Is This How We (All) Do It?:
Butler Lies and Ambiguity Through a Broader Lens

Megan French[1], Madeline E. Smith[2], Jeremy Birnholtz[2], Jeff T. Hancock[1]

Communication Department

Cornell University, Ithaca, NY 14853

Department of Communication Studies

Northwestern University, Evanston, IL 60208

mrf239@cornell.edu, jeremyb@northwestern.edu, jth34@cornell.edu, madsesmith@u.northwestern.edu

ABSTRACT

The ubiquity of mobile devices has resulted in more opportunities to interact with more people than ever before. Given a finite capacity for interaction with others, people commonly manage their availability by limiting others' access to them. Prior work has demonstrated the importance of doing so in a relationally sensitive way and identified the butler lie, in which deception is used to manage availability, as a common linguistic strategy. Two key limitations of existing exploratory work, however, are limited samples of primarily students and a focus on media properties in understanding ambiguity that enables butler lies to be plausible. This paper aims to address these issues via a broad field study of deception and butler lies using a novel message-sampling method employed via a custom mobile app. Results show clear evidence of butler lies occurring in a broader population, with some gender differences; and urge adoption of a multi-level framework for understanding ambiguity that also includes private information and infrastructure-level attributes of interaction media.

Author Keywords

Texting; butler lies; availability; deception; CMC

ACM Classification Keywords

H.5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces

INTRODUCTION AND BACKGROUND

One of the most basic features of mobile phones – allowing interaction with nearly anybody at nearly any time – is both a benefit and drawback. On one hand, social coordination is easier and people can stay in touch in novel ways [7]. On the other hand, constant connection can overwhelm users with interaction opportunities and interruptions. Mobile phone users must maintain a delicate balance between the desire to avoid interaction and the possible consequences of offending their contacts [2]. To maintain this balance, people use relationally sensitive strategies, such as what

have been called "butler lies," for managing availability [3].

Several studies of butler lies have investigated their use and perception [reviewed in 2] but our understanding of this everyday behavior remains constrained in important ways. First, prior work has been exploratory in nature, relying primarily on student participants whose media use and social coordination may be different from other adults' [2].

Second, drawing on Aoki & Woodruff [1], we know that butler lies and other lies online are plausible because they reference ambiguous information that is not verifiable by the recipient (e.g., [3]). Prior work on butler lies has focused mainly on ambiguities related to media themselves, extending work by [5]. For example, the fact that mobile phone users are physically separated makes a lie about one's current location difficult to verify, whereas this lie is essentially impossible when partners are co-present. Literature on deception, however, suggests several other kinds of information that is unverifiable, particularly lies based on private information (e.g., one's feelings) [4]. A clearer picture of how people use private information in butler lies will help us understand how people draw on ambiguity to manage availability.

These open issues are of interest to the CHI community in that butler lies, and availability management strategies more generally, can influence the design of tools to provide awareness and availability information such as interruptions, read notifications, and typing-in-progress indicators. By improving our understanding of how people manage availability, we can design interaction tools that better support people's needs and behavior; and extend our theoretical understanding of how people strategically use media and language to advance relational goals.

Does Everybody Really Do it?

A consequence of prior research on butler lies being exploratory is that all participants in these studies were students at US universities. Given documented differences in deception, texting and other online communication behavior [4, 9], it is important to explore more broadly how these messages are used by different generational and demographic groups. We therefore asked:

RQ1: What is the frequency and nature of butler lying behavior in a broader population?

Ambiguous Ambiguities

Secondly, Birnholtz et al. [3] argued that butler lies are constructed by drawing on the ambiguities inherent in media, such as SMS mobile texting. These ambiguities do not refer to the content of the lie, but rather what makes the lie uncertain enough to be plausible. They presented a taxonomy of ambiguities used in butler lies that largely stem from properties of the communication medium, including: timing (e.g., when something happened), location (e.g., where a person is texting from), and activity (e.g., what the sender is doing). While these ambiguities clearly appear in butler lies, there is arguably an important omission in this taxonomy. Prior work on deception shows that lies are often about private states or information known only to the liar, such as beliefs (e.g., "I really want to see you"), preferences (e.g., "that party will be great"), and intentions (e.g., "yep I'm planning on going") [4].

This omission matters because the ambiguity of private information is distinct from ambiguity related to media. Deception about private information is plausible even face-to-face, as with "white lies." Such lies, when told in mediated interactions, may remain plausible even after the conversation is over because the information is more difficult to verify. Better understanding and capturing this distinction can shed light on how people manage availability strategically. Given the common use of private information in deception, we would expect to see evidence of this in a larger sample. We asked:

RQ2: What types of private information are the subject of butler lies and what are the sources of ambiguity?

THE PRESENT STUDY

To these questions, we introduce a novel message-sampling method using a custom-developed texting application that collected participants' text messages over seven days, and asked at send time if each outgoing message was deceptive.

Participants. We recruited people to participate in a study on text messaging via e-mail listservs, Craigslist ads in 10 cities, and a web-based recruitment system. There were 248 participants, of whom 47 were removed for insufficient activity (< 10 messages) or attempting to participate twice. The remaining 201 included 92 males and 103 females (six did not report gender), and were aged 19 to 63 (*M*=31, *SD*=8.9). Participants had, on average, 7.5 years of texting experience. Sixty-six percent indicated they were non-students, 6% part-time students, and 28% full-time students. The racial/ethnic background of participants was: White (51%), African American (19%), Asian (14%), Hispanic (9%), and Other (7%). Participants were located throughout the United States: 40% in the midwest, 22% in the west, 20% in the northeast, and 18% in the south.

Message Sampling. We developed Butler Messaging, a custom Android app that was substituted for participants' ordinary SMS app during the study. The app interfaced with our secure database, to which each sent or received message was automatically saved. To protect privacy, common proper names were removed and phone numbers were one-way hashed. To allow for send-time identification of deception, a pop-up dialog appeared when send was clicked that asked if the message was deceptive. Deceptiveness information was stored in the database, but not stored on the participant's phone or shared with contacts.

Procedure. After consenting, participants were instructed to install Butler Messaging, import their existing messages, and make it their default messaging app for seven days. Afterwards, participants were completed an online post-study questionnaire and received instructions to uninstall the app. The questionnaire presented them with specific deceptive messages they had sent, and asked them: 1) if they recognized each recipient, 2) their relationship to each contact (e.g. friend, family, etc.), and 3) closeness to each contact. For each deceptive message, participants were asked to explain why it was deceptive.

Data Analysis. Two coders independently coded all deceptive messages for butler content, jocularity (e.g., jokes or sarcasm) and messages accidentally marked as deceptive, using a prior coding scheme [3]. Next, butler lies were coded for the type of ambiguity that enabled plausibility [3]. Ambiguity types from prior work included activity (what the sender was doing), location (where the sender was) and time (when something occurred). Following RQ2, we identified two additional sources of ambiguity: private information and infrastructure, described further below. Of 401 butler lies, 40 could not be coded for ambiguity because participants did not provide enough detail. Inter-coder reliability for ambiguity was acceptable (77% agreement, Cohen's *kappa*=0.65).

RESULTS

In all, participants sent 12,587 and received 29,050 messages, exchanged with 1,661 distinct contacts.

Did Everybody Do It?

Our first research question was about the frequency of butler lies in a sample that extended beyond a predominantly student demographic. We first looked at the prevalence of general deception in our data. Of the 12,587 sent messages, 769 messages were deceptive, with a mean general deception rate of 8% (*SD*=13%). Of these deceptive messages, on average, 56% (*SD*=37%) were identified as butler lies. It should be noted that the rate of butler lies in this study is distinctly higher than past studies, which have averaged around 39% [3, 8].

Deception Across Demographics

We wondered if, in a broader sample with more messages, there were differences across demographic categories for deception or butler lie rates [see Table 1]. Looking first at gender differences using a t-test, we see that that men told significantly more lies in their text messages overall than women [*t*(128)=2.78*p*<0.01]. For butler lies data, however, the reverse was true. Women told significantly more butler

lies on average than men [χ^2= 5.3, $p<0.05$; Kruskal-Wallis test was used due to bimodal distribution of butler lies]. Using ANOVA, we examined differences across previously under-represented racial or ethnic groups. The difference in rates for butler lies across ethnic groups approached significance [$F(4,131)=2.1$, $p=0.08$].

	Deception Rate		Butler Lie Rate	
Gender	*M (SD)*	*N*	*M (SD)*	*N*
Male	0.11 (0.18)**	92	0.48 (0.35)*	63
Female	0.06 (0.08)**	103	0.58 (0.36)*	71
Age				
Under 25	0.09 (0.11)	57	0.55 (0.35)	46
25-30	0.09 (0.16)	58	0.50 (0.35)	37
31-40	0.10 (0.16)	52	0.50 (0.37)	37
Above 40	0.05 (0.10)	29	0.60 (0.44)	17
Ethnicity				
African American	0.08 (0.15)	38	0.61 (0.39)	26
Asian	0.12 (0.17)	29	0.50 (0.37)	20
Hispanic	0.10 (0.18)	18	0.26 (0.32)	12
White	0.08 (0.12)	103	0.55 (0.36)	71
Other	0.06 (0.04)	8	0.55 (0.30)	7
Student				
Full Time Student	0.08 (0.10)	55	0.54 (0.38)	44
Non-Student	0.09 (0.15)	144	0.57 (0.36)	93

Note: For butler lies only participants who told at least one lie were included, ** $p<0.01$, * $p<0.05$

Table 1. Demographic breakdown of average rate of deceptive messages and butler lies.

We believe this was driven in part by the lower rate found for Hispanic participants. However given the small sample of Hispanics, we hesitate to draw any conclusions about this marginally significant relationship.

Despite a large number of non-students, comparing across age groups, using ANOVA, and status as full-time student, using a t-test, findings revealed no significant differences for deception or butler lies.

Understanding Ambiguity
Our second research question addressed the nature of private information and the kinds of ambiguity that people drew on to make their butler lies plausible. To compare the frequency of different types of ambiguity, we ran a mixed model with the five ambiguity categories as a fixed factor and participants as a random factor, revealing significant differences [$F(4,535)=41.2$, $p<0.001$]. Private information was used most (45%), followed by activity (26%), time (15%), location (5%), and infrastructure (4%) [see Figure 1]. Pairwise comparisons, using a Least Significant Difference correction ($p<.05$), showed that private information, activity, location, and time all significantly differed from each other. There was no significant difference between location and infrastructure.

Private Information
What is most striking in these results is that the most frequent category was private information. We saw that
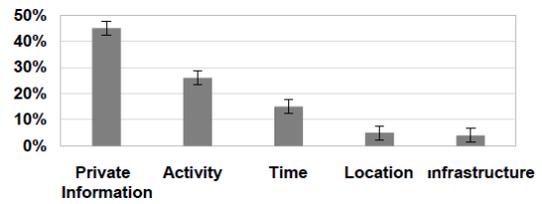

Figure 1. Estimated mean proportion of each participant's butler lies drawing on ambiguity types.

participants regularly drew on private information that was not verifiable, either because it was private to the individual (e.g., intent, feelings) or because it is not visible or normatively acceptable to verify (e.g., money troubles).

Participants drew on ambiguity from private information when being deceptive about their true feelings or intentions, such as a participant who said she was "still interested!" in a social event even when she indicated in the post-study questionnaire that she had known she was not. The same was true for knowledge of future plans. One participant – regarding his availability for later activities– said "not sure where ill be. Maybe," but told us he had knew then he would not be available.

Another use of private information was to ask questions where the answer was already known, but the question served to support conversational goals. One participant, for example, said "you working today…?", but indicated to us they already knew the recipient was not working. This is interesting, in that it does not contain false information. The participant likely rated it as deceptive because the question falsely implies that the recipient's work status is unknown.

Infrastructure
Another novel form of ambiguity we saw here was related to infrastructure. Where prior work did not draw a distinction between attributes of a medium and attributes of the underlying infrastructure, we believe the distinction captures an important phenomenon. Mobile phones have batteries that can fail, for example. Those batteries are an attribute of the phone itself, not an attribute of texting. The ambiguity stems not from the medium, but the functioning of the device or its infrastructure.

While butler lies drawing on this type of ambiguity were not frequent (see Figure 1), considering this allows us to better capture how people take of advantage of device characteristics to manage availability within different contexts. We saw people blame both the infrastructure, as with a participant who falsely said "Hey, sorry -- I was having all kinds of issues with texts going through yesterday…"; as well as the device, as with another that said "the phone keeps freezing" even when it had not.

DISCUSSION
This work has several implications for our understanding of and ability to support availability management. First, we presented Butler Messenger, a message sampling method

that allowed for in-the-moment data collection across geographic regions.

Second, in using our message-sampling method we show a clear difference in the rate of butler lies as compared with prior studies. This raises the question of whether this is due to our experiential sampling rather than retrospective survey method (e.g., [3]); or due to an actual change in texting behavior. Given the relatively short time since prior work, smaller sample and other prior methodological limitations, we argue that the present study presents a more accurate depiction of deception and butler lies in texting.

In addition to overall higher rates of butler lies, we found that gender differences in rates of deception and butler lies. Men lied at almost twice the rate that women lied. However, a higher proportion of women's deceptive messages were butler lies. There were no age, ethnicity, or student status differences for either of our dependent measures. This was surprising both in that one might expect group differences and in that it stands in contrast to work suggesting that students lie at a higher rate [4]. Our result provides some evidence that deception as a strategy for availability management is common across demographics.

Third, we build on prior work by illustrating two distinct ways in which people use ambiguity in crafting plausible butler lies. To our surprise, we found that our participants drew on ambiguity related to private information more frequently than any other type of ambiguity. This suggests that, even when interactions are mediated in ways that provide additional ambiguity, people still frequently use private information strategically, perhaps because of the difficulty that such lies could be detected even face-to-face.

We also saw participants draw on ambiguity related to infrastructure. This distinction is useful because many common communication media (e.g., Skype, WhatsApp, SMS, etc.) run on different types of devices and depend on infrastructures with varying technical properties. Where past research has focused on media features [3, 5], our work considers the reliability and attributes of the devices and infrastructures that underlie these features.

A more complete taxonomy of ambiguity types can inform designers about the potential impacts of making certain information visible to interaction partners. We urge researchers and designers to consider the multiple levels at which ambiguity occurred here in availability management and deception, including private information, details obscured by media properties, and attributes of infrastructure. With a multi-level approach to ambiguity, we can better understand the relationship between media (e.g. GlancePhone [6]) and availability management. One issue this highlights is the way that available/busy status indicators conflate private information with media features that serve to provide or obscure this information.

Limitations and Future Work

While butler lies and deception in texting were examined on a larger scale than prior studies [2, 3, 8], there are limitations to consider. Despite a broader sample, the participants in this study are still not representative of the US population. Our methodological and theoretical contributions provide a foundation for future work.

We also acknowledge the possibility that asking people about deception during the study may have altered their texting behavior. We do not believe this was the case, however, as we saw a noticeable difference only for butler lies, and not for deception overall. In addition, the overall average rate of lying was not significantly different between the first and second half of the study (M=0.08, SD=0.01; M=0.09, SD=0.01, p>0.10). An alternative approach for future work to address this issue would be to have the app query participants about a range of topics at send time (e.g., "Did this message contain humor?"), asking about deception only for a random subset.

REFERENCES
1. Aoki, P. & Woodruff, A. Making spaces for stories: Ambiguity in the design of personal communication. *Proc. ACM CHI* (2005), 181-190.

2. Birnholtz, J., Reynolds, L., Smith, M. & Hancock, J. (2013). "Everyone has to do it:" A join action approach to managing social inattention. *Computers in Human Behavior, 29,* 2230-2238.

3. Birnholtz, J., Guillory, J., Hancock, J., & Bazarova, N. "on my way": Deceptive texting and interpersonal awareness narratives. *Proc. ACM CSCW (2010)*, 1-4.

4. DePaulo, B., Kashy, D., Kirkendol, S., Wyer, & Epstein, J. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979-995.

5. Hancock, J., Thom-Santelli, J., & Ritchie, T. Deception and design: The impact of communication technology on lying behavior. *Proc. ACM CHI (2004)*, 129-134.

6. Harper, R. & Taylor, S. Glancephone – an exploration of human expression. *Proc. ACM MobileHCI (2009),* 24-32.

7. Ling, R. (2004). *The Mobile Connection: The Cell Phone's Impact on Society*. San Francisco: Morgan Kaufmann.

8. Reynolds, L. Smith, M, Birnholtz, J., & Hancock, J. Butler lies from both sides: Actions and perception of unavailability management in texting. *Proc. ACM CSCW (2013*), 769-778.

9. Zickuhr, K., & Smith, A. (2012). Digital differences. *Pew Internet & American Life Project*, *13*.