a purely textual context. What is, perhaps surprising, and interesting is that subjects are even more likely to interpret an implicature.

Table 5: Log Linear Significance

	Df	Deviance	Resid.	Df Resid.	Dev	Pr(>Chi)
NULL				15	28.6821	
Graphics	1	0.0423	14	28.6398	0.8371346	
DF	1	0.0423	13	28.5976	0.8371346	
Privacy	1	0.0047	12	28.5929	0.9453725	***
Implicature	1	13.3274	11	15.2655	0.0002616	
Graphics:DF	1	0.1155	10	15.1501	0.7340215	
Graphics:Privacy	1	0.0419	9	15.1082	0.8378711	
DF:Privacy	1	0.2248	8	14.8834	0.6353792	
Graphics:Implicature	1	7.3723	7	7.5110	0.0066235	**
DF:Implicature	1	0.0361	6	7.4750	0.8493662	
Privacy:Implicature	1	0.0960	5	7.3790	0.7566902	
Graphics:DF:Privacy	1	0.1101	4	7.2689	0.7400369	
Graphics:DF:Implicature	1	1.8719	3	5.3970	0.1712625	
Graphics:Privacy:Implicature	1	1.0591	2	4.3379	0.3034217	
DF:Privacy:Implicature	1	4.2340	1	0.1039	0.0396209	*
Significance codes: o *** 0.001** 0.01* 0.05 0.1 1						

Table 6: 1B Results

	Other	Implicature
Control	79	22
Feedback	95	4

There is also a significant difference for implicature where deontic force occurs with pictures ($p < .05$). There is no firm basis for understanding why this may be the case. I speculate that this may be because subjects might ordinarily expect pictures to be displayed and thus might imagine that they were being asked if they wished to block them. For example, some email clients do not load pictures automatically, but explicitly ask users whether they would like to trust loading pictures for a particular sender.

Finally, there are some limitations to this analysis. For linguistic studies such as this, mixed model effects are often desired. A potential issue with my analysis is that it commits what H. H. Clark (1973) terms the “language-as-fixed-effect” fallacy. This means that the results do not generalize beyond the sample studied here. It is easy to understand how important this is in light that simply by changing the words “accept” / “block” to “on” / “off” we saw a tangible effect. This may well relate to the nature of deontic force implied by “accept” / “block”, but it could also be to other reasons. For example, it is possible that lexical and semantic features of other pairs such as “yes” / “no” may also affect findings.

Increasingly, language researchers are moving to mixed effects models which simply model fixed and random effects in a linear form. Random effects may be added for variables specific to the data sample such as effects relating to subjects, words, and other items. Such an analysis is not substantially more difficult than the one presented here. However, the log linear approach is adequate for the purpose of this pilot study.

5.10.2 Experiment 1B Results

Table 6 below presents raw frequency counts for each of the two conditions (see also, Figure 29).

As before, because the IV is categorical (binary), a non-parametric statistics is required. Using a Pearson’s chi-square

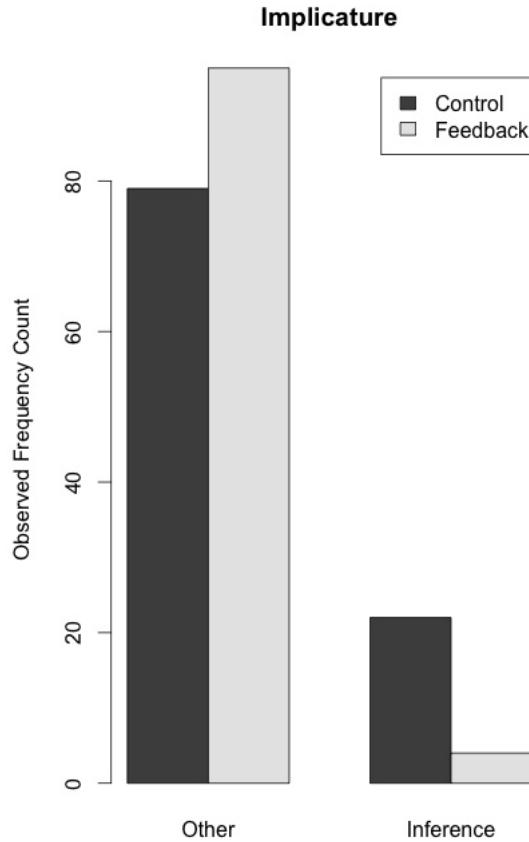


Figure 29: Implicature Per Condition

test with a Yates' continuity correction, subjects in the feedback condition showed significantly less implicature than in the control condition, $\chi^2 (1, N=200) = 12.39$, $p < .001$. The odds of implicature are 6.6 times more likely when there is no feedback given.

Of interest is that 30 participants in the control group and 33 in the treatment groups chose to “accept” cookies (and only 2 chose to “block”; [Table 7](#)).

Click results were different than observed in the pre-pilot. Possibly, some workers believed that that must accept cookies

Table 7: Click Actions

	Accept	Reject	Learn More	Dismiss
Control	30	0	1	1
Feedback	33	2		

in order to participate. Or workers may have believed that the survey wouldn't work properly without accepting cookies. This indicates a potential bias for cooperation. This is also evidenced by one of the subject's comments:

I didn't figure, at first, that the thing about allowing cookies was part of the survey; I figured the survey was based in the UK and was trying to comply with the new UK laws about cookies. (I always allow cookies, because disallowing them creates more problems than allowing them does.)

Several comments were along the lines of this one:

Wow, this definitely made me question whether I knew anything about cookies and ad tracking or not. The questions seemed super simple, but I realized I was making a lot of vague assumptions when you asked more specific questions.

It is also useful to consider whether an implicature might result in harm. By harm, I mean whether the user infers there are no ad cookies. Non-harmful inference would occur in a situation where a participant believed there were ad cookies and also answered "I chose allow" or "I didn't choose block" yet took no action on the cookie banner — inferring passive consent through inaction. On the other hand, an inference that is potentially harmful falls under the situation where the participant did not believe there were ad tracking cookies and responded, "I didn't choose allow" or "I chose block" yet took no action on the cookie banner — inferring non-consent through inaction.

Under this interpretation of harm, approximately 10% of those in the control group made harmful inferences while only 4% in the treatment group. These 4% made an incorrect inference "I didn't choose allow" despite very visible feedback to the contrary.

6

STUDY TWO: HYPERDEIXIS IN ADVERTISEMENTS

How do we understand interaction with objects on the Internet? This study considers the role of reference in hyper-linked advertisements. Intentionally misleading or not, “opt-out” mechanisms for behavioral advertisements are not intuitive. I believe the cause for confusion is intimately aligned with how we process and understand linguistic signs.

6.1 MOTIVATION

In 2009, under pressure from the FTC, industry advertisers formed an alliance called the Digital Advertising Alliance (DAA). The purpose was to establish self-regulatory principles for online behavioral advertising (AAAA, ANA, BBB, DMA, & IAB, 2009). These principles advocate transparency for:

[...] clearly disclosing and informing consumers about data collation and user practices with online behavioral advertising [...] Compliance with the Principle will result in new links and disclosures on the web page or advertisement where online behavioral advertising occurs. (AAAA et al., 2009, p. 2)

Further:

Links to consumer notices will be clear, prominent, and conveniently located... Such enhanced notice will be provided at the time of such collection and use, through common wording and a link/icon that consumers will come to recognize... One option for providing this... is for an entity to attach a uniform link/icon and wording to each advertisement that it serves. Click this link/icon will provide a disclosure from the entity in the form of an expanded text scroll, a disclosure window, or a separate page. (AAAA et al., 2009, p. 5)



Figure 30: AdChoices Icon

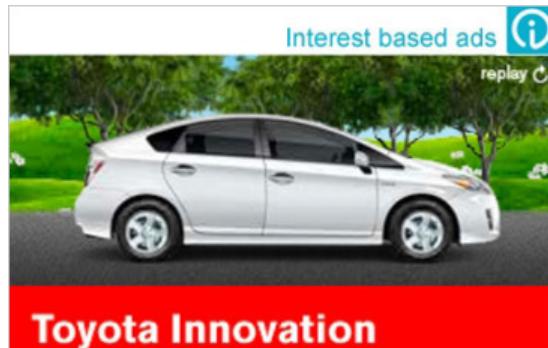


Figure 31: Original AdChoices Design

This section considers whether DAA self-regulation has designed an effective means for "enhanced notice".

6.1.1 *Advertiser Self-Regulation and AdChoices*

DAA self-regulation has come under repeated fire. The implementation of enhanced notice is the AdChoices icon ([Figure 66](#)). [Komanduri, Shay, Norcie, and Ur \(2012\)](#) examined DAA principles and came up with a list of ten requirements appropriate for compliance checks. Included was a check of whether participating advertisers complied with enhanced notice. One problem is that it is nearly impossible to tell whether a particular ad is behavioral or not. However, they report that according to industry estimates of the time, about 80% of advertisements encountered are behavioral. Based on this statistic, [Komanduri et al. \(2012\)](#) found serious compliance problem including infrequent compliance with "enhanced notice".

Separately, [Mayer \(2011\)](#) manually inspected pages from 449 domestic websites from the Alexa list of top 500 U.S. websites for third-party ads. Only 11.3% included an AdChoices icon in or around the ad.

Mayer also showed [Figure 31](#) and [Figure 32](#) on his blog at the Center for Internet and Society at Stanford Law School: [Mayer \(2011\)](#):

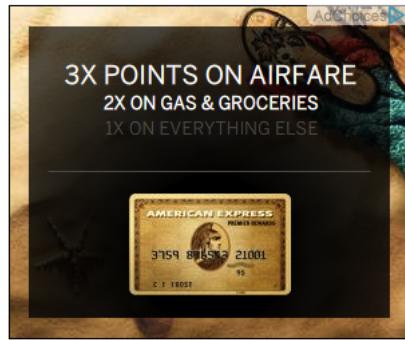


Figure 32: Final AdChoices Design

Note the differences. According to current DAA Creative Guidelines (DAA, 2011), icon only size with container is specified at 19x15 pixels with a rounded left corner radius and 70% opacity. When embedded in an image, it should be in the upper right corner and there should be no space between it and the ad corner.

6.1.2 Effectiveness of the AdChoices Icon

In October of 2012, Evidon served about two billion AdChoices privacy notices daily (Otlacan, 2011). This includes delivery across 40 countries and 38 languages. Earlier in 2011, Evidon reported 500,000 “opt-outs” as a result of presenting over 50 billion ads (Dunaway, 2011). However,

[...] those are not "user opt outs" but "company opt outs." For example, if a browser hits an advertiser's opt-out page on Evidon that has 10 ad tech companies listed total, if all are opted out, that counts as 10 opt-outs, even though it's just one browser. (Dunaway, 2011)

Disregarding this, Dunaway (2011) reports the rate of “opt-out” at 0.0046% and Evidon (2011) reports 0.0035%.

Tribal Fusion (2012) further reports from the period of January 2012 to mid-May 2012, the average daily activity of icon activity was 0.009% (clicks/impression). Of these, 0.003% were users who had set “opt-out” cookies in their browsers.

Evidon (2011) notes that “engagement” in opt-out campaign is influenced by the prominence of the icon — a white icon container on a black ad — results in the highest click rates.

They also note that "transparency doesn't foster opt-out". 1 in 700,000 people who see the icon actually opt-out (about 0.0001%). Furthermore, they calculate that about 1 in 5 people who make it to the opt-out page actually opt-out.

It's important to note that though AdChoices was introduced in 2010, its rollout to comply with European law was compliancy by 30 June, 2012 ([Crimtan.com, n.d.](#)).

6.1.3 *AdChoices and Confusion*

In a study conducted by [Blase Ur \(2012\)](#), 48 participants were asked about their familiarity with the DAA advertising icon and older "Power I" icon. First, they were presented enlarged icons and tag lines with context. 41 responded they had never seen the icons. Later, presented in context next to advertisements, 25 participants still stated then had never before seen the icons while 8 more were unsure. Only 5 participants thought the icon was intended to provide information about OBA.

This report highlights many usability issues surrounding OBA, to include notice and choice mechanisms. However, they did not address the cause of why AdChoices appears so ineffective.

6.1.4 *What Makes a Good Ad Design?*

While there is controversy about the effectiveness of "enhanced notice", the advertising industry is actually very good at design. In an online Banner Ad Design Web clinic ([McGlaughlin, 2011](#)), Flint McGlaughlin describes three key objectives that banner ads must accomplish in order to drive "maximum conversion".

1. **Attract interest.** McGlaughlin gives examples from A/B market tests manipulating size, shape, color, motion, and position toward ad effectiveness.
2. **Generate interest.** In this, Mclaughlin and others (for example, posts concerning banner ad effectiveness on <http://retargeter.com>) focus on the need to understand where targeted consumers might be in the buying process. This principle speaks primarily to tailoring the message to an intended audience.

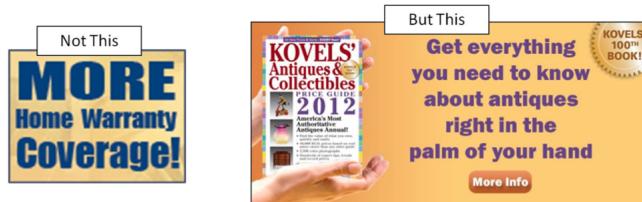


Figure 33: Example "Ask for the Click" (McGlaughlin, 2011)

3. **"Ask for the Click."** Finally, McGlaughlin emphasizes the need for a "call for action". While Internet users implicitly understand that ads are representative of the target site to which they are linked, by placing the image of a button in an ad, McGlaughlin stresses that clarity of the message is increased.

In Figure 33 from McGlaughlin (2011) and associated audio transcript below, *it is clear to see that while advertisers believe that, despite an understanding that ads convey an implicit message to click, conveying this explicitly is vital.*

Not this, here is an ad [left]. It is animated. They have got motion. Never mind the ugly color. Never mind that it screams the entire time you are looking. Never mind that you have to watch it for a long time before you get the message. Just consider the fact that even once you have watched the whole message, you don't know what to do.

To this, now this ad on the right shows...clearly, we could optimize this way better, but let's just look at this to learn from it. I see a picture of what I am getting, and because it is in a set of hands, I can actually get a sense of the size, and I can imagine it. Never promise a download or any other incentive that you don't help the audience visualize or imagine. Secondly, in the cover and in the key content of the cover, there is appeal already built for somebody in this space, and it is a well-known brand, and then, here it says, get everything you need to know about antiques right in the palm of your hand.

That is also quite helpful in understanding what the offer is, but these people have identified that at this stage, the person may not be ready to buy the

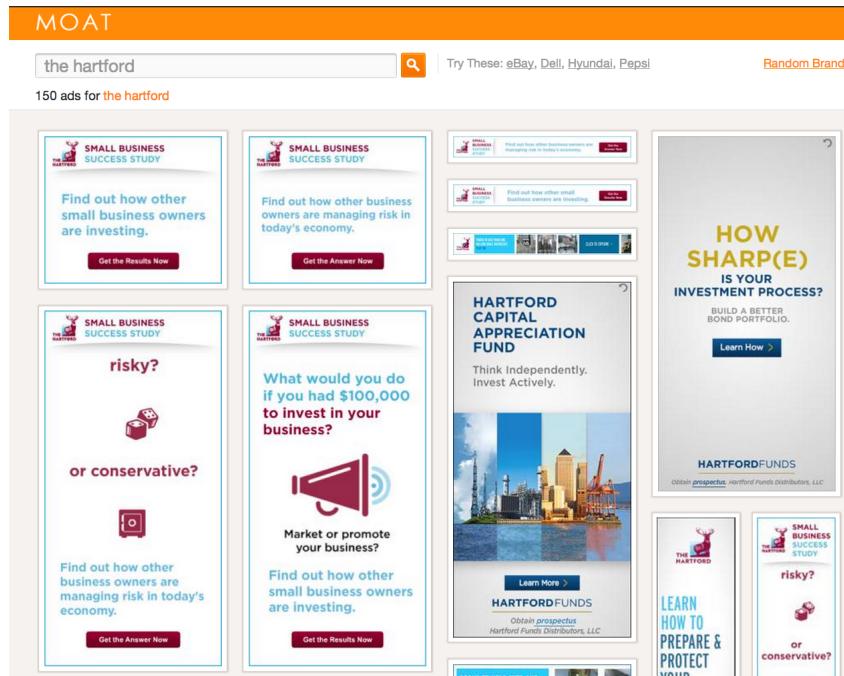


Figure 34: Screen Capture from <http://moat.com>

book. They may want more info. That is the difference between saying a button 'Buy It Now'. You know, 'Get Your Copy' and instead 'Learn More'. (McGlaughlin, 2011)

Finally, by regarding current ad images on <http://moat.com> (Figure 34), it is evident that iconic buttons in image-based ads are highly desirable. This is the case, whether these are flash-based ads (buttons are functional) or simple image-based ads (buttons merely iconic).

6.2 RELATED RESEARCH

Advertising is an intensely competitive business, so it is not surprising that specific details about the effectiveness of the AdChoices icon are difficult to ascertain. Research presented below is directed toward iconicity and indexicality of ads and their parts. This relates to the notion of a "call-to-action".

6.2.1 Salience

Foundational to the representational use of an icon in an ad is whether or not it is perceived in a composite image. Low-level visual properties have a potentially large effect on

users. Hoffman and Singh (1997) find that the saliency of a part depends on at least three factors: size relative to the whole object, degree to which it protrudes, and strength of its boundaries. They find that these factors influence visual processes determining the choice of figure and ground. Elazary and Itti (2008) furthermore find that the selection of interesting objects in a scene is also largely constrained by low-level visual properties. In 76% of images studied, one or more of the top three salient locations fell on an outlined object.

By contrast, top-down models of attention concern the effects of objects and semantic relevance. In studies of task salience, Hegarty, Canham, and Fabrikant (2010) find that good display design facilitates performance by highlighting visual features representing task-relevant information.

While visual saliency is concerned with how well an object stands out from other objects, cognitive saliency is concerned with the relative importance of information. Schmid (2007) describes cognitive salience as concepts directly activated and loaded into current working memory or via spreading activation where one concept activates another (e.g., "dog" activates "bark", "fur", "poodle", etc.). But he also makes reference to ontological saliency, described as a more stable property of objects. He gives the example of a photograph where dog is more salient than the field in the background. Cognitive saliency seems to refer, in part to the effect of visual properties in foregrounding and backgrounding, but also of the ontological saliency of objects, in general.

Cognitive salience is important to understanding the referential use of language in production and comprehension. Gundel, Hedberg, and Zacharski (1993, 2012) account for the lexical expression of referring forms via two kinds of information: conceptual information about a speaker's intended referent, and procedural information about the assumed cognitive status of that referent in the mind of the addressee. That is, the more accessible an entity is to an addressee, the less information a referring expression needs to contain to be correctly understood.

So, for example, imagine the scene where two people are in a room where there is a basket of apples on a table. Person A notices person B looking at the apples and says,

"Are you hungry? Please take one."

The apples are present and, in fact, are uniquely identifiable. If they had not been, Person A might have said,

"Are you hungry? Would you like an apple?"

The apples were highly salient so a pronominal form of reference contained enough information for Person B to understand the referent. With respect to visual scenes, **Vogels, Krahmer, and Maes (2011)** note that visually salient referents are more activated in memory and are thus better accessible than less visually salient referents. In controlled experimentation, expressions referring to visually salient entities tended to be more reduced.

6.2.2 *Deixis*

Deixis is a type of linguistic reference that depends on context for understanding. For example, context can be linguistic as in:

1. Julie went to the library. She was there all day.

The pronoun anaphorically points to a referent in previous linguistic context. Deictic reference can also be extra-linguistic.

2. I'll take that one (pointing to item in a glass case).

(**Fillmore, 1971**) distinguishes between several sorts of deixis which he identifies as **gestural**, **symbolic**, and **anaphoric**. A gestural use of a deictic expression is like sentence (2) above where a demonstrative is accompanied by a pointing gestures. It relies on the hearer monitoring some aspect of the physical situation.

Sentence (3) below exemplifies a symbolic deictic expression.

3. This administration doesn't know what its doing.

It's interpretation depends on a hearer knowing about some aspect of the communicative situation. Finally, an anaphoric use of an expression is one which can be interpreted by knowing what other part of discourse expression is coreferential with it. Sentence (1) above exemplifies this.

Deictic reference is particularly problematic in pragmatic understanding because it involves subjective, attentional, intentional, and other context dependent properties (Levinson, 1983).

The difference between deictic & non-deictic conceptions can be understood by analogy. Consider the difference between a sculptured representation of a human figure, set up in the middle of a courtyard, and a photograph of a human figure. The sculpture does not represent any particular observer's-point-of-view, but the photograph does. The photograph does because the camera had to repositioned at a particular place in front of or to the side of, above or below, or on the same level as, the model." (Fillmore, 1971, p. 235)

6.2.3 Indexicality and Knowledge

Semiotician Charles Peirce influenced theories of reference through his work on signs. In Peirce's semiotic theory there are three basic types of signs: *icons*, *symbols*, and *indexes*. Every sign stands for something — some object. What distinguishes them is the relation between the sign and the object. Icons resemble their object and, therefore, depend on perceptual knowledge. Symbols depend on communicative associations between sign and object. And indexes are "physically connected" — they depend on knowledge about the connection in order to relate index to its object. Indices simply direct attention-to and refer to objects but, otherwise, assert nothing (Atkin, 2005; H. H. Clark, 2003). However, pre-requisite for understanding the referent of an index is *knowing the relation between the index and its object*.

6.2.4 Demonstrative Reference

In study of demonstrative reference, Kaplan (1980) characterized demonstration as directing intention, where the demonstration itself does not determine the referent. What he meant is that the demonstration is simply an "externalization" of intent Kaplan (1980, p. 589). The demonstratum — pointing — has no intrinsic meaning: recognizing speaker intent is key.

In the case of a hypertext link, visual cues (e.g., color of digital ink, mouse hand icon) indicate the presence of a link. This link is indexical to some other body of text. After brief introduction, users understand that this relationship as indexical. But even when visual cues are not available, users may still infer an indexical relation.

In order for deixis to be effective, speakers communicate using signals with the intention to produce awareness. In this rests a coordination problem. Lewis (1969) describes two ways in which indexicals may be established through convention: precedence and salience. He gives the example where he comes upon a patch of quicksand and wants to warn others of its presence. He puts up a scarecrow up to its chest hoping that whoever sees it will understand. He intends to produce awareness, but there is no established convention for communicating quicksand. This he hopes is communicated through force of salience. And through precedence. And at some point in time, as a signal becomes more salient, its indexicality is established through convention.

6.2.5 *Hyperdeixis*

How do Internet users know advertisements are indexical? Often, there are perceptual cues. Mouse rollovers change the iconic representation of a mouse cursor to a pointer. Often visual cues are available, such as text color. More important, advertisements are known to be indexical through convention.

But what about complex images or graphics where the user can perceive no discernible change in the cursor nor other physical attributes to cue off? For example, would you expect the Figure 35 to be indexical, or the parts themselves to be indexical?

In the pre-pilot described in the next section, users did not agree on indexicality. Some thought the image linked to one unique location, and others thought each part of the image linked to a unique location. (In fact, the latter were correct, but the group was divided evenly.)

In light of the discussion above, features of advertising images I am most interested in examining are:

1. **Indexicality** - By convention, users know that to click on an image-based advertisement is to direct them to the advertiser's website.



Figure 35: Complex Image



Figure 36: Complex Advertisement

2. **Iconicity** - Buttons iconically communicate that an action may be performed and, depending on the context, have an indexical relation with a target webpage.
3. **Brand symbols** - Brands are indicative of particular companies and may also be indexical in an online context.

A single advertisement can contain all of these attributes in combination. Some advertisements can become quite complex to the point of containing multiple microinteractions.

As in the case of the complex image in [Figure 36](#), visual and perceptual cues (e.g., mouse cursor change to iconic hand) may not help users identify either indexicals or referents.

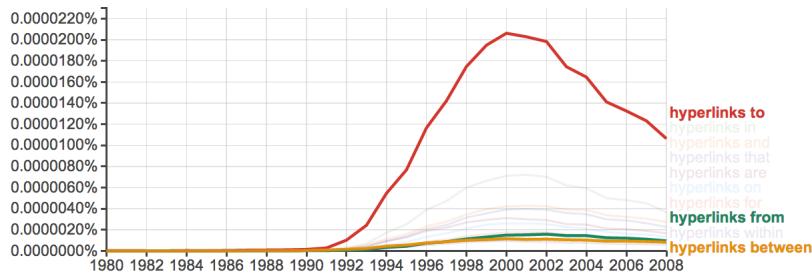


Figure 37: Google Ngram Viewer (hyperlinks *)

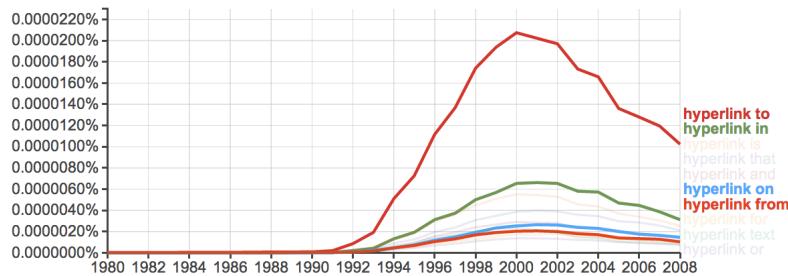


Figure 38: Google Ngram Viewer (hyperlink *)

While Loehr (1997) refers to hypertext deixis as anchored links between text on a webpage to some other text, I refer more generally to such links as examples of *hyperdeixis*. Arguably, user interaction online has become sophisticated to the point where one could dream up any sort of interaction and make it happen. For example, there is no intrinsic reason that hyperlinked text couldn't point to multiple objects. It would not be difficult to create this effect using simple JavaScript: a user mouses over text and a multi-option select box pops into view.

However, there is ample corpus-based linguistic evidence to suggest that we conceive of hyperlinks as "pointers". When we talk about hyperlinks, we use spatial metaphors. In Figure 37 and Figure 38, it's easy to plot the use and distribution of the word "hyperlink" by its usage in modern media. Hyperlinks go "to" and "from" locations and are things "between" objects. They are also found "in" or "on" web text.

In some sense, we are wired to understand the world this way. Our conceptual system has properties grounded in perception — both physically and socially (Lakoff, 2008). Thought and reason are not simply linguistic abstractions but come from an intimate understanding the physical world around us.

6.2.6 *Usability of Hyperlinked Icons*

There appears to be little relevant research (if any) on the topic of hyperlinked images. However, there exists some on the topic of icons. H.-I. Cheng and Patterson (2007) studied the use of iconic hyperlinks on e-commerce websites. Not surprisingly, they found that easily identifiable icons increase system usability. Those that are ambiguous affect the rate of task performance. Wiedenbeck (1999) specifically examined the effect of combining text labels with icons. For a learning-based task, performance was poorest on icon-only interfaces, however, once learned, perception of ease of use was higher.

The pre-pilot discussed in the next section explores how people understand indexicality in the context of advertisements.

6.3 PRE-PILOT

The purpose of this pre-pilot was to better understand how users perceive and interact with advertisements composed of graphical and iconic elements.

6.3.1 *Method and Procedure*

Between the dates of 7–14 October 2013, I created an AMT assignment for 10 workers to annotate a series of approximately 25 images. For this task, I paid \$0.75.

There were two groups of images representing two tasks. The first four image annotations were a test to see whether there might be a tangible effect of a “call-to-action” button icon on an ad. That is, would the presence of an iconic button influence where a person clicked on an ad. For these images, participants were given a screen shot of a digital news article and asked “Please find the [Ford] Ad and click on it” (substituting the brand name as appropriate).

The 21 remaining images comprised the second task which explored participant perception of image advertisements, including those containing the AdChoices icon. Workers were presented ads in isolation and asked to click on “clickable things”.

Before the second set of images was presented, I provided a short training session to familiarize annotators with the task.

Let's change gears now.
 This is a practice image.
 If you saw this in a newspaper, how many things do you think would be "clickable"? Things are clickable if they do something or take you to another webpage.
 In fact, the whole image, but nothing inside, is clickable. Please select '1' below to indicate one thing is clickable.

1 2 3 4 5 6

Now click on the image. To continue, you should have one sticky note on this image. Don't worry if its a little offset from where you clicked.



>>

Survey Completion 0% 100%

Figure 39: Training Task 1

For the first training task (Figure 39), annotators were asked simply to click on the ad. By clicking, a tiny sticky note was left behind to remind the annotator where they clicked. Annotators were also asked to count the number of clickable things (things that do something or take you somewhere). For the first image, annotators were instructed to select "1".

For the second training task, a typical sign-in /registration box was presented. Participants were instructed that text links and buttons from forms were clickable, but that text entry fields were not (Figure 40).

Finally, in the third training task, participants were told that there were multiple things one could click on, but there were only three links (Figure 41).

After annotators completed all three training tasks, they were presented images to annotate on their own. A typical image annotation task is shown in Figure 42.

6.3.2 Results and Discussion

Results from the first four heat map visualizations appear in Figure 43, Figure 44, Figure 45, Figure 46.

In each of the four images, one or more participants clicked directly on the iconic button link, despite implicit knowledge

This is a second practice image.

In the image below, two things are clickable.

The input boxes are not "clickable" since if you type in them they do not take you to a different webpage. But clicking "forgotten password" and "sign in" do cause you to go to different page views.

Select "2" below to indicate two things are "clickable".

- 1 2 3 4 5 6

Now click on the image. You should have two sticky notes on this image. You can click on sticky notes to delete.

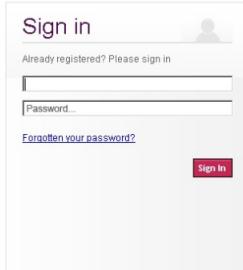


Figure 40: Training Task 2

Last practice image. There are three places this image will take you. But you could say that there are either three or six things that are clickable. Select '3' to say that there are three "links".

- 1 2 3 4 5 6

Click in all of the places that will take you to another webpage.

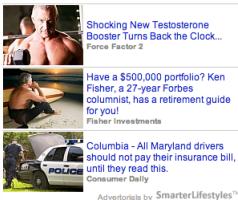


Figure 41: Training Task 3

Links?

- 1 2 3 4 5 6

Click



Figure 42: Image Annotation Task

6.3 PRE-PILOT



Figure 43: Ford Ad



Figure 44: Cars.com Ad

6.3 PRE-PILOT



Figure 45: Traveler's Ad



Figure 46: Visa Ad



Figure 47: Results for Complex Ad

that the ad itself was hyperlinked to the advertiser's landing page. In addition, it appears that brands iconic in shape and style as buttons, also elicited clicks. Note the distinctive shape of the Ford logo in the first image and cars.com logo in the second. However, it could also be said that most clicked on the center of an ad, rather than on a button or iconic brand image.

[Figure 47](#) shows the result for the example task 2 image above:

Roughly half of the participants saw this ad as containing only one link. Others appeared to see fake buttons as links.

[Figure 48](#) is another example from task 2. Here you can see that half of the annotators clicked on the Twitter icon.

Finally, [Figure 49](#) is an ad with the AdChoices icon. No one clicked on the AdChoices icon. But 8 of 10 clicked on the "See the Phones" call-to-action. And several each clicked on the Motorola, Droid, and Verizon brands at the bottom of the image.

By examining task 2 results, I learned that my task was too complex and it was possible participants may have misunderstood what was asked of them. However, the pre-pilot also proved the feasibility of studying user perception of the indexicality of advertisements. This simple qualitative study



Figure 48: Results for McDonald's Ad with Twitter Icon

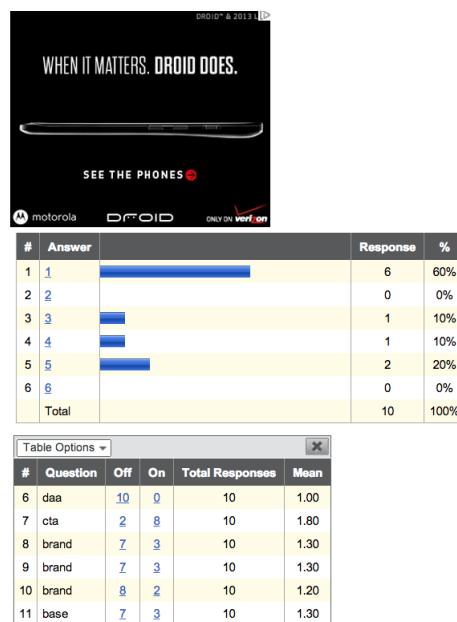


Figure 49: Android Ad with AdChoices Icon

provided insight valuable to the design and implementation of controlled experiments.

6.4 AIMS

In this pilot, I consider indexical reference in hyperlinked advertisements in the context of task-based interaction. In experiment 2A I look at how knowledge contributes to perception of an indexical link on a brand icon. In experiment 2B, I look at iconic buttons on advertisements to see if the presence of a button affects where a user clicks on an ad.

6.5 EXPERIMENTAL DESIGN

Study 1A considers the proposition that iconic elements of advertisements are perceived as indexical if people know what the icon represents, it contrasts semantically with the advertisement, and it signals communicative intent (for example, as a “call-to-action”). This breaks down into three study different hypotheses.

Hypothesis 2A1

Known icons, embedded in advertisements, with brands contrastive to the background advertisement, are perceived differently than icons that semantically match the brand of the primary advertisement.

A McDonald’s icon on a McDonald’s ad should not be seen as indexical: merely iconic. However, an Apple icon on a McDonald’s site might be considered indexical.

Hypothesis 2A2

Known icons, embedded in advertisements, are perceived differently than those that are unknown.

This hypothesis predicts that icons which are unknown may be treated differently than those that are known. That is, a fake brand may be perceived differently than the known Apple brand. Likewise, the AdChoices icon, if unknown, should be treated as an unknown brand.

Hypothesis 2A3

An icon that signals action is perceived differently than an icon that does not signal action.

No sign of comet after pass around sun: scientists

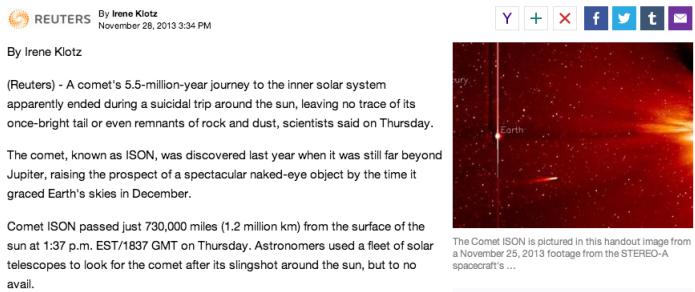


Figure 50: Twitter and FaceBook Icons

A Twitter logo (as appears in Figure 50) is iconic and known to serve one of several functions: "tweet" or "go to Twitter" (to tweet/follow). Likewise, a plain FaceBook icon (also in Figure 50) may be perceived to function as "share" or "go to FaceBook" (to share/follow).¹ That people know these things about Twitter and FaceBook is a function of their relative popularity online. Hypothesis 2A3 predicts that icons, known or unknown, will be perceived differently if they signal action such as share/follow/post/goto, etc.

Experiment 2A is a single factor posttest-only design. There is one **independent variable with 5 levels**: "same brand", "different brand", "unknown brand", "known signal", and "AdChoices". There is one **dependent variable**: perception of an indexical reference. Five groups are required. The control group sees a McDonald's ad with an icon McDonald's logo in the upper right hand corner. Other groups see a different logo of the same size in the same place: Apple, Unknown, FaceBook, and AdChoices. Apple and FaceBook logos were chosen since they have among the highest global brand recognition (Interbrand, 2013).

Hypothesis 2B

Iconic graphical elements influence the click behavior by affecting where people click.

Experiment 2B is a single factor posttest-only design with three levels. The control and two treatment groups are presented with a newspaper article containing an ad. The control group receives an ad with no "call-to-action", while one

¹ These observations are made on the basis of a short AMT survey where I asked 40 turkers what they thought these icons meant in the context of the Yahoo news article in Figure 50.

treatment group sees the same ad but with a textual “call-to-action”, and the other treatment group sees the same ad with an iconic button containing the same “call-to-action” text as the first treatment group. The **independent variable** is “icon” and **dependent variable** is click location (on the “call-to-action” or elsewhere).

Data for Experiment 2A and 2B were collected in separate events. Following presentation of a single task, survey data was collected.

6.6 METHOD

These two experimental designs follow the general procedure outlined in Chapter Four.

6.6.1 *Experiment 2A Method*

Settings and Participants

Using Amazon’s Mechanical Turk, 303 workers were paid \$0.15 to participate in one of two conditions of the single-factor design previously described. Participants were assigned randomly via Qualtrics block randomizer. Results were collected during the period of 7–14 October 2014. The task was expected to take approximately three minutes, though workers were allowed up to 10 minutes to complete.

Procedure and Materials

The general procedure follows that described in Chapter Four. [Figure 51](#) gives a more detailed graphical depiction of flow.

Following presentation of instructions and consent form, participants were each presented one question, as illustrated in the exemplar in [Figure 52](#), followed by survey questions. Participants were instructed to “Please click on things you think are links.”

The image used was taken from an online article at <http://yahoo.com>, while the advertisement from <http://moat.com>. Alterations were accomplished using PhotoShop to create five variants. Logos were restricted to a size of approximately 20px x18px to approximate the 19px x15px (icon only with container) size specified by [DAA \(2011\)](#). However, all logos including the AdChoices icon were presented with 100% opac-

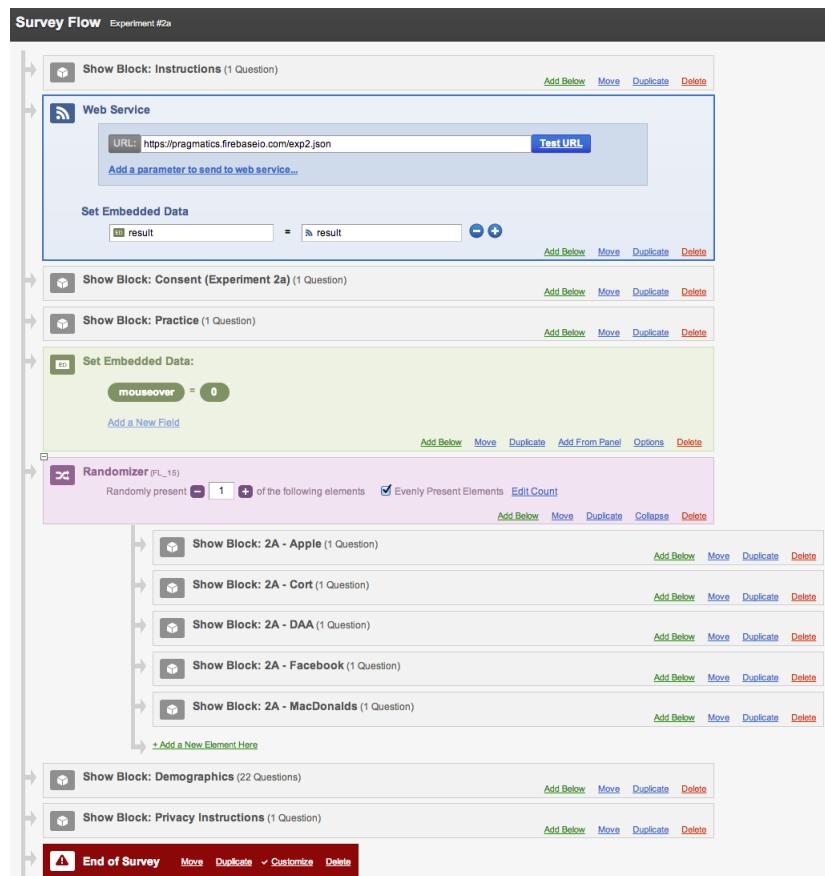


Figure 51: Experiment 2A Flow

The Mentalist's Owain Yeoman Marries Gigi Yallouz

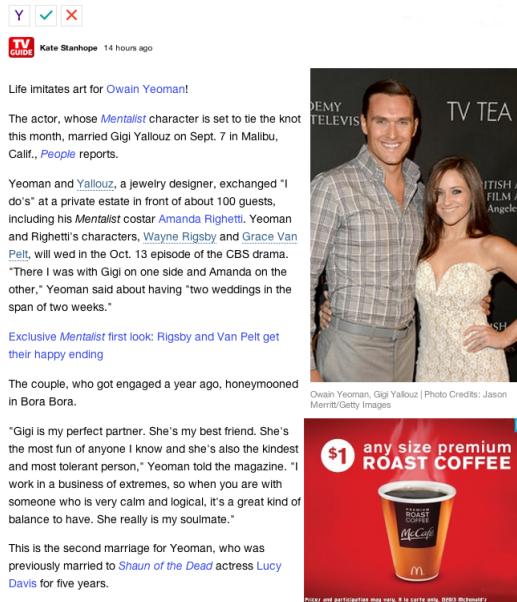


Figure 52: McDonald's Ad with AdChoices Icon

ity on a white background, except for the Apple logo which was unaltered as white on a black background.

Instrumentation

In addition to features provided by AMT and Qualtrics, my scripts performed the following:

1. Disable preview using a custom CSS class blocking Qualtrics controls
2. Check worker hash against a Firebase web service and request worker to return HIT if the hash is in the exclusion list
3. Add worker hash to Firebase
4. Submit results to AMT on completion of the Qualtrics survey
5. Hide Qualtrics hotspot regions from mouseovers
6. Create "sticky" notes on clicks
7. Observe whether mouseovers were detected over treatment icons

The last three capabilities were developed as custom code for this particular experiment. Qualtrics provides two image annotation capabilities, both of which are limited in some respect.

The heat map annotation provided by Qualtrics used in the pre-pilot, required that the number of clicks be pre-specified by the experimenter. This is not desirable for tasks in which participants will click an arbitrary number of times.

Instead, I used the hotspot annotation capability provided by Qualtrics. This let me create image zones to detect where a user clicks. Unfortunately, Qualtrics zones become visible when a participant mouses over a region. For the purposes of the pre-pilot and Experiment 2, I modified the Qualtrics hotspot capability in two ways: 1) never show zones; and 2) clicking creates a small sticky note so that the participant can see some evidence of click feedback at the location of the click.

6.6.2 *Experiment 2B Method*

Settings and Participants

Using Amazon's Mechanical Turk, 202 workers were paid \$0.15 to participate in one of two conditions of the single factor design previously described. Participants were assigned randomly via Qualtrics block randomizer. Results were collected between the days of 7–14 October 2013. The task was expected to take approximately three minutes, though workers were allowed up to 10 minutes to complete.

Procedure and Materials

The general procedure follows that described in Chapter Four. [Figure 53](#) gives a more detailed graphical depiction of flow.

Following presentation of instructions and consent form, participants were presented a screen shot from a *The New York Times* article with my ad substituting for the same-sized block ad in which an ad originally appeared. I used the Ford ad that had been previously used in the pre-pilot (same as figure [Figure 43](#)). Participants were instructed, "pretend you are interested in getting a hybrid car. Find the Ford ad and click on it."

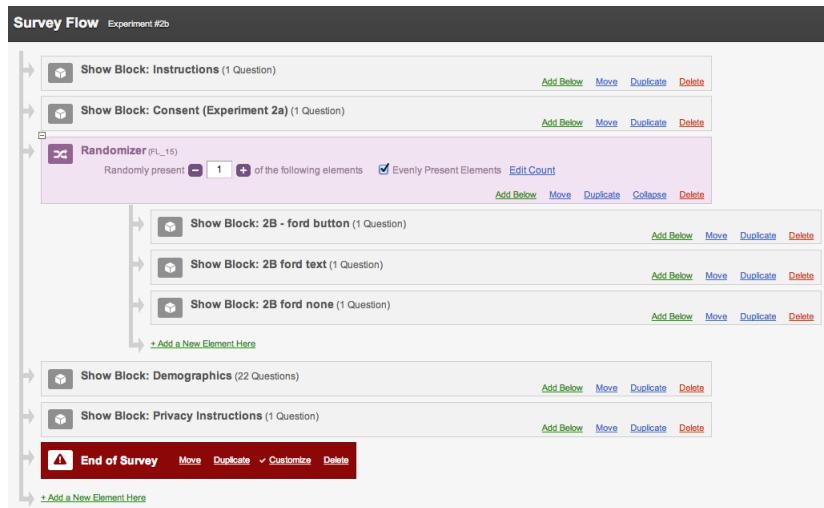


Figure 53: Experiment 2B Flow

Instrumentation

In addition to features provided by AMT and Qualtrics, my scripts performed the following:

1. Disable preview using a custom CSS class blocking Qualtrics controls
2. Add worker hash to Firebase
3. Submit results to AMT on completion of the Qualtrics survey

In this experiment, I used the heat map capability of Qualtrics since only one click was requested.

6.7 DATA COLLECTED

Using G* Power 3 (Berger, 2012) chi square goodness of fit test, for both experiments I estimated a sample size of approximately 220 was necessary in order to detect a medium effect with power of 0.95. All surveys initiated were completed: there were no known drop-outs.

Once all assignments had been completed, (as indicated on my AMT requester dashboard), I downloaded data as a single CSV (comma separated values) file from the Qualtrics website. Data was organized such that each participant's data was on its own row.

Table 8: 2A Raw Clicks

Clicks	Total	Icon
1	60	Apple
2	59	Unknown
2	58	AdChoices
10	52	FaceBook
0	59	MacDonald's

6.8 RESULTS

6.8.1 Experiment 2A Results

Because the IV is categorical (binary), a non-parametric statistics is required. Strikingly, there are very few clicks on icons at all.

9% (27 of 303) of participants said they recognized the AdChoices icon, though when questioned what it meant, only 5.6% were correct. Of the two that clicked on the AdChoices icon, one was correct and the other mistakenly believed it would link to the advertiser's page. Of 303 participants, there were 21 mouseover events over icons — 3 (including the two that clicked) over the AdChoices icon. This means, that had the AdChoices icon been animated on mouseover, one additional person might have been alerted that the icon was meaningful.

Given relatively little data, it is difficult to come to any strong conclusion. Results suggest that the effect size was considerably smaller than anticipated. In the pre-pilot, approximately 50% (5 of 10) clicked the Twitter icon. In this study, only 19% of those presented with the FaceBook icon clicked on it. As a follow-up, I added a condition that simply enlarged the icon to approximately twice the size (30px x 28px). This had no effect: only 17% (10/58) presented with the larger FaceBook icon clicked on it.

I also considered that it might be possible that, despite studies showing little difference in quality of annotation with price variation, amount of pay might be a factor. However, I ran another 200 subjects at a higher pay rate (25 cents for a 2-minute task). This also caused no change.

Table 9: 2B Results

	Button	Logo	Other
None	0	2	64
Text	0	1	66
Button	1	1	67

It is still possible, that the small number of click observations could reflect omission effects of the paid volunteer ([Rush et al., 1978](#)). This remains un-tested.

Finally, the pre-collection focused on images in isolation. This experiment placed ads in context. It is quite possible this accounts for differences seen. Typically, linguistic judgment tasks are presented under isolated conditions where external context plays no role. However, this experiment afforded the opportunity to observe participant behavior under relatively natural conditions and was useful for this reason, if no other.

Though I created additional stimuli, due to time constraints, they were not used in this experiment. As with Experiment 1, advertisements as stimuli (and, indeed, the background newspaper itself) introduce random effects. Thus, the results presented here do not necessarily generalize across all advertisements or embedded contexts.

Committee: I will re-run this in isolation over the Holidays. I may also test both Twitter & FB icons.

6.8.2 *Experiment 2B Results*

[Table 9](#) presents raw frequency counts for each of the three conditions.

As before, because the IV is categorical (binary), a non-parametric statistics is required. Again, some cell counts were quite low and, overall, unbalanced — so I used a Fisher’s Exact Test. The p-value was equal to 1 indicating no significant difference between conditions.

As in 2A, results differed somewhat from the pre-pilot. The difference here was the title of task and payment amount, plus slight variation in the instructions. It is possible that the effect size here was simply too small to detect. A/B testing of ads generally involve a large number of subjects such that it might be possible to detect actual effects on behavior where none in this size study was seen.

This task was also presented in the context of a news article and not in isolation, as had been in the pre-pilot. As with

Committee: I will re-run this in isolation over the Holidays.

2A, I ran a follow-up study offering 25 cents for a shorter task. There was no effect on performance. It is possible that a larger effect size might be observed if the task were presented in isolation, rather than in context.

STUDY THREE: MULTIPARTY DISCOURSE IN THE BROWSER

Why do users feel OBA is “creepy”? As advertisers say, OBA simply tailors ads so that users receive more relevant advertising. Granted, this involves collecting data about online activity, but consumers willingly give data to first-party sites all of the time. What makes this feel different?

7.1 OBA IS “CREEPY”

Blase Ur (2012) conducted semi-structured interviews with 48 non-technical users about their attitudes and understanding of OBA. As they note, these participants are not representative of the general Internet population. The researchers were interested in collecting data that would reveal mental models of “lay people” to better understand attitudes and behaviors.

Here are a selection of participant comments from this study after they had viewed an instructional video on OBA:

"It is a little creepy... because I feel that I should get to decide what is going in and out of my computer." (Blase Ur, 2012, p. 6)

"It makes me feel very insecure. Like if this is what people can figure out about me, then what else can they get off my computer?" (Blase Ur, 2012, p. 6)

"I guess I would be more willing to do it if I had a firmer understanding of how everything worked." (Blase Ur, 2012, p. 7)

"I don't think I really noticed it... but it definitely is kind of creepy when you think about it." (Blase Ur, 2012, p. 7)

"It's kind of a creepy thought that you are being followed and monitored." (Blase Ur, 2012, p. 7)

One of the subjects,

relating a story about how she was searching for furniture the previous night and was confused when

her advertisements started to feature those items....

'It's scary. It makes me nervous. I was thinking about it last night when I was searching for stuff. Like I thought how do they know all this, how do they keep track of this, how do they do this?'
 (Blase Ur, 2012, p. 7)

"It makes me want to go home and delete all my cookies, but then I know that's not gonna help much. It makes me mad." (Blase Ur, 2012, p. 7)

The lay person appears to have a strong emotional response to OBA. Can a theory of social communication lend insight into why?

7.2 RELATED PRIVACY RESEARCH

What makes people willing to share sensitive information online? Behavioral economists Acquisti, John, and Loewenstein (2012) find that disclosure operates under certain principles espoused by Kahneman and Tversky (1979) and A. Tversky and Kahneman (1981).

In an analysis of decision under risk, A. Tversky and Kahneman (1981) find that people make choices by weighing their decisions in light of choices and perceived outcomes of competing choices. Choices are comparative:

They also find that choices depend on a starting reference point:

An essential feature of the present theory is that the carriers of value are changes in wealth or welfare, rather than final states. This assumption is compatible with basic principles of perception and judgment. Our perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes. When we respond to attributes such as brightness, loudness, or temperature, the past and present context of experience defines an adaptation level, or reference point, and stimuli are perceived in relation to this reference point. (Kahneman & Tversky, 1979, p. 277)

Acquisti et al. (2012) focus on the notion of a changing reference point in order to see what sorts of signals might

affect a respondents willingness to disclose sensitive information. They do so by measuring people's propensity to admit to engaging in sensitive (embarrassing, unethical, and illegal behavior) in a series of surveys.

They find evidence that people will admit to having engaged in sensitive behaviors when they are lead to believe that others have admitted in similar behaviors (*herding effect*). They also examine the effect of question order on disclosure (*ordering effect*). And indeed it does. Participants in the increasing conditional were less likely to admit behaviors than in the condition where questions decreased started as very sensitive and then decreased in sensitivity. Both experiments support [Kahneman and Tversky \(1979\)](#) and also give insight into how and why people disclose information online.

In a study related to [Acquisti et al. \(2012\)](#), [John, Acquisti, and Loewenstein \(2009\)](#) manipulated the saliency of privacy to see whether this would affect propensity to disclose. When saliency of privacy was high, willingness to admit to engaging in sensitive behaviors decreased. And when subjects were distracted from privacy concerns, their propensity to disclose was increased.

Timing also has an effect on behavior. [Egelman, Tsai, Cranor, and Acquisti \(2009\)](#) studied the effect of the presentation of privacy information in the context of purchase decisions. By controlling the placement, timing, and privacy level on several consumer websites during a purchase task, participants reacted differently for the purchase of privacy-sensitive items than for items with minimal privacy concerns. In four conditions, different privacy icons were used to study both when and how icon presence affects behavior. [Egelman et al. \(2009\)](#) found that 1) the presence of privacy indicators influenced purchase decisions, and 2) online shoppers who were less privacy-aware, paid significantly more for privacy when privacy indicators were presented *before* visiting websites than *after* they arrived at the website.

Clearly, people are willingly disclose sensitive information online every day. And they are also sensitive to the presence of visual indicators that indicate degree of privacy. In the next section, I show that behavior is also affected by knowledge of the presence of external participants in interaction.

7.3 RELATED LINGUISTIC RESEARCH

7.3.1 *The Stateful Web*

When Tim Berners-Lee first implemented the Hypertext Transfer Protocol (HTTP) in 1990, it was to facilitate communication between a client and a server on the Internet (Cern, n.d.). HTTP functioned as a protocol to facilitate communication much as the telephone: it provided the means for a client to request information from a server via a small set of verbs such as “GET” or “POST”. As a protocol, HTTP is stateless and does not require the server to track state between connections. This means, if a user requests a webpage that includes a number of embedded assets (e.g., images), the web server does not have to know anything about these assets nor does it have to track the status of those requests. In response to a request, a server generally replies with a status such as OK plus information about what sort of content is to be transferred followed by that content. In fact, an HTTP request and response look something like this:

```
GET /
Accept-Encoding: gzip
User-Agent: Rested 2.3 (Macintosh; Mac OS X 10.8.0; en_US)
---blank line---
```

And a response looks something like this:

```
HTTP/1.1 200 OK
Server: gws
Set-Cookie: PREF=ID=00a5929d47bf3487:FF=0:TM=1350849438:LM
=1350849438:S=YYHqxKmLxevdwvbo; expires=Tue, 21-Oct-2014
19:57:18 GMT; path=/; domain=.google.com, NID=65=hS0s8
SoRDwLN1gqEDi3Mx_rjCNUBddfp2ou0mMn40GQBBV9
VEejtx0uKKaIDFt5TMyrBs0ZGBZ3BH-449m2GtPjqxVsYXjuN96
Dgxa0ybcNHf0jXwe2R9t6G05z3Hqhd; expires=Mon, 22-Apr-2013
19:57:18 GMT; path=/; domain=.google.com; HttpOnly
Content-Type: text/html; charset=ISO-8859-1
Transfer-Encoding: Identity
P3P: CP="This is not a P3P policy! See http://www.google.com
/support/accounts/bin/answer.py?hl=en&answer=151657 for
more info."
Date: Sun, 21 Oct 2012 19:57:18 GMT
X-Frame-Options: SAMEORIGIN
X-XSS-Protection: 1; mode=block
Cache-Control: private, max-age=0
Expires: -1
```

---blank line---
---HTML entity here---

In order to accommodate stateful applications — applications that remember information about previous transactions, a small piece of data called a cookie can be attached to an HTTP request and deposited in the client web browser. This concept, envisioned by Lou Montulli of Netscape in 1994, was invented to facilitate a simple sort of memory (J. Schwartz, 2001). In so doing, it also enabled the sense of dialogue between server and client. Originally, cookies were intended to support basic dialogue between user and website publisher for simple transactions such as those needed to support online shopping (e.g., a cookie can be used to keep track of a user's session to include items in a shopping cart).

Though the purpose of cookies was to support the need for state between HTTP transactions, this very simple concept fostered economic and social revolution on the Web: it fundamentally changed interaction on the web from *private* to *public*. Content providers found that sites could be supported by advertising revenue simply because an advertiser with an advertising server could use cookies to decide what ad to return in response to an HTTP request. Thus, in any web page transaction, tailored advertiser content could be presented alongside publisher content.

7.3.2 A Model of MultiParty Interaction

The client-server interaction described above is a sort of computer-mediated communication between humans. HTTP as a machine-based communications protocol enables two parties to exchange messages. Because messages are typically simple HTTP requests with textual or other media content as a response, there is no sense of shared social context involved between a user and web provider and such dialogues are not been of linguistic interest. On the other hand, linguistic analyses of machine mediated human conversation that occur in chat rooms, texting, weblogs, etc. are a vibrant area of study in peer-review venues such as language@internet¹.

Though traditionally, the term discourse has been associated with written and spoken communication, early sociolinguists such as Hymes (1974), Gumperz (1982), Goffman (1981),

¹ <http://www.languageatinternet.org/>

Garfinkel (1967) and others have extended our understanding of discourse causing us to re-consider boundaries between linguistic and non-linguistic phenomenon in social discourse.

As such, the main objects of study in sociolinguist approaches to discourse analysis are *communicative signs and their patterning* (Gumperz, 1982). Dell Hymes argues that it is not language that is fundamental in discourse, but society.

One must take as context a community, or network of persons, investigating its communicative activities as a whole, so that any use of channel and code takes its place as part of the resources upon which the members draw. (Hymes, 1974, p. 4)

Furthermore,

It is [...] not linguistics, but ethnography, not language, but communication, which must provide the frame of reference within which the place of language and culture is to be assessed. (Hymes, 1974, p. 4)

When considering Internet communication, it matters little whether particular interactions are "conversational" — whether responding on a forum, or requesting an article from the New York Times, or "liking" a photo, or requesting an arbitrary web page. What matters is the context; what matters is that **web interaction supports a paradigm of multi-party communication.**

The web is now a very public medium in which interactions between particular persons — or between a user and publisher — occur in a virtual space occupied by many other simultaneous participants. Though web interaction may feel very personal while sitting at ones desk at home, it is as public as conversing in a cocktail party. Moreover, interaction with a particular web site can feel very much as a private dialogue where one meets again and again to continue interacting as if no absence has occurred between sessions.²

² The hidden nature of personalization is a fascinating, related area of interest. In 2011, Eli Pariser wrote a book on the topic entitled *The Filter Bubble: What the Internet is Hiding from You* in which he describes how data mining algorithms are largely responsible for a very personalized experience on the web that influences everything from search results to whose activity is seen on Facebook.

7.3.3 Modeling the Hearer

Though conversation has more generally been described in dyadic (two-party) models, several notable sociolinguists have touched upon multi-party conversation and distinguished between hearer / listener roles (Bell, 1984; H. H. Clark, 1996; H. H. Clark & Carlson, 1982; H. H. Clark & Schaefer, 1987, 1992; Goffman, 1981; Hymes, 1974; Levinson, 1988; Schober & Clark, 1989). Dynel (2010) has attempted to reconcile these various classifications into a unified hearer model. In part, the need to do so stems from the analyses of new forms of discourse, such as radio, television dramas, etc.

Goffman (1981) first divided hearers into two categories: ratified and unratified.³ Ratified hearers are those "entitled to listen to the speaker" whereas unratified are non-official participants, such as bystanders, who are present but not addressed. Ratified participants include the speaker, addressee and other "official hearers" (e.g., "third parties")⁴ who are expected to follow the conversation but who are not addressed. Bell (1984) and H. H. Clark and Schaefer (1992) further distinguish unratified hearers as "bystanders" and "eavesdroppers". The primary distinction between "bystander" and "eavesdropper" is that the *speaker is aware of the presence of bystander and capacity to overhear*. (H. H. Clark & Schaefer, 1992; Dynel, 2010). See Figure 54 for a graphical representation of Dynel's (2011) synthesized model.

The distinction between bystanders and eavesdroppers serves an important distinction with respect to how speakers communicate. When the speaker is aware of a bystander, she can adjust her utterance in a number of ways. In dealing with overhearers, speakers estimate how much the overheard can infer; then design their utterances, accordingly (H. H. Clark & Schaefer, 1992).

With respect to a shared common ground, there are two sorts of information that the speaker must consider. *Open information* is information that overheard believes or could readily guess to be in the common ground. *Closed information* is information that the overheard doesn't believe, and could

³ Levinson (1983), following Fillmore (1971), refers to authorized speakers and authorized recipients.

⁴ Note overlapping terminology with "third party" used in a legal sense. Superficially, these appear to have the same meaning. However, behaviorally, third parties in online tracking act may across a range of hearer / listener roles.

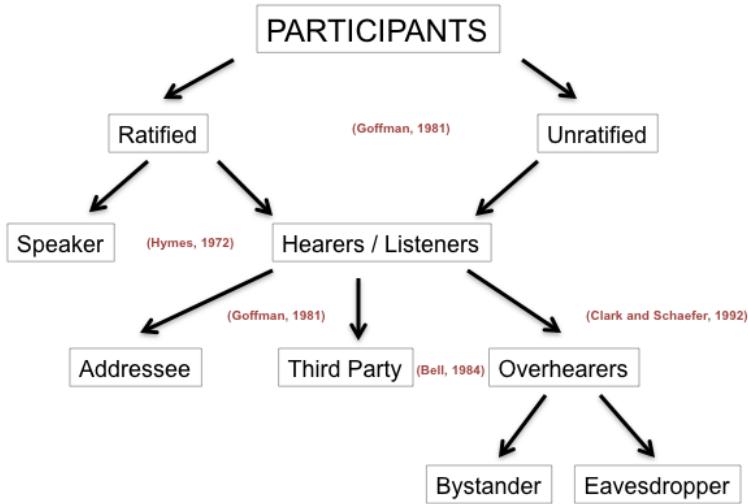


Figure 54: Modeling Hearers in Discourse (Adapted from Dynel, 2010)

not readily guess to be in the common ground (H. H. Clark & Schaefer, 1992). It is the discrepancy between these two sorts of information that may be exploited by the speaker to affect what an overhearer (bystander or eavesdropper) might conjecture (H. H. Clark & Schaefer, 1992).

There are four main attitudes that a speaker can take toward overhearers: **indifference** (the speaker doesn't care whether the overhearer understands what she is saying or not), **disclosure** (the speaker tries to provide enough information that the overhearer can infer the right meaning), **concealment** (the speaker tries to deprive the overhearer of enough information to correctly infer what is meant — e.g., you-know-who did you-know-what to whom), and **disguisement** (the speaker attempts to conceal her meaning from the observer while also deliberately mis-representing that meaning; (H. H. Clark & Schaefer, 1992)).

Thus, **audience design** can be divided roughly into addressee design, third party design, and overhearer design (H. H. Clark & Carlson, 1982). It's worth noting that participant roles are constantly negotiated between interlocutors according to regular (e.g., turn-taking) procedures. And as

Dynel (2010) notes, one role may be performed by a number of individuals simultaneously.

Online tracking represents a sort of discourse where the primary participants are a user interacting with a particular web site — which has a participant role represented by a company or organization. Other hearers are present. Perhaps, ratified and perhaps not. When overhearers are not ratified, their understanding is not grounded. What they understand is conjecture. Sometimes what they may understand may be the same as a ratified hearer, and sometimes not. In an experiment conducted by Schober and Clark (1989), ratified hearers were able to match understanding with the speaker with a very high degree of success. But despite the speaker and hearer deliberately obfuscating their speech in the presence of an overhearer, overhearers were still successful some of the time. More importantly, overhearers were also sometimes wrong.

7.4 AIMS

In this pilot, I consider whether experiment participants view advertisers as bystanders or eavesdroppers. To do this, I compare behavior across two conditions: subjects in the control condition answer a series of sensitive questions. In the treatment conditions, subjects perform the same task but are presented with visible evidence of advertisers as bystanders. This follows the general experiment design of (Acquisti et al., 2012) but poses a new dependent variable for investigation: *visual presence*.

7.5 EXPERIMENTAL DESIGN

This study has one experimental hypothesis:

Hypothesis 3

People answering sensitive questions will behave differently in the presence of unratified participants than when simply notified in advance about the possibility of unratified participants.

In this **between group, single-factor, repeated measures** posttest-only design, the control group is asked to respond to a survey containing questions relating to sensitive activities

(using questions drawn from (Acquisti et al., 2012)). Participants are notified in a privacy statement that there may be ad trackers on the site. The treatment group receives exactly the same notification and survey. However, tracker presence is indicated in a visual display throughout the participant's session.

There is one **independent variable**: "visual presence". The **dependent variable** is the propensity to respond in affirmative to engaging in specific behaviors. This represents the willingness to divulge sensitive information under circumstances which may appear to be more or less private.

Following presentation of stimuli, survey data is collected for both experimental designs.

7.6 METHOD

This experimental design follow the general procedure outlined in Chapter Four.

7.6.1 *Settings and Participants*

Using Amazon's Mechanical Turk, 213 workers were paid \$0.15 to participate in one of two conditions of the single-factor design previously described. Participants were assigned randomly via Qualtrics block randomizer. Results were collected during the period of 7–14 October 2014. The task was expected to take approximately three minutes, though workers were allowed up to 10 minutes to complete.

7.6.2 *Procedure and Materials*

The general procedure follows that described in Chapter Four. [Figure 55](#) is a more detailed graphical depiction of flow.

Following presentation of instructions and consent form, participants were each presented the following six question in the order given.

1. Have you bounced a check?
2. Have you cheated on a tax return?
3. Have you made a false or even somewhat inflated insurance claim?

The screenshot displays the Survey Flow interface for Experiment #3. At the top, there is a "Web Service" section with a URL input field containing "https://pragmatics.firebaseio.com/exp3.json" and a "Test URL" button. Below this is a "Set Embedded Data" section where "result" is mapped to "result". The main content area shows a hierarchical structure of survey blocks:

- Show Block: Instructions** (1 Question)
- Show Block: Consent (Experiment 3)** (2 Questions)
- Randomizer (P1_15)**: Randomly present 1 of the following elements. Options: Evenly Present Elements (checked), Edit Count.
- Randomizer (P1_19)**: Randomly present 1 of the following elements. Options: Evenly Present Elements.
- Show Block: Control** (7 Questions)
- Randomizer (P1_20)**: Randomly present 1 of the following elements. Options: Evenly Present Elements.
- Show Block: Treatment** (7 Questions)
- Show Block: Demographics** (20 Questions)
- Show Block: Privacy Instructions** (1 Question)
- End of Survey**

Each block has "Add Below", "Move", "Duplicate", "Collapse", and "Delete" buttons. There are also "Add a New Element Here" links between some blocks.

Figure 55: Experiment 3 Flow

4. While an adult, have you had sexual desires for a minor?
5. Have you had sex with the current husband, wife, or partner of a friend?
6. Have you fantasized about having violent, non-consensual sex with someone?

Possible answers were:

1. Never
2. Once or twice
3. Sometimes
4. Frequently

As in [Acquisti et al. \(2012\)](#), propensity to disclose is indicated by a positive answer. All positive answers are combined into a single score. Non-answers (skipped questions) count as non-disclosure. Thus, the dependent variable is binary.

7.6.3 *Instrumentation*

In addition to features provided by AMT and Qualtrics, my scripts performed the following:

1. Disable preview using a custom CSS class blocking Qualtrics controls
2. Check worker hash against a FireBase web service and request worker to return HIT if the hash is in the exclusion list
3. Add worker hash to FireBase
4. Submit results to AMT on completion of the Qualtrics survey

No other instrumentation was required for this experimental design.

7.7 DATA COLLECTED

Using a G* Power 3 f-test for linear multiple regression with a fixed model, I estimated a sample size of approximately 89 was necessary in order to detect a medium effect with power of 0.95. All surveys initiated were completed: there were no known drop-outs.

Once all assignments had been completed, (as indicated on my AMT requester dashboard), I downloaded data as a single CSV (comma separated values) file from the Qualtrics website. Data was organized such that each participant's data is on its own row.

7.8 RESULTS

There is some difference in the demographic profile between the Experiment 3 sample and the [Acquisti et al. \(2012\)](#) sample. That study reported a mean age of 40, 45% male, and 82% caucasian. This study was 63% male, 82% caucasian, 68% under 35, and, likely, with lower income levels (given NYT has a subscription fee for regular readers): 53% reportedly make below 30K. Regardless, comparing control group responses from [Acquisti et al. \(2012\)](#) with control group responses here, propensity to disclose was very similar.

Interestingly, the percentage of male respondents for Experiment 3 was higher than the aggregate population studied across all experiments in this dissertation (63% versus 53%; see [Appendix D](#)). This suggests that some population of workers, including females, elected not to participate in this study. I will come back to this point in a bit.

[Table 10](#) below presents averages across all six questions. A simple chi-square test reveals no difference between groups ($p = 1$).⁵

There are two factors that may account for observed differences between this study and [Acquisti et al. \(2012\)](#). First, the population sampled here is surprisingly sensitive to privacy issues. Moreover, AMT as a platform is not seen as anonymous — as the [Acquisti et al. \(2012\)](#) survey may have been perceived. The mere fact that I presented an IRB form with names and contact information affected discourse context.

⁵ This is with and without a bonferroni adjustment for repeated questions ($p = .046$)

Table 10: Raw Results (Aggregate Mean)

	Control	Treatment
Never	41.33	42.17
Once or twice	6.83	7.00
Sometimes	2.33	1.33
Frequently	.50	.50

Second, answering specific questions in the [Acquisti et al. \(2012\)](#) survey was completely voluntary; ‘no answer’ accounted for between roughly 10–15% for each question asked. This was calculated in the metric for propensity to disclose.

Selectional differences in this study are reflected in a couple of ways.

- **Population Bias.** There is no way to know if the people who elected not to participate in Experiment 3 did so because they did not want to see “sensitive” questions or because they were concerned with privacy. Regardless, there is a population bias — as reflected by the relative gender disparity between Experiment 3 and aggregate demographics across all five studies.
- **Obligation.** No one in this experiment tried *not* to answer a question. This may relate to the nature of the obligation workers feel to requesters as paid subjects.

Over the course of the past two years, tremendous change has occurred on the Internet. Savvy users, like those found on AMT, are well aware that what they say and do may be monitored. One of my participants noted in comment, “nothing is private online now due to the NSA.” Events have conspired to raise the saliency of privacy issues on the Internet for even casual Internet users. Despite this, my survey reveals perceptual differences for feelings of privacy on the Internet whether connected at home or in a public setting. [Figure 56](#) represents answers to the following two questions:

1. How private do you feel on the Internet when you are using your computer (laptop, or tablet) in public. Public means something like a coffee shop, library, school, etc.
2. How private do you feel on the Internet when you are using your computer (laptop, or tablet) at home?

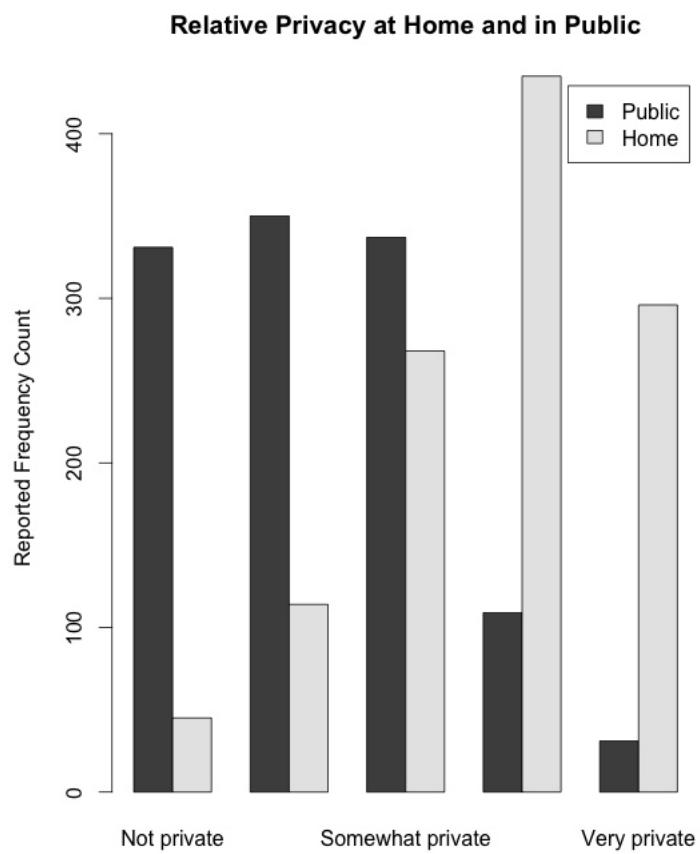


Figure 56: Feelings of Privacy on the Internet in Public and at Home

It's clear that physical setting, at least, lends some expectation of privacy whether one is on the Internet on a personal machine at home or in public. Though visual saliency was not a factor in propensity to disclose in this experiment, users have an expectation of privacy when they are not in a public setting. Given the Internet is now "public", there exists some mis-match of expectation that is not accounted for in Internet browser design.

8

DISCUSSION

This thesis presents the argument that discourse understanding 1) shapes how people understand and interact with graphical user interfaces on the Internet; and 2) how such knowledge might be used to manipulate context, affecting comprehension and, therefore, behavior.

By analyzing GUI microinteractions in the way that we analyze linguistic discourse, it may be possible to explain some sorts of mis-understandings and even present solutions for improving user interaction design. This chapter summarizes results of experiments from the previous three chapters in context of related work in cognitive science.

While qualitative methods are extremely important for guiding a research program, it is through quantitative methods that causality is determined. While Experiments 1 and 2 focused on user comprehension, Experiment 3 concerned how knowledge about the presence of third-parties affects production. The third experiment was problematic in several respects.

8.1 SUMMARY OF FINDINGS

Table 11 summarizes results for all experiments in this dissertation.

Table 11: Summary Results

1A Design	2x2x2 between groups design where the control group is presented a set of textual expressions and asked to answer questions about their meaning. Treatment groups are presented with textual expressions or a dialog box expressing the same set of choices and asked the same questions. Confounding factors include both a privacy (knowledge) bias and also deontic force.
-----------	---

8.1 SUMMARY OF FINDINGS

1A Results	Significant difference between text and graphics for implicature ($p<.01$). Participants are even more likely to interpret an implicature in the graphical conditions than in the textual condition. Also, a significant interaction in the condition where deontic force accompanies "pictures" ($p<.05$). When there is no privacy bias (less knowledge), the conditional implicature has larger effect.
1B Design	Between groups design where two groups are presented with a cookie banner and later asked about whether or not they believe the website placed "cookies" in their browser. The treatment group is presented with feedback about the consequence of their action (or non-action) following presentation of the banner.
1B Results	Significant difference in understanding with presentation of feedback ($p<.001$). With no feedback, the odds of implicature are 6.6 times more likely.
2A Design	Between groups design where five groups are presented an advertisement in the context of a webpage and asked to identify elements with hyperlinks. Treatment groups are presented an advertisement with embedded icons at four levels (known icon different company, unknown icon, known CTA, DAA opt-out) while the control group is presented with an embedded image from the same company as the advertiser.
2A Results	Significant difference only for the FaceBook icon. At this sample size, no difference detected between AdChoices icon and known and unknown icons. Possibly, the sample size was too small for detection of a difference. A much larger effect was seen in the pre-pilot where advertisements were presented in isolation (not embedded in news page context).
2B Design	Between groups design where three groups are presented a digital news page and asked to click on a specific advertisement. Treatment groups are presented an advertisement at two levels (iconic button, textual CTA) while the control group is presented with an advertisement containing no CTA.
2B Results	No difference between groups. The presence of an iconic button does not affect where a user targets a click. Possibly, the effect size is too small to see.
3 Design	Between groups repeated measures design where a control group is asked to respond to a survey containing sensitive questions. The treatment group is presented a visual indicator of ad tracker presence during presentation of each question.

3 Results	No difference between groups. It does not appear that a visual presence indicator for advertiser activity had any affect on user behavior with regard to propensity to disclose sensitive, personal information. It is likely that participants in both groups are equally aware of "bystanders". The population on AMT was not an ideal population for this study due to very high privacy awareness and properties of AMT, in general.
-----------	--

8.2 DISCUSSION OF RESULTS

Pragmatic theory distinguishes between several types of factors affecting comprehension and reasoning:

1. *Social dimensions* of discourse (e.g., cooperative principle; intent; norms; participant roles)
2. *Knowledge* (e.g., subject /content matter; deictic relations)
3. *Saliency* (e.g., cognitive saliency of referents; task saliency)

Attitudes and beliefs also affect reasoning as described by Kahneman and Tversky (1984), though these do not fall under the purview of pragmatic theory.

Experiment 1 concerned pragmatic reasoning where subjects reasoned over both textual and mixed-modal content. Not only were mixed-modal representations subject to the same sorts of pragmatic inferences as purely textual representations, small changes in language and content were shown to have significant effect in understanding and, ultimately, behavior. Furthermore, immediate feedback was seen to affect comprehension.

Experiment 2 primarily considered the role of knowledge (icon referent) in the inference of an indexical relation. By manipulating an icon embedded in a graphical ad, I looked at whether knowledge about that icon (against a background context) played a role in its interpretation. In the pilot, it seemed this was the case, but in the actual experiment, any effect was too small to see. The primary difference was that in Experiments 2A and 2B, the ad image was embedded in the larger context of a news article while in the pre-pilot, the ad was considered in isolation.

Finally, Experiment 3 concerned whether manipulation of visual presence (and, therefore, participant role) would affect responses in a task measuring propensity to disclose sensitive information. In theory, this should be the case. However, I believe elements of task context (including use of AMT) and user knowledge may have affected results.

Below I discuss each experiment in more detail.

8.2.1 Experiment 1: Modal Dialog Boxes

Assuming a common set of principles driving understanding of both multimodal and textual content, people should be as sensitive to implicature in multimodal messages as linguistic messages. In fact, it appears not only is this so, but to a higher degree. This has clear implication for design: *user choices made using non-forced choice modal dialog boxes may be mis-understood.*

Why might people mis-understand the meaning and consequence of choice?

Evans (2003) asserts evidence for two types of thought processes: *heuristic* and *analytic* (also referred to as System 1 and System 2, as well as, Type 1 and Type 2 in Kahneman, 2011; Manktelow, 2012, respectively). Much of this work is based on studies of deductive reasoning. For example, the Wason selection task (Figure 57; Wason, 1960) asked people to select cards which falsify the statement in (a). Subjects were shown 4 cards and told that each card had a letter on one side and a number on the other. The correct answer is A and 7, though only 10–20% of people answered correctly (Wason, 1960). However, when given task (b) and told they play the role of a police officer checking people drinking in a bar, about 75% were successful (Griggs & Cox, 1982).

These studies and others, point to a belief bias where belief is seen to affect reasoning about logical arguments. Similarly, a matching bias is seen in more abstract tasks; a matching bias is one in which one tends to select answers which contain lexical content which matches content over which one is reasoning (e.g., selection of A and 3 as answers in the Wason task mentioned earlier). Such biases provide support for a heuristic, autonomous system for reasoning which takes into account knowledge and belief. Other processes relating to the social dimensions of language use, also appear to be automatic (discussed in the next section).

There are several theories of reasoning that aim to explain this sort of phenomenon: P. Cheng and Holyoak (1985) propose domain specific knowledge structures (pragmatic reasoning schemes) which are remembered and applied from previous experience to new situations; (Johnson-Laird, 1983) argues that people create mental models of assertions and rely on these to reason; and, (Cosmides, 1989) proposes a “social contract” algorithm which is specialized to assess costs and benefits in social exchange.

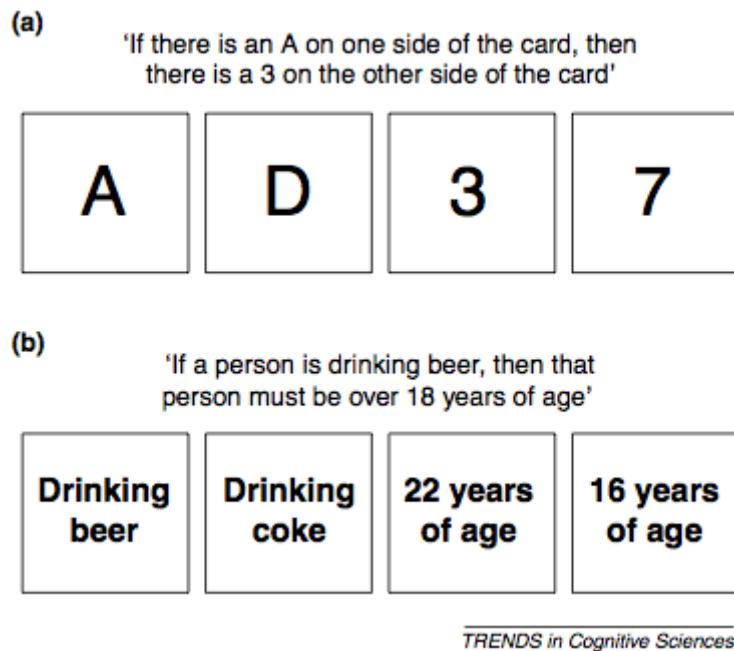


Figure 57: Wason Selection Task (Image credit: [Evans, 2003](#))

Whether through discourse models, pragmatic reasoning schemas (PRS), or other heuristics, people are adept at reasoning over discourse with little apparent effort or thought. In all conditions, subjects were subject to implicature (see [Appendix F](#) for raw results). What is surprising is how sensitive implicature is to small changes context. Across eight conditions, interpretation of implicature varied from 38% (text, pictures, on/off) to 84% (graphics, cookies, on/off).

Experiment 1B was concerned with how people interpret the question presented in Experiment 1A — but in the context of decision-making. In Experiment 1A, subjects were asked what they thought ‘cancel’ or ‘x’ meant. They had time to reflect and reason.

In Experiment 1B, subjects were asked to think about what they thought their choice meant immediately after making a decision. Surprisingly, a great number of subjects attended to the dialog box and chose to respond (30–35%). This is likely a higher percentage than what might be expected of the population of users on the Internet as a whole. Only 22% of those in the control condition made an inference about the presence of cookies based on their choice. Of those that didn’t, 45% of those that thought there were cookies answered “there are probably always cookies” and 72% of those who didn’t think there were cookies said they used an ad blocker. The large number of turkers who believed there were always

cookies suggests a belief bias. And the number that used ad blockers strengthens this: turkers are very knowledgeable of privacy issues.

Of those that did interpret an implicature, it's possible a matching bias was in effect. It is also possible subjects shortcut reasoning using a simple schema or model. Either way, subjects may have reasoned:

- "I didn't select NO and I didn't select YES, therefore neither NO nor YES."
- Or, perhaps, "I didn't select NO, therefore YES."
- Or, similarly, "I didn't select YES, therefore NO."

Why one would select one or the other of the latter two is a good question. Either way, they made a choice without understanding the consequence.

The number of people making an inference about the presence or absence of cookies based on choice was lower than expected. The high "participation" for choice seems likely due to the nature of the task environment: workers may be likely to cooperate and take action on prompts because they have agreed to do work for pay.

Regardless, there were far fewer inferences made when feedback was provided. *Without immediate feedback, odds were 6.6 times higher that subjects understood an implicature to be true.*

8.2.2 Experiment 2: Embedded Icons in Graphical Ads

Experiment 2 was also about pragmatic reasoning, though not the sort associated with logical or semantic reasoning. In the pre-pilot to Experiment 2, participants examined advertisements and were asked to click wherever they thought there might be a hyperlink. Hyperlinks associate an object on a page or in a scene with other content. When there are no visual cues, such a relation largely derives from knowledge people have about the meaning of an object and to what it might refer. Even so, judgements seems to be context sensitive.

In Experiments 2A and 2B, participants were given a task in context — advertisements were embedded in a news page. From Experiment 2A, it seems that an icon embedded in a graphical advertisement is not likely to be considered indexical, regardless of what that icon represents — except in the case of the FaceBook icon which is an explicit "call-to-action"

(like /share) associated with all sorts of different content on the Internet. I had hypothesized that several factors might be at play: 1) icon salience by a contrastive relation (visually /semantically) with the background ad; 2) knowledge that the icon is indexical to some brand; and, 3) recognition of a signal or “call to action” for clicking. In 2A, the only icon which was considered indexical was the FaceBook icon, exhibiting all three properties. It is possible that estimating for a medium effect size for an ad in context, was not enough to see whether (1) and (2) made a difference, independently.

Experiment 2B attempted to determine whether explicit indexicality affected behavior. Would an ad with a CTA that had the form of a button invite a user to click directly on the button? In the pre-pilot, this appeared to be the case, but there was no effect in the embedded context of Experiment 2B. Again, it is possible that the embedding context (ad within a news page) had an effect, potentially reducing the effect size. It is also possible that we may see no effect with more savvy Internet users but would see one with those who are less savvy.

Regardless, these insights were useful in understanding why the AdChoices icon may not be effective for conveying intent to users. Size has less of an effect than one would imagine. More important, *embedded graphics in ads may not be considered indexical by default*. At least not today. As users become accustomed to certain associations and patterns, they may well adapt.

8.2.3 *Experiment 3: Third-Party Bystanders Versus Eavesdroppers*

Experiment 3 attempted to address whether participants might change their behavior depending on whether they thought advertisers were actively observing them (i.e., the difference between non-ratified bystanders versus non-ratified, eavesdroppers). If the treatment group behaved differently, this would lend support to the notion that participants aware of being monitored, changed their behavior in response. However, it appears that both groups were equally cognizant of the potential for third-party monitoring.

Two years ago, turkers, along with most people on the Internet, had little knowledge of OBA. In a recent survey conducted by Lease et al. (2013), at least some AMT workers were aware of the potential for leakage of personal identity via workerId: worker identity is not completely private and can be found in web searches. The survey in this dissertation

revealed that more than 90% of turkers knew what cookies were, and what's more, nearly 75% used ad blockers. Though most of them felt "more private" on the Internet at home than when sitting in a public setting, they seem to have a heightened sensitivity toward privacy issues which may have affected the outcome of this experiment.

Video gaming site owner, Niero Gonzalez was surprised to find that nearly half of his 3 million visitors a month blocked ads (Hill, 2013). From the same source, PageFair says that ad blocking is growing at a rate of 43% per year. Given how knowledgeable my sample was about Internet privacy and tracking technology, few would consider their communications anonymous. Moreover, given the contractual relation between the requester and worker on AMT, only those willing to share sensitive information would be willing to accept such a HIT. Quite possibly, this rendered the question of bystanders versus eavesdroppers irrelevant.

Blase Ur (2012) deliberately recruited non-technical people who were little aware of tracking on the Internet. By contrast, turkers are very aware of tracking. Perhaps, results would have been different two years ago. In Experiment 3, the sample studied was likely not representative of the larger population of Internet users. From this perspective, AMT was not a suitable recruitment platform. However, turkers may reflect what we will see from the majority of users in time. And, good or bad, what feels creepy now, may not after longer familiarity.

8.2.4 *Use of AMT*

Only recently have researchers in pragmatics and sociolinguistics adopted controlled experimentation (and on platforms such as Mechanical Turk) for the study of discourse phenomena. Anand and Andrews (2011) note that pragmatic phenomena are difficult to study because of the many contextual parameters that affect understanding. Crowdsourcing platforms such as AMT now make it possible to vary context parametrically without excessive expense and time. In a study of scalar implicature, Anand and Andrews (2011) found that simply changing the task instructions to "quality control" altered the proportion of responses in an implicature-oriented situation. Though such methods look useful for testing UI design comprehension, there is much to learn about how to best do this.

8.2.5 Potential Population Bias

In both methods and presentation of experiment results, I've noted attributes or characteristics of the AMT population that makes it difficult to typify them as characteristic, non-privacy aware Internet users. Other potential issues, regarding use in cognitive and other experiments were also noted. Like other researchers, I performed tests varying payment looking for quality differences across collection events. However, differences between results in Experiment 2 pre-pilot and Experiments 1 & 2 suggest that there may be another factor.

Rush et al. (1978) noted that subjects that volunteer for pay tend to commit more errors of omission on selective attention tasks. They suggested that paid subjects might working more for task completion than success. In experiment 2B, I offered a 3-minute task for 15 cents. Even when I added collection events increasing payment to 25 cents, and shortening the task by omitting demographic questions, I saw no difference in results. Yet, corresponding samples from the pre-pilot at 75 cents (15–20 minute task) obtained different results. It also possible some other variable might be in play. Perhaps, some workers (e.g., those focused on pay) simply prefer more attractive tasks (easiest; highest reward to time ratio). These tasks are consumed quickly leaving “less attractive” tasks for other turkers to tackle. This question remains future work.

8.3 COGNITIVE ARCHITECTURE

As central goal of this dissertation is to provide support for the notion that user interaction with graphical user interfaces involves discourse processes, a basic discussion of the cognitive architecture facilitating language understanding is essential.

8.3.1 Situation Models

Language, perception, and action are closely aligned (Lakoff, 2008). In recent years there has been much support from brain imaging. For example,

- Activity in pre-motor cortex suggests a causal link for the understanding of action verbs (Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011);
- High manipulability words evoke greater activation in the motor cortex (Madan & Singhal, 2012);

- Perspective in pictures has an effect of learning (de Nooijer, van Gog, Paas, & Zwaan, 2013);
- Evidence of the activation of modality specific representations during discourse show that mental imagery occurs during discourse comprehension.(Kurby & Zacks, 2013)

Furthermore, theories of language comprehension suggest that text comprehension involves the constructions of a mental representation of events described in text (Kintsch & van Dijk, 1978; van Dijk, 1990; Zwaan & Radvansky, 2002).

In semantics, a propositional representation conveys logical meaning using lexical and grammatical features. Depictive representations, on the other hand, do not encode such relations symbolically. Propositional and depictive representations make different sorts of information explicit and accessible (Kosslyn, Thompson, & Ganis, 2006).

Kosslyn et al. (2006) make the argument that visual imagery evokes many of the same processing mechanisms in visual perception. While sensory visual perception is driven by sensory input, visual perception makes use of stored information.

Mental representations form the building block of situation (viz. discourse) models (Zwaan & Radvansky, 2002, 2013). Zwaan and Radvansky (2013) is careful to distinguish these from the notion of schemas such as learned stereotypical “scripts” (e.g., greeting rituals) in interaction. Individual elements of a script have mental representations, and schemata themselves may be used in the construction of models.

Discourse models are necessary for explaining language processing. Discourse understanding requires readers to integrate and recall information across spans of sentences. Moreover, discourse models account for how representations are constructed across modalities. Gernsbacher, Varner, and Faust (2013); Zwaan (1999) found a correlation in comprehension across three modalities (written, auditory, and visual) suggesting that mental representations are independent of modality and used in the integration of verbal and visual information.

8.3.2 Interactive Alignment and Priming

In a theory of *interactive alignment*, Pickering and Garrod (2003) attempt to account for dialogue as the more basic framework for language comprehension and production over

monologue. They argued that linguistic representations become aligned at many levels, as a result of an automatic processes. This is accomplished by “dialogue routines” which simplify language production and comprehension by short-circuiting decision-making processes.

Essential to successful dialogue, is alignment of situation models. Doing so, eliminates the need for speakers to actively maintain a listener model. [Pickering and Garrod \(2003\)](#) give evidence of this as alignment in coordinating referring expressions in dialogue through a model of common ground ([Brennan & Clark, 1996](#); [H. H. Clark & Wilkes-Gibbs, 1986b](#); [Wilkes-Gibbs & Clark, 1992](#)). They also give examples of alignment across other levels of abstraction: articulation ([Bard, Anderson, Sotillo, Aylett, & Newlands, 2000](#)), syntax ([Brani-gan, Pickering, McLean, & Cleland, 2007](#)), and comprehension.

Though the goal of monologue (spoken or written) is not aligned representations, readers draw inferences on the basis of knowledge of the writer — and the writer must infer what the listener has inferred. Only through interaction, do processes of feedback help interlocutors recognize and correct mis-understanding ([H. H. Clark, 1996](#)).

Citing [Levelt \(1993\)](#) and [Dell \(1986\)](#), [Pickering and Garrod \(2003\)](#) point out that it is not only advantageous to avoid levels of representation in comprehension but also in production. Noting the very repetitive nature of dialogue, they discuss the implication of “routines” in alignment processes. Through shared routines (repetitive expressions, lexical items, etc.) language production is eased by the prior activation of relevant lexical and syntactic representations. A prediction of their account is that word frequency effects are reduced by accessibility in dialogue contexts. That is, less frequent meaning becomes more accessible than more frequent meaning in the local context of dialogue. Such affects are facilitated through priming ([Bargh, 2006](#)). “Language perceivers implicitly learn the statistical regularities in their linguistic input, and they use this prior experience to guide comprehension of subsequent language” ([MacDonald, 2013](#)).

This is relevant not only to the acquisition of lexical expressions, but also the acquisition of graphical symbols. “Through the grounding process, interactive graphical communication enables participants to develop symbolic representations from what started out as primarily iconic representations” ([Garrod, Fay, Lee, Oberlander, & MacLeod, 2007](#), p. 4). Experiments 1 and 2 of this dissertation concern mental representations

derived from both linguistic and graphical elements. Moreover, composite components such as modal dialog boxes and graphical advertisements have become routinized by “fluent” Internet users. As (Bargh, 2006) notes,

... the study of language comprehension and production has provided social cognition with highly useful models that have enabled us, over the years, to discover ‘new’ and important social psychological phenomena. Given this stellar track record it might be the case that the underlying mechanisms of language production and of social behavior production are one and the same. (Bargh, 2006, p. 16)

And, particularly relevant to new work presented here:

"One might speculate that language pragmatics on the one hand and goal directed activity in social contexts on the other are derived from a common set of rules whose aim is to enable members of complex social organizations to interact." (Girotto & Politzer, 1990, p. 105)

8.4 THEMES FOR FUTURE AREAS OF STUDY

Each of the experiments described here present opportunities for future, interesting follow-up studies. However, I will present them in the context of several broad themes.

8.4.1 *Micro-Interactional Design Patterns*

Interacting with graphical user interfaces (GUIs) sometimes feels conversational and sometimes not. A dialogue box that asks a yes-no question doesn’t seem that much different from a verbal yes-no question. But file selection prompts don’t feel as natural.

When we speak in our own language about everyday matters, we speak effortlessly with little thought about how to produce an utterance and little thought about how to understand one. It’s a bit different with written language. We endure years at school learning to read and write properly, following style guides which we memorize and practice endlessly until we know how to recognize and avoid passive constructions — as well as generate them with practiced ease. Reading and writing text is just not as intuitive as speaking

conversationally and most people require years of practice to achieve any competency.

When GUIs sprang into public awareness in the 70s and 80s, it became possible for people to interact with a computer without having to learn seemingly arbitrary complex commands typed into a terminal window. GUIs made sense. You could point and push buttons. An icon was a symbol that stood for something meaningful like a “program”. And there was plain old English language mixed in with icons and graphics (since GUIs originated in the United States). GUIs felt natural.

During all of the hype of GUI-based interfaces, software developers learned to adopt the notion of software patterns. The idea was to document shared knowledge of best practices for solving common or typical problems in software design. A software pattern is a fantastic concept for anyone learning to write software. Writing software and designing UIs has much in common with learning how to read and write text. It requires a lot of practice before you become any good at it. Software design patterns help new programmers speed up the process so they don’t have to figure out how to solve every problem encountered through trial-and-error.

When desktop computers became really popular — and web interaction more so in the 2000s, some software engineers began to specialize in “front-end” development. Anyone learning web interaction might consult the Yahoo UI design pattern library while developing a new web application. In fact, Yahoo’s purpose for its design pattern library was to solve a business problem. They wanted a way to communicate standards across development teams in order to increase “consistency, predictability, and usability” across their site — and their brand (Leacock, Malone, & Wheeler, 2005).

Human Computer Interface (HCI) guidelines and design patterns are intended to capture specific problems, examples, usage, rationales, supporting research, standards, etc. Patterns range across stylistic conventions (e.g., page headers and footers), attentional mechanisms such as animation, navigation and organization, layout, common functions such as registration or login, and even more complex patterns such as social sharing and feedback.

In addition to useful, everyday patterns, the notion of *anti-patterns* and *dark patterns*¹ document practice in common use which may be ineffective (anti-patterns) or ethically questionable (dark patterns) patterns.

¹ <http://darkpatterns.org>



Figure 58: Cat in the Hat

When we learn language, some concepts and patterns become **entrenched** with frequent use. So, for example, the phrase “I don’t know” is not something most of us have to think about before uttering. The grammatical construction I (subject) + do (1st person aux verb) + negation + know (infinitive) is not something you have to think about assembling before you say it. Such forms become **routinized** through repeated activation and use — such that both *production* and *understanding* require less cognitive effort and is more automated (Pickering & Garrod, 2003).

This sort of automatization is not limited to conversational speech. Routinization happens at many levels of production and understanding. For example, when you see a familiar word like “key”, pronunciation is automatic. It also interacts with syntactic chunks such that when someone says, “cat in the _”, you anticipate “hat” (Pickering & Garrod, 2003).

When you see a tall hat with red and white stripes as in [Figure 58](#), you may immediately think cat-in-the-hat, as well as “cat”, “Dr. Seuss”, and any number of related concepts. According to (Pickering & Garrod, 2003), **priming** may occur at different levels including lexical, syntactic, semantic, and situation.

Routinization isn’t limited to long-term memory. Suppose you are having a conversation with a friend. You are talking about a movie of which neither of you can remember the name. So you say, “that movie with Harrison Ford”. It would not be surprising if your friend then referred to the same movie as, “the Harrison Ford movie”. You can routinize a reference to something during the course of an interaction in order to communicate more easily.

Priming is also applicable to GUI patterns. When the GUI presents a dialog box like the one below, it is familiar. It offers a choice (1) or (2). Typically, the choice is binary — cancel or

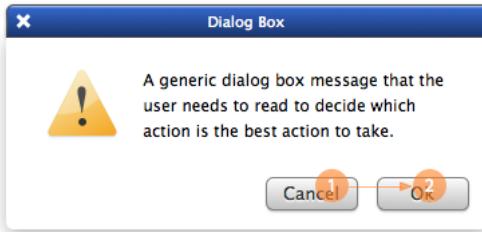


Figure 59: Image credit: <http://uxmovement.com/buttons/why-ok-buttons-in-dialog-boxes-work-best-on-the-right/>



Figure 60: One Choice and Ignore?

accept; yes or no; permit or deny, etc. You read the text and make a choice.

Once you've seen [Figure 59](#), you don't have to ponder over similar dialog boxes each time you encounter one. In fact, the more often you see and recognize a design pattern, the more entrenched it becomes.

What is the difference between an interaction design pattern and linguistic pattern? Well one obvious difference is while we use language all day long, only a few of us know how to produce UIs. And when people interact with UIs, they aren't interacting directly with the designer. So the designer doesn't get direct (or continuous) feedback on how well the user understands the interaction. Production is not directly linked to comprehension in a realtime feedback loop like face-to-face dialog; it is a bit more like interacting with monologue text. This means that patterns are not aligned and refined in the same way.

Interaction design patterns help improve communication, but because they are easily modified by the designer with no direct feedback about how such changes affect a user's comprehension — there is the possibility for non-obvious error.

In fact, there might be a lot of information packed into a dialog box. Dialog boxes don't have to be simple binary choices. The UI designer can make a dialog box for any purpose. [Figure 60](#) illustrates a very simple one where explicit choice is omitted.

[Figure 61](#) illustrates an example of where the designer decided that a form could be a dialog box. What happens if I don't fill something out right?



Figure 61: Dialog Box or Form?



Figure 62: Select File Dialog

[Figure 62](#) is the familiar open document dialog. Sometimes you can open more than one document but there is no way to know without trying it.

[Figure 63](#) is complex dialog that combines the basic cancel, accept pattern with other buttons and choices. From experience, I expect that I can do a bunch of things and then choose “OK” or “Cancel” when I’m done. But it requires a bit more thought on my part — and also a bit of trust where if I spend a lot of time on this and then mess something up, I don’t know what will happen.

[Figure 64](#) is an example of a dialog box where I worry that if I click on the hyperlink that I don’t know if I’m still in this dialog or I’m sent off on a wild goose chase.

Is it more confusing if I move the “cancel” button in [Figure 65](#)? Will the user even see the “cancel” button? Will they become confused because the design conflicts with expectation?

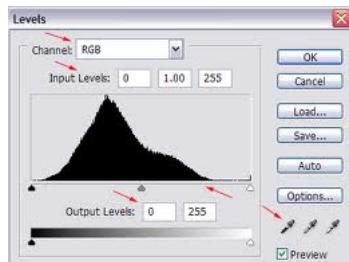


Figure 63: Complex Dialog Box

8.4 THEMES FOR FUTURE AREAS OF STUDY



Figure 64: Dialog Box with Hyperlink

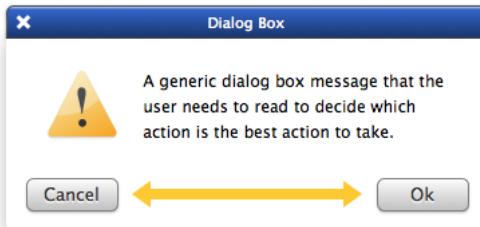


Figure 65: Button Layout

Considering the examples above, it's easy to see how UI and software designers are able to easily break such a design pattern by:

- Packaging the information differently (adding more to a screen or component, for example)
- Altering text on labels
- Altering the position of a component so it is not in a familiar position
- Creating a semantic mis-match between text and button labels

In a study of scalar implicature, [Geurts and Pouscoulous \(2009\)](#) compared an inference task with a verification task and found that there was a positive response bias in the inference task not seen in the verification task. Under what situations might a modal dialog be better stated as a verification dialog than an interrogative?² One direct follow-up to Experiment 1 would be a follow-up to see if a verification dialog box reduces unwanted implicature.

Common design patterns are easily broken. Any designer or developer with a text editor can easily make changes without understanding the impact on comprehension. Producing comprehensible GUIs requires as much practice as writing

² However, ([Clifton & Dube, 2010](#)) argue that the [Geurts and Pouscoulous \(2009\)](#) task (and corresponding interpretation of results) is flawed.

clear, well-structured text. Design patterns arguably serve as a cognitive aid, boosting mechanisms supporting routinization and automatization of understanding. But as designers we need to be sensitive to the effect of alteration, no matter how benign the change might seem. We learn such sensitivity writing prose. It is time to take a hard look at micro-interactional design patterns and question what we think we know.

8.4.2 *Interactive Ads*

There is also ample room to continue exploration of deixis in graphical advertisements, as well. A next step would be to design follow-ups where advertisements are presented in isolation, in a manner similar to the pre-pilot for Experiment 2. Though Flash-based advertisements are becoming more rare, HTML5 is sufficiently richly interactive to expect ads will become increasingly interactive over time. And as we saw in Chapter 3, cross-media interactive advertising is already upon us.

How will such advertisements be perceived by users over time? Goldstein, McAfee, and Suri (2013) used AMT to study the cost of “annoying ads” on subjects. Prior research indicates, while there may be beneficial effects to small amounts of animation, too much may have a detrimental impact on ad effectiveness (Buscher, Dumais, & Cutrell, 2010; Dreze & Hussherr, 2003; Goldfarb & Tucker, 2011). Goldstein et al. (2013) studied the effect of annoying ads in an email sorting task. They believe that there may be a real cost beyond annoyance.

A key observation made in Experiment 2 was that embedded icons did not trigger inference of indexicality while in the context of an online news article. “Pointing” through animation might be effective for getting a user’s attention, but this needs to be presented in such a way to avoid “annoyance”. Regardless, micro-interactive graphics, including advertisements, are a rich new area for study. As hinted in Chapter 3, the rise of both behavioral advertising and interactive advertising in tandem will likely lead to new way to persuade and manipulate.

8.4.3 *Display Adaptation*

Technology outpaces our ability to adapt standards for interaction. If the AdChoices icon was difficult to see and under-



Figure 66: AdChoices Icon in iPhone Browser

stand in a desktop browser, what does it look like in a mobile browser?

Try as I might — I could not select the AdChoices icon in [Figure 66](#) nor was there an obvious way to expand the advertisements size.

What about the iconic buttons advertisers so like? Do they look out of place on the display in [Figure 67](#)?

Finally, its important to note that visual cues that we typically rely on for signaling an element on the page is interactive are beginning to disappear from touch-interactive displays. [Figure 68](#) is a page from the Washington Post application on the Apple iPad.

Nowhere do you see blue text indicating textual hyperlinks. Once you dive into an article, such links are present. But presentation patterns in mobile devices are shifting expectations over time.

8.4.4 Social Microinteractions

Experiment 3 suggests two avenues of future research. First, while it may not be possible to study the effects of eavesdropping in AMT, studying how people intentionally tailor messages for overhearers seems practicable. For example, [Egelman et al. \(2009\)](#) found that users adapted their purchasing behavior to the presence of privacy indicators when purchasing privacy-sensitive items.

8.4 THEMES FOR FUTURE AREAS OF STUDY

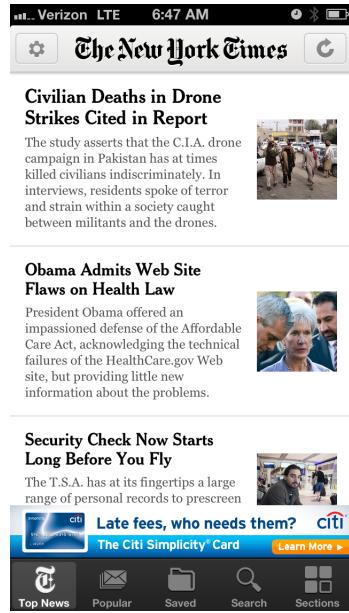


Figure 67: Iconic Button in iPhone Browser

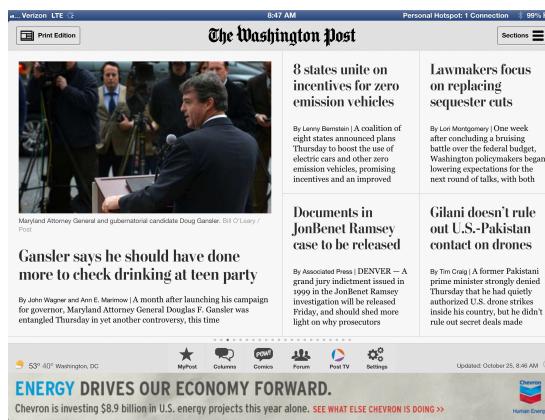


Figure 68: Washington Post on the iPad

What if consumers were given the choice to receive micro-payments in return for sharing information about what they purchased? What if, for example, a consumer, while in an Amazon shopping cart, had the ability to choose from three levels of micro-payment in return for sharing information about purchases. The smallest micropayment might be used to advertise only that he shopped for a “drugstore item” on Amazon. A second level might convey that he shopped for a “hair product for men”. Finally, a third higher-priced level might convey brand such as “Touch of Grey”. Would that consumer adjust levels on the basis of privacy-sensitivity and price?

This is an example of social sharing not yet seen on the Internet. Yet it is a plausible interaction. People are becoming more accustomed to tailoring messages to audiences on FaceBook, Google Plus and other social platforms. The challenge is that, while we may think this gives us finer-grained control over privacy, this may not be true. [Stutzman, Capra, and Thompson \(2010\)](#) found that while Facebook users have over time exhibited increasing privacy-seeking behavior (decreasing the amount of personal data shared publicly), the amount and scope of personal information they have revealed privately to other connected profiles has increased over time. And, as a result, so have disclosers to “silent listeners” — third-party apps, and advertisers. While the impression is that users have greater and finer-grained control over privacy, the reality is that much of this is illusion.

A second potential avenue for increasing social control over privacy is discussed below. Potentially, more transformative, is the need to redesign browser technology to reflect the nature of Internet browsing as inherently social — that is, while it may be evident we are “conversing” with a particular website, third party participants may be unseen and present. Embracing a physical metaphor of spatial awareness may engender new styles of interaction for users seeking greater control over privacy.

8.4.5 *Socially Aware Browsers*

Technology has made it possible to engage in large-scale surveillance in the Internet and piece together little bits of behavioral data in the browser in order to create fairly ac-

curate consumer profiles.³ If one party (the publisher) has given consent for monitoring, U.S. law does not consider this as eavesdropping. Nonetheless, users find such monitoring “creepy” and invasive. When ads follow us across sites, it feels like eavesdropping. While browsing the Internet may feel anonymous and private, it is not. Despite lukewarm attempts by the advertising industry to self-regulate, the burden is on the user rather than the publishers and advertisers to opt-out of tracking. Furthermore, in the presence of a do not track signal from the user’s browser, advertisers still intend to track, though they *may not* display targeted advertising. Of course the potential harm is that any information collected this way may be wrong. Users have no way to know what is collected, what is right or wrong, or who may use this data and for what purpose.

A fundamental problem is this: despite bombardment by privacy policies that may mention “affiliates” and third-party trackers, while users are in the process of interacting online, they are not perceptually aware of such participants — whether they are present are not. Overhearers are hidden inside of an HTTP request. Undoubtedly, the legal system will spend long years grappling with out-of-date privacy and data protection laws. In the meanwhile, we have an opportunity to change the current paradigm by making browsers more socially aware. That is, by acknowledging that interaction in a web browser is inherently multi-party, we can change the user experience to make transparent the social role of all participants present. The challenge is not to disrupt or annoy the user, but to keep such social awareness in the periphery. Social awareness should operate in real time and on a continuous basis. Perhaps, a good metaphor is a social network in the sense of Facebook or Google+. Showing the user who is present, who is a “friend”, who “friends” of “friends” may be, and blocking un-ratified participants should be as natural as turning one’s head to see who is close and lowering one’s voice for better privacy. This does not solve the myriad of privacy problems associated with online behavioral tracking, but might at least afford users the chance to select an appropriate communication strategy for overhearers.

³ Most recently, it appears that political parties are using these techniques, and purchasing consumer data, in order to better target voters (Duhigg, 2012b).

8.5 CONCLUSION

The goal of this dissertation was to prove the following:

1. Interaction with graphical users interfaces involve discourse processing affecting comprehension, and ultimately behavior;
2. Online behavioral advertising participates in a new form of advertising discourse where user's beliefs are affected *during the course of interaction*; and,
3. The use of quantitative methods lend insight into subtle problems in the design of microinteractions.

This dissertation has demonstrated that graphical user interface design can not only cause users to make faulty pragmatic inferences, but that interaction with such interfaces fundamentally exploits properties of the same cognitive architecture underlying linguistic competence. To my knowledge this is the first work to suggest how users of graphical user interfaces might be manipulated through small changes in context that affect understanding. If intentional, this represents what Sperber and Wilson (1986) might call a form of *covert communication*. Though interaction designers may be trained to recognize and understand psychological processes in user interaction, I believe equal attention is merited for the understanding of discourse processes. Both play a role in social cognition.

In addition, two other points are worthy of note.

First, it is not uncommon that technical specifications of formal standards — such as those that drove “Do Not Track” and the AdChoices icon — are developed with a focus on policy over user comprehension. It is in our best interest, as designers and users, to put such specifications under the microscope to see how they stand up in practice. Not doing so has the potential to rob these specifications of their desired effect.

Second, though advertisers have long been known as masters of manipulation of basic psychological processes, much less said about how they also manipulate meaning and understanding. Effects are arguably more subtle and difficult to see. There is rich opportunity to study so-called “dark patterns”, in addition to more commonly used interaction design patterns, to identify potential levers and controls relating to comprehension. By doing so, we are in a better position to

educate users online about how to interpret the myriad of messages and interactions to which they are exposed.



PARTICIPANT CONSENT FORM

I. INTRODUCTION/PURPOSE:

I am being asked to participate in a research study. The purpose is to research a topic in user interface design. My involvement in this study will begin when I agree to participate and will continue until the survey has completed.

II. PROCEDURES:

As a participant in this study, I will be asked to complete an online survey launch to an external website from Amazon Mechanical Turk (AMT). My participation in this study is voluntary and I am free to withdraw or discontinue participation at any time.

My participation in this study will last for approximately 3 minutes.

III. RISKS AND BENEFITS:

IV. My participation in this study may cause some discomfiture. I can stop at any point. If I complete the survey, I will receive monetary compensation for my time.

V. CONFIDENTIALITY:

My data in this study is anonymized. I will NOT be identified by AMT workerID. Amazon will not have access to survey data stored on Qualtrics servers. Qualtrics will not have access to my workerId nor IP address. No personally identifiable data is collected or stored. By signing this form, I allow the research study investigator to make these records available to the University of Baltimore Institutional Review Board (IRB) and regulatory agencies as required to do so by law. Fully anonymized data may be made available to other researchers after the study is complete.

VI. SPONSOR OF THE RESEARCH:

This research study is for a doctoral dissertation.

VII. COMPENSATION/COSTS:

I will be paid 15 cents for my participation.

VIII. CONTACTS AND QUESTIONS:

The principal investigator(s), Lisa Harper and Dr. Kathryn Summers has offered to and has answered any and all questions regarding my participation in this research study. If I have any further questions, I can contact them at lisa.harper@ubalt.edu. For questions about rights as a participant in this research study, contact the UB IRB Chair: Eric Easton, Chair, University of Baltimore Institutional Review Board, 410-837-4874, eeaston@ubalt.edu.

Certification:

I have read and understood the above information:

Yes

No

>>

B

AMT TERMS OF SERVICE

8/16/13

Amazon Mechanical Turk - Participation Agreement



Amazon Mechanical Turk Participation Agreement

Last updated: November 1, 2012

Welcome to the Amazon Mechanical Turk services platform.

BY REGISTERING FOR AND USING THE SITE, YOU CERTIFY THAT (1) YOU ARE AT LEAST 18 YEARS OLD; (2) YOU HAVE THE AUTHORITY TO ENTER INTO THIS AGREEMENT AND BIND YOURSELF OR THE COMPANY YOU REPRESENT; (3) YOU AUTHORIZE THE ELECTRONIC TRANSFER OF FUNDS TO YOUR BANK ACCOUNT IN ACCORDANCE WITH SECTION 4 OF THIS PARTICIPATION AGREEMENT; AND (4) YOU AGREE TO BE BOUND BY ALL TERMS AND CONDITIONS OF THIS AGREEMENT, INCLUDING THE TERMS AND CONDITIONS OF THE PAYMENT SERVICE DESCRIBED IN SECTION 4 AND ALL APPLICABLE POLICIES, PROCEDURES AND GUIDELINES. This Participation Agreement (the "Agreement") is between you and Amazon Mechanical Turk (as defined below) and governs your and Amazon Mechanical Turk's respective rights and obligations with respect to your offering for sale, selling, requesting, purchasing, and/or providing Services (defined below) on or through the Site (as defined below).

For purposes of this Agreement, (a) "Amazon Mechanical Turk," "we," "us" or "our" means Amazon Mechanical Turk, Inc., a Delaware Corporation, (b) "Site" means the Amazon Mechanical Turk web site located at mturk.amazon.com, requester.mturk.com, www.mturk.com and any successor website thereto, provided by us to you through the service platform on the Site, (c) "Services" means any service that you sell, offer to sell, request, purchase, and/or provide on or through the Site, (d) "Affiliate" means any entity controlled by, in control of, or under common control with Amazon Mechanical Turk, (e) "Requester" means you, if you use the Site to request a Provider perform Services, (f) "Provider" means you, if you use the Site to perform Services for a Requester, (g) "Amazon Account" means any customer account that you have established with a website owned or controlled by Amazon or its Affiliates or on behalf of third parties, including without limitation those websites currently located at <http://www.amazon.com>, <http://www.amazon.ca>, <http://www.amazon.co.jp>, and any successor or replacement websites.

This Agreement consists of the terms and conditions set forth in this document together with all applicable policies, procedures and guidelines that appear on the Site from time to time (collectively, the "Policies" which are hereby incorporated by this reference into, and made part of, this Agreement). Amazon Mechanical Turk reserves the right to change the Policies at any time and for any reason. Any changes will be effective upon posting of the Agreement or Policies on the Site, and may be made without any other notice of any kind. You are at all times responsible for reading and understanding each version of this Agreement and the Policies. YOUR CONTINUED USE OF THE SITE FOLLOWING AMAZON MECHANICAL TURK'S POSTING OF ANY CHANGES WILL CONSTITUTE YOUR ACCEPTANCE OF SUCH CHANGES. IF YOU DO NOT AGREE TO ANY CHANGES TO THIS AGREEMENT (INCLUDING TO ANY OF THE POLICIES INCORPORATED HEREIN), DO NOT CONTINUE TO USE THE SITE.

1. Registration.

a. **Registration.** When you register with the Site, you will be asked to provide us with, at a minimum, your name, a valid email address, your phone number, and your physical address. Providers may also be asked to provide certain tax information at registration or afterwards. You agree to provide us with true and accurate information, and to update that information to the extent it changes in any way. When registering or updating your information, you will not impersonate any person or use a name that you are not legally authorized to use.

You may register with the Site either by (i) using your existing Amazon Account or (ii) creating a new Amazon Account. If you do not have an existing Amazon Account at the time you register with the Site, an Amazon Account on the Amazon.com website located at <http://www.amazon.com> (hereinafter, "Amazon.com") will be automatically and concurrently established in your name with the same e-mail address and password you provide under your Amazon Account. Your Amazon Account will be governed by the [Conditions of Use](#) and [Privacy Notice](#) applicable to Amazon.com, as well as the [Amazon Mechanical Turk Privacy Notice](#). You may not use multiple Amazon Accounts to register with Mechanical Turk. Your Amazon Account username must not suggest affiliation with Amazon, Amazon Mechanical Turk, or any third party unless that third party specifically gave you permission to do so.

b. **Passwords and Account Use.** You are solely responsible for maintaining the secrecy and security of your password. You may not disclose your password to any third party (other than third parties authorized by you to use your account) and are solely responsible for any use of or action taken under your password on the Site. If your password is compromised, you must change your password. You may not permit any other person to perform Services as Provider using your Amazon Account. Additionally, if you are using the Site as a Provider, you may not use different Amazon Accounts to perform Services.

2. **Amazon Mechanical Turk's Role.** Amazon Mechanical Turk provides a venue for third-party Requesters and third-party Providers to enter into and complete transactions. Amazon Mechanical Turk and its Affiliates are not involved in the transaction between Requesters and Providers. As a result, we have no control over the quality, safety, or legality of the Services, the ability of Providers to provide the Services to Requesters' satisfaction, or the ability of Requesters to pay for Services. We are not responsible for the actions of any Requester or Provider. We do not conduct any screening or other verification with respect to Requesters or Providers, nor do we provide any recommendations. As a Requester or a Provider, you use the Site at your own risk.

3. Your Use of the Site

a. **Requesters in General.** Upon completion of Services to Requesters' reasonable satisfaction, Requesters must pay Providers for their Services. As a Requester, you agree that upon your approval of the Services performed by a Provider, payment will be remitted to the Provider automatically (as described in Section 4 below). After you have approved the applicable Services, you are not entitled to any refund of your payment for such Services. If a Requester is not reasonable satisfied with the Services, the Requester may reject the Services. As a Requester, you will be charged a fee for your use of Amazon Mechanical Turk in connection with each request for Services. The fees for using Amazon Mechanical Turk are contained in the Policies for the applicable fee associated with your level of access pursuant to this Agreement. All fees are in U.S. dollars unless stated otherwise.

The Amazon Mechanical Turk fees may vary in the future. You agree to pay the amounts set forth in the Amazon Mechanical Turk Fees from time to time on the terms set forth herein and therein, and to check the fees and terms each time you use the Site. You acknowledge that, while Providers are agreeing to perform Services for you as independent contractors and not employees, repeated and frequent performance of Services by the same Provider on your behalf could result in reclassification of that employment status. If you have any questions about your obligations to comply with local laws and regulations pursuant to Section 6, you should seek independent legal advice. To the extent you receive any contact or personal information regarding any Provider who has performed Services for you, such information may only be used as necessary for you to comply with applicable laws and for no other purpose whatsoever. Further, you agree that you will only accept work product from Providers that has been submitted through the Site.

b. **Providers in General.** You may only register once with Mechanical Turk as a Provider. Providers may perform Services for any Requester in accordance with the specifications submitted by the Requester. However, if the Services do not meet the Requester's reasonable satisfaction, the Requester may reject the Services and repeat the specific request. As a Provider, the Requester for whom you provide Services is your client, and as

<https://www.mturk.com/mturk/conditionsofuse>

1/4

C

SAMPLE AMT ASSIGNMENT DEFINITION

```
title = Language study and demographic survey (3 min)
description: one task and then just a survey
keywords:survey, language

# how much you'll pay each subject
reward:.15

# how many subjects do you want
assignments:30

#####
## HIT Timing Properties
#####

# 60*10, 10 mins to finish a suvey
assignmentduration:600

# 60*60*24*2, 2 day to keep on mturk
hitlifetime:172800

# 10 seconds to auto approve the response
#
autoapprovaldelay:10

#####
## Qualification Properties
#####

# user must have an approval rate of 90% or greater
qualification.1:0000000000000000L0
qualification.comparator.1:greaterthan
qualification.value.1:90
qualification.private.1:false

# user must be in the United States
qualification.2:000000000000000071
qualification.comparator.2:equalto
qualification.locale.2:US
qualification.private.2:true
```

D

SURVEY RESPONSE

Table 12: Aggregate Demographic Profile and Survey Response

Category	Experiment 3	Total
Age		
18-24	29%	28%
25-34	44%	45%
35-49	17%	20%
50-64	10%	7%
64+	0%	0%
Gender		
Male	63%	53%
Female	37%	47%
Ethnic Identity		
American Indian / Native American	0%	1%
Asian or Pacific Islander	10%	10%
Black / African American	3%	7%
Hispanic or Latin American	5%	7%
White / Caucasian	82%	73%
Near Eastern or Arabic	0%	0%
Other	2%	2%
Native Language		
English	99%	97%
Other	1%	3%
English		
A little English	0%	0%
Some English	0%	0%
Fluent English	0%	60%
Near-native English	100%	40%

Education		
Some High School	1%	1%
High School Graduate	17%	13%
Some College or Associate Degree	39%	40%
College Degree	35%	37%
Post-graduate Degree	8%	9%
Income		
Under 20,000	30%	34%
20,000 - 30,000	25%	19%
30,000 - 40,000	13%	15%
40,000 - 50,000	13%	10%
50,000+	20%	22%
IT Job		
Yes	21%	20%
No	79%	80%
Internet Usage		
Fewer than 4 hours per week	1%	2%
4-10 hours per week	13%	11%
10-25 hours per week	41%	30%
25+ hours per week	45%	57%
Shop Online		
Never	1%	1%
Rarely	20%	17%
Sometimes	59%	30%
Often	21%	52%
Sense of Privacy in Public		
Not private	31%	29%
A Little private	34%	30%
Somewhat private	23%	29%
Private	11%	9%
Very private	1%	3%

Sense of Privacy at Home		
Not private	5%	4%
A Little private	15%	10%
Somewhat private	23%	23%
Private	34%	37%
Very private	24%	26%
Importance of Online Privacy		
Not much of an issue	8%	7%
Somewhat important	50%	51%
Really important	42%	42%
Steps to Protect Privacy		
Don't know how to protect	12%	8%
Know how but not consistent	38%	49%
Know how and take measures	50%	43%
Know what a tracking cookie is		
Yes	90%	91%
No	10%	9%
DNT means...		
Do not show targeted advertising	17%	20%
Do not track across sites	44%	54%
Do not track on this site	49%	47%
Do not collect information	39%	48%
Do not store information	38%	43%
Would Turn off Tracking if Easy?		
Yes	88%	91%

No	12%	9%
Browser Configured Opt-Out		
Yes	54%	55%
No	46%	45%
Use a Browser Plu- gin for Privacy Pro- tection		
Yes	70%	73%
No	30%	27%
Total	100	1158

E

EXPERIMENT 1 CONDITIONS

You go to a website which presents the following choices:

- Cookies on
- Cookies off
- Cancel

You pick '**cancel**'. Assuming the website respects your wishes, what does this mean?

- The website will have cookies because I didn't select off.
- The website won't have cookies because I didn't select on.
- The website will have cookies or not.
- The website will neither have cookies or not.

Figure 69: Text, No Deontic Force, Cookie Condition

You go to a website which presents the following choices:

- Allow cookies
- Block cookies
- Cancel

You pick '**cancel**'. Assuming the website respects your wishes, what does this mean?

- The website will allow cookies because I didn't select block.
- The website won't allow cookies because I didn't select allow.
- The website will allow cookies or block them.
- The website will neither allow cookies nor block them.

Figure 70: Text, Deontic Force, Cookie Condition

You go to a website which presents the following choices:

- Pictures on
- Pictures off
- Cancel

You pick '**cancel**'. Assuming the website respects your wishes, what does this mean?

- The website will have pictures because I didn't select off.
- The website won't have pictures, because I didn't select on.
- The website will have pictures or not.
- The website will neither have pictures or not.

Figure 71: Text, No Deontic Force, Picture Condition

You go to a website which presents the following choices:

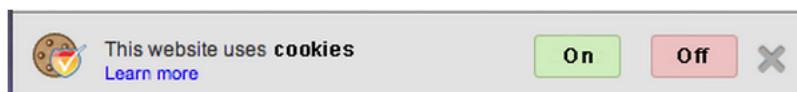
- Allow pictures
- Block pictures
- Cancel

You pick '**cancel**'. Assuming the website respects your wishes, what does this mean?

- The website will allow pictures because I didn't select block.
- The website won't allow pictures, because I didn't select allow.
- The website will allow pictures or not.
- The website will neither allow pictures or not.

Figure 72: Text, Deontic Force, Picture Condition

You go to a website which presents the following:



You pick '**x**'. Assuming the website respects your wishes, what does this mean?

- The website will have cookies because I didn't select off.
- The website won't have cookies because I didn't select on.
- The website will have cookies or not.
- The website will neither have cookies or not.

Figure 73: Mixed-Modal, No Deontic Force, Cookie Condition

You go to a website which presents the following:



You pick 'x'. Assuming the website respects your wishes, what does this mean?

- The website will allow cookies because I didn't select block.
- The website won't allow cookies because I didn't select allow.
- The website will allow cookies or not.
- The website will neither allow cookies or not.

Figure 74: Mixed-Modal, Deontic Force, Cookie Condition

You go to a website which presents the following:

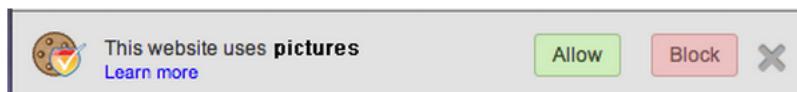


You pick 'x'. Assuming the website respects your wishes, what does this mean?

- The website will have pictures because I didn't select off.
- The website won't have pictures because I didn't select on.
- The website will have pictures or not.
- The website will neither have pictures or not.

Figure 75: Mixed-Modal, No Deontic Force, Picture Condition

You go to a website which presents the following:



You pick 'x'. Assuming the website respects your wishes, what does this mean?

- The website will allow pictures because I didn't select block.
- The website won't allow pictures because I didn't select allow.
- The website will allow pictures or block them.
- The website will neither allow pictures nor block them.

Figure 76: Mixed-Modal, Deontic Force, Picture Condition

EXPERIMENT 1 RAW RESULTS

Table Options ▾				
#	Answer	Response	%	
1	The website will allow cookies because I didn't select block.	7	26%	
2	The website won't allow cookies because I didn't select allow.	6	22%	
3	The website will allow cookies or block them. Create a report subgroup	3	11%	
4	The website will neither allow cookies nor block them.	11	41%	
	Total	27	100%	

Figure 77: Text, Deontic Force, Cookie Condition

#	Answer	Response	%
1	The website will have cookies because I didn't select off.	7	27%
2	The website won't have cookies because I didn't select on.	6	23%
3	The website will have cookies or not.	9	35%
4	The website will neither have cookies or not.	4	15%
	Total	26	100%

Figure 78: Text, No Deontic Force, Cookie Condition

#	Answer	Response	%
1	The website will allow pictures because I didn't select block.	8	31%
2	The website won't allow pictures, because I didn't select allow.	10	38%
3	The website will allow pictures or not.	1	4%
4	The website will neither allow pictures or not.	7	27%
	Total	26	100%

Figure 79: Text, Deontic Force, Picture Condition

#	Answer	Response	%
1	The website will have pictures because I didn't select off.	7	27%
2	The website won't have pictures, because I didn't select on.	5	19%
3	The website will have pictures or not.	7	27%
4	The website will neither have pictures or not.	7	27%
	Total	26	100%

Figure 80: Text, No Deontic Force, Picture Condition

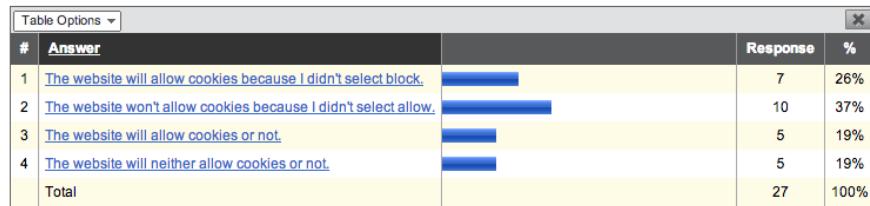


Figure 81: Graphic, Deontic Force, Cookie Condition

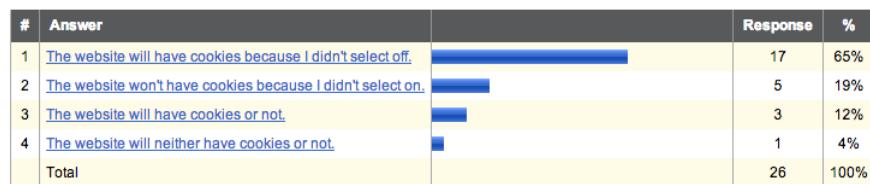


Figure 82: Graphic, No Deontic Force, Cookie Condition

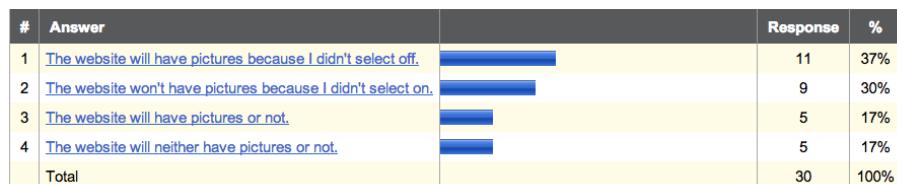


Figure 83: Graphic, No Deontic Force, Picture Condition

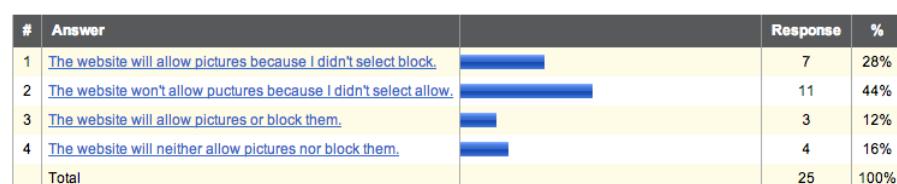


Figure 84: Graphic, Deontic Force, Picture Condition

G

PARTICIPANT SURVEY

Q	Type	Question	Answer Type		Answer 1 18-24	Answer 2 25-34	Answer 3 35-49	Answer 4 50-64	Answer 5 65+	Answer 5 65+
			Radio	Radio						
1	Demographic	How old are you?	Radio	Radio	American Indian / Asian or Pacific Islander	Male	Female	25-34	35-49	Answer 4 50-64
2	Demographic	What is your gender?	Radio	Radio	Asian Indian / Other	Asian or Pacific Islander	American Indian / English	Male	Female	Answer 5 65+
3	Demographic	What is your ethnic identity?	Radio	Radio	Black / African American	Male	Female	Male	Female	White / Caucasian
4	Demographic	What is your native language?	Radio	Radio	Hispanic / Latino	Female	Male	Female	Male	White / Caucasian
5	Demographic	What is your English proficiency?	Radio	Radio	Near-native English	Male	Female	Female	Male	Near-native English
6	Demographic	What is your education?	Radio	Radio	Post-graduate degree	Male	Female	Female	Male	Post-graduate degree
7	Demographic	What is your income level?	Radio	Radio	College graduate degree	Male	Female	Female	Male	College graduate degree
8	Technical Savvy	Do you either have a job in an IT-related field (for example, computer science, web development or similar; are a student in an IT-related area? What is your Internet usage?)	Radio	Yes	40,000-50,000	Male	Female	Female	Male	40,000-50,000
9	Technical Savvy	What is your Internet usage?	Radio	No	30,000-40,000	Male	Female	Female	Male	30,000-40,000
10	Technical Savvy	How much do you shop online?	Radio	Never	25+	Male	Female	Female	Male	25+
11	Privacy	How private do you feel on the Internet when you are using your computer (i.e., laptop, tablet) in public (i.e., a coffee shop, library, school, etc.)	Radio	Not private	Most of my shopping is online	Male	Female	Female	Male	Most of my shopping is online
12	Privacy	How private do you feel on the Internet when you are using your computer (i.e., laptop, tablet) at home?	Radio	Not private	Private	Male	Female	Female	Male	Private
13	Privacy	How important is online privacy to you?	Radio	Occasionally	Very private	Male	Female	Female	Male	Very private
14	Privacy	What steps do you take to protect your privacy?	Radio	A little private	Very private	Male	Female	Female	Male	Very private
15	Privacy	Do you know what web browser tracking cookies are?	Radio	A little private	Mostly private	Male	Female	Female	Male	Mostly private
16	Privacy	Do you know what "do not track" means?	Multiple Select	Do not track me across sites	A really important issue I think about a lot	Male	Female	Female	Male	A really important issue I think about a lot
17	Privacy	If you could easily turn off online tracking, would you?	Radio	Yes	I know how to protect my personal information online but don't consistently do so	Male	Female	Female	Male	I know how to protect my personal information online but don't consistently do so
18	Privacy	Do you opt-out of online tracking in your browser settings?	Radio	No	I know how to protect my personal information online and consistently take the necessary steps to do so	Male	Female	Female	Male	I know how to protect my personal information online and consistently take the necessary steps to do so
19	Privacy	Do you use any browser plug-ins such as adblockers or other privacy enhancing tools on your computer?	Radio	No	Do not collect information on me	Male	Female	Female	Male	Do not collect information on me
					Do not store information on me	Male	Female	Female	Male	Do not store information on me

BIBLIOGRAPHY

- AAAAA, ANA, BBB, DMA, & IAB. (2009, July). *Self-Regulatory Principles for Online Behavioral Advertising*. Retrieved from http://www.iab.net/public_policy/behavioral-advertisingprinciples
- Acohido, B. (2011, February). Most Google, Facebook users fret over privacy. Retrieved from <http://www.gallup.com/poll/146159/facebook-google-users-skew-young-affluent-educated.aspx>
- Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy*, 24–30.
- Acquisti, A., John, L. K., & Loewenstein, G. (2012). The impact of relative standards on the propensity to disclose. *Journal of Marketing Research*, 49(2), 160–174.
- Alonso, O., & Mizzaro, S. (2009). Can we get rid of TREC assessors? Using Mechanical Turk for relevance assessment. *Proceedings of the SIGIR 2009 Workshop on the Future of IR Evaluation*, 15–16.
- Ambati, V., & Vogel, S. (2010). Can crowds build parallel corpora for machine translation systems? In *CSLDAMT '10 Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 62–65).
- Anand, P., & Andrews, C. (2011, July 27). Assessing the pragmatics of experiments with crowdsourcing: The case of scalar implicatures [slides]. In *Crowdsourcing Technologies for Language and Cognition Studies*. Boulder. Retrieved from <http://www.crowdsourcing.org/document/assessing-the-pragmatics-of-experiments-with-crowdsourcing-the-case-of-scalar-implicatures/5558>
- Angwin, J. (2012, Jun 17). Online tracking ramps up. *Wall Street Journal*. Retrieved from <http://online.wsj.com/article/SB10001424052702303836404577472491637833420.html>
- Acquisti, A. (2013). Why privacy matters [video]. *TED Talks*. Retrieved from http://www.ted.com/talks/alessandro_acquisti_why_privacy_matters.html
- Atkin, A. (2005). Peirce on the index and indexical reference. *Transactions of the Charles S. Peirce Society*, 41(1), 161–188.

- Austin, J. L. (1975). *How to Do Things with Words*. Harvard University Press.
- Bailey, B. P., & Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior*, 22(4), 685–708.
- Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2000). Adjusting windows: Balancing information awareness with intrusion. In *Proceedings of the 6th Conference on Human Factors and the Web* (pp. 39–83).
- Bard, Anderson, Sotillo, Aylett, & Newlands. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42(1), 22–22.
- Bargh, J. A. (2006). What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology*, 36(2), 147–168.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54, 569–579.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43(3), 800–813.
- Bell, A. (1984). Language style as audience design. *Language in Society*, 145–204.
- Beller, S. (2008). Deontic reasoning squared. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (p. 2103-2108). Cognitive Science Society.
- Bellman, S., Johnson, E. J., & Lohse, G. L. (2001). On site: to opt-in or opt-out? It depends on the question. *Communications of the ACM*, 44(2), 25–27.
- Bencherif, S., & Tanaka, K. (1987). Covert forms of communication. In *Autumn Meeting of the Linguistic Society of Great Britain, Bradford, September*.
- Bergen, L., Goodman, N. D., & Levy, R. (2012). That's what she (could have) said: How alternative utterances affect language use. In *Cogsci 2012* (pp. 1–6). Sapporo.
- Berger, D. (2012). *Power analysis: Introduction to power analysis with G*Power 3*. Retrieved from <http://wise.cgu.edu/downloads/Power%20Analysis%20with%20GPower%20120409.pdf>
- Blase Ur, P. G. L. L. F. C. R. S. Y. W. (2012). *Smart, useful, scary, creepy: perceptions of online behavioral advertising* (Tech. Rep. No. CMU-CyLab-12-007).
- Brandt, M. T., & Preston, I. L. (2013). The Federal Trade

- Commission's use of evidence to determine deception. *The Journal of Marketing*, 41, 1–7.
- Branigan, H., Pickering, M. J., McLean, M., & Cleland, A. (2007). Participant role and syntactic alignment in dialogue. *Cognition*, 104, 163–197.
- Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 3(2), 193 – 209.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482–1493.
- Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, 5(6), 673–678.
- Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge University Press.
- Bruner, R. E. (1997). Cookie proposal could hinder online advertising. *Advertising Age*. Retrieved from <http://adage.com/article/news/advertisers-win-debate-cookies/405/>
- Buhrmester, M., Kwang, T., & Gosling, S. (2010). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 1–23.
- Buscher, G., Dumais, S. T., & Cutrell, E. (2010). The good, the bad, and the random: an eye-tracking study of ad quality in web search. In *Proceedings of the 33rd international ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 42–49). ACM.
- Business Week / Harris Poll. (2000, March). *A Growing Threat* (Tech. Rep.).
- Callison-Burch, C., & Dredze, M. (2010). Creating speech and language data with Amazon's Mechanical Turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 1–12).
- Center for Democracy and Technology. (2007). *Ebehavioral advertising: Tracking, targeting, and technology*. Retrieved from <https://www.cdt.org/privacy/20071031consumerprotectionsbehavioral.pdf>
- Center for Democracy and Technology. (2008, July). *An overview of the Federal Wiretap Act, Electronic Communications Privacy Act, and State two-party consent laws of relevance to the NebuAd system and other uses of Internet traffic content from ISPs for behavioral advertising* (Tech. Rep.).

- Cern. (n.d.). The website of the world's first-ever web server. Retrieved from <http://info.cern.ch>
- Chandler, J., Mueller, P., & Paolacci, G. (2013, July). Non-naïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. Accepted for publication at *Behavior Research Methods*.
- Chen, Y., Pavlov, D., & Canny, J. F. (2009, June). Large-scale behavioral targeting. In *KDD '09: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM Request Permissions.
- Cheng, H.-I., & Patterson, P. E. (2007). Iconic hyperlinks on e-commerce websites. *Applied Ergonomics*, 38(1), 65–69.
- Cheng, P., & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*(17), 391–416.
- Chiarella, S. (2013, Feb 11). *Subject: Does mturk have a new policy excluding international workers? [email]*. Tips for Requesters on Mechanical Turk. Retrieved from <http://turkrequesters.blogspot.com/2013/01/the-reasons-why-amazon-mechanical-turk.html>
- Chief Marketer. (1999, January). *DoubleClick, abacus stock drop after merger news*. Retrieved from <http://chiefmarketeer.com/news/doubleclick-abacus-stocks-drop-after-merger-news>
- Clark, B. (1944). *The Advertising Smoke Screen*. Harper Press.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 335–359.
- Clark, H. H. (1996). *Using Language*. Cambridge University Press.
- Clark, H. H. (2001). Conversation: Linguistic Aspects. In N. Smelser & P. Baltes (Eds.), *International Encyclopedia of Social and Behavioral Sciences* (pp. 2744–2747). London: Elsevier.
- Clark, H. H. (2002). Bridging. In R. C. Schank & B. L. Nash-Webber (Eds.), *Theoretical Issues in Natural Language Processing* (pp. 1–6). Association for Computing Machinery.
- Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing. Where Language, Culture, and Cognition Meet* (pp. 243–268). Hillsdale: Lawrence Erlbaum.
- Clark, H. H., & Brennan, S. E. (2005). Grounding in communication. In L. B. Resnick, J. M. Levine, & T. S. D (Eds.), *Perspectives on Socially Shared Cognition* (pp. 1–12). APA Books.
- Clark, H. H., & Carlson, T. B. (1982). Hearers and speech

- Acts. *Language*, 58, 332–373.
- Clark, H. H., & Schaefer, E. F. (1987). Concealing one's meaning from overhearers. *Journal of Memory and Language*, 26(2), 209–225.
- Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13(2), 259–294.
- Clark, H. H., & Schaefer, E. F. (1992). Dealing with overhearers. In *Arenas of Language Use*. University of Chicago Press, Chicago.
- Clark, H. H., & Wilkes-Gibbs, D. (1986a). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Clark, H. H., & Wilkes-Gibbs, D. (1986b, September). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Clifton, C., Jr., & Dube, C. (2010). Embedded implicatures observed: a comment on Geurts and Pousoulous (2009). *Semantics and Pragmatics*, 3.
- Constine, J. (2012, Aug 15). Facebook's new retargeted ads performing "very well", adds partners to run them. *Tech Crunch*. Retrieved from <http://techcrunch.com/2012/08/25/facebook-exchange-retargeting/>
- Cook, G. (1989). *Discourse*. Oxford University Press.
- Cook, G. (2001). *The Discourse of Advertising*. Routledge.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31(3), 187–276.
- Cranor, L. (2012). Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *Journal of Telecommunications and High Technology*.
- Crimtan.com. (n.d.). Cookie Deadline Sees AdChoices Icon Rolled Out in Ireland. Retrieved from <http://crimtan.com/cookie-deadline-sees-adchoices-icon-rolled-out-in-ireland/>
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013, March). Evaluating amazon's mechanical turk as a tool for experimental behavioral research. *PLoS ONE*, 8(3), e57410.
- DAA. (2011, June). *Implementing the DAA icon - An ad creative primer*. Retrieved from http://www.iab.net/media/file/DAA_Icon_Ad_Marker_Creative_Guidelines_10-31-2013.pdf
- Degen, J., & Tanenhaus, M. K. (2011, June 2 – 4). A constraint-based approach to scalar implicature processing. In *Poster presented at experimental pragmatics 2011* (Vol. 51, pp. 437–457).
- Dell, G. S. G. (1986, July). A spreading-activation theory

- of retrieval in sentence production. *Psychological Review*, 93(3), 283–321.
- Denis, M. (1997). The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers Psychologie Cognitive*, 16(4), 409–458.
- de Nooijer, J. A., van Gog, T., Paas, F., & Zwaan, R. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left- or right-handed first-person perspectives. *Psychological Science*, 24.
- Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of Marketing Research*, 146–160.
- Donnellan, K. S. (1966). Definite descriptions. *The Philosophical Review*, 75(3), 281–304.
- D’Orazio, D. (2013, Oct 10). Microsoft said to work on technology to replace cookies, track across Windows, Bing, and Xbox. *The Verge*, 1–2. Retrieved from <http://www.theverge.com/2013/10/10/4823944/microsoft-reportedly-working-to-replace-cookies-ad-tracking>
- Dreze, & Hussherr. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing*, 17(4), 16–16.
- Duggal, J. (2012, May). *Promise Unfulfilled? Lessons from the Revolution (Quantcast)* (Tech. Rep.).
- Duhigg, C. a. (2012a, Feb 16). How Companies Learn your Secrets. *New York Times Magazine*. Retrieved from http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?_r=1&pagewanted=all
- Duhigg, C. a. (2012b). *The Power of Habit*. Random House.
- Dunaway, G. (2011). Evidon reports 500,000 OBA opt-outs over 50 billion impressions. Retrieved from <http://www.adotas.com/2011/08/evidon-reports-500000-oba-opt-outs-over-50-billion-impressions/>
- Dynel, M. (2010). Not hearing things - Hearer/listener categories in polylogues. *MediAzioni*, 9.
- Dynel, M. (2011). On "Revolutionary Road": A proposal for extending the Gricean model of communication to cover multiple hearers. *Lodz Papers in Pragmatics*, 6(2), 283–304.
- Eberhard, K. M. K., Spivey-Knowlton, M. J. M., Sedivy, J. C. J., & Tanenhaus, M. K. M. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24(6), 409–436.
- Eckersley, P. (2010). How unique is your web browser? In

- M. Atallah & N. Hopper (Eds.), *Privacy and Enhancing Technologies* (pp. 1–18).
- Egelman, S., Tsai, J., Cranor, L. F., & Acquisti, A. (2009, April 4 - 9). Timing is everything? The effects of timing and placement of online privacy indicators. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009* (pp. 319–328).
- Eichelberger, L. (n.d.). *The cookie controversy: privacy issues*. Retrieved from <http://www.cookiecentral.com/ccstory/cc4.htm>
- Elazary, L., & Itti, L. (2008). Interesting objects are visually salient. *Journal of Vision*, 8(3), 3–3.
- Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454–459.
- Evidon. (2011). *Consumer interactions with ad notice* (Tech. Rep.).
- Farahat, A., & Bailey, M. C. (2012, April). How effective is targeted advertising? In *WWW '12: Proceedings of the 21st International Conference on World Wide Web*. ACM.
- Federal Trade Commission. (2010, December). *FTC staff issues privacy report, offers framework for consumers, businesses, and policymakers* (Tech. Rep.).
- Federal Trade Commission. (2012). *Protecting Consumer Privacy in an Era of Rapid Change—A Proposed Framework for Businesses and Policymakers* (Tech. Rep.).
- Fernback, J., & Papacharissi, Z. (2007). Online privacy as legal safeguard: the relationship among consumer, online portal, and privacy policies. *New Media & Society*, 9(5), 715–734.
- Festinger, L. (1964). *Conflict, Decision, and Dissonance*. Stanford University Press.
- Fielding, R. (2012a, September). *Apache does not tolerate deliberate abuse of open standards [Apache / httpd change request]*. Retrieved from <https://github.com/apache/httpd/commit/a381ff35fa4d50a5f7b9f64300dfd98859dee8d0>
- Fielding, R. (2012b, September 12). *Re: Intermediaries interfering with DNT decision making. [Online forum comment]*. Retrieved from <http://lists.w3.org/Archives/Public/public-tracking/2012Sep/0165.html>
- Fillenbaum, S. (1975). IF: Some uses. *Psychological Research*, 37, 245–260.
- Fillmore, C. (1971). Santa Cruz lectures on deixis. In *Reproduced by the Indiana University Linguistics Club 1975* (pp. 1–89).
- Fowler, A. (2011, November 2). Do not track adoption in

- Firefox Mobile is 3x higher than desktop. Retrieved from <http://blog.mozilla.com/privacy/2011/11/02/do-not-track-adoption-in-firefox-mobile-is-3x-higher-than-desktop/>
- Fowler, A. (2013a, Feb 25). *Firefox getting smarter about third-party cookies*. Retrieved from <http://blog.mozilla.org/privacy/2013/02/25/firefox-getting-smarter-about-third-party-cookies/>
- Fowler, A. (2013b, May 3). Mozilla's new Do Not Track dashboard: Firefox users continue to seek out and enable DNT. Retrieved from <http://blog.mozilla.org/privacy/2013/05/03/mozillas-new-do-not-track-dashboard-firefox-users-continue-to-seek-out-and-enable-dnt/>
- Fulton, S. (2012, February). *Dr. Cranor on "do not track" & the improbability of complete privacy*. Retrieved from <http://www.readwriteweb.com/enterprise/2012/02/dr-cranor-on-do-not-track-and.php>
- Garfinkel, H. (1967). Studies in Ethnomethodology. *Blackwell Publishing Ltd.*
- Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, 31, 961 – 987.
- Geis, M. L. (1982). *The Language of Television Advertising*. Academic Press.
- Geis, M. L. (1995). *Speech Acts and Conversational Interaction*. Cambridge: Cambridge University Press.
- Geis, M. L., & Zwicky, A. M. (1971). On invited inferences. *Linguistic Inquiry*, 2, 561-566.
- Gernsbacher, M., Varner, K., & Faust, M. (2013). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 430–445.
- Geurts, B., & Pousoulous, N. (2009). Embedded implicatures?!? *Semantics and Pragmatics*, 2.
- Gibson, E., & Piantadosi, S. (2011). Using mechanical turk to obtain and analyze english acceptability judgments. *Language and Linguistics Compass*, 5, 509 – 524.
- Girocco, V., & Politzer, G. (1990). Conversational and world knowledge constraints on deductive reasoning. *Cognitive Biases*, 87–107.
- Goel, V. (2013, November 15). Facebook reasserts posts can be used advertise. *The New York Times*. Retrieved from <http://www.nytimes.com/2013/11/16/technology/>

- facebook-amends-privacy-policies.html?_r=0
- Goffman, E. (1981). *Forms of Talk*. Philadelphia. University of Pennsylvania Press.
- Goldfarb, A., & Tucker, C. (2011). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, 30(3), 389–404.
- Goldstein, D. G., McAfee, R. P., & Suri, S. (2013). The cost of annoying ads. In *Proceedings of the 22nd International Conference on World Wide Web* (p. 459-470).
- Gomez, J., Pinnick, T., & Soltani, A. (2009, Jun 1). Know Privacy. Retrieved from <http://knowprivacy.org/>
- Google, I. (2010). Google milestones. *Corporate Information*. Retrieved from <https://www.google.com/intl/en/about/company/history/>
- Gormley, M. R., Gerber, A., Harper, M., & Dredze, M. (2010, June). Non-expert correction of automatically generated relation annotations. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 204–207). Los Angeles: Association for Computational Linguistics. Retrieved from <http://www.aclweb.org/anthology/W10-0732>
- Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (2012). *Handbook of Discourse Processes*. Routledge.
- Grice, H. P. (1975a). Logic and conversation. In Cole (Ed.), *Syntax and Semantics 3: Speech Acts* (pp. 41–58). Elsevier.
- Grice, H. P. (1975b). Logic and conversation. In Cole et al. (Ed.), *Syntax and Semantics 3: Speech Acts* (pp. 41–58). 1975.
- Grice, H. P., & White, A. R. (1961). The causal theory of perception. In *Proceedings of the Aristotelian Society* (pp. 121–153).
- Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73(3), 407–420.
- Grosz, B., Pollack, M., & Sidner, C. (1989). Discourse. In M. Posner (Ed.), *Foundations of Cognitive Science* (pp. 1–33). MIT Press.
- Grosz, B., & Sidner, C. (1986). Attentions, intentions, and the structure of discourse. *Computational Linguistics*, 12(3), 175–204.
- Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge University Press.
- Gumperz, J. J., & Hymes, D. (1972). *Directions in Sociolinguistics; the Ethnography of Communication*. New York: Holt, Rinehart, and Winston.

- Gumperz, J. J., & Hymes, D. H. (1964). The ethnography of communication. *American Anthropologist*, 66.
- Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274–307.
- Gundel, J., Hedberg, N. N., & Zacharski, R. R. (2012). Under-specification of cognitive status in reference production: some empirical predictions. *Topics in Cognitive Science*, 4(2), 249–268.
- Halliday, M. (1977). *Text and Semantic Choice in Social Contexts* (T. A. van Dijk & J. Petofi, Eds.). Berlin: Walter de Gruyter.
- Harris, R. J. (1983). *Information Processing Research in Advertising*. Lawrence Erlbaum Associates.
- Hastak, M., & Culnan, M. J. (2010, January). Future of privacy forum: Online behavioral advertising “icon” study (Tech. Rep.).
- Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 37–53.
- Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Unpublished doctoral dissertation, University of Massachusetts.
- Helft, M., & Miller, C. (2011, January). Privacy law is outrun by the web. *New York Times*. Retrieved from http://www.nytimes.com/2011/01/10/technology/10privacy.html?_r=1&hp
- Hill, K. (2012, October). Microsoft is losing in a bitter battle to protect Internet users' privacy. *Forbes Magazine*.
- Hill, K. (2013, August). Use Of ad blocking is on the rise. *Forbes Magazine*. Retrieved from <http://www.forbes.com/sites/kashmirhill/2013/08/21/use-of-ad-blocking-is-on-the-rise/>
- Hobbs, J. (1985). *On the coherence and structure of discourse* (Tech. Rep. No. Report No. CSLI-85-37).
- Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition*, 63(1), 29–78.
- Hoofnagle, C. J., King, J., Li, S., & Turow, J. (2010, April). *How different are young adults from older adults when it comes to information privacy attitudes and policies?* (Tech. Rep.).
- Hoofnagle, C. J., Soltani, A., Good, N., Wambach, D., & Ayenson, M. (2012). Behavioral advertising: The offer you cannot refuse. *Harvard Law & Policy Review* 273. UC

- Berkely Public Law Research Paper No. 2137601.* Retrieved from <http://ssrn.com/abstract=2137601>
- Hoofnagle, C. J., Urban, J. M., & Li, S. (2012). Privacy and modern advertising: Most us internet users want do not track to stop collection of data about their online activities. *Amsterdam Privacy Conference*, 1–27.
- Horn, L. R. (2006, September). Implicature. In L. R. Horn & W. G (Eds.), *The Handbook of Pragmatics* (pp. 1–46).
- Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine*.
- Huang, Y. (2007). *Pragmatics*. Oxford University Press, USA.
- Hymes, D. (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. University of Pennsylvania Press.
- IAB. (2010, November). *IAB Legislative and Regulatory Affairs Update March 2010* (Tech. Rep.). Retrieved from http://www.iab.net/media/file/Legislative_Update_March_2010-2.pdf
- IAB. (2012, April). *Internet Advertising Revenue Report: 2011 Full Year Results* (Tech. Rep.).
- In re: DoubleClick Inc. (2001). *Privacy Litigation* 154 F. Supp. 2d 497.
- Interbrand. (2013). *Interbrand Best Global Brands 2013*. Retrieved from <http://interbrand.com/en/best-global-brands/2013/Best-Global-Brands-2013.aspx>
- The Internet*. (1995, January). Retrieved from <https://www.youtube.com/watch?v=XluovrUA6Bk>
- Ipeirotis, P. (2010a). Analyzing the Amazon Mechanical Turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students*, 17(2).
- Ipeirotis, P. (2010b). Demographics of Mechanical Turk. *NYU Working Paper*.
- Janis, I., & Mann, L. (1977). *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment*. Free Press.
- Jefferson, G. (1972). Side sequences. In D. Sadrow (Ed.), *Studies in Social Interaction* (pp. 294–338). New York: Free Press.
- John, L. K., Acquisti, A., & Loewenstein, G. (2009). The Best of Strangers: Context-dependent willingness to divulge personal information. Retrieved from <http://ssrn.com/abstract=1430482>
- Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). Defaults, framing and privacy: Why opting in-opting out. *Marketing Letters*, 13(1), 5–15.
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Harvard University Press.

- Johnson-Laird, P. N. (1989). Mental models. In M. Posner (Ed.), *Foundations of Cognitive Science* (pp. 469–499). MIT Press.
- Johnson-Laird, P. N. (2007). Reasoning without logic. In T. Myers, K. Brown, & B. McGonigle (Eds.), *Reasoning and Discourse Processes* (pp. 13–50). Academic Press.
- Johnson-Laird, P. N., & Garnham, A. (1980). Descriptions and discourse models. *Linguistics and Philosophy*(3), 371–393.
- Jupiter Research. (2010, January). *Behavioral Targeting and the Purchase Funnel Opportunity* (Tech. Rep.).
- Kahan, D. (2013, 10). Fooled twice, shame on who? problems with mechanical turk samples. *Cultural Cognition Project at Yale Law School*. Retrieved from <http://www.culturalcognition.net/blog/2013/7/10/fooled-twice-shame-on-who-problems-with-mechanical-turk-stud.html>
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.
- Kahneman, D., & Tversky, A. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review*, 90.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 1–11.
- Kamp, H., & Reyle, U. (1993). *From Discourse to Logic*. Springer.
- Kaplan, D. (1980). Demonstratives.
- Kazienko, P., & Adamski, M. (2007, January). Adrosa-adaptive personalization of web advertising. *Information Sciences: an International Journal*, 177(11), 1–34.
- Kelly, H. (2013, October 11). Why your face might appear in google ads, and how to stop it. CNN. Retrieved from <http://www.cnn.com/2013/10/11/tech/social-media/google-plus-ads-profiles/>
- Kessler, S. (2010, November). Online behavior tracking and privacy: 7 worst case scenarios. Retrieved from <http://mashable.com/2010/11/03/behavior-tracking-privacy/>
- Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3), 427–457.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394.

- Klopfer, P. H., & Rubenstein, D. I. (1977). The concept privacy and its biological basis. *Journal of Social Issues*, 33(3), 52–65.
- Komanduri, S., Shay, R., Norcie, G., & Ur, B. (2012). Ad-Choices? Compliance with online behavioral advertising notice and choice requirements. *CMU-CyLab-11-005*, 1–22.
- Kosslyn, S., Thompson, W., & Ganis, G. (2006). *The Case for Mental Imagery*. Oxford University Press.
- Kripke, S. (1979). Speaker's reference and semantic reference. In Kiteley & Garfield (Eds.), *Meaning and Truth* (pp. 162–188). New York: Paragon House.
- Krishnamurthy, B., Naryshkin, K., & Wills, C. E. (2011). Privacy leakage vs. protection measures: The growing disconnect. *Computer Communications Review*, 40, 112–117.
- Krishnamurthy, B., & Wills, C. E. (2009). On the leakage of personally identifiable information via online social networks. *Web 2.0 Security and Privacy Workshop*, 7–12.
- Kristol, D. M. (2001, November). HTTP Cookies: Standards, privacy, and politics. *Transactions on Internet Technology (TOIT)*, 1(2). Retrieved from <http://arxiv.org/abs/cs.SE/0105018>
- Krux. (2012, Jun). *State of Data Collection on the Web* (Tech. Rep.). Retrieved from http://www.krux.com/pro/broadcasts/krux_news/krux_CIS_2012/
- Kurby, C. A., & Zacks, J. M. (2013). The activation of modality-specific representations during discourse processing. *Brain and Language*, 126, 338 – 349.
- Lai, Y. L., & Hui, K. L. (2006). Internet opt-in and opt-out: Investigating the roles of frames, defaults and privacy concerns. In *Proceedings of the 2006 ACM SIGMIS CPR Conference on Computer Personnel Research: Forty Four Years of Computer Personnel Research: Achievements, Challenges & the Future* (pp. 253–263).
- Lakoff, G. (2008). *Women, Fire, and Dangerous Things*. University of Chicago Press.
- Lanier, J. (2006). Digital maoism. *The Edge*, 183.
- Lanier, J. (2010). *You Are Not a Gadget*. Random House LLC.
- Leacock, M., Malone, E., & Wheeler, C. (2005). Implementing a pattern library in the real world: A Yahoo! case study. *ASIS&T IA Summit*, 1–10.
- Lease, M., Hullman, J., & Bigham, J. (2013). Mechanical Turk is not anonymous. *Social Science Research Network*.
- Leech, G. N. (1966). *English in Advertising*. Longmans.

- Leech, G. N. (1983). *Principles of Pragmatics*. Longman Pub Group.
- Leon, P. G., Cranshaw, J., Cranor, L. F., Graves, J., Hastak, M., Ur, B., & Xu, G. (2012). *What do online behavioral advertising privacy disclosures communicate to users?* (Tech. Rep. No. CMU-CyLab-12-008).
- Leon, P. G., Ur, B., Shay, R., Wang, Y., Balebako, R., & Cranor, L. (2012). *Why Johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising* (Tech. Rep. No. CMU-CyLab-11-017).
- Levelt, W. J. M. (1993). *Speaking*. MIT Press.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Levinson, S. C. (1988). Putting linguistics on a proper footing: Explorations in Goffman's concepts of participation. In P. Drew & A. Woottton (Eds.), *Goffman: Exploring the Interaction Order* (pp. 161–227). Oxford: Polity Press.
- Levinson, S. C. (2000). *Presumptive Meanings*. The MIT Press.
- Lewin, K. (1951). *Field Theory in Social Science*. Harper & Row.
- Lewis, D. (1969). *Convention: A Philosophical Study*. Harvard University Press.
- Locke, J. L. (2010). *Eavesdropping: An Intimate History*. OUP Oxford.
- Loehr, D. (1997, July). Hypertext and deixis. In *Referring phenomena '97: Referring phenomena in a multimedia context and their computational treatment*.
- MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4(226), 1–16.
- Madan, C. R., & Singhal, A. (2012). Consciousness and cognition. *Consciousness and Cognition*, 21(3), 1563–1570.
- Manktelow, K. (2012). *Thinking and Reasoning: An Introduction to the Psychology of Reason, Judgment and Decision Making*. Psychology Press.
- Marcus, S., & Rips, L. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior*, 18, 199–223.
- Markovits, H., & C, L. (1990). Pragmatic reasoning schemas for conditional promises: Context and representation. In J. P. Caverni, J. M. Fabre, & M. Gonzales (Eds.), *Cognitive Biases* (pp. 183–192). Elsevier.
- Marks, J., & Reiter, E. (1990). Avoiding unwanted conversational implicatures in text and graphics. In *AAAI-90* (pp. 1–7). Harvard University, Center for Research in Computing Technology, Aiken Computation Laboratory.

- Mason, H. (2012, July). Machine learning for hackers. In *Devs Love Bacon*. Retrieved from <http://www.hilarymason.com/presentations-2/devs-love-bacon-everything-you-need-to-know-about-machine-learning-in-30-minutes-or-less/>
- Mason, W., & Suri, S. (2011, June). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, 44(1), 1–23.
- Mason, W., & Watts, D. J. (2009, June). Financial incentives and the "performance of crowds". In *HCOMP '09: Proceedings of the ACM SIGKDD Workshop on Human Computation*. ACM.
- Mayer, J. (2011, August). *Tracking the trackers: The adChoices icon*. Retrieved from <http://cyberlaw.stanford.edu/node/6714>
- Mayer, J. (2012a, September 12). *Adobe's lead DNT negotiator blocked the most popular browser in the most popular server. PR: hey, it was on his own time. Wow that's lame. [Tweet]*. Retrieved from <https://twitter.com/jonathanmayer/status/244300727726510081>
- Mayer, J. (2012b, September 7). *Adobe tries to circumvent the W3C Do Not Track consensus process, black-list Internet Explorer 10 in Apache. Nice try. [Tweet]*. Retrieved from <https://twitter.com/jonathanmayer/status/244174917615108096>
- Mayer, J., & Mitchell, J. C. (2012, May). Third-party web tracking: Policy and technology. In *SP '12: Proceedings of the 2012 IEEE Symposium on Security and Privacy*. IEEE Computer Society.
- McDonald, A. M. (2010). Cookie confusion: do browser interfaces undermine understanding? *CHI Extended Abstracts*, 4393–4398.
- McDonald, A. M. (2011, April). User perceptions of do not track. In *W3C Workshop on Web Tracking and User Privacy*. Princeton.
- McDonald, A. M., & Cranor, L. (2010). Beliefs and behaviors: Internet users' understanding of behavioral advertising. *TPRC*.
- McDonald, A. M., & Cranor, L. F. (2009). An Empirical Study of How People Perceive Online Behavioral Advertising (CMU-CyLab-09-015).
- McDonald, A. M., & Peha, J. (2011). Track gap: Policy implications of user expectations for the 'do not track' internet privacy feature. *TPRC*.

- McDonald, A. M., Reeder, R., Kelley, P. G., & Cranor, L. (2009). A comparative study of online privacy policies and formats. In *Privacy Enhancing Technologies* (pp. 37–55).
- McGlaughlin, F. (2011, December). Banner Ad Design. In *Marketingexperiments.com* (pp. 1–40).
- Milne, G. R., & Culnan, M. J. (2004). Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing, 18*(3), 15–29.
- Mozilla. (n.d.). *Mozilla at a glance*. Retrieved from <http://blog.mozilla.org/press/ataglance/>
- Munro, R., & Tily, H. (2011, August). The start of the art: An introduction to crowdsourcing technologies for language and cognition studies. In *Workshop in Crowdsourcing Technologies for Language and Cognition Studies* (pp. 1–10). Boulder.
- Narayanan, A. (2011, July 28). There is no such thing as anonymous online tracking. In *Stanford Law School Center for Internet and Society Blog*. Retrieved from <http://cyberlaw.stanford.edu/blog/2011/07/there-no-such-thing-anonymous-online-tracking>
- Narayanan, A., & Shmatikov, V. (2010). Myths and fallacies of "personally identifiable information". *Communications of the ACM, 53*(6), 24.
- Nowak, S., & Rüger, S. (2010). How reliable are annotations via crowdsourcing. In *MIR '10 Proceedings of the International Conference on Multimedia Information Retrieval* (pp. 557 – 566). New York: ACM.
- Oberlander, J. (1995). Grice for graphics: pragmatic implicature in network diagrams. *Information Design Journal, 8*(2), 163–179.
- O'Connell, J. (2011, Dec 14). The Forrester Wave: Demand-side Platforms, Q4 2011 (Tech. Rep.).
- Official Journal of the EU. (2009, December). *Directive 2009/136/EC of the European Parliament and of the Council*. Retrieved from <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:0036:en:PDF>
- Ohm, P. (2010, August). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *57 UCLA Law Review 1701*, 1–77.
- OMB. (2007, May 22). Safeguarding against and responding to the breach of personally identifiable information [Memorandum for the heads of executive departments and agencies]. M-06-16. , 1–22.

- Otlacan, O. (2011). Evidon passes 50 billion impressions served, 80 million unique users. Retrieved from <http://www.adoperationsonline.com/2011/08/11/evidon-passes-50-billion-impressions-served-80-million-unique-users/>
- Packard, V. O. (1957). *The Hidden Persuaders. An Introduction to the Techniques of Mass-persuasion Through the Unconscious.* David McKay Co., Inc.
- Pandey, S., Aly, M., Bagherjeiran, A., Hatch, A., Ciccolo, P., Ratnaparkhi, A., & Zinkevich, M. (2011). Learning to target: What works for behavioral targeting. In *ACM Conference on Information and Knowledge Management (CIKM)*.
- Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5), 411–419.
- Pickering, M. J., & Garrod, S. (2003). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169 – 189.
- Pole, A. (2010, October). How target gets the most of its guest data. In *Predictive Analytics World*. Retrieved from <http://rmportal.performedia.com/node/1373>
- Politzer, G. (1981). Differences in interpretation of implication. *The American Journal of Psychology*, 461–477.
- Reips, U.-D. (2006). Standards for Internet-based experimenting. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, 49(4), 243–256.
- Reitman, R. (2012, September). A deep dive into facebook and datalogix: What's actually getting shared and how you can opt-out. In *Electronic Freedom Foundation Blog*. Retrieved from <https://www.eff.org/deeplinks/2012/09/deep-dive-facebook-and-datalogix-whats-actually-getting-shared-and-how-you-can-opt>
- Roesner, F., Kohno, T., & Wetherall, D. (2012, April). Detecting and defending against third-party tracking on the web. In *NSDI'12: Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation*. USENIX Association.
- Rosnow, R. L., & Rosenthal, R. (1976). The volunteer subject revisited. *Australian Journal of Psychology*, 28, 97-108.
- Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., & Tomlinson, B. (2010, April). Who are the crowdworkers?: Shifting demographics in mechanical turk. *CHI EA '10: CHI '10 Extended Abstracts on Human Factors in Computing Systems*.

- Rush, M. C., Phillips, J. S., & Panek, P. E. (1978). Subject recruitment bias: The paid volunteer subject. *Perceptual and Motor Skills*, 47(2), 443–449.
- Russell, B. (1905). On denoting. *Mind*.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.
- Saffer, D. (2013). *Microinteractions*. O'Reilly Media.
- Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361–382.
- Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 289–327.
- Schmid, H.-J. (2007). Entrenchment, salience, and basic levels. *Handbook of Cognitive Linguistics*, 117–138.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21(2), 211–232.
- Schwartz, A. (2009, November). *Looking Back at P3P: Lessons for the Future* (Tech. Rep.). Retrieved from https://www.cdt.org/files/pdfs/P3P_Retro_Final_0.pdf
- Schwartz, J. (2001, September 4). Giving web a memory cost its users privacy. *New York Times*. Retrieved from <http://www.nytimes.com/2001/09/04/business/giving-web-a-memory-cost-its-users-privacy.html>
- Scott, W. (1904, January). The Psychology of Advertising. *Atlantic Magazine*. Retrieved from <http://www.theatlantic.com/magazine/archive/1904/01/the-psychology-of-advertising/303465/>
- Searle, J. R. (1969). *Speech Acts*. Cambridge University Press.
- Searle, J. R. (1971). What is a speech act. *The Philosophy of Language*, 3, 1965–1996.
- Sesto, R., & Frankel, J. (2008, September). *How deep packet inspection changed the privacy debate* (Tech. Rep.). Retrieved from <http://www.bingham.com/Publications/Files/2008/09/How-Deep-Packet-Inspection-Changed-the-Privacy-Debate>
- Shankland, S. (2012). Apache web software overrides ie10 do-not-track setting. In *CNET, Internet and Media*. Retrieved from http://news.cnet.com/8301-1023_3-57508351-93/apache-web-software-overrides-ie10-do-not-track-setting/
- Sherman, E. (2008, September). Privacy Policies are great for PhDs. *CBS News Moneywatch*.

- Simms, M. (2000, January). *National opinion fact sheet: Internet use 1998-1999* (Tech. Rep.).
- Snow, R., O'Connor, B., Jurafsky, D., & Ng, A. Y. (2008). Cheap and fast—but is it good? Evaluating non-expert annotations for natural language tasks. , 254–263.
- Solove, D. J. (2008). *Understanding Privacy*. Harvard University Press.
- Soltani, A., Canty, S., Mayo, Q., Thomas, L., & Hoofnagle, C. J. (2009, August). *Flash Cookies and Privacy* (Tech. Rep.).
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition* (2nd ed.). Cambridge: Blackwell Publishers, Ltd.
- Sprouse, J. (2010). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43(1), 155–167.
- Stalnaker, R. (1999). *Context and Content*. Clarendon Press.
- Statcounter. (2012). *Global stats: Top 5 browsers on Aug 2012*. Retrieved from <http://gs.statcounter.com/#browser-ww-monthly-201208-201208-bar>
- Stenning, K. (2002). *Seeing Reason*. OUP Oxford.
- Stenning, K., Lascarides, A., & Calder, J. (2006). *Introduction to Cognition And Communication*. Bradford Books.
- Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19(1), 97–140.
- Stenning, K., & Van Lambalgen, M. (2008). *Human Reasoning and Cognitive Science*. MIT Press.
- Stiller, A., Goodman, N. D., & Frank, M. C. (2011, Jul). Ad-hoc scalar implicature in adults and children. In *Proceedings of the 33rd annual meeting of the cognitive science society*.
- Stutzman, F., Capra, R., & Thompson, J. (2010). Factors mediating disclosure in social network sites. *Computers in Human Behavior*, 27(1), 590–598.
- Tanaka, K. (1999). *Advertising Language*. Routledge.
- Tanenhaus, M. K. M., Spivey-Knowlton, M. J. M., Eberhard, K. M. K., & Sedivy, J. C. J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634.
- Thaler, R., Sunstein, C., & Balz, J. (2010). Choice architecture. Available at SSRN 1583509.
- Tietze, M. I., Winterboer, A., & Moore, J. (2009, March). The effect of linguistic devices in information presentation messages on comprehension and recall. In *Proceedings of the 12th European Workshop on Natural Language Generation*. Association for Computational Linguistics (pp. 114–

- 117).
- Title 18, Part 1, Chapter 119 §2510(12). (n.d.).
- Tribal Fusion. (2012). *AdChoices Our experience*. Retrieved from <http://www.iabuk.net/sites/default/files/Tribal%20Fusion.pdf>
- TRUSTe. (2012). *TRUSTe 2012 UK consumer data privacy study* (Tech. Rep.).
- Turow, J. (2012, November). The non-transparency of key online words. In *NTIA Privacy Multistakeholder Process Regarding Mobile Application Transparency* (pp. 1–10).
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.
- Tversky, B. (2004). Semantics, syntax, and pragmatics of graphics. In K. Holmqvist & Y. Ericsson (Eds.), *Language and Visualisation* (pp. 141–158). Lund: Lund University Press.
- Tversky, B. (2010). Visualizing thought. *Topics in Cognitive Science*, 3(3), 499–535.
- Tversky, B., & Lee, P. (1999). Pictorial and verbal tools for conveying routes. *COSIT '99 Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*, 51–54.
- Ulanoff, L. (2012, October 2). There really is a smartphone inside EW Magazine. *Mashable*. Retrieved from <http://mashable.com/2012/10/02/ew-has-smartphone-inside/>
- United States Code Title 15 §45*. (n.d.).
- van Dijk, T. A. (1990). Social cognition and discourse. *Handbook of Language and Social Psychology*, 163–183.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press.
- Vestergaard, T., & Schrøder, K. (1985a). *The Language of Advertising*. Blackwell Publishers, Ltd.
- Vestergaard, T., & Schrøder, K. (1985b). *The Language of Advertising*. Oxford: Blackwell Publishers, Ltd.
- Vogels, J., Krahmer, E., & Maes, A. (2011). How visual saliency affects referent accessibility. In L. Carlson, C. Hoelscher, & Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3128–2122). Austin.
- W3C TPWG. (n.d.). Retrieved from <http://www.w3.org/2011/tracking-protection/>

- W3C TPWG. (2011). *Tracking Preference Expression (DNT)*. Retrieved from <http://www.w3.org/2011/tracking-protection/drafts/tracking-dnt.html>
- W3C TPWG. (2012, June). *Discussion*. Retrieved from <http://lists.w3.org/Archives/Public/public-tracking/2012Jun/0257.html>
- Wason, P. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology*, 12(3), 1–7.
- Wiedenbeck, S. (1999). The use of icons and labels in an end user application program: an empirical study of learning and retention. *Behaviour & Information Technology*, 18(2), 68–82.
- Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, 31, 183–194.
- Willems, R. M. R., Labruna, L. L., D'Esposito, M. M., Ivry, R. R., & Casasanto, D. D. (2011). A functional role for the motor system in language understanding: evidence from theta-burst transcranial magnetic stimulation. *Psychological Science*, 22(7), 849–854.
- Woods, D. (2012, March). bitly's Hilary Mason on "What is A Data Scientist?". *Forbes Magazine*. Retrieved from <http://www.forbes.com/sites/danwoods/2012/03/08/hilary-mason-what-is-a-data-scientist/>
- Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., & Chen, Z. (2009, February). How much can behavioral targeting help online advertising? In *The 18th International Conference* (pp. 261–270). New York, New York, USA: ACM Press.
- Zwaan, R. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, 28, 81–88.
- Zwaan, R., & Radvansky, G. A. (2002). Situation models and themes. *Thematics: Interdisciplinary Studies*, 35–53.
- Zwaan, R., & Radvansky, G. A. (2013). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 1–24.