9-2015

Deception as forgery: The role of reference information in honesty and deceit

Timothy John Luke
Graduate Center, City University of New York

DECEPTION AS FORGERY: THE ROLE OF REFERENCE INFORMATION IN HONESTY AND DECEIT

by

Timothy John Luke

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the Degree of Doctor of Philosophy, The City University of New York

2015

This manuscript has been read and accepted by the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Maria Hartwig

_____ _____

Date Chair of Examining Committee

Maureen O'Connor

_____ _____

Date Executive Officer

Angela Crossman

Saul Kassin

Charles F. Bond Jr.

Marc-André Reinhard
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

DECEPTION AS FORGERY: THE ROLE OF REFERENCE INFORMATION IN HONESTY AND DECEIT

by

Timothy John Luke

Advisor: Maria Hartwig

Using concepts derived from cybernetics, self-presentation theory, and research on human self-regulation, I develop a cybernetic perspective of deception and self-presentation. In this perspective, human communication, both honest and deceptive, is controlled by feedback mechanisms. I report two studies designed to test the basic prediction derived from the cybernetic framework that deceivers are able to better emulate truth-tellers when they have access to relevant reference information about the way truth-tellers behave. Each study manipulated liars' and truth-tellers' access to reference information in a different manner. In Study 1, some participants viewed video recordings of people being interviewed in a manner highly similar to the way the participants would be imminently interviewed. I hypothesized that the exemplar videos would equip liars with reference information that would enable to them to "counterfeit" a truth-teller's behavior. In Study 2, some participants engaged in activities analogous to the task about which they were to be interviewed. Of those who completed the analogous activities, some were interviewed about those activities, and some were not. I hypothesized that liars who completed the analogous activities and were interviewed about them would gain reference information from their own behavior and be better able to counterfeit truthful responses. The results of Study 1 partially supported the hypotheses. The results of Study 2 failed to support the hypotheses, possibly because processes more complex than expected were

at work. Implications of the results are discussed in relation to the cybernetic framework of self-presentation and deception.

“There seems no reason to believe that we are ever acquainted with other people’s minds, seeing that these are not directly perceived.... All thinking has to start from acquaintance; but it succeeds in thinking *about* many things with which we have no acquaintance.”
Bertrand Russell (1905)

Acknowledgments

First and foremost, I owe thanks to my advisor *Maria Hartwig*. She is an insightful mentor, indispensable intellect, and kindred spirit.

This work would never have been completed without my research assistants, *Johanna Hellgren* and *Benjamin Shamash*. They have earned tremendous gratitude for their reliability, enthusiasm, and ability to dupe undergraduates.

For their patience and attention to detail as coders for this research, I thank *Valerie Gelo*, *Fabiana Alceste*, and *Clarissa Tse*.

I would like to thank my committee, *Angela Crossman*, *Saul Kassin*, *Charlie Bond*, and *Marc Reinhard*. Special thanks are due to Angela and Saul, who have served on the committees for each of the milestones of my doctoral studies. Their insight and knowledge has been invaluable.

The past five years would have been much more tedious and much less enjoyable without my lab-mates. I thank *Laure Brimbal*, for bringing a lot of heart to the lab and for challenging me in unexpected, delightful ways; *Evan Dawson*, for her many years of friendship and for helping me discover a love for science; and *Emily Joseph*, for being a supportive colleague, a partner in mischief, and a loyal friend.

For keeping me grounded, keeping me honest, and frequently keeping me fed, I am grateful to *Kristin Davidoff*.

For his technical savvy, for his willingness to help, and especially for being the kind of person I can talk to for hours without ever being bored, I thank *Chris Gettings*.

For giving me a gratifying life beyond work and for never asking me when I would get a real job, I thank my friends from outside of academic life.

Finally, I want to express my gratitude to my mother, *Judy*, and my brother, *Tom*, for their continued and unwavering support. They have, for a quarter of a century, endured my ramblings and idiosyncrasies with grace.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1: Introduction

HAMLET. ...Will you play upon
this pipe?
GUILDENSTERN. My lord, I cannot.
HAMLET. I pray you.
GUILDENSTERN. Believe me, I cannot.
HAMLET. I do beseech you.
GUILDENSTERN. I know no touch of it, my lord.
HAMLET. 'Tis as easy as lying: govern these ventages with
your lingers and thumb, give it breath with your
mouth, and it will discourse most eloquent music.
Look you, these are the stops.
GUILDENSTERN. But these cannot I command to any utterance of
harmony; I have not the skill.
HAMLET. Why, look you now, how unworthy a thing you make of
me! You would play upon me; you would seem to know
my stops; you would pluck out the heart of my
mystery; you would sound me from my lowest note to
the top of my compass: and there is much music,
excellent voice, in this little organ; yet cannot
you make it speak. 'Sblood, do you think I am
easier to be played on than a pipe? Call me what
instrument you will, though you can fret me, yet you
cannot play upon me.
(Shakespeare, *Hamlet*)

Guildenstern admits he cannot play a recorder, Hamlet observes, yet he has the audacity to attempt to deceive – to play – a human, a being so much more complex than the simple pipe. Contrary to Hamlet's impassioned complaint, humans in fact deceive each other with great effectiveness. Humans are poor at detecting deceit (Bond & DePaulo, 2006). Humans lie fairly regularly and remain undiscovered most of the time (DePaulo et al., 1996). Though Hamlet's indignation about the apparent inconsistency may be misplaced, he calls attention to an important characteristic of human behavior: Humans are beings of immense complexity, yet we "play upon" each other.

To effectively play each other, humans would benefit from access to the beliefs and perceptions of others. Other minds, however, are beyond direct observation. This limitation does

not seem to deter us from drawing inferences – often quite accurate – about other minds (see Epley & Waytz, 2009). Hamlet says these things with irony, but despite the difficulties, we do seem to know each other's stops; we do pluck out the heart of each other's mystery. What he scornfully dismisses – the ease of lying – impels the question of how we are able to so effectively influence something as complicated and mysterious as a human being.

The effectiveness not only of deception but other social behavior suggests that there exists accessible information about the happenings of other minds, despite the impossibility of direct observation. Without such information, it would not be quite difficult for humans to interact so smoothly – or to know what words and actions will make for convincing deceit. Some of this information may be conveyed unintentionally, but much of it is purposefully transmitted. That is, in order to facilitate social interactions, people intentionally send messages about their internal states. Others observe these intentional and unintentional messages and draw inferences from them. My concern here is how humans transmit and receive such information in order to pursue their social goals, with a special focus on deception – the transmission of false information.

Definitions

Following previous deception researchers (e.g., DePaulo et al., 2003; Zuckerman, DePaulo, & Rosenthal, 1981), I define deception as an act intended to mislead others. Communication predicated on mistaken beliefs is not considered deception under this definition. In line with DePaulo and her colleagues (2003), I use the terms lying and deception interchangeably, although some scholars have drawn distinctions between the two (e.g., Bok, 1978). In other contexts, it is useful to distinguish between deception (as a broader term) and lies

(as a more specific kind of deception), but for the purposes of this report, the two terms will be used as if they are synonymous.

It is also important to discuss what is meant by an "act" of deception. Any given act can be construed in many ways (see Vallacher & Wegner, 2012). For the purposes of the present theory, the term "deceptive act" encompasses a collection of behaviors that communicate false information and are aimed at pursuing a given goal. Thus, under these terms, a single deceptive sentence (e.g., "No, I didn't break the vase.") is part of a deceptive act but typically does not constitute the complete act. Rather, the complete act includes all the behaviors enacted by the actor in order to facilitate the intended false impression (e.g., facial expressions, gestures, and follow-up statements supporting the impression).

Throughout this paper, I use the term "actor" or "social actor" to refer to person who is interacting with another person or group of persons, whom I call an "audience." I use these terms as a matter of convenience, to clearly identify the person in whose communicative acts I am primarily interested (i.e., the actor) and the recipient of those communications (i.e., the audience), but in reality, both parties qualify as both actor and audience, as they are both acting in ways that influence each other.

Introduction to the Theoretical Framework

The perspective that informs the work presented here is built on three broad theoretical foundations: a theory of the transfer and use of information, related theory of human goal pursuit, and theory of how humans manage information conveyed to others. Cybernetics, a theory of how information is transferred and how it is used to control actions, serves as the first foundation. A cybernetic conceptualization of human self-regulation and goal pursuit serves as the second

foundation. The self-presentational perspective, which emphasizes that people behave in ways that strategically convey information about themselves, serves as the final foundation.

This section will be followed by a broad review of research and theory of self-presentation and deception. Thereafter, relevant cybernetic concepts will be introduced, followed by their broader application to human behavior and their more specific importance to self-presentation and deception. Applying cybernetic concepts to the analysis of self-presentation is not novel in itself (cf. Bozeman & Kacmar, 1997; Cody & McLaughlin, 1990), but I am familiar with no work that has dealt thoroughly with the specific implications of cybernetic mechanisms (e.g., feedback control) for self-presentation and deception.

Additionally, though there have been previous theories which have discussed the relationship between deception and self-presentation and theories that have emphasized the socially interactive context in which deception typically occurs, the present theory differs from previous theories of deception in its emphasis on the goals for which deception is enacted, how these goals influence communicators' use of social information and self-presentational behavior, and the mechanisms by which social goals are pursued. For example, Interpersonal Deception Theory (Buller & Burgoon, 1996) makes specific predictions about how social interactions can influence deceptive behavior, but it does not propose psychological mechanisms underpinning those predictions (see DePaulo, Ansfield, & Bell, 1996). In contrast, the present theory is explicit about the hypothetical structures and processes through which its predictions operate (namely hierarchically organized feedback control). Though the finer details of the present theory will, of course, be expatiated below, its basic premises can be summarized as follows:

Adaptive goal pursuit depends on gathering information about the present situation and acting in order to change the present state to a desired state[1]. Observable attributes and behaviors carry information, from which observers draw inferences about situations and persons. Observers can then compare the inferred present state to their desired state. This comparison guides the observers' behavior in the pursuit of their goals.

Humans live in a highly social world, and they are not ignorant of the fact that others gather information from them. Thus, they attempt to control the transmission of information about themselves (and others) through self-presentation. Much of the time, this control of information is cooperative and involves the transmission of accurate, though perhaps edited, information. Sometimes, however, people deceive, and in doing so, they attempt to transmit false information to others. This attempted control of information – whether honest or deceptive – is part of the larger enterprise of human goal pursuit, and therefore, the same principles that apply to goal pursuit as a whole apply to self-presentation and to deception.

The cybernetic structures and processes that underpin goal pursuit create a dynamic relationship between a social actor and that actor's audience such that information from both parties is transferred and received. Social actors use information from their audience's behavior, as well as past knowledge, to infer what kind of self-presentation is likely to be adaptive. This inference sets a standard for behavior and guides the actor's behavior, to the extent that the actor compares his or her performance to the standard. Two general predictions emerge from this reasoning. First, social actors will adjust their self-presentations in accordance with the information that is available about the audience. This first prediction is not unique to the present theory and is consistent with past research and theory on self-presentation (Schlencker, 2012).

[1] I use the term "present state" to refer the current perceived situation at a given moment. I use the term "desired state" or "desired outcome" to refer to the mental representation of a situation that a person is attempting to bring about.

Second, as valid information about an audience's beliefs and goals becomes increasingly available, the more optimal a given social actor's self-presentation will be. This pattern is qualified by the extent to which the social actor (1) receives and incorporates that information into his or her internal standards for behavior and (2) compares the present state to the desired outcome, in line with the standard for behavior.

In the present perspective, deception is regarded as a type of self-presentation in which a social actor knowingly transmits false information. Deviating from the truth typically involves contrasting the perceived reality with a preferred alternative. Thus, as a general principle, compared to honesty, deception involves (1) greater attention devoted to comparison of the present state to the desired outcome and (2) more precise standards of behavior. That is, generally speaking, when a person intends to deceive rather than tell the truth, he or she more actively compares the present state, in which an intended impression has not yet been created, to an imagined desired state, in which the audience has successfully formed the impression.

As a function of their higher degree of deliberateness and more specified standards for behavior, deceivers tend to respond differently to difficulties in self-presentations than truth-tellers. More specifically, liars may think they are less likely to be believed and, to compensate, exert more effort in their self-presentation and devote greater attention to detecting information relevant to the deceptive self-presentation. There are, however, exceptions to these predicted patterns, which will be discussed below. Moreover, there are numerous other predictions that can be drawn from the reasoning of the present theory. The general predictions above as well as other, more nuanced, hypotheses will be elaborated in the sections that follow. Finally, I present two studies which serve as a first step in empirically investigating predictions that emerge from

this theoretical perspective. Although this theoretical perspective has numerous and diverse implications, here I focus on those that are relevant to the reported studies.

Self-Presentation

Liars and truth-tellers, as Goffman (1959) wrote, "both must take care to enliven their performances with appropriate expressions, exclude from their performances expressions that might discredit the impression being fostered, and take care lest the audience impute unintended meanings" (p.73). Self-presentation comprises people's attempts to control others' perceptions of them. Such attempted control occurs ubiquitously in human communication – whether it is honest or deceptive (Schlenker, 2012). Self-presentation is not, however, an end in itself. Rather, self-presentation is enacted to pursue a purpose – to serve a goal. An instructor teaching a class does not tell jokes as an end in itself but rather to make a lecture more interesting and to be perceived more positively by students. Additionally, because self-presentation is inherently social, it entrains a transaction between the actor and the audience. As Schlenker (2012) writes, "[self-presentation] is not purely an expression of self, purely a role-played response to situational pressures, or purely conformity to the identity expectations of salient others" (p.548). The actor and the audience jointly determine what kinds of self-presentations are desirable.

It is important for the theoretical perspective developed here to consider what is necessary for an identity claim to be desirable. Below, I review this aspect of self-presentation. Later, self-presentation will be discussed in relation to cybernetic theory.

Desirable Identities

In the transactional view of self-presentation (Schlenker, 2012, 1985; Schlenker & Weigold, 1992), there are two features that make an identity claim desirable. First, a claim should be *beneficial*; it should facilitate other goals of the claimant. Second, an actor cannot

simply make whatever claims suit his or her objectives; the claim must be *believable*. A desirable claim must be able to be justified by evidence available to the audience.

There is ample evidence that people prefer to present themselves in a manner that is beneficial to their interests (see Baumeister, 1982; Schlenker & Weigold, 1992). People are generally motivated to view themselves positively (Alicke & Sedikedes, 2011; Sedikides, 2012), and when presenting themselves, they will often claim positive attributes (Baumeister, 1982; Jones, 1990, 1964) or associate with others who possess admirable traits (Cialdini, Finch, De Nicholas, 1990). Self-promotion is not beneficial in all social contexts, however. Under such conditions, people will occasionally claim negative attributes, for example, when an impression of modesty would be advantageous or when presenting oneself as weak might be useful in soliciting assistance (Baumeister & Scher, 1988; Jones & Pittman, 1982; Schlenker & Weigold, 1992).

In the same way that running is not invariably the best method of transit depending on the terrain and the distance, as Schlenker and Weigold (1992) state, "[t]here is no single best image to cultivate in optimizing power or achieving goals" (p.143). In other words, the social situation determines what identity claims are potentially beneficial. Again, self-presentation represents a transaction between the actor and the audience; self-presentations tend to be beneficial when they facilitate not only the actor's goals but also those of the audience. For instance, portraying oneself as especially competent can be beneficial in a job interview (C. A. Higgins & Judge, 2004), but self-promotion can backfire when it makes the actor appear self-absorbed (Schlenker, Pontari, & Christopher, 2001). Similarly, under some circumstances, offering excuses for errors can be effective in avoiding social penalties (Snyder & R. L. Higgins, 1988), but excuses can also make the actor appear reckless or dismissive of others' concerns (Pontari, Schlenker, &

Christopher 2002; Schlenker, Pontari, & Christopher, 2001). Self-protective and self-promoting performances cease to be beneficial – and become counterproductive – when they suggest traits that audiences find unappealing. In sum, people attempt to adjust their identity claims to the social context in order to facilitate their goals.

In order to be beneficial, an identity claim must also be believable. The failure to make believable claims can have deleterious consequences on social relationships when others discover that one's identity claims are false. To be believable, one must make claims that are justifiable with the information available to the audience. There is evidence from research that people strive to make their self-presentations believable (see Schlenker, 2012; Schlenker & Weigold, 1992). For instance, when the audience is aware of unflattering information about the actor, he or she will often attempt to self-promote in domains about which the audience has no information (Baumeister & Jones, 1978). Sometimes, people will strategically alter the situation to make a beneficial claim more believable. It is difficult, for example, to believably claim that one is competent when one is also performing poorly. In such cases, people will sometimes sabotage their own efforts in order to provide an explanation for the quality of their performance when they anticipate they might do poorly (Baumeister & Scher, 1988; Berglas & Jones, 1978). For example, an athlete might exaggerate the severity of a minor injury (e.g., limping visibly or writhing dramatically) as a means of publicly justifying poor performance. People will also occasionally intentionally perform more poorly than they are in fact able to, in order to reduce the audience's expectations (Baumeister et al, 1985).

Although all self-presentation involves editing of reality (as perceived by the actor), far from all self-presentation is deceptive. Making claims that are inaccurate or implausible can be damaging for social relationships in the long term, even if such presentations are successful in

the short term. For this reason, people will generally adopt self-presentations that are reasonably accurate, though positively slanted (see Sedikedes, 2012). However, there are times when reality – even edited reality – is not beneficial to one's goals. In such cases, one might distort, conceal, or fabricate a fictitious version of reality that does facilitate one's goals. When deceiving, one is still motivated to adopt a desirable identity – one that is beneficial and believable. Deception, therefore, represents a special case of self-presentation in which an actor knowingly misrepresents reality.

Deception

In the self-presentational perspective of deception, people deceive for the same or highly similar reasons that people enact other self-presentational behavior (viz. to pursue social goals). In this way, deception is mundane. Deception, as defined here, is a class of self-presentation in which the claim being made is not merely an edited version of the truth, but is known by the communicator to be misleading. For the purposes of developing the present theory, it is important to consider the reasons for which people deceive and the psychological and behavioral differences between lying and telling the truth.

Nearly everyone lies on a fairly regular basis (Kashy & DePaulo, 1996). Most lies are relatively minor, and people frequently lie about their feelings, thoughts, opinions, and achievements in order to leverage material and psychological benefits (DePaulo, Kashy, Kirkendol, & Wyer, 1996; Kashy & DePaulo, 1996). In the same way that honest self-presentations are often enacted to the benefit of others in addition to the self (Schlenker, 2012), not all lies are selfish; some lies are told to spare people's feelings and to maintain smooth relationships (DePaulo, Kashy, Kirkendol, & Wyer, 1996).

Differences between lying and truth-telling. Despite the widespread stereotype of liars being fidgety and gaze-aversive (Global Deception Research Team, 2006), because truth-tellers and liars are both actively engaged in some level of self-presentational activity at all times, the differences in their behavior tend to be minute (DePaulo et al., 2003). The smallness of these visible differences makes the task of detecting deceit particularly difficult (Hartwig & Bond, 2011). DePaulo and her colleagues (2003) have argued that any behaviors that differentiated truth-tellers from liars are not cues to deception *per se*; rather, they are cues to the underlying processes of deception. In a recent meta-analysis, Hartwig and Bond (2014) found that the detectability of lies remains stable across many situational variables (e.g., level of motivation, content of deception). One possible explanation for this stability is that both liars and truth-tellers are affected in similar ways by the situational factors they examined (Hartwig & Bond, 2014; Vrij, 2006). If this is the case, within a given situation, differences between honest and deceptive behavior would tend to be small. In what ways, then, do lies and truths differ psychologically, if at all?

In their discussion of the self-presentational perspective of deception, DePaulo and her colleagues (2003) succinctly described what they termed the *deception discrepancy*: "[The] vast diversity of lies is united by a singe identity claim: the claim of honesty.... The important difference between the truth teller's claim to honesty and the liar's is that the liar's claim is illegitimate" (p.77). DePaulo and her colleagues (2003) identified two implications of the deception discrepancy that led to predictions about how the behavior of liars differs from truth-tellers'. Specifically, (1) liars embrace their self-presentations to a lesser extent than truth-tellers, and (2) liars are more deliberate in their self-presentations than truth-tellers. From the implications of the deception discrepancy, DePaulo and her colleagues (2003) argue, it is

possible to predict what behavioral differences between liars and truth-tellers may exist – small as those differences may be.

The deception discrepancy changes the conditions under which liars and truth-tellers pursue their goals. These differences imply there are self-regulatory differences between lying and truth-telling (cf. Hartwig, Granhag, Strömwall, & Doering, 2010). The tasks involved in knowingly misrepresenting reality and in benignly editing reality overlap substantially, but there are challenges that are more or less likely to appear in one kind of distortion of reality compared to the other. The present theory applies the cybernetic framework in an attempt to explicate these hypothetical differences between lies and truths.

Cybernetics and Self-Regulation

The term *cybernetics* was coined by Norbert Wiener (1948/1961) to describe "the field of control and communication... in the machine or in the animal" (p.11). The theoretical framework of cybernetics is highly abstract and is purposed to describe how communication, feedback, and control mechanisms function regardless of the specific content of the messages or means by which information is transferred. Cybernetics is not only concerned with the transmission of information but also the ways in which systems can use information in order to pursue intended states – that is, the way systems act purposefully. Cybernetics attempts to describe how systems receive information about their environment, use that information to guide their behavior, and produce output to influence their environment.

Because the theory of cybernetics is abstract and largely free of content, it has applications across a wide range of disciplines, including engineering (e.g., Clark, 1996) and the social sciences (e.g., Cody & McLaughlin, 1990). Psychologists have applied cybernetic concepts to human goal pursuit (e.g., Carver, 1979; Carver & Scheier, 2011, 1998, 1981; Miller,

Galanter, & Pribram, 1960; Powers, 1973a, 1973b, 1998; Powers, Clark, & McFarland, 1960a, 1960b). In the present analysis, I rely heavily on the work of psychological theorists who have used cybernetics to understand the processes through which humans work toward their goals.

The cybernetic concept of feedback is particularly important to the present theory and to the studies presented here. Below I review the concept of feedback in relation to human self-regulation. I then apply the concept to an analysis of self-presentation and deception.

Feedback

Feedback occurs when there exists "circularity of action between the parts of a dynamic system" (Ashby, 1956, p. 53). The cybernetic concept of feedback is broadly applicable to physical machinery as well as human behavior (Ashby, 1956; Wiener, 1948). Feedback forms a loop of action in which the behavior of a system is adjusted in order to reduce discrepancy between the present state and a reference value (or reference standard). Stated differently, a feedback loop is a system for controlling behavior in order to pursue and maintain particular conditions.

A feedback loop comprises four components: (1) an input function, (2) a reference value, (3) a comparator, (4) and an output function. Input entails the receipt of information and sensing the current state of relevant variables. The comparator detects discrepancies between the input and the reference value, which is a desired outcome or standard. Output is the behavior of the system, which is purposed to alter the values of the relevant variables and bring them closer to the reference value. These components work together to sense and alter the environment of the feedback system. Feedback systems that are organized to reduce discrepancies with a reference value (i.e., negative feedback loops) are sometimes known as purposive loops, as they facilitate movement toward goals. Figure 1 illustrates the basic structure of a negative feedback loop. In

order for a feedback loop to function properly, it must be possible to sense discrepancies with the reference value. That is, there must exist information about deviations from the state intended by the feedback loop, and the system must be able to detect that information. Additionally, the feedback system must be capable of producing output to influence variables relevant to the intended state.

The thermostat is a classic example of a discrepancy-reducing negative feedback system. The thermostat is set to a particular temperature (i.e., the reference value). The device senses the present temperature (i.e., the input function) and determines whether it is different from the reference temperature (i.e., the comparator). If the temperature does not match the reference temperature, the thermostat activates a heater to adjust the temperature (i.e., the output function). If there is no discrepancy, the heater is not activated. Some events in the room (e.g., people entering or leaving the room, opening doors and windows) may alter the temperature. These events, called disturbances, produce discrepancies between the current and desired temperature. The thermostat continues to sense the present temperature, comparing it to the reference temperature, and activating the heater if any discrepancy exists. The feedback loop of the thermostat thus attempts to keep the temperature of the room at the desired level.

Applications in human self-regulation. Feedback loops can be applied to explain goal-directed human behavior. Carver and Scheier (2011, 1998, 1981) have argued that people's goals specify the reference values of feedback mechanisms that regulate their behavior. In their analysis, they draw a distinction between approach goals and avoidance goals. An approach goal is one that involves pursuing an objective – reducing the discrepancy between the present state and a desired state. Approach goals thus entail negative feedback loops. However, in Carver and Scheier's analysis, avoidance goals involve increasing distance (literal or figurative) from an

outcome one wishes to avoid – increasing the discrepancy between the present state and an undesired state. Avoidance goals thus entail positive feedback loops.

A positive feedback loop is an "open loop" in which certain kinds of input trigger the system to generate output purposed to enlarge the discrepancy with an undesired reference value. In such an open loop, however, the state of the system goes unchecked following the output. This lack of subsequent comparison creates a theoretical problem: Once a positive feedback loop is initiated, it should continue to act until something outside the system intervenes to interrupt the loop. Carver and Scheier (1998) resolve this problem by arguing that for avoidance goals, positive feedback loops are governed by negative feedback loops (viz. approach goals), which keep them from running out of control. This theoretical resolution has received criticism from E. T. Higgins (2012), who takes issue with the asymmetrical conceptualization of approach and avoidance goals as well as the cumbersome resolution of the positive feedback loop problem.

Similar to Carver and Scheier's conceptualization of approach and avoidance goals, Higgins (2012, 1997) argues that goals can have different regulatory foci, either oriented toward promotion or prevention. People tend to pursue promotion-focused goals by actively seeking matches with the desired outcome. In contrast, people tend to pursue prevention-focused goals by actively avoiding mismatches with the desired state or outcome.

Although Higgins does not take a control theory or cybernetic perspective as Carver and Scheier do, Higgins's notion of prevention-focused goals seems compatible with an alternative cybernetic conceptualization of such goals: It is possible that avoidance/prevention-focused goals are pursued with a single negative feedback loop with a reference value of stability or security. Mismatches with the reference value are actively avoided.

For example, a person who perceives the security of her home (the reference value) to be threatened by an impending hurricane (input indicating discrepancy with reference value of security) may nail wooden boards over the windows and hoard a supply of fresh water (output designed to reduce the discrepancy), but she will cease to make emergency preparations once she perceives that her actions have sufficiently ameliorated the storm's threat to her safety (comparison of current state with reference value). Such behavior would be organized with a focus on prevention: to avoid damage to her home and to avoid depletion of a water supply.

For the purposes of the present perspective, it makes little or no difference whether the avoidance/prevention mechanism is governed by a single feedback loop or by two loops (one positive and one negative). Although the issue of the control mechanism underpinning such goals remains open, in the interest of parsimony, I will assume that avoidance/prevention goals are controlled by negative feedback loops. This conceptualization avoids the cumbersome necessity of invoking open loops (for a presentation of a general feedback control theory of human behavior that does not use open loops, see, e.g., Powers, Clark, & McFarland, 1960a, 1960b).

Problems of control. Feedback control systems are rarely, if ever, perfect in actual practice. Control can vary in the extent to which it is "sloppy or "tight" (Carver & Scheier, 1998). The self-regulation of human behavior especially, due to its many complexities, is frequently loosely controlled and imperfect. Problems of control can be localized to any part of a feedback loop. That is, problems can exist with the input, the comparator, the reference value, or the output. In this section I describe some of the many potential problems that may beset feedback systems, but this summary is not necessarily exhaustive (for a more thorough review of feedback systems, see, e.g., Åström & Murray, 2012).

Input. It is possible for input to be (1) discontinuous, (2) ambiguous, or both. If input is discontinuous (viz. input occurs in intervals, rather than continuously; there is lag time between inputs), then a feedback system may continue to produce output beyond the extent to which it is appropriate. This can result in overshooting the target. To overcome this problem, a feedback loop would have to be "designed" to account for this lag time, producing a measured amount of output calibrated to the lag time (for general discussions of this problem, see Åström & Murray, 2012; Clark, 1996).

Input can also be ambiguous. That is, input does not necessarily contain an optimal amount of information about progress toward the goal. It is not always possible to draw accurate inferences from available information (Brunswik, 1943; Tolman & Brunswik, 1935). Under some conditions, it may be possible to elicit more diagnostic input (e.g., an instructor asking probing questions of her students to infer whether they have understood the lesson), but this is not always tenable. In the case of perceiving other minds, for example, it is impossible to gather direct information about relevant variables (e.g., others' perceptions), so one must rely on other signals and draw inferences about what is occurring in the minds of others (e.g., you see a person smile warmly as he greets you and interpret it to mean that he approves of you).

Comparator. A feedback loop can vary in the precision with which it can detect discrepancies with the reference value. As Carver and Scheier (1998) note, a feedback loop capable of accurate error detection is able to react with relatively precise output aimed at reducing discrepancies. A loop with poorer error detection will be more variable and less precise in its responses to discrepancies.

In human self-regulation, the extent to which people are actively comparing the present state to a desired reference value can vary (Oettingen, 2012). Oettingen and her colleagues have

led a research program aimed at examining the self-regulatory consequences of actively comparing the present state to a desired state. Generally, these studies involve a direct manipulation of the comparison process (i.e., participants are instructed to compare present and desired states), but Oettingen (2012) argues that in the real world, there are probably signals that trigger people to engage the comparison process naturally (e.g., negative mood).

This program of research on the comparison process has found that, broadly, self-regulation is more effective when the comparison process is more engaged and thus can more accurately detect errors. For example, people are better able to detect obstacles to goal-pursuit when they are actively comparing the present state to a reference value (Kappes et al, 2013; Oettingen, Hönig, & Gollwitzer, 2000). Moreover, people better account for the likelihood that their efforts will succeed when the comparator is more engaged (Oettingen, 2000; Oettingen, Pak, & Schnetter, 2001).

Reference value. The reference value with which a feedback loop attempts to reduce discrepancies can vary in the extent to which it is well-defined. When reference values are poorly defined, effective self-regulation becomes more difficult. In these cases, it may not be clear what output must be produced in order to reduce discrepancies, and discrepancies between the present state and a desired future may not be detected. Thus, better defined reference values are generally useful for self-regulation. For example, if a student is preparing for an exam, she benefits from knowing the specific content of the exam. Knowing which lecture notes and textbook chapters to review would help guide pursuit of the goal of succeeding on the exam.

When people fail to imagine desired future end-states, they fail to produce instrumental means of pursuing their goals (Oettingen, Hönig, & Gollwitzer, 2000; Oettingen, Pak, & Schnetter, 2001). One way of making relevant standards more accessible is by directing attention

toward the self (Carver & Scheier, 1981). This activation of relevant standards facilitates the process of comparison between the present state and the reference value, generally making self-regulation more effective (Carver, 1979; Carver, Antoni, & Scheier, 1985; Carver, Peterson, Follansbee, & Scheier, 1983; Carver & Scheier, 1998, 1981; Scheier & Carver, 1983).

Output. Many problems with output can be construed as consequences of problems with other parts of the feedback system. If there is substantial lag time between inputs, for example, then the output of the system may be miscalibrated. It is, however, also possible for the other parts of a feedback loop to be functioning optimally while the output is not. Such cases can occur when it is not possible for the system to produce well-adjusted output. For example, consider a person who does not speak the language of the people to whom he must communicate a complex message. Even if he knows precisely what information must be communicated, he cannot produce speech (viz. output) that will facilitate the others' understanding.

Feedback control in self-presentation and deception

In the sections below, I discuss the application of feedback control to the analysis of self-presentation and deception. I begin by developing a general model of how feedback may function in self-presentation. After the general model has been introduced, discussions of how feedback pertains to deception follow.

Beneficial identities and the interdependence of actor and audience. I propose that self-presentation, like other goal-directed behavior, is controlled by feedback loops. Specifically, this feedback system is designed to establish and maintain a desirable identity. In order to effectively construct a beneficial and believable identity, a social actor must have a mental representation of the identity being claimed. The intended identity defines the reference value against which the present situation can be compared in a feedback loop. The output of the

feedback system is the behavior of the actor, which is intended to make others perceive him or her in the desired way. The behavior of the audience and the social environment produce the input in this feedback system. This behavior indicates – more or less accurately – the audience's perceptions of the actor. Self-presentation is not unidirectional, however; all parties of an interaction engage in self-presentation. Therefore, self-presentation entails the linkage of two (or more) feedback loops. The output of each feedback loop contributes to the input for the other(s). Figure 2 illustrates a generalized sketch of these linked self-presentational feedback systems.

How precisely, however, do people know what identity they should cultivate? In other words, how do people know what is desirable in a given context? Knowing what is desirable is necessary for specifying the reference value of an actor's feedback system. As previously discussed, what is desirable in one context is not necessarily desirable in another. Therefore, it is critical that an effective social actor collects information about what self-presentations will be beneficial and believable in whatever specific situation she finds herself. The goals and intentions of the others in the interaction can inform a social actor what is desirable.

Perceiving what is desirable. Drawing inferences about others' intentions and goals, however, is not merely a matter of perceiving observable behavior. Rather, inferences about intentions are inferences about other minds. The states and structures of others' minds are strictly unobservable, yet humans quite readily draw inferences about what is happening in each other's minds, and these perceptions of other minds guide people's causal understanding of events (Epley & Waytz, 2009; Wegner, 2002). That is, people make sense of observable behavior by inferring that the cause of the observed behavior was a mind – that the behavior was enacted in the pursuit of some goal held by that other mind (e.g., Heider & Simmel, 1944). Despite the logical leap from observable behavior to mental states involved in these inferences, mind-reading

in this way can guide behavior adaptively. Knowing others' goals and intentions can help inform people how to act in order to obtain desired outcomes for both themselves and others (e.g., Epley, Caruso, & Bazerman, 2006; Galinsky, Maddux, Gilin, & White, 2008; Trötschel et al., 2011). After all, knowing what is socially desirable requires knowing what other people want.

The quality of input for the feedback system controlling self-presentational behavior of a given social actor, then, is constrained by the extent to which the behavior and observable attributes of the others in the interaction and attributes of the situation provide valid information about the minds of the others in the interaction (viz. their goals, intentions, and beliefs). This information must then be received and accurately interpreted by the actor (cf. Brunswik, 1943). The fact that social interactions do not constantly collapse into utter chaos suggests that humans are able to perform this complex task with some degree of effectiveness.

Two broad sources of such information are the audience's behavior (past and present) and the social situation. Perhaps the most important source of information about what is desirable is the behaviors and traits of the audience of the self-presentation. Goffman (1959) observed that

> [w]hen an individual enters the presence of others, they commonly seek to acquire information about him or to bring into play information about him already possessed.... Informed in these ways, the others will know how best to act in order to call forth a desired response from him. (p.13)

Generally, an identity claim will be beneficial if it advances the goals of both the actor and the audience (Schlenker, 2012; Schlenker & Weigold, 1992). That is, if a performance facilitates the goals of both the actor and audience, there will be an incentive for the audience to cooperate with the actor's self-presentation.

In order to determine what identities are beneficial in a given context, people must, then, draw inferences about others' goals and intentions. People must be able to read each other with some degree of accuracy. What a person says and does may indicate his or her objectives and intentions. Indeed, research suggests that people are often adept at drawing inferences about people's knowledge and intentions from their verbal behavior (see, e.g., Goodman & Stuhlmüller, 2013), and people can rapidly draw relatively accurate inferences about intelligence, personality, interpersonal expectancies, and other social information from thin slices of behavior (Ambady & Rosenthal, 1992; Borkenau et al., 2004; Fiske & Taylor, 2013). Moreover, people readily use past information and expectations to guide their interactions with others (Snyder & Stukas, 1999). Though it is certainly the case that the inferences drawn from small amounts of information about others are not unerringly accurate, these inferences are adaptive and help in making interactions smooth (see Jones, 1990; cf. Gigerenzer & Todd, 1999).

In addition to the behavior of the audience, the social situation itself may provide relevant normative information about what an actor's goals are (cf. Aarts & Dijksterhuis, 2003; Jones, 1990; Jonas & Sassenberg, 2006). For example, consider two people on a date. Without any information about the actors' specific behaviors, the norms of the situation imply that the pair will likely be interested in cultivating likeable impressions and assessing each other for romantic compatibility. The actors can use this information to draw inferences about each others' goals and to interpret each other's behavior.

Most, if not all, social behavior involves self-presentation, including the behavior of the audience. It cannot be assumed that the audience's behavior, at face value, provides an undistorted representation of the audience's perceptions. After all, the audience is undoubtedly engaging in self-presentation as well as the actor; their behavior will likewise be edited to service

their goals. Therefore, both audience and actor must assess their performances using information that is not necessarily diagnostic (nor even necessarily intended to be diagnostic) of their intentions and attitudes. This condition imposes limits on the quality of the input of all parties' feedback systems. On a first date, two people who are mutually attracted to each other and motivated to present themselves as maximally appealing may give each other nondiagnostic feedback by, for instance, pretending to be more interested in the story the other person is telling than he or she in fact is.

An audience and an actor would need to know each others' intentions in order to optimally present themselves. A cooperative audience is not likely to outright inquire about the actor's self-presentational purpose, since in most cases, it could be disruptive to the social transaction (for a discussion of such disruptions, see, e.g., Goffman, 1959). Similarly, an actor cannot always gather direct information about an audience's goals and preferences. Thus, people must tacitly draw inferences about others' goals and must act accordingly.

In most cases, an effective social actor likely behaves in a way that makes her intentions reasonably transparent or, at least, does not make an active effort to make her intentions opaque. Indeed, social interactions are made smoother when the intentions of each actor involved are at least partially discernable. For example, acts of ingratiation are most effective when their intention is somewhat transparent (Gordon, 1996). Goffman (1959) noted that people generally assist each other in their self-presentational efforts. Empirical research suggests that this is indeed the case. For example, people alter descriptions of their same-sex friends in order to make better impressions of them for attractive members of the opposite sex (Schlenker and Britt, 1999). People also assist new acquaintances in making good first impressions with others (Schlenker, Lifka, & Wowra, 2004).

In addition to discerning what self-presentations would be beneficial in a given situation, social actors must also determine what self-presentations will be believable. In the same way that cues in the environment provide information about what is beneficial, there may be indications of what performances might be credible to an audience. Some research has demonstrated that deceivers are sensitive, for example, to indicators of how likely they are to be caught. Specifically, research on diving (viz. exaggerating the effect of a challenge or feigning injury) in association football suggests that players are less likely to take a dive when they are physically close to a referee and therefore more likely to be detected (David, Condon, Bywater, Ortiz-Barrientos, & Wilson, 2011).

Adjustment of performances in response to feedback. In summary, people can and do collect information about the intentions of others from their behavior and from the social environment. This information, then, can be used to inform a social actor as to what identities may be desirable. The information collected from persons and environments serves to define the reference value for the feedback system. It is possible that the input collected from other people and from the environment defines and adjusts the reference value through a secondary feedback process (Carver & Scheier, 1998). In addition to primary output purposed to reduce discrepancies, the system would produce secondary output to define and adjust the reference value to better facilitate one's goals. Figure 3 illustrates the action of this secondary process. This process of adjustment of the reference value in response to input is crucial to the hypotheses of the studies presented below.

This model captures the transactional nature of self-presentation and portrays the actor and audience as interconnected. They influence each other's behavior and together determine what is desirable in a given context. Each person in the interaction uses the behavior of the

other(s) (viz. their output) as a cue (viz. input) to his own performance and to what is desirable. The inferred desirable identities define and adjust the reference values of the actor's and audience's feedback systems, and they produce output in order to reduce discrepancies with the intended identity. These outputs, again, provide the other person(s) with information about their performance; the feedback loops continue to operate to reduce discrepancies and to adjust the relevant reference values. In an idealized interaction, each person provides the other(s) with enough information, through his behavior, to infer his goals, intentions, and attitudes. Equipped with this information, each person in the interaction can cooperate to maintain their self-presentations and keep the interaction smooth.

However, when the social situation and the other(s) with whom one is interacting provide insufficient information about what is expected or desired, self-presentation becomes more challenging. These insufficiencies represent poor input for the feedback system and poor information with which to define and adjust reference values. Moreover, it is important to remember that the audience of social behavior is engaged in self-presentation as much as the actor is. The audience's behavior is an edited performance of their actual experience. Because the audience's self-presentational output serves as the input for the actor (and vice versa) the accuracy of the actor's perception of his or her performance is necessarily constrained by the extent to which the audience's behavior represents their actual goals and impressions.

Suboptimal input like this would appear in interactions in which one or more of the interactants is not cooperating (or outright competing) with the social goals of the others or when one or more of the interactants is engaged in deception. Generally, the more the audience's behavior deviates from the truth, the less informative the input for the actor will be. When the input is less informative, the actor is less able to adaptively adjust his or her performance.

However, it may be possible for an actor to receive input from an audience and for the actor to recognize that input is a distorted version of reality. Recognizing the input as distorted, an actor may be able to adjust his or her performance in order to account for the distortion, assuming the actor knows in what way the input is distorted. The ability to "see through" a performance is determined by (1) the extent to which cues to the unedited situation exist in the edited performance and (2) the ability of the actor to detect and use such information (cf. Brunswik, 1943; Tolman & Brunswik, 1935).

To illustrate, imagine a prankster performs a practical joke on her friend, intending the joke to be funny for both parties. It unfolds that the joke deeply offends the friend, who attempts to downplay how upset she is in order to spare the prankster some guilt. The prankster, however, perceives that her friend is actually more distressed than she is openly displaying, and the prankster therefore apologizes profusely. In this hypothetical case, the friend's output (the prankster's input) was enacted in order to achieve the goal of preventing guilt on the part of the prankster. However, her behavior still carried sufficient accurate information (viz. contained enough cues to the friend's intentions and feelings) that the prankster was able to accurately infer her friend's actual experience. Perceiving a discrepancy with the prankster's intended mutual enjoyment of the interaction, the prankster then adjusted her behavior – apologized profusely – to reduce the discrepancy. It is, of course, possible that the prankster could have perceived her friend's behavior as an authentic representation of the situation. This perception could be due to the friend's self-presentation being especially compelling (viz. not containing enough information to infer the reality) or due to the prankster failing to detect cues to the real situation. In such a case, the prankster would not necessarily have perceived as much of a discrepancy with her intentions and may not have apologized so profusely or at all.

Feedback problems. The dynamic interaction of actor and audience can lead to complicated results, especially when parties of an interaction are being deceptive – that is, when they are producing self-presentational output that provides intentionally false information about the situation (cf. Wiener, 1950/1954). In a cybernetic perspective, deception entails the transmission of intentionally false and misleading information to others. As in benign self-presentation, a liar behaves in such a way – generates output – which serves as input for an audience in order to create an intended impression, but that input provides information to the audience which misrepresents reality. The results of this input are various. The audience of a deceptive self-presentation may be duped by the claim, or they may detect the deception. In either case, the audience will draw some inference about the goals and intentions of the deceiver and will then enact some behavior in order to further their own goals. If the lie is undetected, the audience is likely to proceed as if the information were true. If the lie is caught, the audience may confront the liar about his or her deception, for instance, or the audience may allow the false claim to go uncontested, rather than risk the social consequences of a confrontation.

Consider, for instance, a nationalistic parade in an oppressive, totalitarian nation. The people who line the streets for the parade may present themselves as fervently nationalistic and elated to see their dictator – who himself is smiling back and waving his hand proudly at his people. One can imagine several possible realities underpinning this scene. It could be that the dictator is naïve; the general population is miserable and terrified of their leader, but they dare not incur the wrath of the dictator and his army by expressing displeasure. The smiles of the people at the parade fool the dictator into believing they are, in fact, happy to live under his regime. Alternatively, it could be that the people are actually pleased to be living in the nation, and their smiles – and the smile of their leader – are authentic representations of their inner

experiences, and they work together to reinforce their national pride. It could also be that this scene is an utter sham, and no one there is, in fact, pleased: The dictator is fearful and insecure that his empire may fall to shambles but smiles to create the appearance of prosperity, and the general population is disgusted and hateful of their oppressor but they conceal such feelings with falsified joy. In this last alternative, it is possible that one or both parties are being fooled by the other's performance, and it is further possible that no one is fooled and that the charade is intended for another audience entirely: the denizens of other nations who may see photographs of the parade in the media. In all of these possibilities, however, each person in the scene is collecting information from the situation and from other persons about what the others – at least ostensibly – intend and desire. Even in cases in which one or both parties are duped, everyone is acting on the information they have collected in order to pursue their goals.

The general tendency for actors and audiences to make their intentions at least somewhat clear may partially explain why liars are often perceived as uncooperative. Because, by necessity, liars do not want their audiences to accurately infer their goals, they are purposefully unhelpful. Indeed, an impression of uncooperativeness is one of the strongest cues to deception studied in empirical research (DePaulo et al., 2003).

Suboptimal functioning of the feedback system can lead to problems during self-presentation. Self-presentation also becomes more difficult when social actors inadequately compare the present situation to the reference value (no matter how well defined it is). That is, a well-engaged comparator facilitates optimal self-presentation. When attention is not or cannot be directed at monitoring one's performance, self-presentation tends to become less optimal (cf. Vohs, Baumeister, & Ciarocco, 2005). That is, people behave in a manner that creates negative

impressions (e.g., arrogantly, egotistically). I will revisit potential problems with the comparator below, especially in relation to deception.

Issues with the input function are another potential source of difficulty for self-presentation. As mentioned above, input can vary with respect to how frequently information is available and the quality of the information available (cf. DePaulo, 1992). Communicative contexts impose limits on the frequency and quality of input. A face-to-face conversation, for instance, provides fairly immediate input, with verbal, nonverbal, and paralinguistic information. In contrast, email and letters provide less frequent input and only verbal information. Despite the diversity of contexts in which communication (and, therefore, self-presentation) can occur, the feedback processes I have described should apply universally, as there is no assumption about the medium of communication or the length of time over which the communication occurs. The frequency and quality of input for the feedback system do, however, have consequences for self-presentational behavior.

Generally, the more continuous the input, the more quickly the system can adjust to problems and the less likely the system will produce output that misses the target. It is potentially easier, therefore, to detect that you have offended your conversational partner when you are face-to-face, compared to exchanging letters. If you misspeak and offend someone to whom you are speaking, as long as you have detected the offense, you can take immediate action to attempt to rectify the problem. In contrast, if you write something offensive in an early paragraph in a long letter, you may have incurred the anger and disgust of the recipient without a means of detecting what you have done, and the remainder of the letter may continue to offend the reader.

Similarly, the better the quality of the information provided by the input, the better one can calibrate one's output. Several features of the people in the interaction and of the situation

constrain the quality of input. The input for each feedback system in the interaction is largely the output of the other parties in the interaction. Therefore, the quality of the input is directly related to the quality of others' output. Even if the other elements of the feedback system are functioning optimally, the effectiveness of a person's self-presentation remains constrained by the range of output he or she is able to produce. People are not free to enact any behavior they desire: They must possess the ability to produce the behavior, and the situation must be conducive (socially and physically) to the production of that behavior. DePaulo (1992) has reviewed the various situational, cultural, practical, and physical constraints imposed on nonverbal self-presentations. The same kinds of constraints exist for verbal behavior. To use a simple example, one cannot produce words that one doesn't know, even if those words would have precisely the desired effect. Individual differences, such as social anxiety, may constrain a person's ability to produce optimal output (cf. Schlenker & Leary, 1985, 1982). Moreover, the norms of the social situation may place limits on what kinds of output are acceptable.

Another potential source of suboptimal input is deceit. When one or more people in a social interaction are behaving deceptively, it may become more difficult for people to calibrate their self-presentations, as the input for social actors' feedback systems becomes less diagnostic of the actual situation. Consistent with this reasoning, research on lies told in everyday life suggests that deceptive interactions are less pleasant and less intimate (DePaulo et al, 1996) and that people who lie in greater frequencies tend to have poorer quality social relationships (Kashy & DePaulo, 1996).

The deception discrepancy. The deception discrepancy is the fundamental difference between lies and truths: Although liars and truth-tellers both assert that their claims are sincere, liars' claims are not. Honest self-presenters do not necessarily present an unadulterated portrayal

of reality, but they differ from liars in that they believe their claims are justified, even if they may be somewhat exaggerated or understated. In contrast, deceivers believe that their claims are unjustified by reality.

Before delving into the specific implications of feedback control for deceptive behavior, it is important to establish what is meant when referring to "differences between lies and truths or liars and truth-tellers." When I speak of such differences, I am referring to a hypothetical comparison of the thoughts, feelings, and behaviors of a person who is making a deceptive identity claim to those of the same person making a summarily identical – but honest – claim. That is, I am imagining a comparison of a liar and an analogous truth-teller – the only difference between them being the deception discrepancy; that one of them is making a false claim and the other is making a true claim.

As previously mentioned, DePaulo and her colleagues (2003) inferred two core implications from the deception discrepancy: (1) Liars less convincingly embrace their claims compared to truth-tellers, and (2) liars are more deliberate in their self-presentations than truth-tellers. I agree that these implications flow from the deception discrepancy. Below, I interpret these two implications using the feedback control framework. Both of these implications entail differences in the manner in which liars and truth-tellers pursue their goals. It is assumed that liars and truth-tellers have substantively identical goals; that is, they both act toward the same desired outcome (i.e., to create a particular impression or belief in their audience), but the way in which they attempt to arrive at that outcome differs.

Liars embrace their claims less convincingly. Liars often lack the knowledge and experience necessary to make compelling claims. In the perspective presented here, this entails an inability to specify reference values for their goal pursuit that match the values a truth-teller

might have if making an analogous claim. The output of the feedback system – even if it is perfectly calibrated – will be oriented toward reducing discrepancies with a reference value that is relatively less compelling. The observable result is that liars make statements that are less rich with detail, are less plausible, and are shorter (DePaulo et al., 2003).

This difference between lies and truths is mediated by the difference between liars' and truth-tellers' knowledge of the content of the deceit. If a deceiver does not have knowledge that a truth-teller would have, he may not be able to simulate that knowledge. As the content of the deceit becomes more richly sophisticated, compelling deception becomes more challenging for those naive of such details. Knowledge about identities can be quite complexly structured, detailed, and dynamic (Markus & Wurf, 1987; Showers & Zeigler-Hill, 2012). It is, therefore, probably generally easier to claim that one has had an episodic experience one has not (e.g., pretending that one attended a seminar, when one slept late and missed it) than to claim that one is a different person (e.g., pretending to be a graduate student when one is in fact an undergraduate). Stated differently, if the deceiver is less familiar with the content of the deception, the deceiver's account will likely contain less detail and be less plausible than the account of an analogous truth-teller.

Differences between honest and deceptive claims may appear but not be detected if the audience of such claims is likewise ignorant of the nuance the deceiver is unable to emulate. If, however, the audience is knowledgeable about the domain of the deception, then there are more constraints on what is believable. These constraints correspond to what DePaulo and Bell (1996) called the *defensibility postulate*, which will be addressed in further detail later.

Liars also have different emotional investments than truth-tellers. In brief, because liars are often motivated by different goals than truth-tellers, their emotional contingencies differ from

truth-tellers. The issue of emotions related to deception will be dealt with in greater detail in a section below.

Liars are more deliberate in their self-presentations. Many researchers have hypothesized that deception is generally more effortful than telling the truth (e.g., Zuckerman, DePaulo, & Rosenthal, 1981; Vrij, 2006). The present perspective is consistent with these previous analyses, but it conceptualizes the increased cognitive effort as part of a feedback process. People lie, in part, because they do not believe the truth will suffice to facilitate their goals. Thus, because lying almost necessarily involves some comparison between the present state of the world and a desired state, liars tend to be more deliberate than truth-tellers. For this reason, the comparator of the feedback system tends to be more highly engaged during deception compared to truth-telling. Greater engagement of the comparator leads to better optimized output for goal pursuit – or at least more precisely-formed output.

If it is the case that lies generally entail greater amounts of mental comparison than truth-telling, liars should be generally more attentive to obstacles to the pursuit to their goals (Kappes, Singmann, & Oettingen, 2012; Kappes et al, 2013). Reference values, therefore, tend to be well-defined during deception, even if they are not elaborated with the same compelling degree of detail a truth-teller might have. Liars specify what they want to tell or what they do not want to tell; truth-tellers are less frequently so specific. For truth-tellers, reference values tend to be like impressionistic images of well-furnished and decorated rooms – imprecise, but rich. For liars, reference values tend to be like photographs of Spartan chambers, furnished with only the essentials – precise, but spare.

Consistent with this analysis, research has demonstrated that although lies tend to be less detailed than truths (as noted above), lies tend to contain fewer ordinary imperfections (DePaulo

et al., 2003): Compared to truth-tellers, liars admit to faulty memory less often, make fewer spontaneous corrections, and pursue fewer tangents. Moreover, when they know they are going to be questioned, liars are more prone to carefully plan out their verbal and nonverbal behavior compared to truth-tellers, but they tend to keep their statements short, simple, and to-the-point (Hartwig, Granhag, & Strömwall, 2007; Hartwig, Granhag, Strömwall, & Doering, 2010; Strömwall, Hartwig, & Granhag, 2006). Moreover, liars are likely to be more attentive to cues to their performance than truth-tellers (cf. Hartwig, Granhag, Strömwall, & Doering, 2010). That is, because the comparator tends to be more engaged for liars, they are especially attentive to input that can inform them about the quality of their performance.

This higher sensitivity to input relevant to their performance may lead liars to act in more tightly controlled and deliberate ways than truth-tellers, especially in response to input that suggests they may not be believed (e.g., signs that the audience is suspicious). Research has been consistent with this idea. Vrij (2006) interviewed liars and truth-tellers in a mock transgression paradigm. After an information-gathering phase, the interviewer expressed suspicion that the participant was concealing something, before returning to a non-accusatory, information-gathering style. Vrij (2006) found that liars' nonverbal behavior was more rigid (i.e., fewer smiles, fewer hand movements) than truth-tellers' behavior, but these differences were strongest in the last phase of the interview. These results suggest that liars may not have taken their credibility for granted and maintained a even more heightened level of control over their behavior, beginning when the interviewer expressed suspicion. The suspicion may have mustered their attention, and this increased effort carried over into the remaining phases of the interview.

In the present framework, these results may represent liars' reference values as being more well-defined (and differently defined) than truth-tellers', and therefore liars sense discrepancies with their reference values in ways that truth-tellers do not. That is, because liars' reference values are better defined than truth-tellers', there are more ways in which a liar's performance, by their own perception, can "go wrong." They more actively consider whether or not they are being viewed as credible. As a result, liars are prone to attempt to compensate with tighter control than truth-tellers typically exert.

The present analysis suggests that these differences in deliberateness depend on differences in the extent to which the comparison process is engaged for lies and truths. There are, however, situations in which truth-tellers are disposed to be or are prescribed to be deliberate. In situations in which liars' and truth-tellers' comparators are similarly engaged, cues to deliberateness should not differentiate lies and truths or should do so to a lesser degree. In the delivery of a prepared speech, for example, people may be deliberate, whether they are lying or telling the truth. Liars' and truth-tellers' comparison processes are likely to be highly engaged, as they carefully deliver their remarks and attempt to keep the script. To use another example, consider a first date between a potential romantic couple. Both parties are likely to be highly alert to self-presentational concerns as they want to make a positive impression on the other. Even if a person engages in deception to manage his or her impression here, cues to deliberateness may not betray the deceit, since both liars and truth-tellers are deliberate in this context. I know of no empirical data that directly supports or refutes this prediction, but this hypothesis follows from the present analysis. However, the study by Vrij (2006) described above indirectly supports the hypothesis: When under apparent suspicion by the interviewer, both honest and deceptive participants behaved similarly deliberately (e.g., making more eye contact

with the interviewer). Differences in apparent deliberateness emerged when the interview transitioned toward an information gathering focus again, as truth-tellers seemed to throttle back their control while liars maintained a high degree of control. It may be that both liars and truth-tellers make an adjustment of reference value while the interviewer appears actively suspicious, insofar as they both recognize the potential benefits of behaving more deliberately in those circumstances. When the style of the interview moved toward an information-gathering approach, however, the reference value for liars may have remained specified such that they were highly deliberate. This may be because liars were looking for clear input indicating that the interviewer was no longer suspicious of them. Truth-tellers, in contrast, shifted their behavior in response to the more suspicious interview style, but they remained less deliberate than liars throughout the entire interview. It is possible that they were, on the whole, less concerned with their credibility (cf Kassin, 2005) and therefore only relatively mildly adjusted their behavior in response to the interviewer. It is possible that overt removal of suspicion may have signaled to liars that it would be appropriate to loosen control (viz. readjust the reference value). Vrij (2006) did not manipulate whether the interviewer explicitly removed suspicion from the participant; the interviewer consistently simply transitioned into an information gathering approach without overt removal of suspicion, so this idea remains empirically untested.

It should be noted, however, that the suspicion introduced by the interviewer in the study by Vrij (2006) was relatively benign (i.e., “Your reactions make me think you are hiding something.”). With stronger accusations (e.g., confident declarations that the subject is lying), it is possible to induce behavior in truth-tellers that appears defensive and dishonest (Kassin, Goldstein, & Savistsky, 2003). In such cases, the accusatory context may determine a person’s

behavior to a greater extent than whether they are truthful or deceptive (cf. Hartwig & Bond, 2014).

There may also be cases in which the comparison process is not especially engaged while lying. One such case may occur when social norms are especially strong. When social norms prescribe a particular response (regardless of its truth), people may deceive in a relatively mindless way. If a friend or romantic partner asks about his or her physical appearance (e.g., "Do these pants make me look fat?"), a dutiful partner may perfunctorily respond with a positive comment (e.g., "No, you look great.") regardless of the answer's veracity and without much or any deliberation (cf. Bargh & Williams, 2006). In situations such as these, a person may lie so automatically that there is little or no comparison between the present state and the intended state. Under conditions in which the comparison process tends not to be engaged for both liars and truth-tellers, truthful and deceptive behavior should not differ with respect to deliberateness.

Generally speaking, liars' relatively high level of engagement of the comparison process is adaptive (cf. Oettingen, 2012). Failure to or inability to engage the comparison processes is likely to lead liars to perform suboptimally. If a person enacts deceptive behavior in order to pursue a goal but does not adequately assess the present situation (i.e., the attitudes, knowledge, and goals of the audience), he may lie in a manner that ignores important constraints on what self-presentations are likely to be effective and what is believable. Truth-tellers may commit similar self-presentational errors when their comparators are not well-engaged (Vohs, Baumeister, & Ciarocco, 2005).

Attention, effort, and cognitive load. In line with the argument that deception tends to be more deliberate, many deception theorists have argued that deception is more cognitively effortful than telling the truth (e.g., Vrij, Fisher, Mann, & Leal, 2006; Vrij, Granhag, Mann, &

Leal, 2011; Vrij & Granhag, 2012; Zuckerman, DePaulo, & Rosenthal, 1981). A growing body of empirical literature has demonstrated that imposing cognitive load impairs liars' ability to effectively deceive (Vrij et al, 2008; Vrij et al, 2011), although some emerging work has found that occasionally load can impair truth-tellers' ability to effectively communicate as well (Lane, Martin, Elliott, & Mennie, 2014). As I have argued previously, though it is generally the case that deception is more effortful and the comparison process is more fully engaged during deception than in non-deceptive self-presentation, there may be situations in which deception occurs automatically and without much cognitive effort.

Cybernetic analyses typically assume that a system has enough energy to continue operation (see Wiener, 1950/1954, 1948/1961), but there is nothing inherently incompatible between limited resource models and cybernetic control models. The two most relevant limited resource models related to self-regulation are the strength model of self-control (Baumeister, Vohs, & Tice, 2007) and the working memory framework (Baddeley, 2003). Broadly, in both of these models, there are limits to the attentional resources a person can devote to monitoring and altering their behavior in order to effectively pursue their goals (Hofmann, Friese, Schmeichel, & Baddeley, 2011).

The strength model of self-regulation construes the self as a resource that is depleted by acts of will (e.g., resisting temptation, planning activities carefully). Bauer and Baumeister (2011) have stated that the strength model of self-regulation becomes relevant when the feedback system has detected a discrepancy with the reference value. As far as I am aware, however, the relationship between the strength model and feedback mechanisms has never been investigated empirically. I am also unaware of any thorough theoretical treatment of the relationship between

these two models of self-regulation. Similarly, working memory and cognitive load have not, to the best of my knowledge, been explored in relation to feedback mechanisms.

Although there is a dearth of data to settle the matter, I disagree with Bauer and Baumeister's (2011) assertion that limitations on attentional resources become relevant for self-regulation after the detection of a discrepancy. Rather, I find it more plausible that limitations on resources affect the comparison process, both before and after the detection of discrepancies. The amount of attention a person can muster constrains his or her ability to detect discrepancies in addition to the ability to calibrate output to reduce discrepancies that are detected. Stated differently, attentional resources are needed both to detect discrepancies and to cope with them.

Thus, the imposition of cognitive load on social actors is likely to impair their ability to monitor their performance and to produce output that advances their social goals. In many cases, imposing load does not significantly impair truth-tellers, whose social task requires less attention – that is, a less engaged comparator. The present analysis predicts that there is an interactive effect such that cognitive load should impair the performances of people whose social tasks are attentionally demanding, whether the person is lying or not. In the feedback perspective, cognitive load impairs the comparator. Social actors whose comparators are impaired are likely to have difficulty monitoring their performance (i.e., getting a sense of whether their audience is forming the desired impression), producing well-calibrated output (i.e., behaving in a way that is likely to influence the audience in the intended way), and adjusting the reference value when appropriate (i.e., noticing important attributes of the audience that may inform the actor what kinds of performances are likely to be useful).

In cases in which deception would be more cognitively taxing than telling the truth, impairment of the comparator should differentiate liars and truth-tellers. Behaviors associated

with increased cognitive activity, such as speech errors and hesitations, should manifest themselves to a greater degree in liars than in truth-tellers (Vrij et al, 2008; Vrij et al, 2011). Moreover, if deceivers are less able to adjust their reference values, they may not be as able to account for factors that should constrain their performances. They may not readily account for what is believable; they may ignore the audience's knowledge and thus make statements that are implausible. I know of no data that speak directly to that prediction, but it follows from the present analysis.

There are notable exceptions to the general principle that deception is more cognitively demanding than telling the truth. For instance, there is some evidence that planning and preparation can ameliorate some of the behavioral signs of cognitive effort (DePaulo et al., 2003). Planning can make it more difficult for lie catchers to detect deceit, but it can also make both liars and truth-tellers appear more deceptive, probably because planned communication, whether honest or deceptive, seems more deliberate (Bond & DePaulo, 2006). It is also worth considering situations in which telling the truth may be more cognitively demanding than lying. In such cases, the behaviors discussed above would distinguish between liars and truth-tellers in the opposite way.

It is also possible that, to the extent that engagement of the comparison process entails the consumption of attentional resources, when people are more deliberate in their self-presentations, the depletion of resources may have consequences for their subsequent behavior (cf. Bauer & Baumeister, 2011; Baumeister, Vohs, & Tice, 2007). For example, it may be the case that when liars (or truth-tellers) are especially engaged, the depletion of energy induces them to seek out later opportunities to restore their resources (Gailliot & Baumeister, 2007; Gailliot et al., 2007). This prediction is related to a secondary hypothesis of the studies presented below.

The defensibility postulate. DePaulo and Bell (1996) argued that when people lie, as much as possible, they attempt to devise deceptions that can be defended as honest or, at worst, cannot be clearly established as deceptive. The advantage of creating such defensible lies is clear: One can avoid social penalties for lying if one eludes discovery, and if it is virtually impossible to discover that one has lied, then one can almost surely avoid being penalized. Liars' consideration for what deceptions are defensible is analogous to the more general self-presentational principle of concern for believability (Schlenker, 2012).

Interpreted in the framework of feedback control, it follows from the defensibility postulate that, in the same way that an honest social actor would use information about the audience to adjust his or her reference values, deceivers use information about the audience to determine what lies are defensible. As mentioned above, deceivers are more likely to be attentive to potential obstacles to their self-presentations (Kappes, Singmann, & Oettingen, 2012; Kappes et al, 2013). Liars are therefore likely to be alert to potential threats to the defensibility of deceptive self-presentations. To account for these threats, the secondary process in control of adjusting the reference values for the feedback system is probably particularly well-engaged for liars, compared to truth-tellers. Stated differently, in addition to the comparator being particularly engaged to detect discrepancies with the reference value for liars, the comparator is also engaged to detect information that indicates that the reference value should be altered. For a liar, it is particularly important to know what the audience knows, as the audience's knowledge directly constrains what claims are believable (cf. Baumeister & Jones, 1978).

Research provides indirect support of this reasoning. There is evidence that liars alter their behavior when they perceive that their audience may be suspicious (Burgoon et al., 1996). Additionally, research on interviewing criminal suspects has demonstrated that deceptive

suspects will adjust their strategies to avoid contradicting the knowledge of the interviewer (Hartwig, Granhag, & Luke, 2014) and will adjust their strategies when the interviewer is particularly alert to contradictions of the facts (Luke, Hartwig, Shamash, & Granhag, 2013). When deceptive suspects are unaware of the interviewer's knowledge, however, their statements often stray far from the truth, and consequently, they make claims that are indefensible, given the evidence. When a deceptive suspect knows or suspects that the interviewer is knowledgeable about his activities, however, he will adjust his story to accommodate this knowledge, creating a more defensible claim (e.g., Hartwig, Granhag, Strömwall, & Vrij, 2005; Luke, Dawson, Hartwig, & Granhag, 2014). Although there is no research of which I am aware that has examined shifts in liars' strategies specifically from the perspective of feedback control, the extant literature is generally consistent with the hypothesis that liars are highly attentive to potential threats to their credibility.

Motivation and expectancy in self-presentation and deception. The importance of success and the expectancy of success are among the variables that appear to influence the difficulty of deception. The meta-analytic results of Bond and DePaulo (2006) indicate that, across studies in deception, when communicators were highly incentivized to be believed, the deception detection accuracy of people judging their statements improved, albeit by a small degree. However, under these higher-stakes conditions, lie catchers tended to view everyone as more deceptive. That is, they were disposed to judge statements as deceptive, even when they were honest (though only when they were able to view their nonverbal behavior). This tendency for deception (and other forms of self-presentation) to be more difficult when it is more important has been termed *motivational impairment* (see DePaulo, LeMay, & Epstein, 1991).

Deception researchers have hypothesized that higher levels of motivation induce liars to monitor and control their behaviors to an extent that appears unnatural (see DePaulo, Kirkendol, Tang, & O'Brien, 1988). Consistent with this reasoning, DePaulo and her colleagues (2003) found that under conditions of high motivation, liars were more tense than truth-tellers, and more motivated liars were more fluent in their speaking than unmotivated liars. However, it seems that when highly motivated, both liars and truth-tellers appear overly controlled (Bond & DePaulo, 2006). Thus, motivational impairment is not an effect endemic only to liars.

There have been few investigations of the effect of expectancy of success on behavior during deception. Deceivers' expectancies of success have been indirectly examined in some studies (DePaulo, Kirkendol, Tang, & O'Brien, 1988; DePaulo, Stone & Lassiter, 1985), but to the best of my knowledge, DePaulo, LeMay, and Epstein (1991) conducted the only study to have directly experimentally manipulated deceivers' expectancy of success. The results of these studies are consistent with the hypothesis that expectancy and motivation interact to influence the amount of effort liars and truth-tellers exert during self-presentation, but it is not entirely clear whether this shift in effort occurs in the direction predicted by the present perspective. For example, based on the perceived sincerity of the communicators' statements, DePaulo, LeMay, and Epstein (1991) speculated that the greatest amount of effort was exerted by truth-tellers when motivation was high and expectancy was low and when motivation was low and expectancy was high. However, DePaulo and her colleagues (1991) did not directly measure exerted effort, so it is difficult to determine the extent of the communicators' engagement in each experimental condition.

The present theory paints a somewhat more complicated picture of motivational impairment than has been previously discussed in the deception literature, though the basic

premise of motivational impairment holds: As engagement with the self-presentational task increases, behavior becomes increasingly tightly controlled. It is important to consider, however, what determines a social actor's engagement with the task and what it means to be engaged with a task.

If a goal is especially important, both liars and truth-tellers may muster their attention toward the task. For both liars and truth-tellers, then, motivation would lead to more goal commitment – and to more engagement of the comparator. Thus, both liars and truth-tellers behavior would be more tightly controlled under conditions of high stakes, but for deceivers, their deception creates an additional source of engagement. At extremely high levels of goal importance, however, it is probably the case that both liars and truth-tellers are mustering virtually all their attentional resources toward the goal (or, at least, as many resources as they can). If, for example, you are tasked with giving a toast at your best friend's wedding, the task is of such great social importance that you are probably giving it your undivided attention, and your behavior is likely to be highly controlled – whether or not you are lying about how much you approve of your friend's choice of spouse. In such cases, the degree to which a communicator's behavior is tightly controlled would not distinguish between truths and lies.

Expectancies of success can interact with goal importance to increase or decrease commitment, and the same relationships between variables that apply to other goals apply to self-presentational goals. When the importance of a self-presentational goal is relatively low, expectancies of success should matter relatively little (e.g., Oettingen, 2000). Greeting a coworker as you pass in the hall requires some self-presentation, but it is typically relatively unimportant. Despite the brevity of a passing greeting, there can be variance in how difficult it is to conduct a smooth interaction. If, for instance, you know the coworker you are passing to be

generally amiable, there is a high likelihood of having a smooth and successful, though brief, exchange of pleasantries. If, however, you know you are passing an insufferable grouch, the likelihood of a smooth exchange is low. In both cases, however, you are unlikely to muster a large amount of effort for the greeting, since the interaction is of such little importance.

As the importance of a self-presentational goal increases, however, expectancies of success become more influential on the exertion of effort, with high expectancies of success encouraging effort and low expectancies discouraging effort. All other things being equal, the effort one is willing to expend will be a function of how difficult one perceives it to be to produce the desired impression and the extent to which the difficulty is surmountable. These two perceptions combine into an expectancy of success, which will then determine the amount of effort. Reconsidering a scenario from above, if you are required to perform a toast at your best friend's wedding, the task is likely to be relatively difficult. You will be in front of an audience of friends and family with diverse attitudes and beliefs, and it will be difficult to propose a toast that is appealing to all of them. If you are not confident in your ability to speak in public (viz. you perceive the difficulty to be practically insurmountable), you may exert moderate effort and resort to a generic, unmemorable, but adequate toast. If, however, you are confident in your skills as a toastmaster (viz. you perceive the task as difficult but surmountable), you may muster a large amount of effort and prepare a witty and memorable toast.

As I mentioned earlier, in cases in which both honest and deceptive self-presenters are mustering virtually all their available resource toward a task, the extent to which their behaviors will appear controlled should not distinguish between them. This implies that, all other things being equal, at moderate and lower levels of importance and expectancy, liars and truth-tellers may be more distinguishable by the extent to which their behavior is tightly controlled.

Chapter 2: The Present Studies

In order to create a successful forgery, a forger must know what the original looks like. The forgery must be similar enough to the object it is emulating so that it is perceived as an original. I propose that deception is analogous to crafting forgeries. Because deceptive communication is, in large part, an imitation of truthful communication, liars must rely on their knowledge of the ways in which truth-tellers behave in order to pass off their messages as truthful. In the same way that a successful forger must create a fake that strongly resembles the original, a successful liar must present him- or her-self in a manner that strongly resembles a truth-teller.

If one considers self-presentation and deception from a cybernetic perspective, it follows that communicators' – especially liars' – self-presentational performances become more optimal as they gain access and make use of information about what kinds of performances are adaptive in a particular context. Stated differently, people are better able to present themselves when they have an idea of what other adaptive performances "look like." Liars, therefore, are better able to lie when they are aware of how truth-tellers act in a given context.

Exposure to such information, in cybernetic terms, permits a social actor to better specify his or her reference value(s) in line with what is adaptive. Actors are therefore able to act consistently with that reference value, in order to better facilitate their social goals. In deception, when liars are aware of how truth-tellers behave, this information can serve as input as to how they should alter their reference value(s). Thus, just as a forger benefits from knowing the properties of an original, information about truthful performances can guide liars' behavior, such that they are able to create better emulations of the truth.

I report two experimental studies here. The present studies provide an important initial test of predictions derived from the cybernetic perspective of self-presentation and deception developed above. The theoretical framework predicts that social actors, especially deceivers, are better able to present themselves when they have access to valid information about the attributes of desirable social performances. These studies will provide initial tests of this central prediction of the cybernetic theory of self-presentation and deception.

Although the concept of feedback I have described above comprises numerous mechanisms, the present studies are designed primarily focused the reference value component of the feedback loop. I take this relatively narrow focus because the notion of reference values is fairly unique among theories of deception. Other theories (e.g., Interpersonal Deception Theory) contain predictions about how liars and truth-tellers respond to input about their performances, but I know of no other theoretical approach to deception that is concerned with the reference information that is used to guide deceptive and truthful performances. The present studies are designed to provide some insight into this relatively novel idea.

Both studies are based on a novel paradigm: the "embarrassed experimenter" paradigm. This section is organized as follows: First, I describe the paradigm and the coding system for the key variables in the two studies. Afterward, I describe the procedure and results of the studies. Each study's description contains a discussion of the results. A general discussion follows in the next section.

The "embarrassed experimenter" paradigm

This new experimental paradigm is purposed to produce a plausible social situation in which the participant is motivated to lie in order to assist an experimenter. The central premise of the paradigm is that a participant can be motivated to lie in an interview in order to apparently

protect an experimenter from embarrassment or social reprisals. A major advantage of the paradigm is its experimental realism; that is, participants are not instructed to lie as part of a study's procedures (at least ostensibly); they are lying for more realistic reasons. Here, I will briefly summarize the core elements of the paradigm. In each of the studies proposed below, the paradigm is modified in order to test particular hypotheses.

Introduction and cover story. Participants were recruited for a study on "memory and impressions of mental tasks." The cover story was important for creating experimental realism. The procedure began with the participant's arrival in the lab. After obtaining informed consent, an experimenter provided the participant with questionnaires to assess individual differences (i.e., regulatory focus and perspective taking), partly in order to facilitate the cover story. After the participant completed these questionnaires, the experimenter explained that the participant would then take part in an activity and then will be interviewed about the experience. The experimenter then led the participant to another room, the door to which was closed.

Veracity manipulation. In the truthful condition, the experimenter opened the door to the room with a key. Participant was led inside the room and took part in a task that was apparently designed to test memory ability. Using a standard deck of playing cards, the experimenter shuffled and placed a number of cards on the surface of a table. The participant was given a few seconds to memorize the values and suits of the cards before the experimenter concealed the cards and asked the participant to recall the values and suits. This was repeated several times, with increasing numbers of cards dealt onto the table. After the task was completed, the experimenter terminated the activity and led the participant to the next stage in the procedure.

In the deceptive condition, the door to the room was locked and the experimenter claimed that he or she did not have the key. The experimenter acted embarrassed and explained that he or she had made several such mistakes in the study's procedure over the last few days, and the lead researcher in charge of the project had been scornfully scrutinizing him or her because of the errors. The experimenter then asked the participant to claim during the subsequent interview that they did in fact complete the task as planned. The experimenter described the procedure for the task so that the participant understood what would have occurred in the room. The experimenter's behavior was scripted, but experimenters were permitted latitude to respond to the participant's behavior and act in a manner that seemed natural. For example, some participants attempted to help the experimenter by bypassing the lock with a credit card. When this occurred, the experimenter encouraged the behavior as if he or she was genuinely interested in gaining access to the room. After obtaining the participant's agreement to "cover" for him or her, the experimenter led the participant to the next stage of the procedure.

The playing card activity served two primary purposes. First, it provided an opportunity to manipulate the veracity of the participants' subsequent claims to the interviewer: Some participants actually completed this activity; others deceptive claimed that they did. Second, the experience was sufficiently rich as to provide the participants with enough content to speak for several minutes. The task and its procedures were elaborate enough to produce many details when described, and the room contained memorable details (e.g., numerous objects on shelves, several pieces of furniture). Because the task was fairly rich in detail, it put liars at a slight disadvantage with regard to information. That is, liars necessarily only had secondhand information about what would have occurred during the memory activity.

Interview. Following the veracity manipulation phase of the procedure, the experimenter guided the participant to a room to meet the interviewer (also allegedly the lead researcher). Upon the participant's arrival, the interviewer explained that he or she was going to ask a series of questions about the activity in which they just participated, in order to assess their memory of the experience. The interviewer followed a series of scripted questions, asking about the procedures of the task, about the participant's impressions of the difficulty of the task, and about the physical space in which the task took place. Participants who have been asked to conceal the experimenter's apparent mistake should attempt to deceive the interviewer. Participants who were not asked to lie had no reason to attempt to deceive the interview and provided honest statements.

The interviews were divided into three phases, each of which pertained to a different aspect of the procedure. In the first phase, participants were asked to describe the procedure for the task in which they took part. In the second phase, participants were asked to describe the physical environment in which the task took place. In the third phase, participants were asked about their impressions of the task (e.g., the easiest and most difficult parts of the task).

Interviews were video recorded for coding. After the interview, the experimenter provided the participant with a questionnaire that probed for suspicion of the true nature of the experiment. The experimenter and interviewer disabused the participant of the deception and conducted a thorough debriefing before dismissing the participant. Following the debriefing, the experimenter offered the participant candy from a dish. There were always fifteen pieces of hard candy in the dish so that the experimenter could later count and record the number of pieces of candy the participant took. The number of pieces of candy the participant took served as an

indirect measure of how depleting the experience was (Gailliot & Baumeister, 2007; Gailliot et al., 2007).

Coding of participant behaviors. After the interviews were conducted and recorded, the behavior of the participants was coded for analysis. The specific behaviors coded in each of these domains were partially adapted from past empirical research on behavioral correlates of deception and the meta-analysis conducted by DePaulo and her colleagues (2003).

The cues to forthcomingness were assessed by trained coders who were blind to experimental condition of the videos and to the hypotheses. Two coders rated 20% of the sample of recordings in order to establish inter-rater reliability. After reliability was established, the coders each rated half of the remainder of sample. Two cues were be coded to assess the how forthcoming a statement is: (1) the amount of time the participant spent talking and (2) the number of details included in the statement. These two variables were adapted from the two most strongest cues to deception in the domain of forthcomingness that DePaulo and her colleagues (2003) meta-analyzed. The two forthcomingness variables were coded for each phase in the interview (i.e., the phases concerning the procedures of the tasks, the description of the environment, and the impressions of the task). Talking time was measured as the percentage of the time (measured in seconds) the participant spent talking to the total amount of time both the interviewer and participant spent talking.

In line with Allwood, Ask, and Granhag (2005), details were measured by counting the number of information units that occurred in the statement. Statements about actors and their actions counted as one unit. For example, the phase ‘‘the experimenter walked in’’ would count as a single unit. An object and a description with one attribute counted as a single unit. The phase ‘‘there was a desk in the corner’’ counted as one unit. When objects are described by multiple

attributes, each additional attribute (beyond the first) counted as an additional information unit. Therefore, the phrase ‘‘there was a cluttered black desk’’ counts as two units. In statements about actors and their actions in which the actor was described with one or more attributes, the act and actor were counted as a single unit, and each attribute counted as an additional unit. The phrase ‘‘the experimenter sat at a white table’’ therefore counted as three units.

Participants’ behavior in the interview was assessed with five additional variables: (1) the *clarity and logic* of the statement, (2) the apparent *deliberateness* of the participant’s statement, (3) the extent to which the participant appeared to be *cooperative*, (4) the extent to which the participant appeared to be *nervous*, and (5) the extent to which the participant appeared to answer *honestly*. In order to measure these variables, the videos of the interviews were shown to observers recruited through Amazon’s Mechanical Turk (MTurk). These observers provided ratings of each of these variables. This procedure is described in greater detail below, in Phase 2 of Study 1.

Study 1

Purpose. Study 1 aimed at testing the hypothesis that deceivers can craft better formed lies when they have exposure to examples of truthful communication in a particular social context. Participants were randomly assigned to be exposed to recorded examples of similar (but not identical) interviews prior to being interviewed themselves. Participants were told that these exemplars may help give them an idea of what to expect when they are interviewed. These exemplars are hypothesized to provide liars with information about the properties of a typical self-presentation in this context (and therefore, the properties of a good lie). Truth-tellers are not predicted to be affected by exposure to the exemplars. Because they have little use for the information, they are unlikely to make use of it.

Study 1 consisted of three stages. In Stage 1, I generated videos of truth-telling participants being interviewed. These videos were then shown as exemplars to some of the participants in Stage 2, which was designed to test the hypotheses concerning participants' behavior in the interview. Finally, in Stage 3, the recordings of the interviews from Stage 2 were shown to observers who judged the veracity of the participants. Each stage is described below.

Study 1 had seven hypotheses:

H_1: I predicted a main effect for veracity, such that, truth-tellers, compared to liars, will produce statements that appear more forthcoming, more cooperative, clearer and more logical, and less deliberate.

H_2: I predicted an interaction between veracity and exemplar exposure. More specifically, liars exposed to example interview recordings will produce statements that appear more forthcoming, more cooperative, clearer and more logical, and less deliberate, compared to statements by liars who have not received exemplar information.

H_3: Because deception tends to be more effortful than honesty, I predicted a main effect for veracity such that liars would self-report having planned and prepared their statements to a greater extent than truth-tellers.

H_4: Because I expected the exemplars to make the task of lying easier, I predicted an interaction between veracity and exemplar exposure such that liars who saw the exemplars would report being less deliberate (i.e., planning and preparing less) compared to liars who did not see the exemplars.

H_5: I predicted a main effect for veracity such that liars, because they had expended more energy than truth-tellers, would take more candy at the conclusion of the study.

H_6:This increased energy consumption should be attenuated when liars' task is made easier by the exemplars. As such, I predicted an interaction such that liars who watched the exemplar videos would take less candy than liars who did not receive the exemplars.

H_7: Because liars who have access to reference information about truth-tellers' behavior should be able to generate more convincing statements than those who do not have such information, I predicted an interaction between veracity and exemplar condition such that observers would be significantly less accurate when judging liars who watched the exemplars, compared to those who did not watch the exemplars.

Stage 1: Production of Exemplars. In order to generate the materials for Stage 2, I produced video recordings of interviews to serve as exemplars. These exemplar interviews were similar, but not identical, to the interviews that took place during the hypothesis testing phase. Participants were undergraduates at a large urban university in the northeastern United States (N = 20). They were given course credit in exchange for their participation.

Stage 1 followed the procedures of the paradigm described above, but during this phase no participants were asked to lie. Thus, this phase produced 20 videos of interviews in which participants truthfully communicate their experiences.

Additionally, the activity about which these participants were interviewed was different from the activity described above. Because I hypothesized that the "active ingredient" of the exemplars is the style of truth-tellers' self-presentations, rather than the content of their statements, it was important to create exemplar interviews that were structurally similar to the ones participants would engage in but did not contain the content they would be attempting to communicate. Specifically, participants were randomly assigned to one of two activities. Both of these activities took place in different rooms, which were similar in size but contained different

furniture and were arranged differently. As can be seen below, these activities resemble the faux memory task described above in that they involve interacting with an experimenter in a structured task that increases in difficulty over time. As in the playing card activity described above, participants' actual performance on these tasks is irrelevant; only the manner in which they describe the experience of the activities is important (e.g., how much detail they include in their account).

One activity involved solving relatively simple math problems shown on cards displayed by the experimenter to the participant. The experimenter showed the participant a math problem printed on a card for a brief period of time. The experimenter then concealed the card and recorded the participant's answer to the problem. The experimenter repeated this process with increasingly difficult problems.

The second activity involved rapidly calculating the total value of sets of several coins. Concealed under a cover, the experimenter placed a number of coins (United States currency) on the surface of a table. The experimenter then lifted the cover for a brief period of time, during which the participant attempted to count the values of the coins. The experimenter then concealed the coins again and recorded the participant's calculated sum. This process was repeated with increasingly difficult sets of coins.

Following completion of one of these tasks, the participants were interviewed in line with the description above. The interviews were video recorded. These recordings were used as exemplars in the second phase of this study.

Stage 2: Primary hypothesis Testing. Study 1 uses a 2 (veracity: truthful vs. deceptive) x 2 (exemplars: no exemplars vs. exemplars) factorial design. Stage 2 of Study 1 followed the embarrassed experimenter paradigm as described above with an additional manipulation: Some

participants watched exemplar videos produced in Stage 1 (i.e., the exemplar condition). These participants watched two randomly selected videos from Stage 1, one video of the math task and one video of the coin task.

Participants in Stage 2 were undergraduates at a large urban university in the northeastern United States (N = 130). Three participants refused to comply with the experimenter's request to lie. In these cases, the experimenter pretended to locate the lost key to the room, and the procedures were continued as if that participant were in the truthful condition. I conducted the statistical analyses with these participants removed, and the results were consistent with analyses conducted with the participants included. All reported analyses include the full sample. The sample was mostly female (64.6% female, 35.4% male) and racially diverse (40.8% Hispanic, 20.0% Black/African-American, 16.2% White, 15.4% Asian, 5.4% mixed, 2.3% other). The average age was M = 20.18 years (SD = 3.91), with a range from 18 to 48. The median age was 19 years. Participants were given course credit in exchange for their participation.

The exemplar manipulation occurred after the veracity manipulation and before the interview. In the exemplar condition, the experimenter will guide the participant to a room where he or she will be seated in front of a computer. The experimenter will then explain that the participant will watch two videos from previous studies in which participants underwent similar procedures. The purpose of watching these videos, the experimenter will explain, is to give the participant a sense of what to expect in the interview. In the no exemplar condition, participants listened to an audio recording of the interview questions, repeated twice. This was done in order to rule out the possibility that any observed effects were the result of merely having been exposed to the interview questions in the exemplar condition. After viewing or listening to the recordings, the experimenter brought the interviewer to the interview room. The interview and

subsequent procedures were conducted as described above. All interviews were video recorded. for coding.

Results.

Coding. In order to assess the forthcomingness of the participants, as described above, two trained coders, who were blind to hypotheses and blind to condition, rated a random sub-sample of 20% of the sample (n = 26) for details and talking time. Inter-rater reliability was adequate for all phases of the interviews, all intra-class correlation coefficients < .72. Once inter-rater reliability had been established, one of the coders rated the remainder of the recordings.

In order to obtain ratings for cooperativeness, clarity and logic, deliberateness, and nervousness the videos of the interviews (from both Study 1 and Study 2) were shown to observers recruited through MTurk. Each of the observers (N = 328) viewed six interviews randomly selected from the samples of Study 1 and Study 2. Observers rated each video on each of the variables on a 101-point scale (0 to 100), presented as a slider with an initial position of 50. Each video was viewed and rated several times (in Study 1, M = 8.34 ratings, SD = 1.75, median = 8). For each variable, I calculated the mean rating. These means were used in subsequent statistical analyses to test the primary hypotheses.

Forthcomingness. I analyzed the measures of talking time and details using separate two-way MANOVAs, using the measures for each phase of the interview as dependent variables. Means and standard deviations for talking time and details are presented in Table 1. To gain an overall measurement of participants' forthcomingness, I also calculated an average talking time, across each of the phases of the interview, for each participant as well as a sum of all the details the participant provided across each of the phases of the interview. Means and standard deviations for average talking time and total details are presented in Table 2.

For talking time, there was a significant main effect for veracity, Wilk's $\Lambda = .750$, F (3, 123) = 13.69, $p < .0001$, partial $\eta^2 = .25$. There was no significant main effect for exemplar condition on talking time, Wilk's $\Lambda = .962$, F (3, 123) = 1.62, $p = .19$. The interaction between veracity and exemplar condition on talking time was not significant, Wilk's $\Lambda = .978$, F (3, 123) = .93, $p = .43$. The multivariate results for details are highly similar. There was a significant main effect for veracity on details, Wilk's $\Lambda = .755$, F (3, 123) = 13.28, $p < .0001$, partial η^2 = .245. There was no significant main effect for exemplar condition on details, Wilk's $\Lambda = .959$, F (3, 123) = 1.76, $p = .16$. The interaction between veracity and exemplar condition on details was not significant, Wilk's $\Lambda = .983$, F (3, 123) = .71, $p = .55$.

To clarify these results, I conducted follow-up univariate ANOVAs. The results for these tests are presented in Table 3. As can be seen in the table, there was a main effect for veracity and for exemplar condition in the second phase of the interview (i.e., the phase concerning the environment) for both talking time and details. In the second phase, truth-tellers spoke more (M = 73.87%, SD = 7.76%) and provided more details (M = 11.79, SD = 2.94) compared to liars, who spoke less (M = 64.06%, SD = 11.69%) and provided fewer details (M = 8.59, SD = 2.98). In the second phase, participants who saw the exemplars spoke more (M = 70.46%, SD = 8.87%) and provided more details (M = 10.60, SD = 3.24) compared to those who did not see the exemplars, who spoke less (M = 66.57%, SD = 12.83%) and provided fewer details (M = 9.48, SD = 3.40).

Although the interactions in the multivariate and univariate ANOVAs were not significant, examining the means and standard deviations for details and talking time (see Tables 1 and 2) suggests there was a consistent pattern across each phase of the interview and for each measure: Truth-tellers who saw the exemplars were approximately as forthcoming as truth-tellers

who did not see the exemplars, but liars who saw the exemplars were consistently more forthcoming than liars who did not. This pattern is clearest when examining participants' average talking time and total details provided (see Table 2 and Figures 4 and 5). Simple effects tests performed on the measures of average talking time and total details are consistent with this pattern, and analyses of the measures of forthcomingness for each phase of the interview are generally consistent with it as well. These results are presented in Table 4. The patterns observed here are consistent with the hypotheses.

Observer ratings of statements. I analyzed the ratings of cooperation, clarity/logic, and deliberateness using a MANOVA, with veracity condition and exemplar condition as fixed factors. Multivariate analyses revealed a significant main effect for veracity, Wilk's $\Lambda = .933$, F (3, 123) = 2.93, $p = .036$, partial $\eta^2 = .067$, a significant main effect for exemplar condition, Wilk's $\Lambda = .940$, F (3, 123) = 2.64, $p = .053$, partial $\eta^2 = .060$, and a significant interaction between veracity and exemplar condition, Wilk's $\Lambda = .929$, F (3, 123) = 3.13, $p = .028$, partial $\eta^2 = .070$. Follow-up univariate ANOVAs were conducted.

The follow up tests suggested that veracity had a significant effect on ratings of cooperation, F (1, 125) = 5.17, $p = .025$, partial $\eta^2 = .040$, such that truth-tellers were rated as more cooperative ($M = 80.25$, $SD = 8.21$) than liars ($M = 76.53$, $SD = 10.15$). Exemplar condition did not have a significant effect on ratings of cooperation, F (1, 125) = 1.69, $p = .196$. There was no significant interaction between veracity and exemplar condition for ratings of cooperation, F (1, 125) = 0.21, $p = .645$. The difference between liars and truth-tellers was in line with the predictions. However the predicted interaction between veracity and exemplar condition was not observed. Means and standard deviations for the cooperation ratings for each condition are presented in Table 5.

Veracity had a significant effect on ratings of clarity/logic, F (1, 125) = 7.83, p = .006, partial η^2 = .059, such that truth-tellers were rated as significantly clearer and more logical (M = 73.53, SD =9.46) than liars (M = 68.63, SD = 10.71). Exemplar condition also had a significant main effect on ratings of clarity/logic, F (1, 125) = 4.35, p = .039, partial η^2 = .034, such that participants who viewed the video exemplars were rated as significantly clearer and more logical (M = 72.61, SD =10.28) than those who did not view the exemplars (M = 69.13, SD = 10.34). There was no significant interaction between veracity and exemplar condition for ratings of clarity/logic, F (1, 125) = .945, p = .333. The difference between liars and truth-tellers was consistent with the hypotheses. However, it was unexpected that exposure to the exemplars would have an effect on both liars and truth-tellers. Means and standard deviations for the clarity/logic ratings for each condition are presented in Table 5.

There was no significant effect of veracity on ratings of deliberateness, F (1, 125) = 1.72, p = .192. There was no significant effect of exemplar condition on ratings of deliberateness, F (1, 125) = 2.72, p = .102. However, there was a significant interaction between veracity and exemplar condition on ratings of deliberateness, F (1, 125) = 7.17, p = .008, partial η^2 = .054. Simple effects tests indicate that among liars, those who saw the exemplar videos were rated as more deliberate (M = 65.23, SD = 11.61) than those who did not (M = 57.70, SD = 9.83), t (69) = 2.95, p = .004, d = .70. Among truth-tellers, there was no significant difference in the ratings of deliberateness between those who saw the exemplars (M = 62.85, SD = 9.41) and those who did not (M = 64.64, SD = 7.67), t (63) = .235, p = .815. Among those participants who did not view the exemplars, truth-tellers were rated as more deliberate than liars, t (64) = 3.15, p = .002, d = .79. Among participants who viewed the exemplar videos, there was no significant difference between liars and truth-tellers, t (61) = .875, p = .385. Truth-tellers who viewed the exemplars

were rated as significantly more deliberate than liars who did not view the exemplars, t (62) = 2.12, p = .038, d = .54. This pattern of results can be seen in Figure 6. In summary, the difference between liars and truth-tellers with respect to their perceived deliberateness was attenuated by exposure to the exemplars. This pattern is consistent with the hypotheses, but the means are in the direction opposite of the predictions. That is, I expected liars to appear more deliberate than truth-tellers; instead, the opposite was true. Means and standard deviations for the deliberateness ratings for each condition are presented in Table 5.

Because I did not have specific predictions for nervousness, the results for nervousness were analyzed using a separate two-way univariate ANOVA. There was no main effect for veracity, F (1, 125) = .006, p = .94, or exemplar condition, F (1, 125) = .001, p = .98. There was a significant disordinal interaction between veracity and exemplar condition, F (1, 125) = 5.43, p = .02, partial η^2 = .042. Truth-tellers who did not see the exemplars (M = 49.04, SD = 11.64) were rated as more nervous than truth-tellers who saw them (M = 44.29, SD = 10.89), although this trend was not significant, t (56) = 1.61, p = .11, d = .41. Truth-tellers who did not see the exemplars were rated no differently than liars who had seen the exemplars (M = 49.27, SD = 12.18), t (63) = .075, p = .94. Truth-tellers who did not see the exemplars were rated as more nervous than liars who did not see the exemplars (M = 44.40, SD = 11.76), but this trend was not statistically significant, t (64) = 1.61, p = .11, d = .40. Truth-tellers who saw the exemplars were rated no differently than liars who had not seen the exemplars, t (62) = .039, p = .97. Truth-tellers who saw the exemplars were rated as less nervous than liars who did not see the exemplars, but this trend was only marginally significant, t (61) = 1.69, p = .096, d = .43. Liars who saw the exemplars were rated as more nervous than liars who did not see the exemplars, but

this trend was only marginally significant, t (61) = 1.72, p = .091, d = .41. The means for nervousness by condition are presented in Table 6 and are plotted in Figure 7.

Self-reported deliberateness. Four self-report items were used to measure self-perceived deliberateness. These four items are included in Appendix A. I analyzed these items using a two-way MANOVA, with veracity and exemplar condition as factors. There was a significant main effect for veracity, Wilk's Λ = .910, F (4, 122) = 3.02, p = .02, partial η^2 = .090. There was no main effect for exemplar condition, Wilk's Λ = .986, F (4, 122) = .44, p = .77. There was a significant interaction between veracity and exemplar condition, Wilk's Λ = .906, F (4, 122) = 3.17, p = .016, partial η^2 = .094. Follow-up univariate ANOVAs were conducted.

For all four items, there were no main effects for veracity or exemplar condition. However, there were significant interactions between veracity and exemplar condition for three of the four items. The results for the ANOVAs are presented in Table 7. Means and standard deviations for each of the items are presented in Table 8. Comparisons of each condition for each item are presented in Table 9. The means for each item are plotted in Figure 8, Figure 9, Figure 10, and Figure 11. As can be seen in the tables and figures, truth-tellers who saw the exemplars generally reported being more deliberate compared to truth-tellers who did not see the exemplars. This pattern was unexpected, but the differences on three of the four items were statistically significant (see Table 9). Additionally, there was a general pattern such that liars who saw the exemplars reported being less deliberate (e.g., planning what they were going to say to a lesser extent) compared to liars who did not see the exemplars. Although this trend for liars was consistent across all four items, none of the differences between liars who saw the exemplars and those who did not were statistically significant (see Table 9).

Energy consumption. As an indirect measure of energy consumption, the experimenter recorded the number of pieces of candy the participant took from the dish after the procedures were complete. One extreme outlier – a participant who took the maximum amount of 15 pieces of candy – was removed from the sample. There was no significant main effect of veracity on the taking of candy, F (1, 123) = .954, p = .331. There was no significant main effect of exemplar condition on the taking of candy, F (1, 123) = .368, p = .545. There was a marginally significant interaction between veracity and exemplar condition, F (1, 123) = 3.07., p = .082, partial η^2 = .034. The means for each condition are plotted in Figure 12.

Simple effects tests revealed that there were marginally significant differences in the expected directions. Among truth-tellers, there was no significant difference between those who saw the exemplars (M = 1.69, SD = 1.44) and those who did not (M = 1.39, SD = 1.56), t (55) = .76, p = .450. Among liars, those who saw the exemplars (M = 1.49, SD = 1.27) took fewer pieces of candy than those who did not see the exemplars (M = 2.11, SD = 1.66), but the statistical significance of this difference was only marginal, t (68) = 1.78, p = .079, d = .42. Liars who did not see the exemplars took more candy than truth-tellers who did not see the exemplars, but the statistical significance of this difference was only marginal, t (64) = 1.83, p = .073, d = .45. Liars who did not see the exemplars did not significantly differ from truth-tellers who saw the exemplars, t (59) = 1.04, p = .303. Liars who saw the exemplars did not differ from truth-tellers who saw the exemplars, t (59) = .54, p = .554, and did not differ from truth-tellers who did not see the exemplars, t (64) = .28, p = .778. To summarize, the general pattern was such that liars who did not see the exemplars took the most candy. The other groups took fewer pieces, and each of the other groups took approximately the same amount.

Stage 3: Deception detection. Stage 3 was designed to test the predictions about the accuracy and bias of the observers. For the purposes of these analyses, I will refer to the participants who are interviewed as "senders". The same observers who provided ratings of cooperativeness, clarity and logic, deliberateness, and nervousness for Stage 2 provided ratings of the extent to which they perceived the senders to be honest.

The measure of perceived honesty was used to assess the deception detection accuracy and judgment bias of the observers. The honesty scale was analyzed as a continuous variable, but in order to calculate accuracy and bias, the scale was used to create variables indicating a binary veracity judgment (0-50 = deception, 51-100 = truth), and this veracity judgment was used to calculate accuracy[2].

Results.

Perceptions of honesty. There was no significant difference in observers' ratings of the honesty of liars (M = 77.22, SD = 8.49) and truth-tellers (M = 78.79, SD = 7.92), F (1, 125) = 1.25, p = .266. There was no significant difference in observers' ratings of the honesty of senders who saw the exemplars (M = 78.67, SD = 8.49) and those who did not (M = 77.22, SD = 8.00), F (1, 125) = 1.26, p = .26. There was no significant interaction between veracity and exemplar condition, F (1, 125) = 1.35, p = .25. Means and standard deviations for perceived honesty for each condition are presented in Table 10.

[2] It is reasonable to ask why accuracy was not calculated as a continuous variable using the honesty scale. Although the honesty scale is a continuous measure, a continuous measure of accuracy could not be meaningfully calculated from the honesty scale. Although each side of the midpoint of the scale meaningfully indicates a judgment that the sender is mostly honest or mostly dishonest, the receiver is not necessarily "more correct" or "more wrong" to have formed a more or less extreme impression of the sender's honesty because I do not have a continuous measure of how objectively honesty the sender was. For example, if a receiver watched a video from a deceptive sender, because it is not clear the extent to which the sender was actually lying, the receiver would not necessarily be more correct to respond "0" on the scale compared to a higher score (e.g. "20"). For this reason, all accuracy analyses were conducted using the dichotomized judgments.

Accuracy and bias. As described above, each observer judged a random selection of six interviews sampled from both Study 1 and Study 2. For Study 1, the 328 observers made 1,098 judgments, but 25 judgments (2.3%) were missing. Thus, observers made a total of 1,073 valid judgments of senders from Study 1.

Observers exhibited a strong truth bias: They judged 89.00% of the senders as truth-tellers. Overall, observers were 46.41% accurate. This accuracy rate was significantly less than chance, $p = .02$. Consistent with the truth bias, observers were significantly more accurate when judging truth-tellers (81.97%) than judging liars (16.75%), χ^2 (1) = 454.96, $p < .0001$, $\varphi = .65$. When judging truth-tellers, observers' accuracy did not significantly differ when judging senders who had not seen the exemplars (81.18%) and those who had seen the exemplars (82.83%), χ^2 (1) = .23, $p = .64$. As expected, when judging liars, observers were significantly more accurate when judging senders who had not seen the exemplars (20.20%) compared to those that had seen the exemplars (13.19%), χ^2 (1) = 5.15, $p = .023$, $\varphi = .094$. Accuracy rates for each condition are presented in Table 11.

Discussion. The results of Study 1 were partially supportive of the predictions. The results frequently demonstrated the predicted differences between liars and truth-tellers. The cybernetic perspective predicts that liars should have benefited from exposure to the exemplars, such that differences between liars and truth-tellers would have been attenuated by having watched the exemplar videos. Study 1 generated several pieces of evidence that are consistent with this theoretical prediction.

Forthcomingness. The pattern of results for both measures of forthcomingness (i.e., talking time and details) was directly consistent with the "counterfeiting" hypothesis. That is, liars who saw the exemplars behaved more similarly to truth-tellers (viz. were more

forthcoming). This suggests that liars, as predicted, adjusted their reference standards in order to behave more similarly to truth-tellers. Importantly, the participants shown in the exemplar videos took part in activities that were different from the activity liars were supposed to have done. Thus, liars could not simply repeat exactly what they heard and saw in the exemplars in order to provide a compelling account. Rather, the exemplars could only provide information about the *properties* of a truthful performance, not its specific *content*. Liars who saw the exemplars were, therefore, better able to adopt the manner of truth-tellers' behavior (i.e., becoming more forthcoming), even if they could not provide statements that were identical to truth-tellers' in terms of content.

Some caution is warranted in drawing the conclusion that the results supported the predictions because, although simple effects tests supported the hypotheses, the interaction between veracity and exemplar condition was not statistically significant. However, the effect sizes for the differences between liars who saw the exemplars and those who did not are moderate to large in size, and the effects are consistent across several of the individual measures of forthcomingness.

Observer ratings and self-reports. I expected that exposure to the exemplars would induce liars to behave more in line with truth-tellers – that is, to appear to be more cooperative, clearer and more logical, and less deliberate. The predictions about ratings of cooperativeness and clarity/logic were not supported. The means of these ratings were in the predicted direction (see Table 5), but the observed differences were not significant. It is not clear why the only predicted significant interaction was observed for ratings of deliberateness. This explanation is speculative, but it is possible that the specific manner in which liars adjusted their behavior in

response to the exemplars (e.g., providing more details) was detected in the ratings of deliberateness but not in the ratings of cooperating or clarity/logic.

In the case of the deliberateness ratings, the predicted pattern emerged, but opposite to my prediction, liars who did not see the exemplars were perceived as less deliberate than truth-tellers. It is not entirely clear why this occurred. One possible explanation is that liars appeared less effortful because they were simply speaking less and providing fewer details, and the exemplars induced them to speak more. Although the direction was opposite what I expected, the pattern was consistent with predictions derived from the cybernetic perspective: The difference between liars and truth-tellers was diminished when liars saw the exemplars.

It is also worth noting that the observer ratings of deliberateness and the self-reports diverge in opposite directions: Observers rated liars who saw the exemplars as appearing to think harder than liars who did not see the exemplars, but the liars themselves reported planning less and feeling less prepared when they saw the exemplars as compared to when they did not. Although there is no direct evidence that this is the case, the observer ratings and the self-reports may measuring different constructs. That is, the self-reports of deliberateness may be measuring participants' perception of how easy the interview was, whereas the observer ratings of deliberateness may be measuring the extent to which it appeared that the participants were exerting effort and were engaged with the interview. It may have been the case, therefore, that liars who saw the exemplar videos perceived the interview as easier, as they had a clearer sense of how they should act in order to create a more compelling performance.

Veracity judgments. Consistent with the predictions, liars who saw the exemplars were judged less accurately compared to liars who did not see the exemplars. Thus, liars exposed to reference information succeeded in eluding detection more frequently than liars who did not

have such information. This result is supportive of the hypothesized "counterfeiting" process. However, this conclusion should be tempered by the fact that the groups did not significantly differ in the extent to which observers rated their honesty on the 101-point scale.

Observers exhibited an unusually strong truth-bias (89% of messages were judged to be truthful, according to the dichotomized measure), one of the highest observed in the deception literature (see Bond & DePaulo, 2006). One likely explanation is that because observers were not given any information about the likelihood that the senders were lying, they generally assumed the senders were being honest (see Levine, 2014). Generally, in lie detection studies, observers are informed that the people whom they are judging might be lying. As a consequence, they are more willing to judge messages as deceptive than they ordinarily would in everyday life, in which people usually take others at their word and in which it often goes against social norms to openly express suspicion that one's interaction partner is lying. Here, because the observers were not told that they might be viewing deceptive messages, they naturally tended to judge the senders as truthful.

Energy consumption. I expected that, because lying is frequently more taxing on mental resources than honesty, liars would seek to compensate for this higher consumption of energy by taking more candy when it was offered to them at the conclusion of the study. Additionally, because I predicted that the exemplars should have made deception easier and less demanding of resources, liars who saw the exemplars should have taken less candy than those who did not see the exemplars. Participants in each condition took candy in a pattern consistent with the predicted pattern of energy consumption, although these effects were only marginally significant. The link between the consumption of self-regulatory resources and blood glucose levels – and thus, taking candy in response to taxing mental processes – has only been explored in a few

studies (Hagger, Wood, Stiff, & Chatzisarantis, 2010). Given the marginally significant effects in the present study, replication will be essential for providing further support for the hypotheses concerning energy consumption and deception.

The effect of reference information on truth-tellers. Several unexpected effects were observed. The exposure to the exemplars unexpectedly seemed to influence the behavior and self-reports of the truth-tellers as well as the liars. This is apparent in the disordinal interactions for the self-reports of deliberateness and observer ratings of nervousness. The cybernetic perspective advanced here is compatible with the finding that reference information affects honest social actors as well as deceptive social actors, but I made no predictions about how exposure to the exemplars might influence truth-tellers in the present study. These patterns nevertheless suggest that liars and truth-tellers both responded to the exemplars, but they did so in ways that differed and sometimes influenced their behavior in diverging directions. Why would it be the case that liars and truth-tellers respond to information differently? The cybernetic perspective would interpret this as a difference of goals. That is, because liars and truth-tellers have different social goals, the meaning and potential usefulness of the reference information differs for each of them. The precise processes underpinning these observed effects are not clear, and the present study was not designed to investigate them. This topic is potentially fertile ground for future research.

Study 2

Purpose. Study 1 investigated a fairly direct method of providing access to relevant reference information. In the real world, reference information is probably rarely so directly available. Rather, people naturally observe others' behavior and their own behavior for information about how to behave. Study 2 aimed at examining whether liars learn to adjust their

reference values by collecting and using information from their own prior truthful performances. In other words, liars may use information from their own past behavior in order to guide the ways in which they deceive. As in Study 1, participants either lied or told the truth about having taken part in a memory activity (henceforth, the *critical activity*). This study manipulated whether or not participants performed activities that are similar to the activity about which they may or may not have been lying (henceforth, *analogous activities*). These analogous activities may serve the same function as the exemplars in Study 1 were hypothesized to serve: They provide information that helps liars to deceive in ways that resemble the truth to a greater extent. As described below, the analogous activities in Study 2 were identical to those activities described in the exemplars in Study 1, so some of the participants in Study 2 should have received, from observing their own behavior, exactly the same information as would be available from the exemplar videos.

It is not necessarily enough, however, that liars have taken part in the analogous activities. The cybernetic perspective of self-presentation and deception predicts a more nuanced pattern. Specifically, in order for liars' reference values to be specified in line with the analogous activities (thus becoming more adaptive), liars must also have a mental representation of how one should report on these activities in an interview. Merely performing these other tasks is not likely to be enough to provide such information. One way to obtain such information is to be interviewed about those analogous activities. In the course of truthfully answering questions about the analogous activities, liars would gain a better understanding of the properties of a performance that is adaptive in that context (e.g., the amount of detail to include, how to structure the description). Liars would thereby gain the necessary information to specify their reference values such that they would be able to provide an account about the critical activity. By

varying whether participants are interviewed about the analogous activities, it was possible to test this more specific prediction.

Study 2 consisted of two stages. Stage 1 was designed to test the primary hypotheses about participants' behavior in the interview. In Stage 2, the recordings of the interviews from Stage 1 were shown to observers who judged the veracity of the participants. Each stage is described below.

Participant behavior. Study 2 had five hypotheses[3]:

H_1: I predicted a main effect for veracity such that, overall, truth-tellers, compared to liars, will produce statements that appear more forthcoming, more cooperative, clearer and more logical, and less deliberate.

H_2: I predicted an interaction between veracity and task/questioning condition such that liars who took part analogous activities *and* truthfully respond to questions about those analogous activities will make statements that are more forthcoming, more cooperative, clearer and more logical, and less effortful than (1) liars who do not take part in the analogous activities or (2) liars who take part in analogous activities but are not asked questions about them.

H_3: As in Study 1, I predicted a main effect for veracity such that that liars would self-report having planned and prepared for the interview, compared to truth-tellers

H_4: I predicted an interaction between veracity and task/questioning condition such that the difference on self-reported deliberateness between liars and truth-tellers would be reduced in the three-task condition, compared to the three-task, limited questioning condition and the single-task condition. Stated differently, I expected liars to find the

[3] Study 1 and Study 2 were both designed before any data were collected, so the predictions for Study 2 do not account for the findings of Study 1.

interview more challenging than truth-tellers but only when they did not have the practice of answering questions about the two analogous activities.

H_5: As in Study 1, I measured the number of pieces of candy the participants took following the procedures in Study 2, and I predicted a main effect for veracity such that liars would take more candy than truth-tellers.

Although I predicted that liars would take more candy than truth-tellers, it was less clear how veracity would interact with the interview/task manipulation. Because performing more tasks may, in itself, demand more energy, it was plausible that those who did three activities (in the three-task and the three-task, limited questioning conditions) would take more candy than those who did only one.

Deception detection. Because being interviewed about the analogous activities should provide reference information for liars to use to make more convincing performances, I expected that accuracy for detecting liars in the three-task condition would be lower compared to the accuracy for detecting liars in the single-task condition and liars in the three-task, limited questioning condition.

Stage 1: Primary hypothesis testing. Study 2 followed the procedures of the embarrassed experimenter paradigm, with additional manipulations prior to the veracity manipulation and during the interview. This study makes use of a 2 (veracity: truthful vs. deceptive) x 3 (tasks/questioning: single task vs. three tasks vs. three tasks, limited questioning). Participants in Study 2 were undergraduates at a large urban university in the northeastern United States (N = 103). Three participants in the deceptive condition refused to comply with the experimenter's request to lie in the interview. As in Study 1, in these cases, the experimenter pretended to locate the lost key to the room, and the procedures were continued as if that

participant were in the truthful condition. Again, as in Study 1, I conducted the statistical analyses with these participants removed, and the results were consistent with analyses conducted with the participants included. All reported analyses include the full sample. However, two participants were non-fluent in English and one participant (in the deceptive condition) admitted he had not completed the critical task, so the data for these participants were discarded. This left a total sample size of $N = 100$. Participants were 72.7% (72) female and 27.3% male (27). One participant refused to provide a gender identification. The sample was 53.1% Hispanic, 20.4% Black, 11.2% White, 6.1% Asian, and 9.1% other races and ethnicities. The average age was $M = 20.69$ years ($SD = 3.31$), with a range from 18 to 35. The median age was 20 years. Participants were given course credit in exchange for their participation.

Prior to the veracity manipulation, participants randomly assigned to the *three tasks* and the *three tasks, limited questioning* conditions took part in two activities that are conceptually similar to the critical activity (i.e., the analogous activities) in addition to the critical activity. Specifically, they took part in the coin-value calculation task and the math problem task that were used in the exemplar stage in Study 1, before taking part in the critical activity. Participants assigned to the *single task* condition did not complete the analogous activities and only completed the critical activity.

During the interview stage, participants in the *three tasks* condition were asked questions about the analogous activities prior to being asked questions about the critical activity. Participants in the *single task* and the *three tasks, limited questioning* condition were only asked questions about the critical activity. Prior to being interviewed, all participants listened to the audio recording of the interview questions. As in Study 1, this procedure was included to eliminate the possibility that any observed effects were due to exposure to the questions in the

three tasks condition. The interviews were video recorded. Coding of the interviews took place largely as described above, but each line of questioning in the interviews (viz. pertaining to different activities) will be coded separately. This separation permitted comparison between conditions that included or did not include more than one line of questioning, and it permitted within-subjects comparisons between the different stages of the interview (for participants in the *three task* condition).

Results.

Coding. Coding of forthcomingness variables (i.e., details and talking time) was conducted in a manner highly similar to the coding for Study 1. However, for participants in the three-task condition, each phase of questioning for each task was coded, resulting in nine ratings of each variable, instead of three. To establish reliability, two trained coders, who were blind to hypotheses and blind to condition, rated a sub-sample of the 20% (n = 20) of the sample of interviews from Study 2. Inter-rater reliability was adequate, all intra-class correlations coefficients > .71. Once adequate reliability had been established, one coder rated the remainder of the videos. Coding by MTurk observers was conducted in the same manner as described for Study 1. However, for participants in the three-task condition, the observers watched a video for each task separately and provided ratings for each task. Each video was viewed and rated several times (M = 8.58 ratings, SD = 2.12, median = 9).

Forthcomingness: Between-subjects (all conditions). It was possible to conduct between-subjects comparisons of the participants' behavior during the interview, comparing the single-task condition to the third task (i.e., critical task) of the three-task and three-task, limited questioning conditions. I analyzed the measures of talking time and details using two two-way MANOVAs, using the measures the three interview phases as dependent variables for each

analysis. Means and standard deviations for talking time and details, by condition, are presented in Table 12.

For talking time, there was a significant multivariate main effect for veracity, Wilk's Λ = .800, F (3, 91) = 7.49, p < .0001, partial η^2 = .20. There was a significant main effect for task/questioning condition, Wilk's Λ = .849, F (6, 182) = 2.56, p = .02, partial η^2 = .079. In contrast to the predictions, there was no significant interaction between veracity and task/questioning condition, Wilk's Λ = .905, F (6, 182) = 1.54, p = .17.

Follow-up univariate ANOVAs revealed a significant main effect for veracity on talking time in the second phase of the interview (i.e., the phase concerning the environment), F (1, 92) = 14.82, p < .0001, partial η^2 = .14, such that truth-tellers spent a significantly greater amount of time talking (M = 72.78%, SD = 9.62%) than liars (M = 63.59%, SD = 14.23%) in that phase. There was no significant effect for veracity on the first phase, F (1, 92) = .12, p = .73, and no significant main effect on the third phase, F (1, 92) = .19, p = .67.

Follow-up univariate ANOVAs revealed a significant main effect for task/questioning condition on the first phase of the interview (i.e., the phase concerning the procedures), F (2, 92) = 4.37, p = .015, partial η^2 = .09. Participants in the single-task condition talked for a significantly shorter amount of time (M = 80.33%, SD = 7.31%) compared to participants in the three-task condition (M = 83.40%, SD = 4.43%), t (62) = 2.00, p = .05, d = .57, and the three-task, limited questioning condition (M = 84.17%, SD = 5.59%), t (66) = 2.43, p = .02, d = .96. The three-task condition and the three-task, limited questioning condition did not significantly differ from each other, t (62) = .60, p = .55. There was no significant effect for task/questioning condition on the second phase, F (2, 92) = .24, p = .79, and no significant main effect on the third phase, F (2, 92) = .39, p = .68.

For details, multivariate analyses indicated there was a significant main effect for veracity, Wilk's $\Lambda = .764$, F (3, 91) = 9.38, $p < .0001$, partial $\eta^2 = .24$. There was no significant main effect for task/questioning condition, Wilk's $\Lambda = .937$, F (6, 182) = 1.00, $p = .42$. Against predictions, there was no interaction between veracity and task/interview condition, Wilk's $\Lambda = .764$, F (6, 182) = .38, $p = .89$. Follow-up univariate ANOVAs revealed that veracity had a marginally significant effect in the first phase of the interview (i.e., the phase concerning the procedures), F (1, 93) = 3.54, $p = .063$, partial $\eta^2 = .037$, such that truth-tellers provided more details (M = 15.51, SD = 6.33) than liars (M = 13.38, SD = 5.65). Veracity had a significant effect in the second phase of the interview (i.e., the phase concerning the environment), F (1, 93) = 25.35, $p > .0001$, partial $\eta^2 = .21$, such that truth-tellers provided more details (M = 16.11, SD = 7.21) than liars (M = 9.74, SD = 4.54).

Forthcomingness: Within-subjects (three-task condition). For participants in the three-task condition, observers rated the participants' behavior separately for each task. As such, I was able to conduct within-subjects comparisons for participants in the three-task condition. To examine the impact of veracity on participants' behavior while answering questions about each of the tasks in the three-task condition, I calculated an average talking time for each participant for each task and a total measure of details provided in each interview phase for each task. I then conducted repeated-measures ANOVAs for average talking time and for total details. Means for average talking time and total details by veracity condition are presented in Table 13.

Overall, talking time did not significantly vary across the three tasks, Wilk's $\Lambda = .947$, F (2, 27) = .76, $p = .48$. However, liars and truth-tellers differed in talking time across the three tasks, Wilk's $\Lambda = .675$, F (2, 27) = 6.49, $p = .005$, partial $\eta^2 = .33$. As can be seen in Figure 13 and Table 13, liars and truth-tellers behaved similarly for the first two tasks, but they diverged on

the third task such that liars spoke significantly less than truth-tellers. A linear within-subjects contrast supports this pattern, F (1, 28) = 8.81, p = .006, partial η^2 = .24. Simple effects tests were generally consistent with this pattern and indicate that there was no significant difference in the amount of time liars and truth-tellers spent talking for the first task, t (28) = .12, p = .90, and second task, t (28) = .43, p =.67, and there was a marginally significant difference between liars and truth-tellers for the third task, t (28) = 1.93, p = .06, d = .71.

Multivariate tests there was a marginally significant trend such that participants different in the number of details they provided for each task, Wilk's Λ = .835, F (2, 27) = 2.67, p = .087, partial η^2 = .17. Additionally, there was a significant interaction between the within-subjects factor (the three tasks) and veracity, such that liars and truth-tellers differed in the number of details they provided for each task, Wilk's Λ = .564, F (2, 27) = 10.44, p < .0001, partial η^2 = .44. As can be seen in Table 13 and Figure 14, liars and truth-tellers provided similar amounts of details in the first two tasks, and they diverged in the third task, with liars providing significantly fewer details than truth-tellers. A linear within-subjects contrast is consistent with this pattern, F (1, 28) = 12.85, p = .001, partial η^2 = .31, Simple effects tests were consistent with this pattern and indicate that there was no significant difference in the amount of details liars and truth-tellers provided for the first task, t (28) = .22, p = .83, and second task, t (28) = .22, p =.83, and there was a significant difference between liars and truth-tellers for the third task, t (28) = 2.73, p = .01, d = 1.05.

*Observer ratings***:** *Between-subjects (all conditions).* I analyzed ratings of cooperation, clarity/logic, and deliberateness, for the third task in all conditions, using a two-way MANOVA, with veracity and task/questioning condition as factors. The analysis revealed a significant main effect for veracity, Wilk's Λ = .892, F (3, 91) = 3.68, p = .015, partial η^2 = .11. There was no

main effect for task/questioning condition, Wilk's $\Lambda = .930$, F (6, 182) = 1.12, $p = .35$. Against predictions, there was no interaction between veracity and task/questioning condition, Wilk's $\Lambda = .979$, F (6, 180) = .33, $p = .92$.

To clarify the main effect of veracity, I conducted follow-up analyses. Veracity had a significant effect on ratings of cooperation, F (1, 93) = 7.96, $p = .006$, partial $\eta^2 = .079$, such that truth-tellers appeared more cooperative ($M = 79.55$, $SD = 6.69$) compared to liars ($M = 74.43$, $SD = 11.31$). Veracity also had a significant effect on ratings of clarity/logic, F (1, 93) = 11.11, $p = .001$, partial $\eta^2 = .107$, such that truth-tellers appeared clearer and more logical ($M = 72.15$, $SD = 9.03$) compared to liars ($M = 64.53$, $SD = 13.44$). Veracity did not a significant effect on ratings of deliberateness, F (1, 93) = 1.51, $p = .22$. Means and standard deviations for each variable by condition are presented in Table 14.

Observer ratings: Within-subjects (three-task condition). Means and standard deviations for each task, by veracity condition are presented in Table 15. A repeated-measures ANOVA indicated that there were no significant within-subjects differences across ratings of cooperation, Wilk's $\Lambda = .964$, F (2, 27) = .50, $p = .61$, and this factor did not significantly interact with veracity, Wilk's $\Lambda = .873$, F (2, 27) = 1.97, $p = .16$.

There were no significant within-subjects differences across ratings of clarity/logic, Wilk's $\Lambda = .903$, F (2, 27) = 1.45, $p = .25$, but this factor interacted marginally significantly with veracity, Wilk's $\Lambda = .812$, F (2, 27) = 3.12, $p = .06$, $\eta^2 = .19$. As can be seen in Figure 15 and Table 15, liars were rated as less clear and logical compared to truth-tellers for the first task; liars and truth-tellers were rated more similarly for the second task; and for the third task, liars were again rated as less clear and logical compared to truth-tellers. A quadratic within-subjects contrast is consistent with this pattern, F (1, 28) = 6.46, $p = .017$, partial $\eta^2 = .19$.

Although there were no significant within-subjects differences for ratings of deliberateness, Wilk's $\Lambda = .993$, $F\ (2, 27) = .10$, $p = .91$, there was a marginally significant interaction, Wilk's $\Lambda = .819$, $F\ (2, 27) = 2.99$, $p = .06$, partial $\eta^2 = .18$, such that ratings of deliberateness varied across interviews in a different pattern for liars and for truth-tellers. As can be seen in Figure 16 and Table 15, liars tended to be rated as decreasingly deliberate across each of the tasks, and truth-tellers tended to be rated as increasingly deliberate across each of the tasks. A linear within-subjects contrast is consistent with this pattern, $F\ (1, 28) = 4.80$, $p = .04$, partial $\eta^2 = .146$.

Self-reported effort and deliberateness. Participants were presented with the same self-report items pertaining to self-perceived deliberateness as in Study 1 (see Appendix A). These four self-report items were analyzed using a two-way MANOVA, with veracity and task/questioning condition as factors. Multivariate analyses suggested that there was no main effect for veracity, Wilk's $\Lambda = .949$, $F\ (4, 90) = 1.21$, $p = .31$, or for task/questioning condition, Wilk's $\Lambda = .897$, $F\ (8, 180) = 1.25$, $p = .27$. There was no significant interaction between veracity and task/questioning condition, Wilk's $\Lambda = .892$, $F\ (8, 180) = 1.32$, $p = .23$. Because no multivariate effects were found, I did not conduct any follow-up analyses for these items. Means and standard deviations for each item are presented in Table 16.

Energy consumption. As in Study 1, as an indirect measure of energy consumption, I measured the number of pieces of candy the participants took at the conclusion of Study 2. There was no main effect for veracity, $F\ (1, 93) = 1.50$, $p = .22$. There was a marginally significant main effect for task/questioning condition, $F\ (2, 93) = 2.74$, $p = .07$, partial $\eta^2 = .056$, such that the three-task, limited questioning condition took the most candy ($M = 2.21$, $SD = 1.61$), followed by the three-task condition ($M = 1.80$, $SD = 1.32$), and the single-task condition ($M =$

1.46, *SD* = 1.04). The single-task condition took significantly less candy than the three-task, limited questioning condition, t (67) = 2.30, p = .02, d = .55. Neither the three-task, limited questioning condition, t (62) = 1.09, p = .28, nor the single-task condition, t (63) = 1.17, p = .25, differed from the three-task condition. Counter to predictions, there was no significant interaction between veracity and task/questioning condition, F (2, 93) = .24, p = .79. Means and standard deviations for each condition are presented in Table 17. A plot of the means is presented in Figure 17.

Stage 2: Deception detection. Data collection for Stage 2 of Study 2 and Stage 3 of Study 1 occurred simultaneously and used the same observers. The analyses for Stage 2 of Study 2 are highly similar to those of Stage 3 of Study 1. Because all participants told the truth about the analogous activities, all analyses were conducted between subjects, concerning the critical activities only.

Results.

Perceptions of honesty. There was no significant main effect for veracity on ratings of honesty, F (1, 93) = 2.30, p = .13. There was no significant main effect for task/questioning condition on ratings of honesty, F (2, 93) = .42, p = .66. There was no significant interaction between veracity and task/questioning condition, F (1, 93) = .23, p = .80. Means and standard deviations for honesty are presented in Table 18.

Accuracy and bias. For Study 2, the 328 observers made 870 judgments, but 14 judgments (1.6%) were missing. Thus, observers made a total of 856 valid judgments of senders from Study 2. As in Study 1, observers were highly biased to judge statements as true: They rated 88.45% of the statements (about the critical activity) in Study 2 to be true. Overall, observers were 53.97% accurate. This accuracy rate was significantly greater than chance, p

= .022. Consistent with the truth bias, observers were significantly more accurate when judging true statements (83.13%) than they were when judging deceptive statements (14.56%), χ^2 (1) = 395.97, $p > .0001$. Accuracy when judging participants in the single-task condition (54.04%), the three-task condition (51.40%), and the three-task, limited questioning condition (56.49%) did not significantly differ, χ^2 (2) = 1.49, $p > .48$. In summary, contrary to predictions, the task/questioning condition did not affect accuracy. Accuracy rates for each condition are presented in Table 19.

Discussion. If we assume that being interviewed about the analogous activities should have had the same effects as the exemplars in Study 1 (as I predicted), the effects observed in Study 1 were not reproduced in Study 2. Some of the predicted differences between liars and truth-tellers emerged (e.g., in ratings of cooperativeness and clarity/logic), but the central predictions of Study 2 were not supported. There was no evidence that being interviewed about the analogous activities influenced liars to behave in a manner more similar to truth-tellers. In fact, directly contrary to predictions, liars and truth-tellers in the three-task condition diverged sharply in the amount of time they spent talking and the amount of details they provided for the critical task (see Figures 13 and 14).

However, there remains some evidence that liars and truth-tellers were engaged in different processes in the course of the interview. Some of the most compelling evidence to this point is the observer ratings of clarity/logic and deliberateness in the three-task condition.

For clarity/logic, liars were less clear and logical than truth-tellers when answering questions about the first task. The difference between liars and truth-tellers decreased for the second task. Their behavior diverged once again for the critical task, as liars were once again less clear and logical than truth-tellers.

For deliberateness, liars and truth-tellers began, in the first task, equivalent in apparent effort, and they began to diverge at the second task, truth-tellers becoming apparently more effortful and liars becoming apparently less effortful. They were most different at the third task, for which participants in the deceptive condition were lying, such that truth-tellers appeared even more effortful and liars appeared even less effortful. However, the differences begin to emerge when participants are being questioned about the second task, about which all participants told the truth.

Why would participants in the truthful and deceptive conditions diverge in behavior when they were both telling the truth at this point? The study's design does not permit strong conclusions to be drawn about this. However, one possible explanation – consistent with the self-regulatory perspective I have taken – is that liars and truth-tellers are engaged in different processes of goal pursuit from the outset of the interview. Although this is admittedly highly speculative, being questioned about the analogous activities may have provided feedback of a kind I did not anticipate. Responding to the questions about the analogous activities may have given liars a signal they could (or should) begin to reduce the apparent effort they were exerting, without any serious cost to (or in fact, to the benefit of) their ability to proceed successfully through the interview. Liars were likely concerned primarily about completing the interview without being detected[4]. As such, liars may have been attentive for signals about how much effort they should exert in order to create a credible impression while also anticipating the fact they could not, in fact, generate responses for the third task that could possibly match those they could generate for the first two. That is, because they anticipated that they could not perfectly

[4] Informally, the experimenters reported to me that participants in the deceptive condition frequently expressed concern, before they were questioned, that they might not succeed in the interview.

counterfeit a truthful performance for the third task, liars may have begun adjusting their performances before being questioned about the critical task.

There are at least two possible advantages to doing this: First, performing differently in the earlier stages of the interview may have been a self-presentational strategy. Doing so would make differences in the third task less remarkable. Stated differently, the low quality of statements for the first two tasks could have been intended to serve as "camouflage" for the low quality in the third task. Second, making lower quality statements for the analogous activities may have altered liars' reference standards for their behavior, in preparation for the performance they anticipated when answering questions for the critical task. Making gradual adjustments to the reference values of the feedback system may demand fewer attentional resources than abrupt changes. This could be the self-regulatory equivalent to coming to a rolling stop, rather than slamming the brakes.

Again, these possibilities are speculative, and this speculation should be tempered by the fact that these patterns did not emerge on other measures (e.g., talking time and details). Moreover, there are methodological limitations that should qualify inferences drawn from the within-subject analyses of the three-task condition: The sequence of the tasks (i.e., the math problems, coin-value counting, and card memorization) was not varied, and it is therefore possible that an order effect may account for some of the results. Ultimately, future research will be needed to clarify these results.

To summarize, being interviewed about the analogous activities did not have the effect I predicted: Liars who were interviewed about the analogous activities did not behave more similarly to truth-tellers than those who did not. However, there is some evidence from the

observer ratings that liars and truth-tellers approached the interview differently, even in parts of the interview during which both truth-tellers and (eventual) liars were being honest.

Chapter 3: General Discussion

The present studies were designed to provide an initial test of some of the key predictions of the cybernetic perspective on self-presentation and deception that I described earlier, specifically those predictions concerning the adjustments of the reference values of feedback systems. Using the analogy of counterfeiting, I predicted that exposure to reference information about truthful behavior would enable liars to adjust their performances to be more in line with truth-tellers' performances.

The results of Study 1 were consistent with the broad hypothesis that self-presentational processes are controlled by feedback mechanisms, and more specifically, the results suggest that both liars and truth-tellers can and do alter their self-presentational behavior in response to reference information. As predicted, on two measures of forthcomingness and on observer ratings of deliberateness, liars behaved more similarly to truth-tellers when they watched the exemplar videos. Moreover, liars who saw the exemplars eluded detection at higher rates than liars who did not see the exemplars. Overall, the results of Study 1 suggests that liars used information from the exemplars to guide their performances to be more in line with truth-tellers' performances, and this counterfeiting appears to translate into less detectable deception.

Although Study 2 found many of the predicted overall differences between liars and truth-tellers, its results were contrary to the central hypotheses: Liars who were interviewed about the analogous activities did not behave in a manner more similarly to truth-tellers. Neither the analogous activities themselves nor being interviewed about them appeared to provide relevant reference information as predicted.

It is not clear why Study 2 failed to find the predicted effect. The analogous activities from Study 2 were the same activities used in the exemplar generation stage of Study 1. Thus,

the same information available to participants through the exemplars in Study 1 should have been available through the participants' own responses to the questions about the analogous activities in Study 2, assuming that participants in Study 2 responded to those questions in a manner similar to the people interviewed in the exemplars. That assumption, however, may not have been the case. The observer ratings of clarity/logic and deliberateness for participants in the three-task condition suggest that liars' behavior while questioned about the analogous activities differed from that of truth-tellers. If that is the case, participants in the three-task condition may not have had access to the reference information I expected that manipulation to provide. If this reasoning is correct, the manner in which people obtain relevant reference information is more complicated than I initially supposed. At minimum, it appears that learning from one's own social interactions is not as simple of a matter as I predicted.

Theoretical Perspectives on Deception

The foregoing theoretical discussion and the present studies were purposed to advance the understanding of deception and self-presentation. Psychological research on deception remains heavily reliant on theoretical frameworks introduced in the 1980s and '90s. Arguably the most influential of these theories have been the cognitive approach (Zuckerman, DePaulo, & Rosenthal, 1981), which describes deception as a more controlled, more physiologically arousing, and more cognitively challenging task than truth-telling, and the self-presentational perspective (DePaulo, 1992; DePaulo et al., 2003), which describes deception as a special case of normal self-presentational behavior. These perspectives are empirically supported and theoretically valuable, and they informed much of my reasoning here. The present theoretical perspective complements those previous perspectives and surmounts some of the limitations in their explanatory scope.

What can the cybernetic perspective contribute beyond what previous theories have? Although other researchers have argued that deception is goal-directed (e.g., Buller & Burgoon, 1996) and related to basic self-regulatory processes (Granhag & Hartwig, 2008; Hartwig, Granhag, Strömwall, & Doering, 2010), I know of no other theory of deception that thoroughly explores how self-regulatory feedback mechanisms operate in deception. Although the present theory shares many predictions with Interpersonal Deception Theory (Buller & Burgoon, 1996), IDT makes no statements about the underpinning psychological mechanisms that explain those predictions (see DePaulo, Ansfield, & Bell, 1996). The cybernetic approach leads to a novel understanding of how liars might obtain reference information to guide their deceptions. In the present studies, some of the predictions of this framework were tested, and Study 1 found support for the prediction that liars can observe truthful behavior in order to craft more compelling lies. Deception researchers have recently tested the effects of exposing liars and truth-tellers to model truthful statements, but the one study (of which I am aware) in which an exemplar manipulation is used is largely atheoretical (Leal et al., 2015).

The cybernetic perspective also offers alternative and more complete explanations for existing empirical findings. For example, consider research in the investigative interviewing literature that has found that deceptive suspects in interviews change their behavior in response to direct disclosures of evidence about their activities (e.g., Hartwig et al., 2005, Granhag et al., 2013; Luke et al., 2013), change their behavior when they become aware that evidence may possibly exist (Luke, Dawson, Hartwig, & Granhag, 2014), and change their behavior when they become aware of the tactics the interviewer is going to use (Luke, Hartwig, Shamash, & Granhag, 2013; Tekin et al., 2015). Granhag and Hartwig (2015, 2008) have plausibly interpreted these findings as deceptive suspects engaging in a strategic, self-regulatory adaptation to threats to

their credibility. More specifically, Granhag and Hartwig (2008) have argued that, in investigative interviews, liars tend to use avoidance-based self-regulatory strategies (i.e., providing short, simple statements and denying involvement), but alerting suspects to evidence prompts them to deviate from these normal strategies (i.e., providing statements that are closer to the truth). The cybernetic perspective would interpret these findings somewhat differently: When evidence is disclosed to a deceptive suspect, he or she uses it as input to adjust the reference value(s) of the feedback system controlling his or her behavior. Evidence disclosures provide information about what kinds of statements are (non)credible. Suspects can and do use this information to guide their behavior.

Indeed, there is close similarity between the results of Study 1 and the results of the research on suspects' responses to evidence disclosure. The consistent finding is that deceptive suspects and truthful suspects behave differently (i.e., contradict the evidence to a greater extent) when they are unaware of the evidence that exists against them, but these differences are reduced when deceptive suspects are made aware of the evidence (for a meta-analytic review of these findings, see Hartwig, Granhag, & Luke, 2014). Similarly, in Study 1, many of the differences between liars and truth-tellers were mitigated when liars were shown the exemplars.

If the cybernetic interpretation of these findings is correct, one would also expect that evidence disclosure, information about the interviewer's tactics, and warnings that evidence may exist would produce effects that are circumscribed to the particular information carried by each of those interventions. This might explain why, for example, Luke, Dawson, Hartwig, and Granhag (2014) found that when deceptive suspects were warned that investigators might have obtained surveillance tapes of their recordings, many deceptive suspects became highly forthcoming with truthful information, but roughly equally as many were extremely withholding

with details. Extremely few liars adopted a moderate strategy of providing only some detail. The specific content of the surveillance recordings was not disclosed, so liars lacked the information necessary to calibrate their performances to the investigator's knowledge: some overcompensated, others undershot. Luke and his colleagues (2014) did not have a clear interpretation of why they observed this bimodal distribution, but the cybernetic perspective offers an explanation.

Further Investigations

It would require a substantial amount of research to thoroughly investigate the theoretical framework developed here. It is beyond the scope of this report to describe outline a research program designed to conduct such a thorough exploration. Rather, I will focus on some recommendations for future developments related to the present studies and their subject of reference information's role in self-presentation and deception.

Although the results of Study 1 were consistent with the key hypotheses about the effect of the exemplar videos, it is important to note some of the study's limitations. The exemplar video manipulation is a fairly direct way of manipulating access to reference information, but it is rather artificial. Outside of the laboratory, it is not likely people frequently encounter video recordings of social interactions like the one in which they are imminently about to participate. Rather, people witness others engaging in social interactions and engage in interactions themselves, and from these events, people likely obtain relevant reference information (cf. Schlenker, 2012). Research has demonstrated that lie-catchers are more credulous in interactive contexts (Bond & DePaulo, 2006), and this may be because liars are better able to detect, from the behavior of their interaction partner and the situation, what kinds of performances are effective. Study 2 attempted to more closely simulate more naturalistic conditions, but it failed to

find evidence that answering questions about analogous activities assisted in creating more compelling statements when questioned about the critical task. Future research should explore alternative methods of providing reference information to social actors, especially methods that more closely emulate what is likely to occur in naturalistic settings. As I described above, people probably gain reference information from the behavior of their audiences. Research should test methods of providing reference information to senders through the behavior of receivers. This sort of input might take many forms, such as expressions of suspicion (cf. Bond & Fahey, 1987), disclosure of evidence (cf. Hartwig et al., 2005), or direct statements about the receiver's goals.

In future studies of feedback processes in deception and self-presentation, it will be critical to more closely examine the processes involved in the use of reference information. Another limitation of Study 1 is that its design did not include a "placebo" condition. Past research has occasionally demonstrated that bogus stimuli can have effects that are similar to "real" stimuli. For example, Levine and his colleagues (2005) found that providing lie-catchers with bogus lie detection training (i.e., lacking content that would plausibly help people detect lies) improved their accuracy at levels similar to training actually designed to improved accuracy. The cybernetic perspective predicts that liars should benefit from context-specific reference information – that is, information about how truth-tellers behave in the specific context in which a person intends to lie. One possible method of testing this hypothesis would be to show participants a "placebo" exemplar that is similar to an authentic exemplar in modality, length, and other superficial attributes but contains information about a substantively different social interaction. If the predictions of the cybernetic perspective are correct, this placebo exemplar should not provide liars with the information they need in order to effectively counterfeit a truth-teller's performance.

The pattern of deception detection rates in Study 1 conformed to the predicted pattern when analyzing the dichotomized veracity judgments, but the predicted pattern was not found when analyzing the 101-point honesty ratings. In order to increase our confidence that the results of Study 1 are not an artifact of the dichotomized scale, future research should reassess the deception detection rates using a less ambiguous metric (e.g., dichotomous judgments, rather than continuous ratings that are later dichotomized). Additionally, it should be reiterated that the observers received no information about the purpose of the studies. It is possible that a different pattern of results would have been found if the observers were informed that some of the senders were lying.

One limitation of the design of Study 2 is that, in the deceptive condition, the questions concerning the analogous activities were always asked after the participant had committed to lying about the critical task. Because of this design element, it is possible that participants' anticipation of lying may have affected their behavior in the interview during questioning for the analogous activities, as I discussed earlier. One way of testing this possibility would be to use a highly similar paradigm but to interview participants about each task directly after the task is completed. Participants would not be asked to lie until after they had already been interviewed about the first two tasks. That way, the participant's responses could not be influenced by the anticipation of lying in a subsequent interview. Their interviews for the analogous activities would be conducted as authentic truth-tellers and not as eventual liars. Their truthful performances (and the reference standards they specified for themselves therewith) may provide them with the information they need to craft more compelling lies later.

There is an alternative hypothesis worth noting here, however: Altering the manipulation in this way may also alter the manner in which participants use the reference information from

their own performances. That is, because they would not yet have formed the intention of lying, they may attend to that information differently than they would have if they had already formed the intention. A person's social goal (i.e., lying or telling the truth) may influence which inputs, of the innumerable amount humans receive at any given time, receive attention. As I discussed above, if a person intends to lie, they should be particularly attentive to reference information that is relevant to their future deceptive performance. This possibility should be tested empirically, but altering the manipulation such that participants are interviewed immediately following the analogous activities may not equip liars with the reference information necessary for them to improve their performances, since they may not be pay attention to relevant aspects of their performances.

An important implication of the cybernetic framework, which was not explored directly in the present studies, is that feedback processes should occur as continuously, as social interactions are occurring. For the purposes of simplicity, the present studies examined how reference information could be introduced prior to a critical social interaction and only explored highly social interactions in which the interviewer followed a script. In the real world, however, interactions are more dynamic, and people may be actively gathering information about their interaction partners as interactions are happening (cf. Buller & Burgoon, 1996). To more thoroughly investigate predictions derived from the cybernetic perspective, it is important to explore more organic interactions and the ways in which social actors may respond to feedback from their interactions partners. One possible way to experimentally manipulate feedback from interaction partners is to have confederates interact with participants and to vary the behavior of the confederates (e.g., having the confederate express skepticism about a possibly deceptive statement). Similar manipulations have been used successfully in the interviewing and

interrogation literature to influence participants' behavior (e.g., Tekin et al, 2015). Future research could make use of such methods to explore the nuances of feedback processes.

Conclusion

The cybernetic perspective of self-presentation and deception has far-reaching implications, only a tiny fraction of which were examined in the present studies. These studies have provided some evidence in support of basic predictions derived from this framework. Moreover, the present studies made use of a novel paradigm that may be useful in future research on deception. The foregoing description of the cybernetic perspective can hopefully serve as inspiration for further research into feedback mechanisms and their role in social interactions. Although the contribution of this report is relatively modest, it is my hope that it plants a seed from which further research and theoretical developments can sprout.

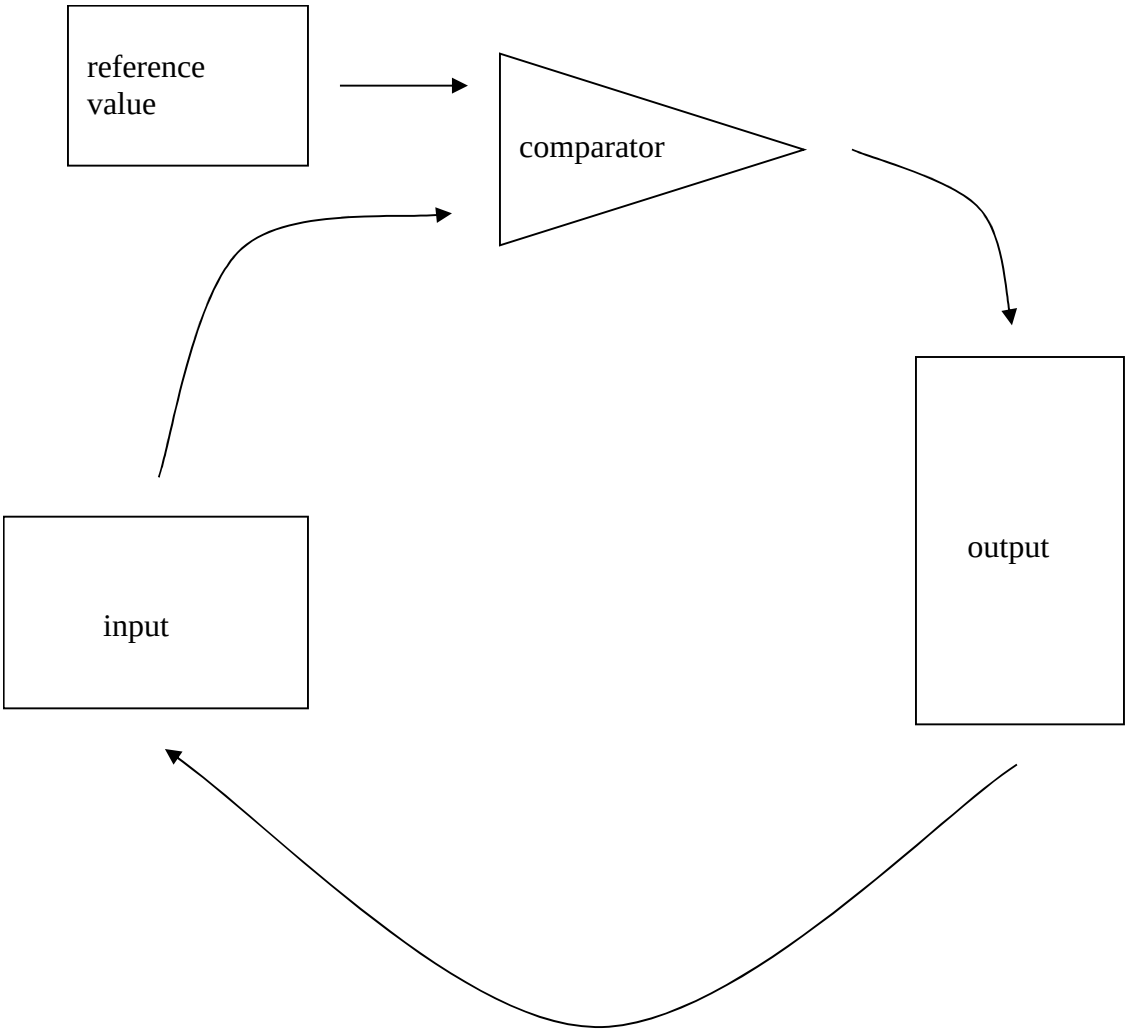

Figure 1. Negative feedback loop.

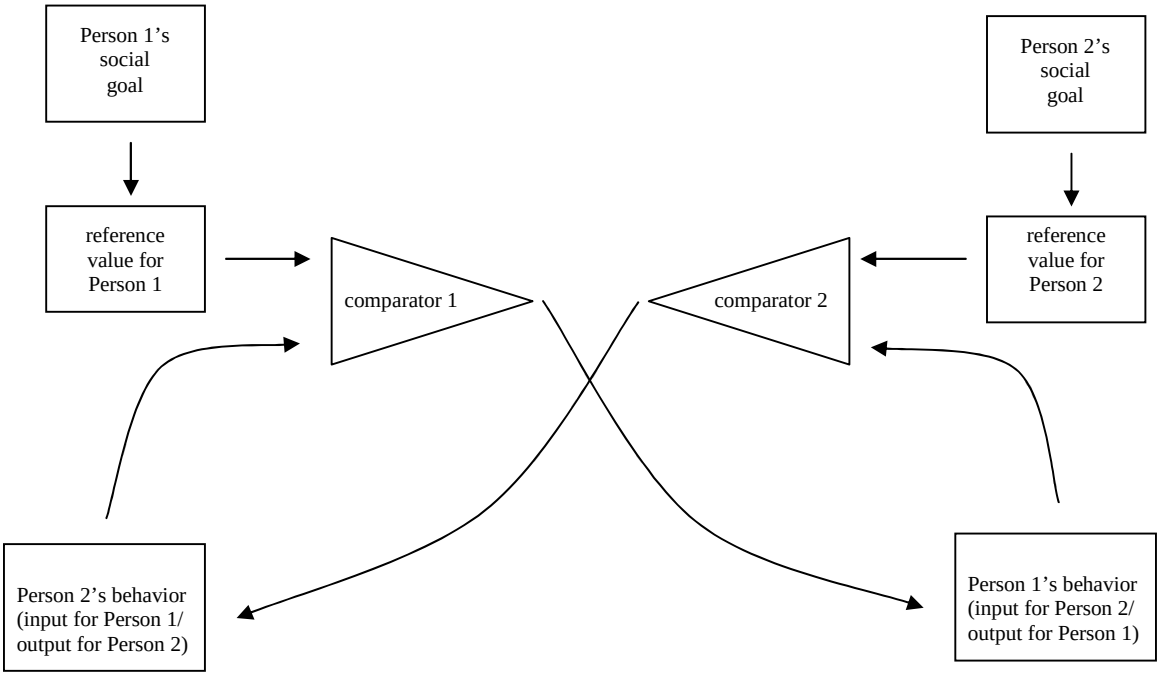

Figure 2. Generalized sketch of self-presentational feedback systems

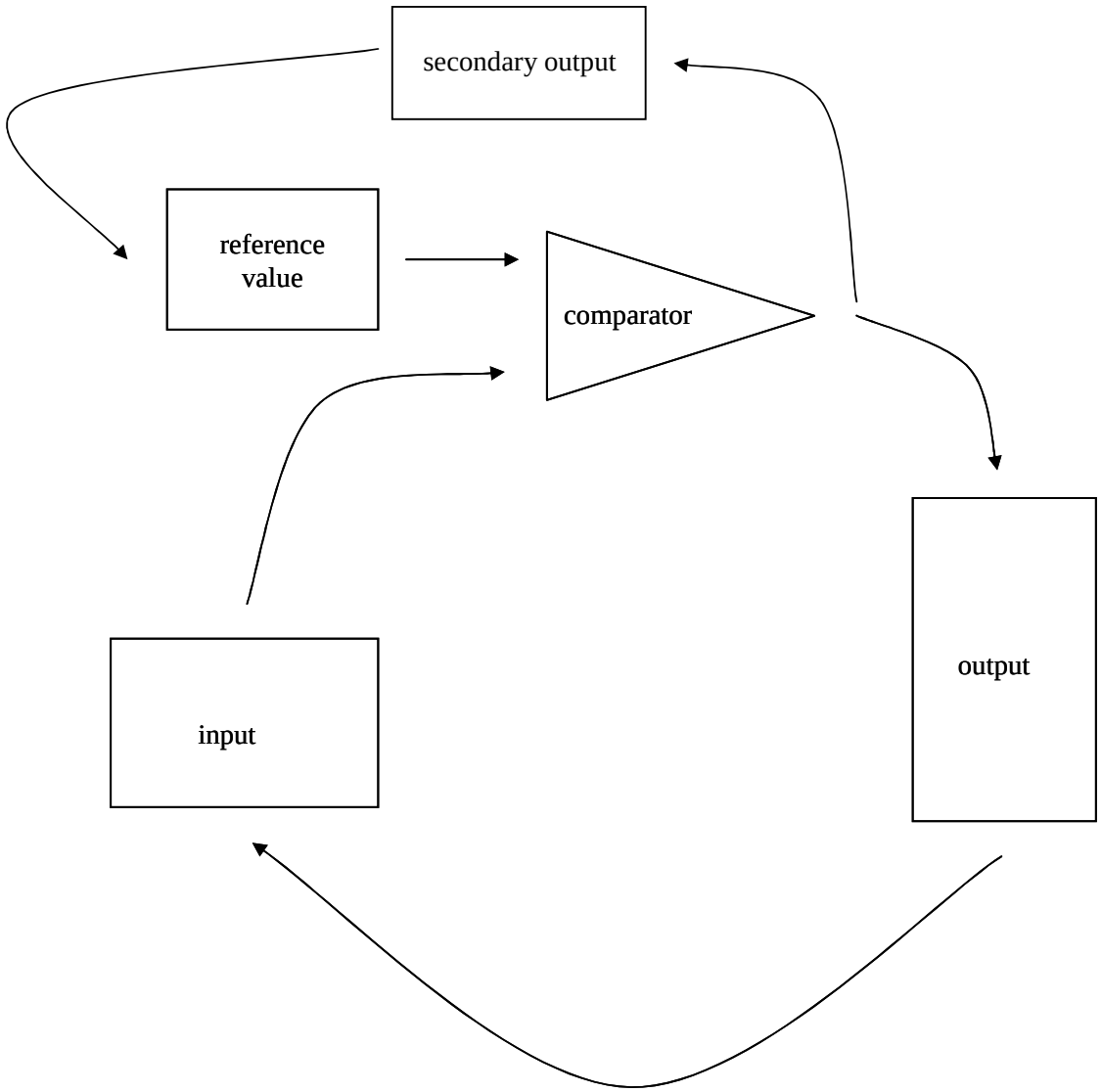

Figure 3. Secondary process adjusting reference value of a feedback loop

Table 1

Means and standard deviations for details and talking time for each phase of the interview, by condition (Study 1)

Talking Time		Mean	SD
Phase 1			
Truthful	No Exemplars	49.65%	12.16%
	Exemplars	50.63%	8.06%
Deceptive	No Exemplars	49.33%	10.43%
	Exemplars	52.44%	11.22%
Phase 2			
Truthful	No Exemplars	73.03%	7.02%
	Exemplars	74.77%	8.51%
Deceptive	No Exemplars	61.19%	14.11%
	Exemplars	67.02%	7.66%
Phase 3			
Truthful	No Exemplars	70.14%	11.85%
	Exemplars	69.14%	8.82%
Deceptive	No Exemplars	67.14%	12.78%
	Exemplars	72.56%	10.51%
Details			
Phase 1			
Truthful	No Exemplars	11.33	4.22
	Exemplars	10.79	2.54
Deceptive	No Exemplars	9.86	3.06
	Exemplars	10.63	3.99
Phase 2			
Truthful	No Exemplars	11.50	2.73
	Exemplars	12.11	3.18
Deceptive	No Exemplars	7.81	2.99
	Exemplars	9.40	2.79
Phase 3			
Truthful	No Exemplars	10.60	4.16
	Exemplars	10.75	3.53
Deceptive	No Exemplars	9.86	3.50
	Exemplars	11.40	3.98

Table 2

Means and standard deviations for average talking time and total details, by condition (Study 1)

Average Talking Time			
Truthful	No Exemplars	64.27%	7.32%
	Exemplars	64.85%	6.35%
Deceptive	No Exemplars	59.22%	10.10%
	Exemplars	64.10%	7.36%
Total Details			
Truthful	No Exemplars	33.43	7.59
	Exemplars	33.64	6.84
Deceptive	No Exemplars	27.53	6.47
	Exemplars	31.43	7.53

Table 3

Univariate ANOVA results for talking time and details (Study 1)

Talking Time		
Veracity	Phase 1	F (1, 125) = .16, p = .69
	Phase 2	F (1, 125) = 31.01, p < .0001, partial η^2 = .20
	Phase 3	F (1, 125) = .01, p = .92
Exemplar	Phase 1	F (1, 125) = 1.18, p = .28
	Phase 2	F (1, 125) = 4.63, p = .03, partial η^2 = .036
	Phase 3	F (1, 125) = 1.25, p = .27
Veracity x Exemplar	Phase 1	F (1, 125) = .32, p = .57
	Phase 2	F (1, 125) = 1.35, p = .25
	Phase 3	F (1, 125) = 2.63, p = .11
Details		
Veracity	Phase 1	F (1, 125) = 1.70, p = .20
	Phase 2	F (1, 125) = 38.31, p < .0001, partial η^2 = .24
	Phase 3	F (1, 125) = .004, p = .95
Exemplar	Phase 1	F (1, 125) = .03, p = .86
	Phase 2	F (1, 125) = 4.53, p = .035, partial η^2 = .035
	Phase 3	F (1, 125) = 1.58, p = .21
Veracity x Exemplar	Phase 1	F (1, 125) = 1.11, p = .30
	Phase 2	F (1, 125) = .911, p = .34
	Phase 3	F (1, 125) = 1.07, p = .30

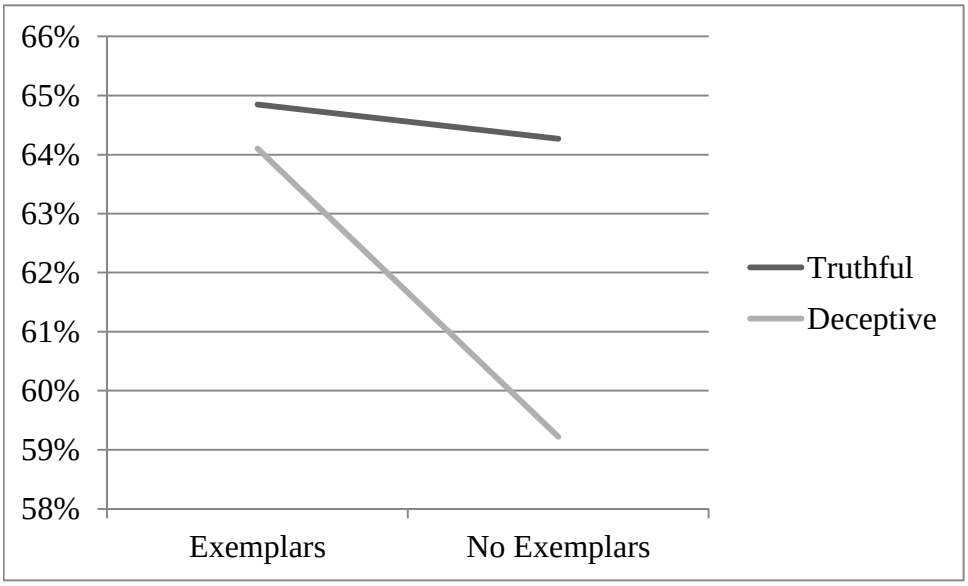

Figure 4. Means plot for average talking time, by condition (Study 1)

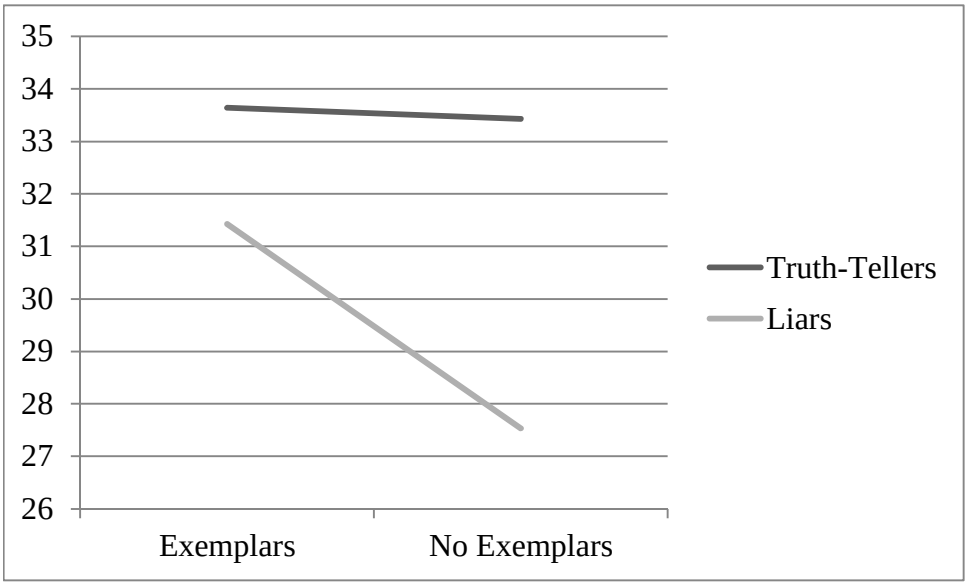

Figure 5. Means plot for total details, by condition (Study 1)

Table 4

Simple effects tests for forthcomingness (Study 1)

Truth-tellers, exemplars vs. truth-tellers, no exemplars	
Average talking time	t (56) = .32, p = .75
Total details	t (56) = .11, p = .91
Talking time, Phase 1	t (56) = .36, p = .72
Talking time, Phase 2	t (56) = .85, p = .40
Talking time, Phase 3	t (56) = .36, p = .72
Details, Phase 1	t (56) = .59, p = .56
Details, Phase 2	t (56) = .78, p = .44
Details, Phase 3	t (56) = .15, p = .88
Liars, exemplars vs. liars, no exemplars	
Average talking time	t (69) = 2.28, p = .026, d = .58
Total details	t (69) = 2.34, p = .022, d = .56
Talking time, Phase 1	t (69) = 1.21, p = .23
Talking time, Phase 2	t (69) = 2.16, p = .035, d = .70
Talking time, Phase 3	t (69) = 1.95, p = .055, d = .53
Details, Phase 1	t (69) = .91, p = .37
Details, Phase 2	t (69) = 2.32, p = .023, d = .55
Details, Phase 3	t (69) = 1.73, p = .088, d = .41
Liars, exemplars vs. truth-tellers, no exemplars	
Average talking time	t (63) = .15, p = .88
Total details	t (63) = 1.07, p = .29
Talking time, Phase 1	t (63) = .96, p = .34
Talking time, Phase 2	t (63) = 3.28, p = .002, d = .80
Talking time, Phase 3	t (63) = .87, p = .39
Details, Phase 1	t (63) = .69, p = .49
Details, Phase 2	t (63) = 3.06, p = .003, d = .76
Details, Phase 3	t (63) = .79, p = .43
Liars, no exemplars vs. truth-tellers, no exemplars	
Average talking time	t (64) = 2.29, p = .026, d = .58
Total details	t (64) = 3.41, p = .001, d = .84
Talking time, Phase 1	t (64) = .12, p = .90
Talking time, Phase 2	t (64) = 4.17, p < .0001, d = 1.08
Talking time, Phase 3	t (64) = .98, p = .33
Details, Phase 1	t (64) = 1.64, p = .11
Details, Phase 2	t (64) = 5.20, p < .0001, d = 1.29
Details, Phase 3	t (64) = .78, p = .44
Liars, exemplars vs. truth-tellers, exemplars	
Average talking time	t (61) = .48, p = .63

Total details	t (61) = 1.21, p = .23
Talking time, Phase 1	t (61) = .72, p = .48
Talking time, Phase 2	t (61) = 3.80, p = .0001, d = .94
Talking time, Phase 3	t (61) = 1.38, p = .17
Details, Phase 1	t (61) = .18, p = .86
Details, Phase 2	t (61) = 3.60, p = .001, d = .91
Details, Phase 3	t (61) = .68, p = .50
Liars, no exemplars vs. truth-tellers, exemplars	
Average talking time	t (62) = 2.58, p = .012, d = .73
Total details	t (62) = 3.66, p = .001, d = .92
Talking time, Phase 1	t (62) = .54, p = .59
Talking time, Phase 2	t (62) = 4.49, p < .0001, d = 1.19
Talking time, Phase 3	t (62) = .71, p = .48
Details, Phase 1	t (62) = 1.29, p = .20
Details, Phase 2	t (62) = 5.55, p < .0001, d = 1.39
Details, Phase 3	t (62) = 1.00, p = .32

Table 5

Ratings of cooperation, clarity/logic, and deliberateness (Study 1)

Cooperation		Mean	SD
Truthful	No exemplars	78.84	8.36
	Exemplars	81.75	7.90
Deceptive	No exemplars	75.84	10.88
	Exemplars	77.23	9.45
Clarity/Logic		Mean	SD
Truthful	No exemplars	70.91	8.60
	Exemplars	76.35	9.67
Deceptive	No exemplars	67.66	11.49
	Exemplars	69.64	9.90
Deliberateness		Mean	SD
Truthful	No exemplars	64.64	7.67
	Exemplars	62.85	9.41
Deceptive	No exemplars	57.70	9.83
	Exemplars	65.23	11.61

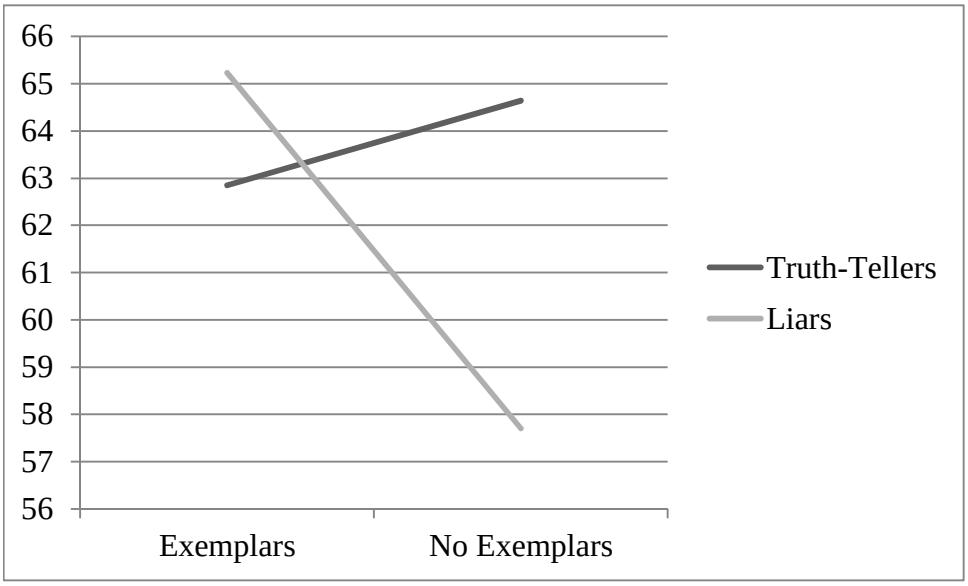

Figure 6. Means plot for ratings of deliberateness, by condition (Study 1)

Table 6

Means for nervousness by condition (Study 1)

		Mean	SD
Truthful	No exemplars	49.05	11.65
	Exemplars	44.29	10.89
Deceptive	No exemplars	44.40	11.76
	Exemplars	49.27	12.18

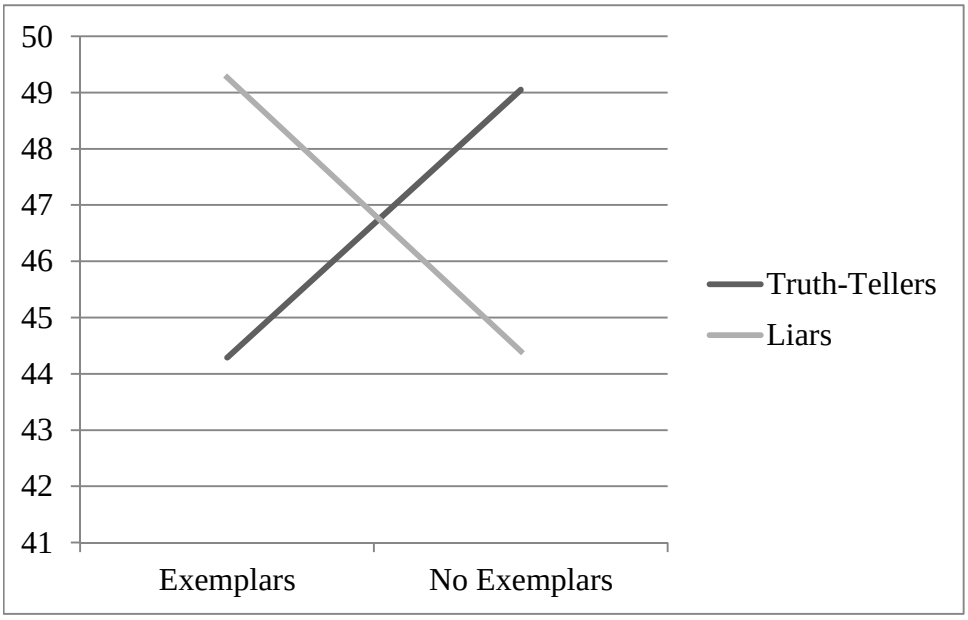

Figure 7. Means plot for nervousness (Study 1)

Table 7

Univariate ANOVA results for self-reported deliberateness (Study 1)

Veracity	Plan what to say	F (1, 125) = 1.61, p = .21
	Plan body language	F (1, 125) = 0.28, p = .60
	Plan to be believable	F (1, 125) = 3.18, p = .08, partial η^2 = .025
	Felt prepared	F (1, 125) = 0.91, p = .34
Exemplar	Plan what to say	F (1, 125) = 0.06, p = .80
	Plan body language	F (1, 125) = 0.52, p = .47
	Plan to be believable	F (1, 125) = 0.23, p = .63
	Felt prepared	F (1, 125) = 0.66, p = .42
Veracity x Exemplar	Plan what to say	F (1, 125) = 0.68, p = .41
	Plan body language	F (1, 125) = 4.27, p = .04, partial η^2 = .033
	Plan to be believable	F (1, 125) = 5.29, p = .02, partial η^2 = .041
	Felt prepared	F (1, 125) = 8.00, p = .005, partial η^2 = .060

Table 8

Means and standard deviations for self-reported deliberateness (Study 1)

Plan what to say		Mean	SD
Truthful	No exemplars	4.13	1.57
	Exemplars	4.30	1.64
Deceptive	No exemplars	4.00	1.62
	Exemplars	3.69	1.75
Plan body language		**Mean**	**SD**
Truthful	No exemplars	2.29	1.35
	Exemplars	3.07	1.84
Deceptive	No exemplars	2.72	1.58
	Exemplars	2.34	1.59
Plan to be believable		**Mean**	**SD**
Truthful	No exemplars	2.84	1.57
	Exemplars	3.70	1.68
Deceptive	No exemplars	4.11	1.79
	Exemplars	3.54	1.93
Feel prepared		**Mean**	**SD**
Truthful	No exemplars	3.58	1.41
	Exemplars	4.56	1.37
Deceptive	No exemplars	4.08	1.48
	Exemplars	3.54	1.72

Table 9

Comparisons of each condition on self-reported deliberateness (Study 1)

Truth-tellers, exemplars vs. truth-tellers, no exemplars	
Plan what to say	t (56) = .40, p = .69
Plan body language	t (56) = 1.87, p = .07, d = .48
Plan to be believable	t (56) = 2.02, p = .05, d = .53
Feel prepared	t (56) = 2.66, p = .01, d = .70
Truth-tellers, exemplars vs. liars, exemplars	
Plan what to say	t (60) = 1.40, p = .17
Plan body language	t (60) = 1.68, p = .10, d = .42
Plan to be believable	t (60) = 0.34, p = .73
Feel prepared	t (60) = 2.51, p = .015, d = 1.15
Truth-tellers, exemplars vs. liars, no exemplars	
Plan what to say	t (61) = 0.72, p = .48
Plan body language	t (61) = 0.82, p = .42
Plan to be believable	t (61) = 0.92, p = .36
Feel prepared	t (61) = 1.29, p = .20
Truth-tellers, no exemplars vs. liars, no exemplars	
Plan what to say	t (65) = .33, p = .74
Plan body language	t (65) = 1.19, p = .23
Plan to be believable	t (65) = 3.07, p = .003, d = .75
Feel prepared	t (65) = 1.42, p = .161
Truth-tellers, no exemplars vs. liars, exemplars	
Plan what to say	t (64) = 1.08, p = .28
Plan body language	t (64) = .14, p = .89
Plan to be believable	t (64) = 1.61, p = .11
Feel prepared	t (64) = .10, p = .92
Liars, exemplars vs. liars, no exemplars	
Plan what to say	t (69) = .79, p = .43
Plan body language	t (69) = 1.01, p = .32
Plan to be believable	t (69) = 1.29, p = .20
Feel prepared	t (69) = 1.42, p = .16

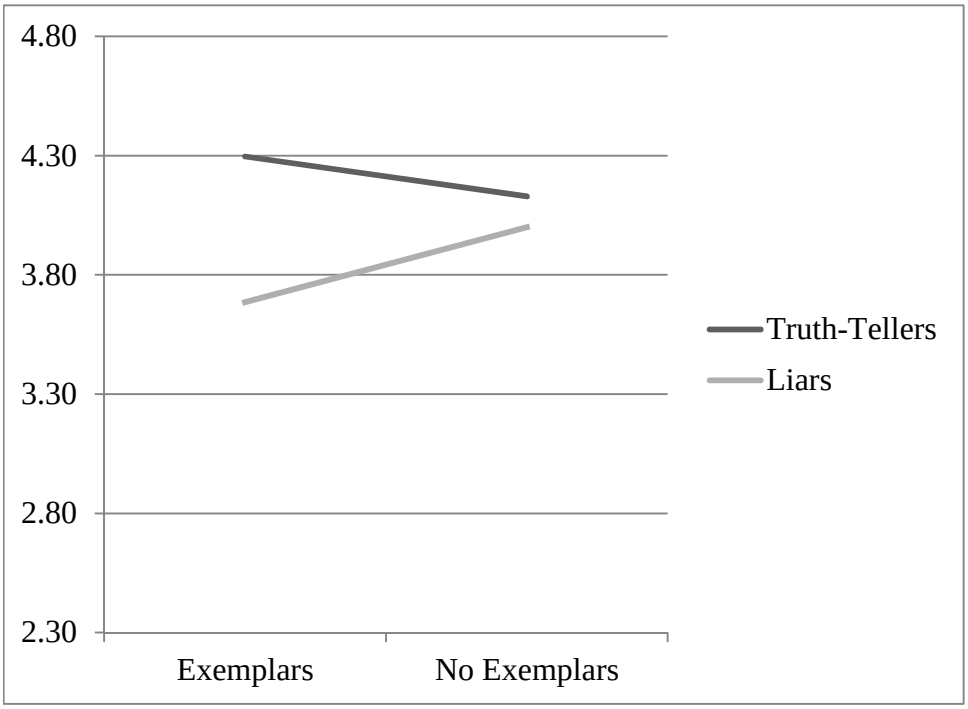

Figure 8. Means plot for the extent to which participants reported that they planned what to say (Study 1)

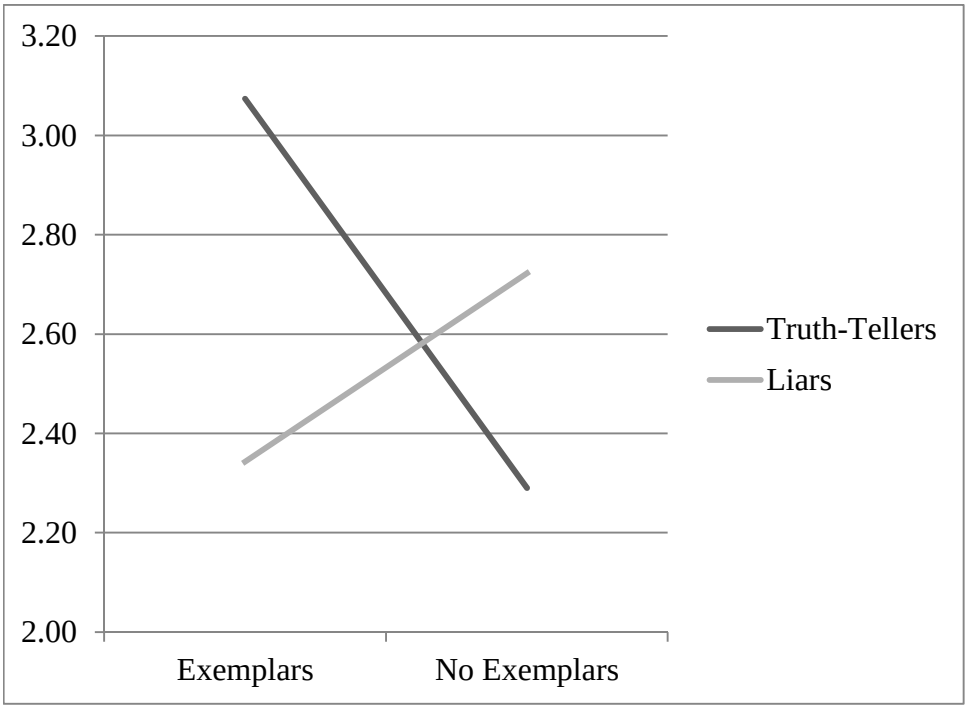

Figure 9. Means plot for the extent to which participants reported that they planned their body language (Study 1)

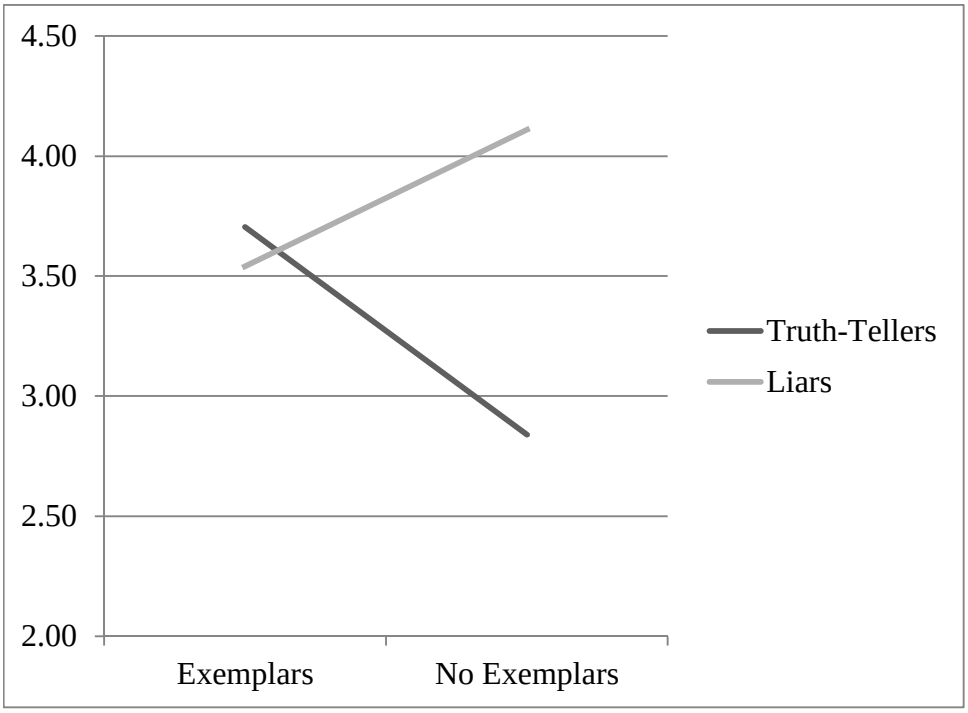

Figure 10. Means plot for the extent to which participants reported that they planned their responses to be believable (Study 1)

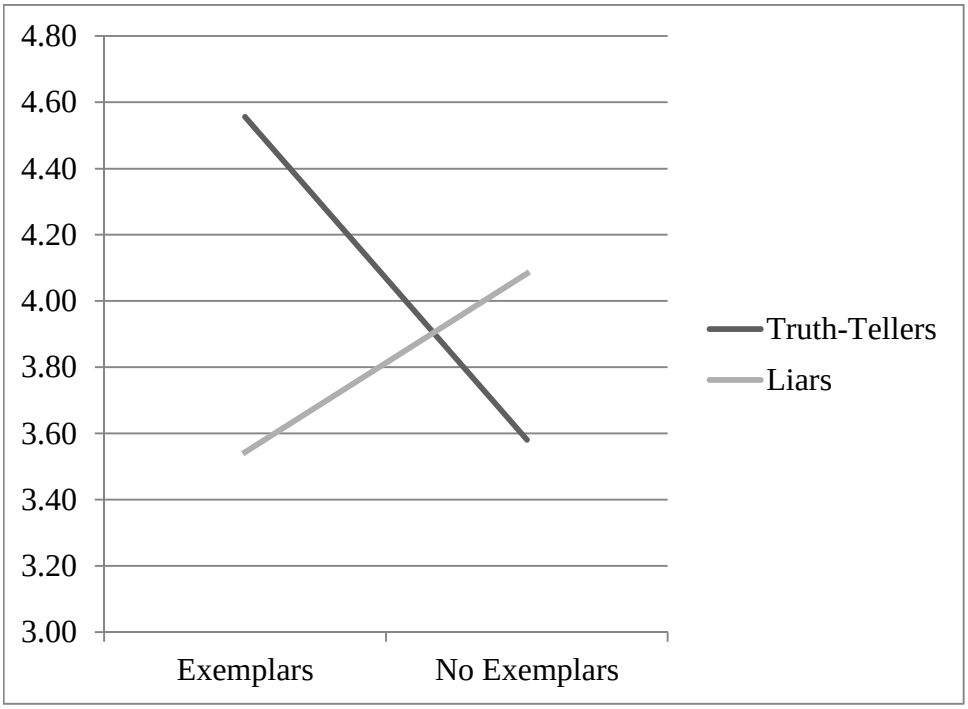

Figure 11. Means plot for the extent to which participants reported that they felt prepared (Study 1)

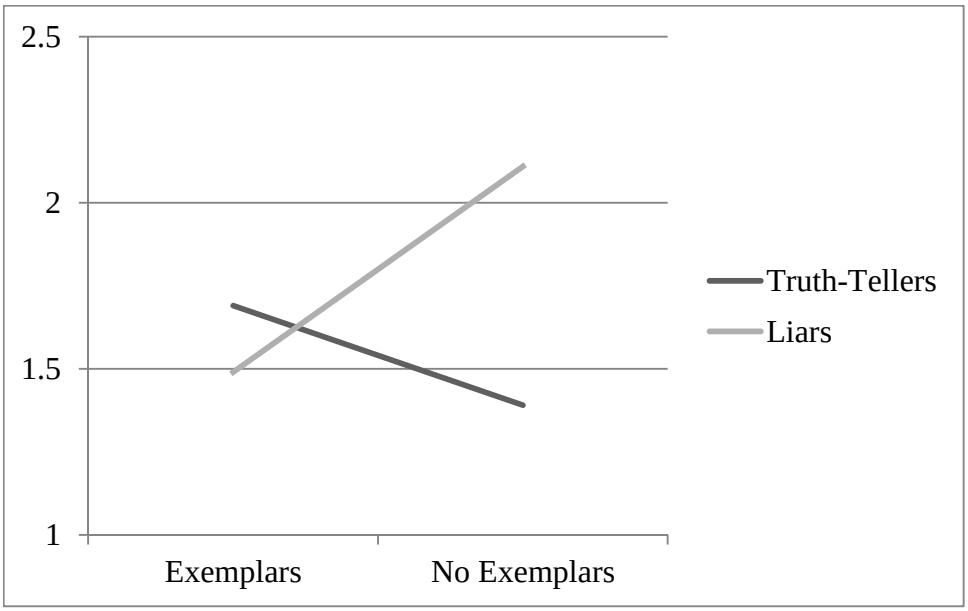

Figure 12. Means plot for the number of pieces of candy taken (Study 1)

Table 10

Ratings of honesty (Study 1)

		Mean	SD
Truthful	No exemplars	77.1862	7.16241
	Exemplars	80.5149	8.44646
Deceptive	No exemplars	77.2521	8.74102
	Exemplars	77.1924	8.35392

Table 11

Deception detection accuracy by condition (Study 1)

	Truthful, no exemplars	Truthful, exemplars	Deceptive, no exemplars	Deceptive, exemplars	Total
Accuracy	81.18%	82.83%	20.20%	13.19%	46.41%

Table 12

Means and standard deviations for talking time and details for each phase of the interview (critical task only), by condition (Study 2)

Talking Time		Mean	SD
Phase 1			
Truthful	Single-Task	80.95%	5.81%
	Three-Task	83.53%	5.06%
	Three-Task, Limited Questioning	82.82%	5.92%
Deceptive	Single-Task	79.45%	9.21%
	Three-Task	83.24%	3.65%
	Three-Task, Limited Questioning	85.88%	4.81%
Phase 2			
Truthful	Single-Task	72.01%	9.41%
	Three-Task	74.50%	7.37%
	Three-Task, Limited Questioning	72.04%	11.70%
Deceptive	Single-Task	66.49%	12.41%
	Three-Task	59.87%	12.89%
	Three-Task, Limited Questioning	64.10%	16.91%
Phase 3			
Truthful	Single-Task	77.08%	8.47%
	Three-Task	72.51%	11.86%
	Three-Task, Limited Questioning	72.19%	11.32%
Deceptive	Single-Task	72.10%	10.98%
	Three-Task	71.74%	11.55%
	Three-Task, Limited Questioning	75.00%	12.83%
Details			
Phase 1			
Truthful	Single-Task	14.50	5.76
	Three-Task	16.59	6.88
	Three-Task, Limited Questioning	16.16	6.25
Deceptive	Single-Task	11.93	5.27
	Three-Task	12.15	3.72
	Three-Task, Limited Questioning	15.80	6.78

Phase 2			
Truthful	Single-Task	16.50	7.67
	Three-Task	15.71	6.55
	Three-Task, Limited Questioning	16.53	7.52
Deceptive	Single-Task	9.64	3.91
	Three-Task	7.85	2.94
	Three-Task, Limited Questioning	11.47	5.68
Phase 3			
Truthful	Single-Task	14.25	7.77
	Three-Task	13.24	7.46
	Three-Task, Limited Questioning	14.47	9.13
Deceptive	Single-Task	11.14	6.15
	Three-Task	10.85	4.49
	Three-Task, Limited Questioning	14.33	9.02

Table 13

Means and standard deviations for average talking time and total details by veracity, three-task condition (Study 2)

Talking time		Mean	SD
Task 1	Truthful	73.81%	7.57%
	Deceptive	74.14%	6.67%
Task 2	Truthful	74.43%	8.42%
	Deceptive	75.71%	7.48%
Task 3	Truthful	76.85%	7.28%
	Deceptive	71.61%	7.49%
Details			
Task 1	Truthful	36.47	12.10
	Deceptive	35.46	12.88
Task 2	Truthful	41.29	19.48
	Deceptive	39.85	15.05
Task 3	Truthful	45.53	17.64
	Deceptive	30.85	9.06

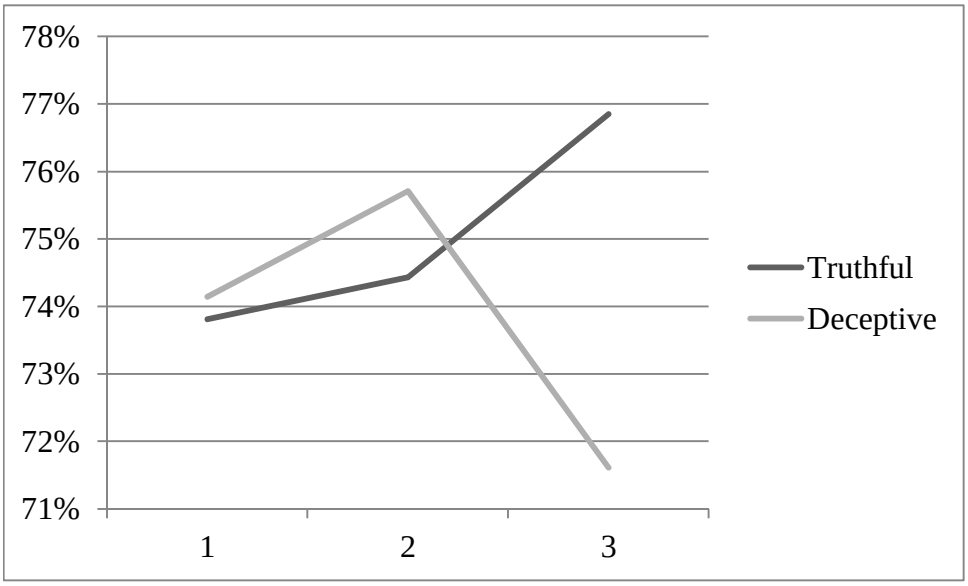

Figure 13. Average talking time for each task, by veracity condition in the three-task condition (Study 2)

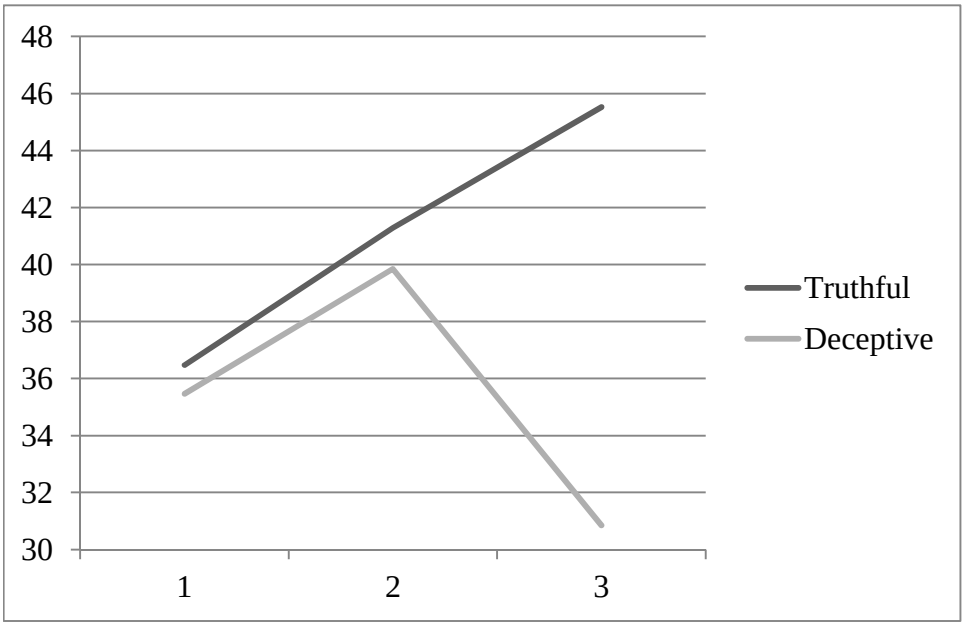

Figure 14. Means plot for total details for each task in the three-task condition, by veracity condition (Study 2)

Table 14

Means and standard deviations for cooperation, clarity/logic, and deliberateness, between-subjects (Study 2)

Cooperation		Mean	SD
Truthful	Single-task	77.65	7.35
	Three-task	80.67	6.06
	Three-task, limited questioning	80.64	6.32
Deceptive	Single-task	73.27	12.31
	Three-task	75.54	10.80
	Three-task, limited questioning	74.55	11.47
Clarity/Logic		Mean	SD
Truthful	Single-task	70.89	7.71
	Three-task	74.14	9.17
	Three-task, limited questioning	71.77	10.35
Deceptive	Single-task	64.63	17.73
	Three-task	67.74	9.63
	Three-task, limited questioning	61.64	11.82
Deliberateness		Mean	SD
Truthful	Single-task	63.96	9.16
	Three-task	67.14	11.84
	Three-task, limited questioning	64.79	7.97
Deceptive	Single-task	61.78	8.71
	Three-task	61.85	10.68
	Three-task, limited questioning	64.37	14.28

Table 15

Means and standard deviations for cooperation, clarity/logic, and deliberateness for each task in the three-task condition, by veracity (Study 2)

Cooperation		Mean	SD
Truthful	Task 1	79.31	7.03
	Task 2	79.31	6.28
	Task 3	80.67	6.06
Deceptive	Task 1	75.70	10.32
	Task 2	77.01	7.94
	Task 3	75.54	10.80
Clarity/Logic		Mean	SD
Truthful	Task 1	71.68	10.29
	Task 2	71.88	7.95
	Task 3	74.14	9.17
Deceptive	Task 1	67.92	8.80
	Task 2	70.59	7.45
	Task 3	67.74	9.63
Deliberateness		Mean	SD
Truthful	Task 1	64.91	11.89
	Task 2	66.93	10.88
	Task 3	67.14	11.84
Deceptive	Task 1	64.99	9.82
	Task 2	63.05	8.35
	Task 3	61.85	10.68

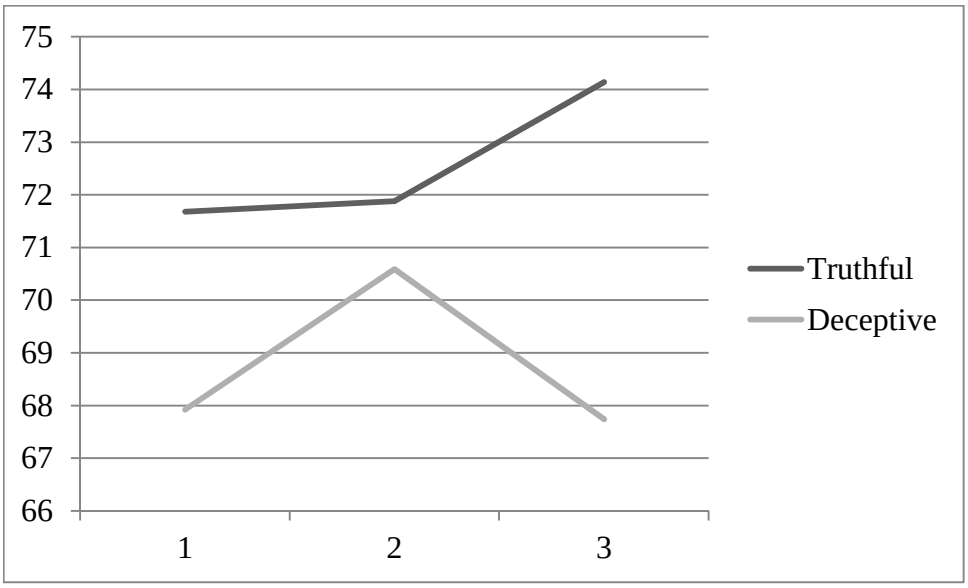

Figure 15. Means plot for ratings of clarity/logic for each task in the three-task condition, by veracity (Study 2)

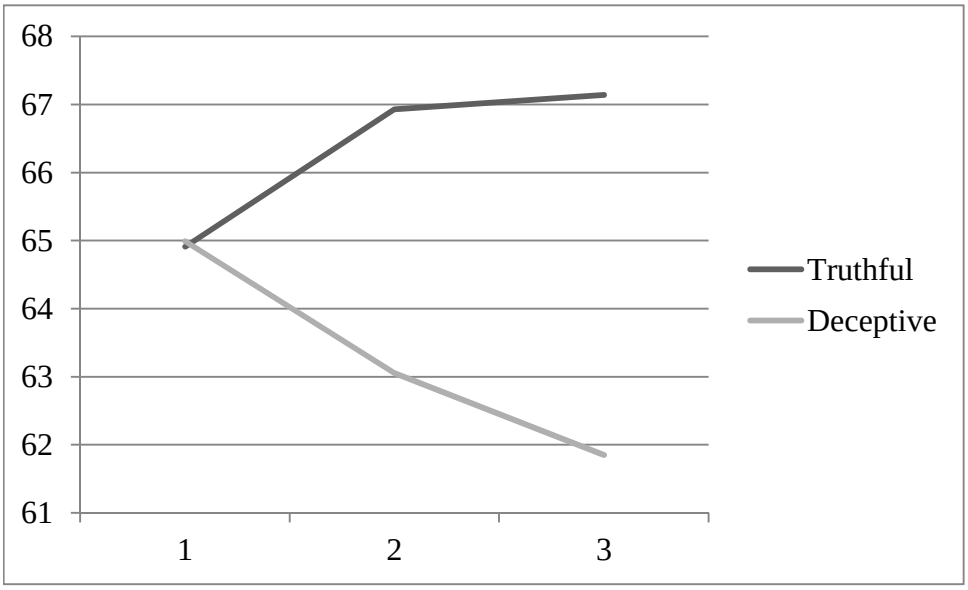

Figure 16. Means plot for ratings of deliberateness for each task in the three-task condition, by veracity (Study 2)

Table 16

Means and standard deviations for self-reported deliberateness (Study 2)

Plan what to say		Mean	SD
Truthful	Single-task	4.19	1.33
	Three-task	2.41	1.50
	Three-task, limited questioning	3.26	1.24
Deceptive	Single-task	3.07	1.59
	Three-task	3.00	1.41
	Three-task, limited questioning	3.20	1.32
Plan body language		Mean	SD
Truthful	Single-task	3.00	1.41
	Three-task	2.00	1.27
	Three-task, limited questioning	2.42	1.39
Deceptive	Single-task	2.64	1.45
	Three-task	2.08	1.19
	Three-task, limited questioning	2.20	1.42
Plan to be believable		Mean	SD
Truthful	Single-task	3.71	1.49
	Three-task	2.35	1.54
	Three-task, limited questioning	2.79	1.55
Deceptive	Single-task	3.50	1.74
	Three-task	2.92	1.26
	Three-task, limited questioning	2.93	1.79
Feel Prepared		Mean	SD
Truthful	Single-task	3.33	1.46
	Three-task	2.00	1.12
	Three-task, limited questioning	3.00	1.29
Deceptive	Single-task	2.64	1.74
	Three-task	2.46	1.27
	Three-task, limited questioning	2.07	1.33

Table 17

Means and standard deviations for number of pieces of candy taken by condition (Study 2)

		Mean	SD
Truthful	Single-task	1.57	1.12
	Three-task	2.06	1.56
	Three-task, limited questioning	2.26	1.76
Deceptive	Single-task	1.29	0.91
	Three-task	1.46	0.88
	Three-task, limited questioning	2.13	1.46

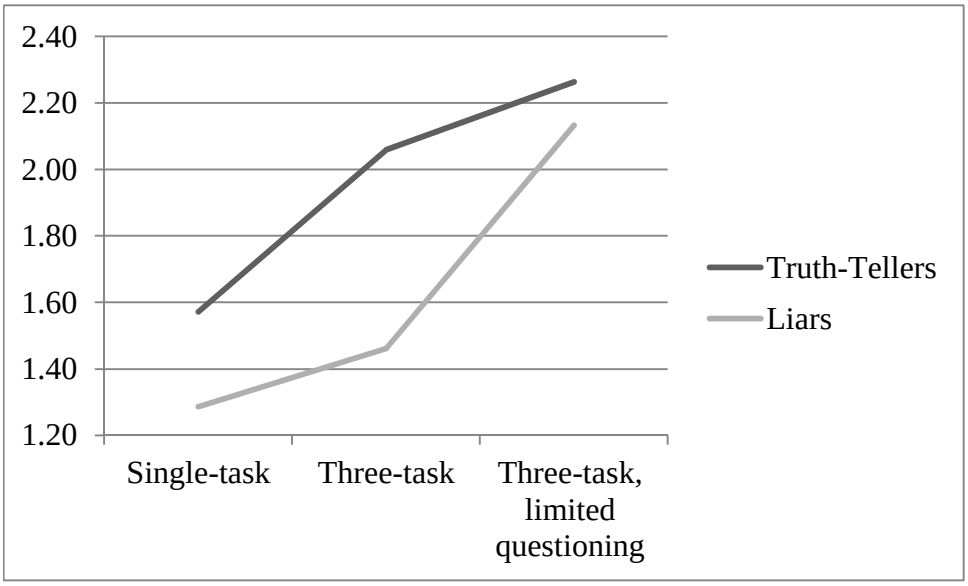

Figure 17. Means plot for number of pieces of candy taken, by condition (Study 2)

Table 18

Means and standard deviations for ratings of honesty (Study 2)

		Mean	SD
Truthful	Single-task	78.50	6.45
	Three-task	79.74	7.73
	Three-task, limited questioning	77.89	6.10
Deceptive	Single-task	74.44	10.49
	Three-task	77.14	8.74
	Three-task, limited questioning	76.64	11.75

Table 19

Deception detection accuracy by condition (Study 2)

	Truthful, Single-Task	Truthful, Three-Task	Truthful, Three-Task, Limited Questioning	Deceptive, Single-Task	Deceptive, Three-Task	Deceptive, Three-Task, Limited Questioning	Total
Accuracy	80.92%	86.99%	82.08%	12.50%	14.29%	16.96%	53.97%

Appendix A

Self-Reported Deliberateness Items

(1) How much did you plan what you were going to say ahead of time?

(2) How much did you plan your body language ahead of time?

(3) To what extent did you plan your responses in order to appear believable?

(4) How much did you feel prepared for the interview?

References

Aarts, H. & Dijksterhuis, A. (2003). The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, *84*, 18-28.

Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, *36*, 715-729.

Albrechtsen, J. S., Meissner, C. A., & Susa, K. J. (2009). Can intuition improve deception detection performance?. *Journal of Experimental Social Psychology*, *45*, 1052-1055.

Alicke, M. D., & Sedikides, C. (Eds.). (2011). *Handbook of self-enhancement and self-protection*. New York: Guilford Press.

Ambady, N. & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*, 256-274.

Arkin, R. M. (1981). Self-presentation styles. In J. T. Tedeschi (Ed.), *Impression management theory and social psychological research* (pp.311-333). San Diego: Academic Press.

Ashby, W. R. (1956). *An introduction to cybernetics.* New York: John Wiley & Sons.

Ashby, W. R. (1952). *Design for a brain.* New York: John Wiley & Sons.

Åström, K. J. & Murray R. M. (2012). *Feedback systems: An introduction for scientists and engineers* (Version v2.11b). Princeton: Princeton University Press.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829-839.

Bargh, J. A., & Williams, E. L. (2006). The automaticity of social life. *Current directions in psychological science*, *15*, 1-4.

Bauer, I. M., & Baumeister, R. F. (2011). Self-regulatory strength. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp.64-82). New York: Guilford Press.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3-26.

Baumeister, R. F., Hamilton, J. C., & Tice, D. M. (1985). Public versus private expectancy of success: Confidence booster or performance pressure? *Journal of Personality and Social Psychology*, *48*, 1447-1457.

Baumeister, R. F., Hutton, D. G., & Tice, D. M. (1989). Cognitive processes during deliberate self-presentation: How self-presenters alter and misinterpret the behavior of their interaction partners. *Journal of Experimental Social Psychology*, *25*, 59-78.

Baumeister, R. F., & Jones, E. E. (1978). When self-presentation is constrained by the target's knowledge: Consistency and compensation. *Journal of Personality and Social Psychology*, *36*, 608-618.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497-595.

Baumeister, R. F., & Scher, S. J. (1988). Self-defeating behavior patterns among normal individuals: review and analysis of common self-destructive tendencies. *Psychological Bulletin*, *104*, 3-22.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*, 351-355.

Bell, K. L., & DePaulo, B. M. (1996). Liking and lying. *Basic and Applied Social Psychology*, *18*, 243-266.

Berglas, S., & Jones, E. E. (1978). Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement. *Personality and Social Psychology Bulletin, 4*, 200-206.
Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Pantheon Books.
Bond, C. F. Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: accuracy and bias. *Psychological Bulletin, 134*, 477-492.
Bond, C. F. Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214-234.
Bond, C. F., Jr. and Fahey, W. E. (1987). False suspicion and the misperception of deceit. *British Journal of Social Psychology, 26*, 41-46.
Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology, 86*, 599-614.
Bozeman, D. P. & Kacmar, K. M. (1997). A cybernetic model of impression management processes in organizations. *Organizational Behavior and Human Decision Processes, 69*, 9-30.
Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological review, 50*, 255-272.
Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*, 203-242.
Buroon, J. K., Bulkr, D. B., Ebesu, A. S., White, C. H., & Rockwell, P. A. (1996). Testing interpersonal deception theory: Effects of suspicion on communication behaviors and perceptions. *Communication Theory, 6*, 243-267.
Button, S. B. (2004). Identity management strategies utilized by lesbian and gay employees A quantitative investigation. *Group & Organization Management, 29*, 470-494.
Carver, C. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition & Emotion, 17*, 241-261.
Carver, C. S. (1979). A cybernetic model of self-attention processes. *Journal of Personality and Social Psychology, 37*, 1251-1281.
Carver, C. S., Antoni, M. H., & Scheier, M. F. (1985). Self-consciousness and self-assessment. *Journal of Personality and Social Psychology, 48*, 117-124.
Carver, C. S., Blaney, P. H., & Scheier, M. F. (1979). Reassertion and giving up: The interactive role of self-directed attention and outcome expectancy. *Journal of Personality and Social Psychology, 37*, 1859-1870.
Carver, C. S., Peterson, L. M., Follansbee, D. J., & Scheier, M. F. (1983). Effects of self-directed attention on performance and persistence among persons high and low in test anxiety. *Cognitive Therapy and Research, 7*, 333-353.
Carver, C. S., & Scheier, M. F. (2011). Self-regulation of action and affect. In K. D. Vohs, R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications,* 2nd edition (pp. 3-21). New York: Guilford Press.
Carver, C. S. & Scheier, M. F. (1998). *On the self-regulation of behavior*. Cambridge: Cambridge University Press.
Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97*, 19-35.
Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.

Cialdini, R. B., Finch, J. F., & De Nicholas, M. E. (1990). Strategic self-presentation: The indirect route. In M. Cody & M. L. McLaughlin (Eds.), *The psychology of tactical communication* (pp.194-206). Bristol, PA: Multilingual Matters.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893-910.

Cheng, C. M., & Chartrand, T. L. (2003). Self-monitoring without awareness: using mimicry as a nonconscious affiliation strategy. *Journal of Personality and Social Psychology, 85*, 1170-1179.

Chrobot-Mason, D., Button, S. B., & DiClementi, J. D. (2001). Sexual identity management strategies: An exploration of antecedents and consequences. *Sex Roles, 45*, 321-336.

Clark, R. N. (1996). *Control system dynamics.* Cambridge: Cambridge University Press.

Cody, M. & McLaughlin, M. L. (1990). *The psychology of tactical communication*. Bristol, PA: Multilingual Matters.

David, G. K., Condon, C. H., Bywater, C. L., Ortiz-Barrientos, D., & Wilson, R. S. (2011). Receivers limit the prevalence of deception in humans: evidence from diving behaviour in soccer players. *PloS one*, *6*(10), e26017.

DePaulo, B. M., Ansfield, M. E., & Bell, K. L. (1996). Theories about deception and paradigms for studying it: A critical appraisal of Buller and Burgoon's lnterpersonal Deception Theory and research. *Communication Theory, 6,* 297-310.

DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: lies are told to those who care. *Journal of Personality and Social Psychology, 71*, 703-716.

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

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12*, 177-202.

DePaulo, B. M., LeMay, C. S., & Epstein, J. A. (1991). Effects of importance of success and expectations for success on effectiveness at deceiving. *Personality and Social Psychology Bulletin, 17*, 14-24.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-118.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Telling ingratiating lies: effects of target sex and target attractiveness on verbal and nonverbal deceptive success. *Journal of Personality and Social Psychology, 48*, 1191-1203.

Ekman, P. (19852009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (revised edition). New York: W. W. Norton & Company.

Epley, N., Caruso, E. M., & Bazerman, M. H. (2006). When perspective taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology, 91*, 872-889.

Epley, N. & Waytz, A. (2009). Mind perception. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology*. New York: Wiley.

Fiske, S. T. & Taylor. S. E. (2013). *Social cognition: From brains to culture,* 2nd edition. Los Angeles: Sage.

Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective - taking and empathy in strategic interactions. *Psychological Science, 19*, 378-384.

Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, *11*, 303-327.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., ... & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*, 325-336.

Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology*, *37*, 60-74.

Gigerenzer, G. & Todd, P. M. (1999). *Simple heuristics that make us smart: Evolution and cognition*. New York: Oxford University Press.

Gollwitzer, P. M., Parks-Stamm, E. J., & Oettingen, G. (2009). Living on the edge: Shifting between nonconscious and conscious goal pursuit. In E. E. Morsella, J. A. Bargh, & P. M. Gollwitzer. Oxford: Oxford University Press.

Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in cognitive science*, *5*, 173-184.

Gordon, R. A. (1996). Impact of ingratiation on judgments and evaluations: A metaanalytic investigation. *Journal of Personality and Social Psychology*, *71*, 54-70.

Granhag, P.A., & Hartwig, M. (2015). The Strategic Use of Evidence (SUE) technique: A conceptual overview. In P.A. Granhag, A. Vrij, & B. Vershuere (Eds.), *Deception detection: Current challenges and new directions*. Chichester: Wiley.

Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law*, *14*, 189-200.

Granhag, P. A., Strömwall, L. A., Willén, R. M., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology*, *18*, 341-355.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101-124.

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychological Bulletin*, *136*, 495-525.

Hartwig, M. & Bond, C. F. Jr. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, *28*, 661-676.

Hartwig, M. & Bond, C. F. Jr. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, *137*, 643-659.

Hartwig, M., Granhag, P. A., & Luke, T. J. (2014). Strategic use of evidence during investigative interviews: The state of the science. In D.C. Raskin, C.R. Honts, & J.C. Kircher (Eds.), *Credibility assessment: Scientific research and applications* (pp. 1-36). Waltham, MA: Academic Press.

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during interrogations. *Psychology, Crime & Law*, *13*, 213–227.

Hartwig, M., Granhag, P. A., Strömwall, L., & Doering, N. (2010). Impression and information management: On the strategic self-regulation of innocent and guilty suspects. *The Open Criminology Journal*, *3*, 10-16.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior, 29*, 469–484. doi:10.1007/s10979-005-5521-x.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 243-259.
Higgins, E. T. (2012). *Beyond pleasure and pain: How motivation works*. New York: Oxford University Press.
Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280-1300. doi:10.1037/0003-066X.52.12.1280
Higgins, C. A., & Judge, T. A. (2004). The effect of applicant influence tactics on recruiter perceptions of fit and hiring recommendations: a field study. *Journal of Applied Psychology, 89*, 622-632.
Hofmann, W., Friese, M., Schmeichel, B. J., & Baddeley, A. D. (2011). Working memory and self-regulation. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp.204-225). New York: Guilford Press.
Hsee, C. K., & Abelson, R. P. (1991). Velocity relation: Satisfaction as a function of the first derivative of outcome over time. *Journal of Personality and Social Psychology, 60*, 341.
Hsee, C. K., Abelson, R. P., & Salovey, P. (1991). The relative weighting of position and velocity in satisfaction. *Psychological Science, 2*, 263-266.
Hsee, C. K., Salovey, P., & Abelson, R. P. (1994). The quasi-acceleration relation: Satisfaction as a function of the change of velocity of outcome over time. *Journal of Experimental Social Psychology, 30*, 96-111.
Jonas, K. J., & Sassenberg, K. (2006). Knowing how to react: Automatic response priming from social categories. *Journal of Personality and Social Psychology, 90*, 709-721.
Jones, E. E. (1990). *Interpersonal perception*. New York: Appleton-Century-Crofts.
Jones, E. E. (1964). *Ingratiation: A social psychological analysis*. New York: Holt.
Jones, E. E., & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. In J. Suis (Ed.), *Psychological Perspectives on the Self* (pp.231-262.). Hillsdale, NJ: Lawrence Erlbaum.
Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist, 60*, 215-228.
Kappes, H. B., Oettingen, G., Mayer, D., & Maglio, S. (2011). Sad mood promotes self-initiated mental contrasting of future and reality. *Emotion, 11*, 1206-1222.
Kappes, A. Singmann, H., & Oettingen, G. (2012). Mental contrasting instigates goal pursuit by linking obstacles of reality with instrumental behavior. *Journal of Experimental Social Psychology, 48*, 811-818.
Kappes, A., Wendt, M., Reinelt, T., & Oettingen, G. (2013). Mental contrasting changes the meaning of reality. *Journal of Experimental Social Psychology, 49*, 797-810.
Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology, 70*, 1037-1051.
Kruglanski, A. W., Bélanger, J. J., Chen, X., Köpetz, C., Pierro, A., & Mannetti, L. (2012). The energetics of motivated cognition: a force-field analysis. *Psychological Review, 119*, 1-20.
Kruglanski, A. W., Köpetz, C., Bélanger, J. J., Chun, W. Y., Orehek, E., & Fishbach, A. (2013). Features of Multifinality. *Personality and Social Psychology Review, 17*, 22-39.
Kruglanski, A. W., Pierro, A., & Sheveland, A. (2011). How many roads lead to Rome? Equifinality set-size and commitment to goals and means. *European Journal of Social Psychology, 41*, 344-352.

Kwang, T., & Swann, W. B. (2010). Do people embrace praise even when they feel unworthy? A review of critical tests of self-enhancement versus self-verification. *Personality and Social Psychology Review*, *14*, 263-280.

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, *14*, 334-339.

Lalwani, A. K., Shrum, L. J., & Chiu, C. Y. (2009). Motivated response styles: the role of cultural values, regulatory focus, and self-consciousness in socially desirable responding. *Journal of Personality and Social Psychology*, *96*, 870-882.

Lane, S., Martin, S., Elliott, E., & Mennie, K. (March, 2014). *The impact of cognitive load and WMC on information elicitation and deception detection.* Paper presented at the annual conference of the American Psychology-Law Society, New Orleans, LA.

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. P. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*, *20*, 129-146.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, *58*, 317-344.

Leary, M. R., Nezlek, J. B., Downs, D., Radford-Davenport, J., Martin, J., & McMullen, A. (1994). Self-presentation in everyday interactions: effects of target familiarity and gender composition. *Journal of Personality and Social Psychology*, *67*, 664-673.

Levine, T. R. (2014). Truth-Default Theory (TDT) A Theory of Human Deception and Deception Detection. *Journal of Language and Social Psychology*. Published online May 24, 2014.

Levine, T. R., Feeley, T. H., McCornack, S. A., Hughes, M., & Harms, C. M. (2005). Testing the effects of nonverbal behavior training on accuracy in deception detection with the inclusion of a bogus training control group. *Western Journal of Communication*, *69*, 203-217.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communications Monographs*, *66*, 125-144.

Luke, T. J., Dawson, E., Hartwig, M., & Granhag, P.A. (2014). How awareness of evidence induces forthcoming counter-interrogation strategies. *Applied Cognitive Psychology*. Available online March 14, 2014.

Luke, T. J., Hartwig, M., Brimbal, L., Chan, G., Jordan, S., Joseph, E., Osborne, J., & Granhag, P. A. (2013). Interviewing to elicit cues to deception: Improving strategic use of evidence with general-to-specific framing of evidence. *Journal of Police and Criminal Psychology*, *28*, 54-62.

Luke, T. J., Hartwig, M., Shamash, B., & Granhag, P.A. (October, 2013). *Suspect counter-interrogation strategies: The effect of prior knowledge of the interviewing techniques.* Paper presented at the High-Value Detainee Interrogation Group Research Symposium, Washington DC.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299-337.

Masip, J., Barba, A., & Herrero, C. (2012). Behaviour analysis interview and common sense: A study with novice and experienced officers. *Psychiatry, Psychology and Law*, *19*, 21-34.

Masip, J., Garrido, E., & Herrero, C. (2009). Heuristic versus systematic processing of information in detecting deception: Questioning the truth bias. *Psychological Reports*, *105*, 11-36.

Masip, J., Herrero, C., Garrido, E., & Barba, A. (2011). Is the behaviour analysis interview just common sense? *Applied Cognitive Psychology*, *25*, 593-604.
Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Holt, Rinehart, & Winston.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
Oettingen, G. (2012). Future thought and behaviour change. *European Review of Social Psychology*, *23*, 1-63.
Oettingen, G. (2000). Expectancy effects on behavior depend on self-regulatory thought. *Social Cognition*, *18*, 101-129.
Oettingen, G., Hönig, G., & Gollwitzer, P. M. (2000). Effective self-regulation of goal attainment. *International Journal of Educational Research*, *33*, 705-732.
Oettingen, G., Mayer, D., Sevincer, A. T., Stephens, E. J., Pak, H. J., & Hagenah, M. (2009). Mental contrasting and goal commitment: The mediating role of energization. *Personality and Social Psychology Bulletin*, *35*, 608-622.
Oettingen, G., Pak, H. J., & Schnetter, K. (2001). Self-regulation of goal-setting: Turning free fantasies about the future into binding goals. *Journal of Personality and Social Psychology*, *80*, 736-753.
Pontari, B. A., Schlenker, B. R., & Christopher, A. N. (2002). Excuses and character: Identifying the problematic aspects of excuses. *Journal of Social and Clinical Psychology*, *21*, 497-516.
Powers, W. T. (2007). Emotions and PCT: A brief overview. In D. Forssell (Ed.), *Perceptual control theory* (2010) (pp.10-12). Hayward, CA: Living Control Systems Publishing.
Powers, W. T. (1998). *Making sense of behavior: The meaning of control*. Montclair, NJ: Benchmark Publications.
Powers, W. T. (1973a). *Behavior: The control of perception*. Chicago: Aldine.
Powers, W. T. (1973b). Feedback: Beyond behaviorism. *Science*, *179*, 351-356.
Powers, W. T., Clark, R. K., & McFarland, R. L. (1960a). A general feedback theory of human behavior: Part I. *Perceptual and Motor Skills*, *11*, 71-88.
Powers, W. T., Clark, R. K., & McFarland, R. L. (1960b). A general feedback theory of human behavior: Part II. *Perceptual and Motor Skills*, *11*, 309-323.
Russell, B. (1905). On Denoting. *Mind*, *14*, 479-493.
Sedikides, C. (2012). Self-protection. In M. R. Leary, J. Tangney (Eds.) , *Handbook of self and identity* (2nd ed.) (pp. 327-353). New York: Guilford Press.
Scheier, M. F., & Carver, C. S. (1983). Self-directed attention and the comparison of self with standards. *Journal of Experimental Social Psychology*, *19*, 205-222.
Schlenker, B. R. (2012). Self-presentation. In M. R. Leary, J. Tangney (Eds.) , *Handbook of self and identity* (2nd ed.) (pp. 542-570). New York: Guilford Press.
Schlenker, B. R. (1985). Identity and self-identification. In B. R. Schlenker, *The self and social life*, 65-99. New York: McGraw-Hill.
Schlenker, B. R. (1975). Self-presentation: managing the impression of consistency when reality interferes with self-enhancement. *Journal of personality and social psychology*, *32*, 1030-1037.
Schlenker, B. R., & Leary, M. R. (1985). Social anxiety and communication about the self. *Journal of Language and Social Psychology*, *4*, 171-192.

Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization model. *Psychological Bulletin*, *92*, 641-669.
Schlenker, B. R., Lifka, A., & Wowra, S. A. (2004). Helping New Acquaintances Make the Right Impression: Balancing Image Concerns of Others and Self. *Self & Identity*, *3*, 191-206.
Schlenker, B. R., Pontari, B. A., & Christopher, A. N. (2001). Excuses and character: Personal and social implications of excuses. *Personality and Social Psychology Review*, *5*, 15-32.
Schlenker, B. R., & Weigold, M. F. (1992). Interpersonal processes involving impression regulation and management. *Annual Review of Psychology*, *43*, 133-168.
Schwarz, N., & Clore, G. L. (2007). Feelings and phenomenal experiences. A Kruglanski & E. T. Higgins (Eds), *Social psychology: Handbook of basic principles* (2nd ed.) (pp.385-407). New York: Guilford.
Schütz, A. (1998). Assertive, offensive, protective, and defensive styles of self-presentation: A taxonomy. *The Journal of Psychology*, *132*, 611-628.
Shah, J. & Higgins, E. T. (1997). Expectancy x value effects: Regulatory focus as determinant of magnitude and direction. *Journal of Personality and Social Psychology*, *73*, 447-458.
Shah, J., Higgins, T., & Friedman, R. S. (1998). Performance incentives and means: how regulatory focus influences goal attainment. *Journal of Personality and Social Psychology, 74*, 285-293.
Shah, J. Y., & Kruglanski, A. W. (2003). When opportunity knocks: bottom-up priming of goals by means and its effects on self-regulation. *Journal of Personality and Social Psychology*, *84*, 1109-1122.
Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology*, *38*, 368-383.
Showers, C. & Zeigler-Hill, V. (2012). Organization of self-knowledge: Features, functions, and flexibility. In M. R. Leary, J. Tangney (Eds.), *Handbook of self and identity* (2nd ed.) (pp. 105-123). New York: Guilford Press.
Snyder, C. R. & Higgins, R. L. (1988). Excuses: their effective role in the negotiation of reality. *Psychological Bulletin*, *104*, 23-35.
Snyder, M. & Stukas Jr, A. A. (1999). Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual Review of Psychology*, *50*, 273-303.
Strömwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behaviour and strategies during a police interrogation. *Psychology, Crime and Law*, *12*, 207-219.
Swann, W. B. & Buhrmester, M. D. (2012). Self-verification: The search for coherence. In M. R. Leary, J. Tangney (Eds.), *Handbook of self and identity* (2nd ed.) (pp. 405-424). New York: Guilford Press.
Tedeschi, J. T. & Lindskold, S. (1976). *Social psychology: Interdependence, interaction, and influence.* New York: Wiley.
Tekin, S., Granhag, P. A., Strömwall, L., Mac Giolla, E., Vrij, A., & Hartwig, M. (2015). Interviewing Strategically to Elicit Admissions From Guilty Suspects. *Law And Human Behavior 39*, 244-252.
ten Brinke, L., Stimson, D., & Carney, D. R. (2014). Some evidence for unconscious lie detection. *Psychological Science*.
Thibodeau, P. H. & Boroditsky, L. (2013). Natural language metaphors covertly influence reasoning. *PloS one*, *8*, e52961.

Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS One*, *6*, e16782.
Tice, D. M., Butler, J. L., Muraven, M. B., & Stillwell, A. M. (1995). When modesty prevails: Differential favorability of self-presentation to friends and strangers. *Journal of Personality and Social Psychology*, *69*, 1120-1138.
Tolman, E. C. & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review*, *42*, 43-77.
Trötschel, R., Hüffmeier, J., Loschelder, D. D., Schwartz, K., & Gollwitzer, P. M. (2011). Perspective taking as a means to overcome motivational barriers in negotiations: when putting oneself into the opponent's shoes helps to walk toward agreements. *Journal of Personality and Social Psychology*, *101*, 771-790.
Vallacher, R. R., & Wegner, D. M. (2012). Action identification theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins, *Handbook of theories of social psychology* (pp.327-348). Thousand Oaks, CA: Sage.
Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, *57*, 660-671.
Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*, 3-15.
Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology*, *88*, 632-657.
Vrij, A. (2006). Challenging interviewees during interviews: The potential effects on lie detection. *Psychology, Crime & Law*, *12*, 193-206.
Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, *24*, 239-263.
Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, *10*, 141-142.
Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, *1*, 110-117.
Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: toward a cognitive lie detection approach. *Current Directions in Psychological Science*, *20*, 28-32.
Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior*, *33*, 159-166.
Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: the benefit of recalling an event in reverse order. *Law and Human Behavior*, *32*, 253-265.
Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: Bradford.
Wegner, D. M., Vallacher, R. R., Kiersted, G. W., & Dizadji, D. (1986). Action identification in the emergence of social behavior. *Social Cognition*, *4*, 18-38.
Wegner, D. M., Vallacher, R. R., Macomber, G., Wood, R., & Arps, K. (1984). The emergence of action. *Journal of Personality and Social Psychology*, *46*, 269-279.
Wiener, N. (1950/1954). *The human use of human beings: Cybernetics and society*, 2nd edition. Boston: Houghton Mifflin.

Wiener, N. (1948/1961). *Cybernetics: Control and communication and the animal and the machine*, 2nd edition. Cambridge, MA: MIT University Press.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 1–60). New York: Academic.