

TESTING THE INTRINSIC BENEFIT MODEL OF THE SIGNALING THEORY

by

Sosuke Okada

Copyright © Sosuke Okada 2020

A Dissertation Submitted to the Faculty of the

SCHOOL OF SOCIOLOGY

In Partial Fulfilment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2020

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

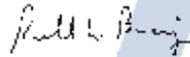
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: **Sosuke Okada**
titled:

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.



Joseph J Galaskiewicz

Date: Jan 12, 2021



Ronald L. Breiger

Date: Jan 12, 2021



Tamar Kugler

Date: Jan 12, 2021

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.



Joseph J Galaskiewicz
Dissertation Committee Chair
School of Sociology

Date: Jan 12, 2021



Acknowledgement

This work was supported by National Science Foundation Doctoral Dissertation Improvement Grant to the Sociology Program, ID#1404371

I like to thank Joseph Galaskiewicz, Ronald Breiger, and Tamar Kugler for their guidance and mentorship throughout the research. I like to thank Charles Noussair for his assistance in conducting the experiment at the Economic Science Laboratory. Finally, I like to thank my parents and an unnamed friend for all the emotional supports they had provide me through the difficult years of my life.

Table of Contents

Abstract p5
Introduction p6
Chapter 1: Theoretical Background, Literature Review p14
Chapter 2: Intrinsic Benefit and the Signaling Theory p31
Chapter 3: Intrinsic Benefit Model: Hypotheses P42
Chapter 4: Experiments #1 and #2 P59
Chapter 5: Experiment#3	----- p72
Chapter 6: Discussion and Conclusion P102
Appendix P125
References P136

Abstract

This study proposes the intrinsic benefit model of the signaling theory for sociology. The signaling theory is a subtheory of the game theory. It was developed independently within Evolutionary Biology and Economics, and it is concerned with the communications under the situations with asymmetrical information. Although the signaling theory have been widely adapted across social science, its influence within Sociology has been limited so far. This study proposes the argument that the signaling theory can achieve the increased relevance within Sociology by focusing on the role of (perceived) intrinsic benefit obtained from the signal production. The focus on the intrinsic benefit would allow the signaling theory to be applied on the broader range of phenomena which are of sociological interests, while at the same time analytically integrating additional social and symbolic contexts of the signals. Based on this argument, the propositions were developed about the role of the signal visibility and the intentionality of the signal. The three experiments were conducted to test the propositions. The two vignette experiments were conducted to test the effect of signal visibility on the signaling of environmental commitment through the purchases of electronic vehicles. A laboratory experiments was conducted to test the effect of the intentionality of the signal on the signaling of trustworthiness through donations. The first experiment gave the strong support to the propositions, whereas the second and the third experiment produced the mixed results. The author suggests that the overall findings are consistent with the main argument underlying the intrinsic benefit model.

Introduction

The signaling theory provides the game-theoretic framework to explain interactions under the condition of asymmetrical information. In its base form, the signaling theory assumes that there are two actors. The sender of the signal possesses the certain quality – such as competence, trustworthiness, and so on – which is not visible to the receiver of the signal. For the receiver of the signal, the potential outcome of her interaction with the sender depends on this invisible quality of the sender; for example, if a professor wishes to hire a student as a research assistant, she would like to ensure that the research assistant is reliable. The sender of the signal, on the other hand, wishes to convince the receiver that he indeed possesses the desirable quality, as the sender then can benefit from the future interaction with the receiver; the student who is looking for the work as a research assistant would like to ensure that the professor thinks he is reliable.

How can the student in the above situation convince the professor that he will be a reliable research assistant? The student, obviously, can tell the professor that he has been known to be reliable and industrious. However, the professor has no reason to believe what the student tells her; she is fully aware that he wants to work as a research assistant, and everyone in his position would be telling her that he is reliable regardless of their actual quality. She has no reason to believe him unless the student can provide actual evidence for his ability. The student, in turn, needs to ensure that the professor thinks she has enough reason to believe that he is a capable assistant. We can imagine many potential signals – in this case, “evidence” – that the student can utilize in this situation. The signaling theory identifying the unifying feature of such signals; the ratio of marginal cost of

producing the signal to the marginal advantage of improved assessment must be an increased function of quality being signaled. Typically, this means that the cost of producing the signal is less for those of high quality than for those of low quality. Applied to the above example, the student needs to find a “signal” for which a reliable and industrious student can produce more easily than a student who is merely pretending to be so could. Perhaps he can provide an example of his previous work as a research assistant. Or he can ask his former teacher to vouch for his ability. A lazy student would have difficult time producing such signals. If we assume that the similar interactions between the professor and the students recur over time, it would be the case that the students who are able to produce these signals are likely to be of higher quality overall than the students who do not; thus, the signal is reliable.

In the above example, I suggested that there is a cost-quality association because it is plausible to assume that high-quality candidates have an easier time (thus, cheaper) producing the signals such as the examples of previous work or the letters of recommendation. This probably is a reasonable assumption to make – nevertheless, there are possible counterarguments. For example, perhaps a student who is apt at cheating can forge any such materials with great ease, though most certainly he would not be considered a reliable worker. Having easy access to such signals might simply be the result of having a well-known mentor, and he may not do so well without this mentor. More importantly, there is no practical methodology to prove that the high-quality students would incur less cost producing these signals than their low-quality counterpart would; as such, it often is difficult to ascertain the cost-quality relationship here. For many social situations where the signaling could be relevant, it often is difficult to establish the cost-quality association

that is plausible. I suggest that this constitutes a notable hurdle to the adaption of the signaling theory within Sociology.

In this study, I propose a novel approach for establishing the cost-quality association within the signaling theory – named the intrinsic benefit model for the purpose of this study. Instead of focusing on the cost of producing a signal, the new approach focuses on the (perceived) intrinsic benefit obtained from producing the signal. If we assume that a certain behavior is commonly associated with the certain individual characteristic so that it is widely understood that those with the high level on the characteristic would obtain more intrinsic benefit from this behavior, the intrinsic benefit obtained from producing the signal (the behavior above) can function as the link to establish cost-quality association. Even if the actual cost to produce the signal is unrelated to the quality of the individual characteristic, the intrinsic benefit obtained from it can differentiate the effective cost of producing the signal between the high-quality individual and the low-quality individual. For example, most people would agree that it is likely that those who are more committed to a political cause would obtain more intrinsic benefit from joining a protest than those who are less committed – this, in turn, can differentiate high-commitment individuals and low-commitment individuals even though there is no difference in the cost of attending the protest.

I argue that this assumption is in line with how people commonly understand the motivations behind many behaviors. It could be especially relevant for qualities that cannot be easily associated with tangible outcomes. It would be relatively easy for the student in our first example to produce a signal that shows he has the necessary technical skills to be

the research assistant – for example, the sample codes from his previous work. But how can someone signal his commitment to a political cause? Someone’s commitment to a political cause does not make it any easier to attend the street demonstration. However, the intrinsic benefit obtained from producing the signal – participating in the demonstration, in this example – can differentiate between the high-quality actors (those committed to the cause) from the low-quality actors (those who are not).

The emphasis on the perceived intrinsic benefit can affect the very dynamics of signaling. In the intrinsic benefit model, the cost-quality association of a signal rests on the perceiver; therefore, any factors which weaken the perceived association between the intrinsic benefit obtainable from the signal production and the quality being signaled could hinder the effectiveness of the signal. This study proposes that the visibility of a signal is one such characteristic of the signal which could affect how the signal is perceived. In the conventional signaling theory, a signal is always assumed to be visible. However, with the focus on the perceived intrinsic benefit, visibility of a signal becomes a double-edged sword; a high visibility signal can inadvertently signal the sender’s motivation for signaling, which, in turn, makes the receiver to perceive that the signal was intended as a signal, and therefore was not produced for its intrinsic benefit.

This study focuses on signaling of one’s pro-social (as in willing to act toward the common – instead of individual – benefit) characteristic. Pro-social characteristics, such as altruism and trustworthiness, are highly desirable characteristics. Indeed, as we have discussed earlier, trustworthiness is inherently tied with the risk under asymmetrical information. However, being a person with pro-social inclination does not make any action

less costly – altruism does not magically make it easier to do something. Therefore, the traditional cost-quality association would not be applicable here. However, it is reasonable to assume that those who are more pro-social would obtain more intrinsic benefit from acting selflessly. Indeed, this is the underlying assumption of “warm glow” model of altruism. Thus, signaling of pro-social characteristic is a natural fit to the intrinsic benefit model of signaling.

I have conducted three experiments. The first two experiments were conducted online, using the vignette design. The experiments focused on the signaling of pro-environmental attitude by purchase of electronic vehicle (EV). The first experiment focused on the perceiver of the signal, evaluating how characteristic of the signal (EV purchased) affects the subjects’ evaluation of the characters in the vignettes. It used the 2x2 design, where the signal cost (the price of the hypothetical EV) and the signal visibility (whether the EV has a distinctive appearance or not) are two variables of interest. The result supported the predictions based on the intrinsic benefit model. The sender of the high cost signal was perceived as more committed to the environment, and the sender of low-visibility signal was perceived as more committed to the environment. The signal visibility had the stronger effect than the signal cost, and the effect of the signal visibility on the perceived environmental commitment was mediated by the perceived signaling motive.

The second experiment focused on the sender of the signal; it used the vignette design, asking the subjects how much they are willing to pay for the hypothetical EV. The visibility of the EV was the treatment condition (high/low), and the subjects’ environmental commitment and signaling motivation were measured as additional

independent variables. The result failed to support the predictions; although both the environmental commitment and signal visibility were positively associated with the amount the subjects were willing to pay, the signaling motive had no statistically significant effects.

The third experiment was conducted in the laboratory, and it was intended to test both the sender and the receiver at the same time in the situation with monetary consequences. The experiment concerned the signaling of trustworthiness based on participants' charitable giving, and it utilized the trust game (including both one-shot and repeated design) to measure trust and trustworthiness. There were three treatment conditions. In the no signaling condition, the subjects were asked about making a charitable donation to a nonprofit before they learned they were going to participate in trust games, and the information about the subjects' donation was unavailable to their partners during the trust games. In the no signaling motivation condition, the subjects were asked about their donation before they learned they were going to participate in trust games, and the information about the subjects' donation was available to their partners during the trust games. In the signaling motivation condition, the subjects were aware that information about their donation would be available in the trust games when they decided on the donation, and the information about the subjects' donation was available to their partners during the trust games.

The results of the third experiment were mixed. As predicted, the donors were more trustworthy than non-donors (i.e., they reciprocated the transfer of funds from the first subject by sharing the profit). However, contrary to the predictions, the first subjects

in the trust game did not treat the donors and non-donors differently – that is, the first subjects were no more likely to transfer to the donors than to non-donors – in the trust games where donation information was available. In turn, the second subjects in the trust game were no more likely to make donation in the signaling motivation condition than in non-signaling motivation condition. These results contradict previous research, which utilized similar designs. Interestingly, the treatment conditions had the strong effect in the frequency of transfer in the repeated trust game. The first subjects in the signaling motivation condition were far less likely to make transfer than the subjects in the other conditions, suggesting that the salient signaling motivation eroded overall level of trust among subjects.

The three experiments based on the intrinsic benefit model produced mixed results. The potential explanations and interpretations for these results will be discussed in the later chapters. Most notably, I suggest that Goffman's notion of strategic interaction – the concept aimed to bridge the game theory and symbolic interactionism – can provide the useful framework to interpret the findings obtained from this study. I argue that the symbolic interactionism has the unique potential to augment – or, rather, fulfil the potential of – the signaling theory, though such approaches have their pitfalls. Finally, several potential avenues for expanding the research around the intrinsic benefit model are identified.

In this first chapter I gave an overview of the theory, methods, and results of the dissertation. In Chapter 2, I present the literature review on the signaling theory focusing on the aspects of the signaling theory most relevant to this research. In Chapter 3, I will

describe the intrinsic benefit model used in this study. In Chapter 4, I describe the first experiment conducted for this research, using the online vignette design to study the subjects' perception of the hypothetical sender of the signal. In Chapter 5, I describe the second experiment which used the online vignette design to ask the subjects about their hypothetical behavior. In Chapter 6, I describe the third experiment, which used the laboratory experiment to test how the donation is used as the signal of trustworthiness. Finally, in Chapter 7, I discuss the findings of the study, and explore the potential future research projects to extend this study.

Chapter 1 Theoretical Background, Literature Review

Signaling theory deals with situations where there is conflicting interests and information asymmetry. The sender of the signal has information about his/her internal characteristic, but the receiver of the signal has no way of directly accessing that information. Although both parties can potentially obtain benefits from the interaction, the sender of the signal also has an incentive to exaggerate her quality. Such conditions could be applicable for many situations in the real world. Even in situations that are mainly cooperative, actors frequently have the incentive to appear better than they really are. The receiver, knowing that the sender has the incentive to exaggerate, would need a reason to believe the signal. If it is less costly for those of high quality to produce a certain signal than those of low quality, the signal can reliably communicate the high quality of the sender; the association between cost and quality can function as a way to ensure the credibility of the signal by making it more costly to produce dishonest signals than to produce honest signals.

a. Economics/Managerial Science

Spence's signaling model came out of a growing interest in asymmetrical information (Spence, 2002). In its most basic form (Spence, 1973), it assumes that there are two types of actors who send a signal – high-quality (high ability) and low-quality (low ability). The sender (prospective employees) knows their own quality, but the receiver (potential employer) does not. The receiver (potential employer) benefits from identifying the qualities of applicants accurately so that they can be assigned to appropriate positions. The receiver has the expectation on how well senders' education actually corresponds their

quality, which they update based on their experience with the previous senders. The senders, on the other hand, all prefer the position for high-quality types, as it will have better pay. Under this condition, education can (though not always) function as a signal of senders' quality when the cost of education is higher for the low-quality actors than for the high-quality senders; it is assumed that those with high quality have a much easier time attaining the higher education than those with low quality.

In Spence's signaling game, there are two equilibria that are of interest. In the separating equilibrium, the type of signal corresponds with the type of sender; all senders with the high-quality type receive education, and no senders with the low-quality type receive education, so that all senders self-select according to their quality. In the pooling equilibrium, every sender acts in the same way (everyone receives education, or everyone does not receive education) regardless of their type; it suggests that the difference in cost of education (between high quality and low quality) is not sufficiently large to induce the self-selection among the senders.

Following Spence, the signaling theory within economics evolved in multiple directions. As with Spence's original model, the signaling theory has been utilized to analyze the behaviors in the labor market (for example, Spence (1976), Tyler, Murnane & Willet (2000), Hopkins (2012)). These applications are largely considered as successful – it is argued that this model provides a better explanation for the association between education and wages than explanations based on human capital (Weiss, 1995).

On the theoretical level, some of the subsequent works on the signaling theory worked on refining the definition of equilibrium (for example, Cho & Kreps (1987)), or

analyzing equilibrium selection, either mathematically (for example, Nöldeke & Samuelson 1997) or experimentally (for example, Brandts & Holt 1992). At the same time, some studies aimed to further elaborate the signaling theory by adding the specific theoretical tweaks: for example, how noisy signals affect the signaling (for example, de Haan, Offerman & Sloof (2011)). Finally, there are some studies which speculate on the potential applications of the signaling theory on phenomena which are less typical in economics. For example, Austen-Smith and Fryer (2005) applied the two-audience (peers and potential employers) signaling model to explain the phenomenon of “acting white.” Because the lack of education is costly (as it hurts one’s prospect for employments), the rejection of educational institution can function as the signal of one’s sociability within a subgroup. I found this study to be somewhat noteworthy, due to how the authors came up with a novel approach to provide an explanation for the cost-quality association.

The economics branch of the signaling theory have found notable success in the fields related to management, particularly in finance. Investment-related decisions, by their nature, are high-stake situations characterized by asymmetrical information. Investors do not possess complete information on businesses they are investing in; therefore, they employ signals to determine the quality of a potential candidate for investment. In turn, sending the right signals of the quality would be of great interest for success of any enterprises. Within this area of research, the most well-studied application of the signaling theory concerns how the insider ownership signals the quality of the business (for example, Filatotchev & Bishop (2002), Bell, Moore & Al-Shammari (2008)). It can be assumed that insider ownership would be more costly for the executives of the low-quality business than for the executives of the high-quality business, as the performance of the

business would directly affect the future value of their stocks. In this sense, insider ownership presents a relatively straightforward application of the signaling theory within managerial science. In addition, there are the researches which explore other potential signals to the investors – such as Courtney, Dutta, & Yi (2016), Ahlers et al (2015), and Eddleston et al (2016) – which tend to rely on cost-quality associations which are not as straightforward. For example, Alsos & Ljunggren (2017), relying on Connelly et al (2011)'s review piece, aimed to apply the signaling theory to explain the role of gender in the investor-entrepreneur relationships. Although Alsos and Ljunggren discuss the characteristics of signals raised by Connelly et al (2011), the cost-quality argument seemed to have disappeared in their argument. Curiously, they wrote *(c)ostly signals are regarded as more credible as they are less likely to be fake*, indicating that they had fallen into the trap of the handicap principle.

Recently, another line of research explored the signaling roles of corporate social responsibility and corporate philanthropy. Given that CSR and corporate philanthropy, by their definition, does not directly produce revenue for organizations, whatever benefits an organization can obtain from such activities should be reputational in nature. There were some studies which aimed to analyze such activities through the signaling theory (for example, Su et al (2016), Xu, Xeng & Chen (2018); for a recent work utilizing a novel dual-audience approach, see DesJardine, Marti & Durand (2020)). These studies typically tend not to emphasize the cost-quality association; indeed, it is unclear if it is even possible to identify the specific quality which is being signaled by CSR or by corporate philanthropy. Typically, such studies use stock prices to measure the effect of signaling; however, I argue that the validity of such applications of signaling theory tends to be weaker, as it is unclear

what specific quality is being signaled, how the said quality results in the increased price of stock, and how they can rule out any confounding characteristics. It is most likely the case that CSR and corporate philanthropy can result in improved reputation, but it is important to note that reputational improvement is a phenomenon which is distinct from successful signaling – reputation, unlike signaling, does not need to correspond with any underlying characteristic. It is plausible that CSR and corporate philanthropy signal certain organizational characteristics. At the same time, I believe that the application of the signaling theory on CSR/corporate philanthropy would benefit from having a more narrow focus, explicitly defining what quality is actually being signaled.

Podolny(2010)'s status signaling model makes an interesting twist on the argument behind the use of signaling to obtain reputational benefit upside down. Instead of suggesting that an organization can obtain reputational benefit from signaling some underlying qualities by CSR, Podolny proposed that the (high) status itself can signal the quality of an organization (and its service/products), because it can be assumed that the high status – or, positive reputation – itself is so valuable for an organization so that it would be more costly for high status organization to do anything which could endanger their reputation; high status organizations have too much to lose from malfeasance (or, getting caught with malfeasance). The underlying argument here likely is older than the signaling theory itself – indeed, the author was always told by his mother during childhood to buy products from established players, because they have more to lose from failing to meet their standards. However, the beauty of Podolny's model of status signaling is that it elegantly marries the two concepts – status and signaling – together while at the same time maintaining the central structure of the signaling theory: the cost-quality association. As

long as the key assumption behind Podolny's model – that the high status organizations have more to lose from damaging its reputation than the low-status organizations do – holds true, it can be inferred from the signaling theory that the high status of an organization signals its quality. Although it can be argued that it may not be true that the high status organization would lose more for failing to uphold its reputation – high status itself can be self-perpetuating (Correll & Ridgeway, 2006), for example, thus is well-protected from harming its reputation – an organization can successfully signal its quality from its high status as long as the audience of its signal *perceives* that the high status organizations have more to lose from damaging its reputation than the low-status organizations do, as the author's mother does.

Instead of focusing on establishing cost-quality associations or elaborating the formal models, the applications of the signaling theory within managerial science tend to focus on integrating characteristics which could affect how a signal is used and how effective it is in real-world situations where signals are being used. The review by Connelly et al (2011) lists many such characteristics, including (un)intentionality of a signal, type of a signal, contradicting signals, and so on. Such characteristics do not necessarily interact with the signaling theory on the theoretical level; it probably is not impossible to develop a formal model based on the signaling theory, which can integrate all of the above characteristics, but the model developments are seldom the main goals of these researches. There is no question that such characteristics do affect the real-world use of the signals within the spheres of phenomena which are of interest within managerial (or more broadly, social) science; developing a formal model of the signaling theory which integrates these characteristics is not strictly necessary for the empirical studies on how such

characteristics affect the practice of signaling. At the same time, I argue that there is potential benefit to be had in developing the signaling theory in such ways so that it can account for additional characteristics which are relevant to the topics being studied, while maintaining the core framework on the signaling theory. Podolny's status signaling model can be seen as an example of such approaches, and I argue that it displays the possibility of extending the theoretical reach of the signaling theory.

b. Evolutionary Biology

Within evolutionary biology, signaling theory emerged as part of the broader interest in animal communication, especially for the purpose of mate selection. In Zahavi's (1975) handicap model, he argued that seemingly unproductive traits such as feathers of peacocks can evolve through sexual selection because they can signal fitness. These traits confer handicaps to the signalers, as the resource (such as carotenoid) to produce the signal (red-colored feathers) can be used in more productive manners (in the case of carotenoid, helping the immune system). In turn, more fit individuals should be able to better absorb the cost incurred by wasting resources on a signal than individuals who are less fit.

Taking the cues from Spence's signaling model, Grafen (1990) further elaborated Zahavi's model by adding the crucial element, that the ratio of marginal cost of advertising to the marginal advantage of improved assessment must be an increased function of quality. If we treat marginal advantage as fixed, it suggests that there should be an inverse

association between quality and cost. This results in the common, simplified model which is roughly equivalent to Spence's signaling model. This reformulation moved the signaling theory away from Zahavi's model, which strictly focused on the cost to produce the signal. Contrary to the commonly held misassumption on signaling theory, signaling cost by itself cannot ensure the reliability of signals (Grose, 2011); a signal whose cost (or benefit) is unrelated to its underlying quality, however costly it is, would not function as a signal at all. Therefore, it was Grafen's elaboration which led to the establishment of signaling theory within evolutionary biology. Note that the positive association between quality and the benefit obtained from successful signaling can also result in reliable signals. However, the applications of this aspect of the signaling theory has been relatively rare (Grose, 2011). This study will focus on the situations where there is an (inverse) association between the cost of producing a signal and the quality being signaled. The implications if the potential benefit-quality association will be discussed in the subsequent chapters.

The development of the signaling theory based on evolutionary biology, as applied to phenomena relevant to social science, primarily took place within the field of evolutionary anthropology. The earliest examples of such approaches were influenced more by Zahavi's handicap model than by Grose's elaboration of the signaling theory. These studies were typically influenced by evolutionary biology, and they tended to be unaware of the works by economists (Grose, 2011). Oftentimes, these studies used the signaling theory as a potential tool to provide rational-choice based explanation to seeming irrational, frivolous behaviors. For example, there are some researches which connect signaling theory and collective action problem by suggesting that frivolous display of generosity can function as signals of fitness or high status (Boone 1998: Smith and Bird,

2000; Griskevicius 2010; Van Vugt and Iredale, 2012). Or, closely following works in evolutionary biology, some researches focus on how some behaviors (for example, seemingly pointless risk-taking behaviors) can be useful for advertising to potential mates (Bird, Smith, and Bird, 2001; Wang and Griskevicius, 2014). Much like Zahavi's handicap model, these arguments rest on the implied cost-quality association: those with more resources – which can be seen as an analogue of fitness in evolutionary biology – can better absorb the cost of seemingly wasteful behaviors. Indeed, we may argue that such arguments share the same theoretical structure as Zahavi's handicap model.

The theoretical origin of such applications of the signaling theory is most evident in their emphasis on the importance of mate selection, where signaling of fitness is of paramount importance. It is reasonable to assume an individual is always motivated to seek a mate who is more fit; considering the definition of fitness within evolutionary biology, this argument almost seems tautological. Although this line of research has made meaningful contributions to evolutionary anthropology, the overt emphasis on costly signaling and mate selection is rather limiting. Many socially consequential instances of signaling could not neatly fall under the umbrellas of costly signaling or mate selection. One may argue that an entrepreneur seeking investment or a potential employer looking for an employee are close enough to mate selection as in they all are high-stake interactions between two parties under asymmetrical information. However, they are distinct from mate selection in that specific patterns of behaviors would not propagate themselves through the process of Darwinian selection; and these social interactions most certainly contain characteristics which are not relevant for mate selection.

It appears that the more recent applications of the signaling theory based on the evolutionary biology tended to be centered around evolutionary psychology, often with the relatively narrow focus on mate selection (for example, van Vugt & Iredale (2013)). Nevertheless, some studies took approaches which are less psychological (evolutionary psychological, to be specific) and more sociological (or social scientific/social psychological). Among the studies of this type, I suggest that the application of the signaling theory made by evolutionary anthropologists of religions is the most relevant for this study. Religions frequently prescribe (or proscribe) certain behaviors which cannot be explained away on utilitarian grounds. Signaling theory provides an attractive tool to explain why certain religious norms exist. This line of research (for example, Sosis & Alcorta (2002), Henrich (2009), Wilkins (2018)) typically argues that certain religious norms can function as the signal of one's sincerity of religious belief, allowing one to distinguish the believers from non-believers. Compared to studies focused on mate-selection, it is difficult to resolve the need for the cost-quality association for the studies on religions. Because of this, I argue that theoretical approaches used by the evolutionary anthropologists to resolve this problem can provide helpful toolkits for the purpose of this study, which we will discuss in later chapters.

Finally, we note the curious convergence between the signaling theory within managerial science and the signaling theory within evolutionary biology; the behaviors which would not allow for the simple explanations based on rational-choice/utilitarian frameworks – the behaviors which could be considered as pro-social (as in they benefit others) but which do not seem to directly benefit the actors themselves, in particular – make the especially suitable topics of analysis utilizing the signaling theory.

c. Sociology

Unlike in Economics, Evolutionary Anthropology, or Managerial Science, the signaling theory failed to garner much attention within Sociology. Several attempts were made to introduce the signaling theory into Sociology. Gambetta's work on the uses of signals in New York underground (2009) likely has been the most successful attempt to introduce a signaling theory into Sociological research projects, though it did not manage to "make a splash". There is one study which used the signaling theory to explain gang recruitment (Densley 2012). Przepiorka, who co-authored with Gambetta an article on the signaling theory, is a sociologist who did several studies using the signaling theory; however, his signaling-related works typically appeared outside of Sociology. Otherwise, I failed to find any significant influence of the signaling theory within Sociology. So far, it appears that these attempts failed to establish the signaling theory as a cohesive theoretical tradition within Sociology, unlike within the other fields of social science.

One example is Camerer (1988), who is a renowned behavioral economist, who proposed the signaling-based model of gift-giving in a prominent sociology journal. I argue that the signaling theory actually provides a superior explanation to the practice of gift-giving than exchange theory can – which is the dominant theory used to explain gift-giving within Sociology (for example, Emerson (1976)). The exchange theory, unlike the signaling theory, assumes that the continued exchanges of the gifts are efficient for both actors involved (note that the exchange theory, like the signaling theory, is based on the notion of rational actors). However, most instances of gift-giving in practice are inefficient – indeed,

they often are designed to be inefficient. Spending money to purchase a gift card of equal value, which only works in specific stores is inefficient compared to simply gifting the money which was used to purchasing the gift card. Yet, directly gifting money is frowned upon in the Western Europe. Even in the East Asia where money-gifting does take place in specific circumstances, the practice of money-gifting typically involves a built-in inefficiency, such as purchasing a specific paper bag to contain the money being gifted. Similarly, receiving Christmas cards would provide no tangible benefits whatsoever – so much that Schilling once described the custom of trading Christmas cards as an example of the suboptimal equilibria which became unavoidable, due to the difficulty of coordinated collective actions to move away from it (1978). As seen here, the practice of gift-exchange does not appear to produce material benefits at all to the parties involved, therefore violating the key assumption of the exchange theory. The signaling theory does not suffer from any such theoretical shortcomings, because the receiver of a signal does not need to obtain any benefit from the signal itself. However, the signaling-based interpretation failed to take off within Sociology, even among the rational-choice inclined sociologists who typically utilize the exchange theory.

The one application of Camerer's model within Sociology is the work by Paik and Woodley (2012) on courting gifts. It is curious that their topic went back to the ancestral home of the signaling theory in mate selection; however, the gifts in their model signal romantic commitment, not mere fitness. This change of focus seems appropriate for Sociology, as a signal within Sociology can certainly mean something more than fitness. Unfortunately, Paik and Woodley do not explain how the cost of gift-giving can be associated with romantic commitment; instead, they merely assume the existence of such

association. Beyond gift-giving, Raub (2014) proposed the model of signaling by hostage posting; although Raub is a sociologist, the structure of his study closely followed the style typical of economics. Indeed, Diego Gambetta, one of the few sociologists who actively promotes the signaling theory, suggested that “the use of signaling theory has been scant and not of uniformly high quality” (2011).

The signaling theory’s relative lack of popularity within Sociology does not mean the idea behind the signaling theory is irrelevant within Sociology. Indeed, I argue that asymmetrical information frequently takes an important, if implied, role in several classic works of Sociology and Social Theory concerned with status reproduction. For example, Veblen’s classic argument on conspicuous consumption (1899) was based on the same assumption as the model of costly signaling first proposed by Zahavi; those with more resources would be better able to absorb the cost of conspicuous consumption, so that the conspicuous consumption can signal one’s wealth. In another classic work, Simmel (1957) proposed the following theory of fashion: elites initiate a fashion trend, and the masses then follow it. Elites then move on to a new fashion, as their old fashion no longer distinguishes them from the masses. This is precisely what happens when signaling cost has no association with quality – the association between a signal and a quality is not sufficient for the signal to be stable, as anyone can easily fake it. If fashion is to act as a reliable signal of elite status, it must be less costly for the elites than for the masses. We can note that those with high status indeed have less difficulty absorbing the cost of high fashion.

Bourdieu’s influential argument on cultural capital in *Distinction* (1984) rests on the similar assumption. The key idea of *Distinction* is that the elites actively use their

sophisticated cultural disposition, or cultural capital, to distance themselves from the masses. Here, cultural capital can be understood as a signal of status. Cultural capital is not only transferred early in one's life, but individuals actively utilize their resources to obtain cultural capital (for themselves, or for their offspring), and it can be argued that such behaviors are consistent with the signaling theory. Thus, the signaling theory is a natural fit for explaining the role of culture in status reproduction.

Although it is evident that the signaling theory is relevant to the core research questions within Sociology, it can be argued such "signaling-like" approaches within Sociology have noticeable weaknesses. These approaches tend to first look at conspicuous wastes of resources and then try to hypothesize "signaling" explanations which show why such wasteful behaviors are in fact "rational". As such, their underlying logics are similar to Zahavi's handicap model, and the subsequent "costly signaling" models within social science. Indeed, Bird and Smith (2005) recognized this focus on cost (and cost alone) as the shared characteristics of what they identified as the forerunners of the signaling theory in classic works on social thought – the list which includes, Thorsten Veblen, Pierre Bourdieu, and Marcel Mauss. Gambetta (2011) criticized such approaches, suggesting that they risk falsely assuming that everything which is wasteful must be a signal. This criticism is similar to the one raised by Grose (2011), who suggested that the applications of the signaling theory within social science tend to overemphasize the importance of cost of producing signal. In the signaling theory, not all costly behaviors are signals and not all signals need to be costly; however, such theoretical nuances are lost when one is preoccupied with the behaviors which seemingly are wasteful.

On the flipside, the preoccupation on seemingly frivolous signals could distract our attentions from the signals that do not appear to be frivolous wastes of resource. Indeed, in the subsequent chapters, I suggest that subtle signals may be more effective in certain situations than conspicuous signals. However, such nuances of signaling would be lost when one focuses on “costly signaling”. Additionally, I argue that “costly signaling” models often lack in theoretical precision; it often is unclear what characteristic is being signaled by the wasteful behavior, and how this behavior signals the said characteristic (as in cost-quality association). Indeed, the literature in managerial science is unclear even about what characteristic is being signaled by corporate philanthropy. Of course, the “signaling-like” theories by Veblen or Bourdieu do not actually specify the characteristics being signaled; rather, it is assumed that the signals are considered to be desirable to their audience for some reasons. We may even say that the omnipresent idea of “fitness” in the evolutionary biology is rather vague when it is applied to social science; it is unclear what kind of “fitness” is being signaled when a member of a hunter-gatherer tribe hosts a feast to the fellow members of the tribe. Julius Caesar was well-known for taking debts to finance many grand feasts for the Roman populace; whatever characteristic he was signaling, it seemed to have worked for him – yet, it seems obvious that he wasn’t signaling his fitness by taking debts. (Note: Gambeta seems to love using historical examples – therefore, I declare that I am merely conforming to the norms for the subfield). With these reasons, it is evident that the “costly signaling” school of the signaling theory is insufficient in explaining the wide range of signaling-related behaviors within Sociology.

d. Discussion

Why would a sociologist want to employ the signaling theory, when so far it has been largely ignored within Sociology – and that its most popular form within sociology (costly signaling) has notable flaws? For answering this question, it is useful to think about the underlying appeal of the signaling theory within social science. Grose (2011) suggested that the main appeal of the signaling theory is strength of its simple and intuitive model. Considering the popularity of prisoner’s dilemma – the prime example of game theoretical models which are relatively simple and easy to understand – within social science, it is reasonable to argue that the relative simplicity of the signaling theory helped its adaptation within social science. However, I argue that the signaling theory within social science does not need to be simple to be effective – indeed, many important social phenomena would not be able to neatly fit into the scope of the basic signaling model unless we are willing to keep the model sufficiently vague (as in the case of costly signaling model).

Instead, I suggest that the signaling theory within social science should strive to suit the theoretical demands, which are unique to social science. Bird and Smith (2005) argued that the main benefit of the signaling theory is its capability to integrate symbolic communication into rational-choice/strategic action framework. Within Sociology, symbolic communications are typically analyzed through their meanings. However, it is also possible that many communications are strategic in nature, as Bird and Smith suggested “communication is aimed at securing social benefits with real material consequences, and these consequences are crucial to an understanding of form, dynamics, and fate of communicative and symbolic behavior” (2005). Following their view, I argue that signaling theory in social science can go beyond simple signaling of resources and fitness – instead, it can offer a model, which can potentially explain a wide range of

communicative behaviors. I suggest that Gambetta's study of the uses of signals in the criminal underworld (2009) illustrates a direction, which the signaling theory can take within Sociology. Gambetta's richly detailed, highly contextualized ethnographic study evokes the symbolic interactionism of Goffman far more than it evokes the rational choice model of Spence. At the same time, Gambetta's work is clearly based on the signaling theory – as it is focused on why the cost of producing a specific signal would be associated with the quality being signaled. However, I suggest that the context-rich, qualitative research alone is not sufficient to help establishing the signaling theory within Sociology; I argue that there is the need to develop a framework around the signaling theory which produces empirically testable predictions, while at the same time is being capable of accounting for the unique demands on social science.

Chapter 3: Intrinsic Benefit and the Signaling Theory

In the previous chapter, I have suggested that, for the signaling theory to be successful within Sociology, there must be a relatively straightforward approach to establish cost-quality associations across the diverse set of phenomena. Traditionally, the costly signaling approach occupied the de-facto standard role as the framework for establishing cost-quality associations. Although this approach of the signaling theory can be effective in certain situations, it has some notable drawbacks, as criticized in Grose (2011), Gambetta (2011) and Barker et al (2019). For the purpose of Sociology, I argue that its most important flaw is that many unobservable qualities which would be beneficial to be signaled – trustworthiness, loyalty, commitment, and authenticity, for example – are not resources that can be “wasted”; therefore, they cannot be accounted by the costly signaling approach. In addition, the costly signaling approach may not provide full picture of the phenomenon even when it is applicable. Following the logics of the costly signaling approach, both frivolous displays of generosity and conspicuous consumptions can achieve the same goal of signaling one’s resource. Yet, social meanings attached to these behaviors are clearly different; the costly signaling by itself would not be able to account for this difference. However, if we wish to apply the signaling theory to more diverse phenomena, cost alone would not be able to offer sufficient explanations.

This project will follow the approach which is used by the evolutionary anthropologists studying religion. They argue that one’s observance of a religion’s costly requirements can signal her commitment to the religion. Those who are more committed to the religion should obtain more intrinsic benefits from observing religious requirements,

making these requirements less costly for them (Sosis and Alcorta, 2003). This, in turn, will increase solidarity among members of the religion. Testing this application of signaling theory, Sosis and Bressler (2003) have found that a higher number of costly requirements within religious communes is associated with greater longevity, but this is not true of secular communes. Similarly, Tan and Vogel (2008) found that religiosity can function as the signal of trustworthiness. As a more generalized form of this argument, Henrich (2009) proposed a model of credibility enhancing display, focusing on signals of sincerity of (verbally displayed) beliefs. According to this model, a belief is more likely to be learned if it is accompanied by an action which appears to be more costly for those who do not actually believe in what they say. For example, religiously prescribed hardships (such as fasting) would be perceived as less costly for those with sincere religious beliefs. Therefore, such behaviors can signal one's sincere religious beliefs. Although Henrich's model is more general in nature, it is focused on what makes beliefs more credible; thus, it is not necessary applicable to all phenomena where signaling theory could be relevant. In addition, Henrich does not state exactly why certain behaviors would be perceived as less costly for those with sincere beliefs.

This project is based on the idea that the same general principle applies to many social phenomena other than religion, as long as there is the perceived association between quality and intrinsic reward. The intrinsic benefit in this study is conceptualized to generalize Henrich's argument, focusing on the missing link in Henrich's argument on why a certain behavior would be perceived as less costly for those with a sincere belief. The crux of Henrich's argument is that there are certain beliefs with which it is perceived that increased sincerity of beliefs would reduce the costliness of certain actions. For example,

those who believe in otherworldly salvation would be more likely to find following religious rules less costly. I propose that intrinsic benefit is what enables this perceived association between beliefs and costliness of actions. One's sincere belief in his/her otherworldly salvation would not directly reduce the physical cost of attending church every week by itself. Instead, what the sincerity of belief does is that it is (perceived to be) likely to produce more positive feeling from the act of attending church every weeks – be it joy, satisfaction, or relief from anxiety. Intrinsic benefit in this model is conceptualized to encapsulate these positive feelings which could be produced from certain actions depending on actors' characteristics. Intrinsic benefit obtained from an action could be seen as increased utility, in a way similar to how some economists conceptually treat “warm glow” – positive feelings obtained from altruistic acts – as increased utility. Benefits from intrinsic benefit, much like benefits from warm glow, are not something tangible; instead, its benefit is assumed to take the form of positive emotions.

At the same time, unlike Henrich's model, the conceptualization of intrinsic benefit makes it possible to extend its application beyond signaling of beliefs, thereby making it applicable to other characteristics. Sincerity of beliefs is not the only thing which can enable certain actions to produce positive emotion. It is possible that positive emotion is produced by an underlying trait or characteristic, instead of being produced by a sincere belief. This conceptualization of intrinsic benefit enables us to extend signaling theory into various characteristics, traits, and attitudes of the producer of a signal. For example, it can be assumed that those who are politically committed would obtain more intrinsic benefit from attending rallies, and those who have sophisticated taste would obtain more intrinsic benefit from high culture. In addition, the conceptualization of intrinsic benefit allows to

treat “warm glow” from an altruistic act as a product of pro-social personality (or, increased social preference); in turn, an altruistic act can function as a signal of prosocial personality. As previously discussed, the applications of the signaling model in social science typically focus on the signals of the resources or wealth. Although signaling of the resource can be important in many situations, not every signal can be explained as a signal of resources or wealth. There are many other characteristics that people would want to signal, such as loyalty, trustworthiness or competence. The intrinsic reward model can provide the simple explanation why some cultural practices can be used as the signals of these characteristics. This can allow signaling theory to be applied to a wider range of phenomena; any cultural acts where there is the widely agreed upon association between an internal characteristic and an intrinsic benefit, the signaling theory can be applied to.

In addition, the intrinsic benefit model will allow for the incorporation of more in-depth interactions than a simple model of costly signaling would. The effect of signaling in the intrinsic benefit model depends on the perception of the perceiver; therefore, it can potentially incorporate sociological elements such as interpretations or shared meanings. While the model in Henrich (2009) is specifically based on the perceived, instead of actual, association between the cost and the quality, Henrich did not explicitly discuss the consequences of the shift from the actual association to the perceived association. However, I suggest that this shift can have drastic effects on the signaling theory. The perceived association, by its nature, is fluid and dependent on the social contexts surrounding the signal. Therefore, the focus on the perceived association allows us to bring the social contexts into the core of the signaling theory. This can potentially complicate the theoretical clarity of the signaling theory. At the same time, the social context of a signal

inherently contains the symbolic elements. Bird and Smith (2006) argued that the main advantage of the signaling theory is that it can bridge between semiotic and materialistic frameworks. The intrinsic benefit model, by its nature, integrates symbolic elements into the signaling theory. In a sense, the intrinsic benefit model tends to see signaling similar to the processes of strategic interaction, as discussed by Goffman (1970) – where symbolic communications are strategically utilized to obtain tangible benefits while employing tools which are available through social contexts. Signaling in social world – unlike with the assumption behind the signaling in evolutionary biology – is inherently strategic, as the actors actively engage in managing the signal they produce, and in interpreting the signal they receive; the intrinsic benefit model can provide a tool to capture the strategic dynamics of signaling. As we have seen here, the intrinsic benefit model can extend both the breadth and the depth of the sociological applications of signaling theory.

Intrinsic Benefit Within Sociological Signaling Theory

How can the emphasis on the intrinsic benefit resulting from the production of the signal improve the sociological applications of the signaling theory? For the most obvious example, we can go back to the hypothetical comparison between the billionaire who bought an expensive yacht and the billionaire who donated the same amount of money (as the yacht purchase) to a charity. From the perspective of costly signaling, the two signals appear to be nearly identical; neither signal would be likely to directly benefit the billionaires, and they both waste the equal amount of financial resource. However, seeing through the lenses of intrinsic benefit, the two signals are clearly distinct from each other.

It is plausible that those who purchase an expensive yacht would be perceived as being motivated for the showing off their wealth, or their preference toward the lifestyle associated with extravagant spending on the yacht. On the other hand, those who donate to a charity would be perceived as being motivated by the commitment to the cause behind the charity, or by their desire to be seen as being committed to the same cause. In turn, the two signals, despite having the equivalent monetary cost, would evoke the very different perceptions on the perceivers. Similarly, although Bourdieu's notion of cultural capital can be understood as a form of signaling, his key idea around the cultural capital – the cultural capital is signified both by the resource which is required to acquire the certain cultural practice and by the shared understanding that the certain cultural practices are more “worthy”. I suggest that the signaling theory can provide a more suitable fit to the notion of cultural capital incorporating the intrinsic benefit obtained from the signal production to the emphasis on the cost of the signal production. For example, a piece of highblow art frequently requires the certain background knowledge to be appreciated, such as the knowledge of the Greek/Roman/Biblical mythology in case of classic (non-modern) European arts. If this background knowledge typically is only held by the members of the elite class, then the ability to obtain the intrinsic benefit from this highblow art would also be contingent on having the elite background. Therefore, the appreciation of highblow arts can signal one's elite upbringing even though one's upbringing would not directly affect the physical cost of going to the museum.

The intrinsic benefit also is a natural fit for applying the signaling theory to the practice of gift-giving. Oftentimes, the most socially valued gifts are not necessarily the most expensive ones. Indeed, so-called “thoughtful” gifts – the gifts which were made with

the careful considerations of the receivers' preferences – seem to be as appreciated. It is plausible to assume that the givers who are more committed to the well-being of the receivers would make more enjoyment out of preparing such gifts. Indeed, it is likely this association is much stronger for “thoughtful” gifts than for “expensive” gifts; I suggest that together with the intrinsic benefit, the signaling theory can provide the satisfactory account of the actual practice of gift-giving.

Additionally, the signaling theory, together with the emphasis on the intrinsic benefit, can be employed to explain various other behaviors which were not traditionally explained by the signaling theory. For example, Willer, Kuwabara, and Macy (2009) argued that people willingly enforce norms that they do not actually agree with, so that they can appear to agree on the norm. Although they do not refer to signaling theory, it is possible to explain their argument in terms of signaling theory. Enforcing a norm is much more costly than simply claiming to agree on the same norm. At the same time, those who are highly committed to a norm should obtain considerable intrinsic benefit from enforcing the norm. Therefore, norm enforcement can function as a signal of one's commitment to the norm. More recently, Kim (2017) had made the argument which is identical to Willer, Kuwabara, and Macy under the context of managerial science; unlike Willer, Kuwabara and Macy, Kim had made an explicit reference to the signaling theory. Through the intrinsic benefit, the signaling theory can also explain how members of disadvantaged groups, like low-income youth, tend to practice cultural acts that are oppositional to the mainstream society, as reported in Willis (1977) and others – this also as the motivation behind the model by Austen-Smith and Fryer (2005). Similarly, some have applied the signaling theory to risk-taking behaviors, as seen in Farthing (2005).

I suggest that the signaling of authenticity provides another promising avenue for the signaling theory focused on the intrinsic benefit. Authenticity often is a highly desirable quality, especially for organizations and individuals who deal with cultural products. Authenticity is the quality which strictly resides in the perceiver of the signal – there is no objective criteria for “authenticity” – and which is inherently ambiguous. Money is frequently perceived as the anathema to authenticity; because of this, it would be difficult to try to signal authenticity through spending monetary resources. The past studies have noted that the organizations sometimes make frivolous displays of efforts – particularly in the areas closely relevant to their organizational identity – to signal their authenticity (Carrol and Swaminathan, 2000; Beverland, 2005). Typically, these attempts are explained in terms of strategic identity claims. However, these attempts can also be considered as signals. For example, the wineries that Beverland studied engaged in many costly, seemingly pointless practices – for example, selecting trees used for making their barrels – to display their commitment to authentic wine-making (2005). Again, we can argue that it is widely perceived that those with high level of commitment for authentic wine-making could derive more intrinsic benefit from the seemingly wasteful practices. Thus, their behaviors are consistent with the signaling model. Similarly, Negro, Hannan, and Fassiotto (2015) explicitly applied the signaling theory to winemaking. They argued that wine made by winemakers belonging to “Biodynamic” category are rated higher in non-blind tasting tests than wine made by winemakers belonging to “organic farming” category, because the cost of entry to the former category is higher than the cost of entry to the latter category. Although they suggest there is cost-quality association because “Biodynamics” category requires more discipline, I suggest that the same argument is equally applicable to “organic

farming” category. However, according to Negro, Hannan, and Fassiotto, “Biodynamic” winemaking includes elements which seem to be irrational and esoteric – for example, “preparations” used for Biodynamic winemaking includes bizarre items such as “Yellow flowers buried sheathed in a stag’s bladder” (2015). If anything, these elements resemble religious rituals more than rational business practices; thus, I argue that the intrinsic benefit model could provide the strong explanation for their findings. It is plausible that the esoteric practices could invoke stronger association with the intrinsic benefit than seemingly rational practices; the perceiver would assume that, if winemakers are willing to perform intricate and (seemingly) nonsensical rituals, they must be truly committed to the art of winemaking. Therefore, I believe that their findings are consistent with the intrinsic benefit model.

Organizations, of course, are not the only entities would strive to signal their authenticity through making frivolous efforts. For example, Hietanen and Rokka (2015) described the efforts put by DJs in “Dubstep” music subculture to maintain their claim of authenticity; for example, DJs they interviewed claimed that they scraped off track listings from the records they use in DJing. Although it is possible to argue that such efforts are the simple acts of self-handicapping (as in the handicap model of signaling), I suggest that the self-handicapping provides an insufficient explanation; much like the biodynamics winemakers in Negro, Hannan, and Fassiotto (2015), DJs in Hietanen and Rokka (2015) made the frivolous efforts in such ways so that there would be reasonable association between the identity being signaled and the intrinsic benefit obtained from producing the signal. I argue that the signaling theory, together with the emphasis on the intrinsic benefit obtained from the signal production, can show how the ritualistic behaviors such as

scrapping off track listings – the signals which are neither utilitarian nor entirely pointless – can serve as the potent signals of one’s authentic commitment to a cultural identity.

Conclusion

As we have discussed so far, this study aims to both develop the signaling model based on the association between an internal characteristic and an intrinsic benefit from a cultural act. There are criticisms that researchers using the signaling theory, both in biology and in social science, tends to focus too much of the model itself instead of model verification (Grose, 2011). Following the tradition of sociological social psychology which emphasizes theory-testing using experiments, this study intends to verify the intrinsic benefit model through controlled experiments.

For verifying a model based on the signaling theory, it is useful to discuss how evolutionary biologists employ and test models from the signaling theory. Biologists typically verify their model by empirically testing two key elements of the signaling model: the association between the signal and the quality of the sender, and the association between the signal and the receiver’s assessment of the sender (Searcy and Nowicki, 2005). For example, since carotenoid in birds can be used both for improving immunity or for pigmentation, more fit birds can better handle the cost of pigmentation than less fit ones (Searcy and Nowicki, 2005). This signal influences mating success: female birds of many species prefer males with red coloration, therefore red coloration is accepted as the signal of fitness by its receivers (for example, (Hill 1991)). Here, I suggest that a similar approach can be used to test signaling theory within social science. This study will test the prediction

based on the intrinsic benefit model on both sides of the signaling theory – the association between the quality of the sender and the signal produced by the sender, and the association between the signal and the receiver’s association of the sender.

A particular issue of note for testing the intrinsic benefit model is that intrinsic benefit itself is not directly measurable. Indeed, the intrinsic benefit in this study can be understood as the blanket terminology for any psychological reward which propels individuals to make a costly cultural act which is unlikely to provide much material benefits. Therefore, it would be impossible to directly measure intrinsic benefit as the singular observable variable. Instead, this study aims to manage this issue by generating and testing the predictions which are unique to the intrinsic benefit model. In particular, the intrinsic benefit model is based on the idea that the dynamics of signaling is driven partly by the perceived association between internal characteristic and intrinsic benefit – therefore, the factors which can affect the perceived association can affect the dynamics of signaling. The present study will focus on signal visibility as such factor – it will experimentally test the predications made by the intrinsic benefit model, including the ones centered around the signal visibility.

Chapter 3 Intrinsic Benefit Model: Hypotheses

In this chapter I will outline in detail the intrinsic benefit model of signaling. I will review the assumptions upon which it is based, derive the propositions based on the assumptions, and offer several hypotheses that can be tested empirically.

Assumptions, Scope Conditions, Propositions

Theoretically, the model of cultural signaling is based on two components; one for the receiver of the signal, and another for the sender of the signal. Both components share the same assumptions and the scope conditions.

- Assumption#1a: there are two actors – p, the sender of the signal, and q, the receiver of the signal
- Assumption#1b: p's signal has its intended audience, A
- Assumption#1c: a is a subset of A who is able to recognize the signal – a/A is the proportion of the audience who can recognize the signal
- Assumption#1d: q is a member of a
- Assumption#2: p possesses an internal characteristic I, and the level of I cannot be directly observed by others
- Assumption#3a: S is a cultural act that can be performed by p, with the cost of C
- Assumption#3b: The meaning of S is shared across the society to a reasonable level
- Scope condition#1a: both p and q are self-interested
- Scope condition#1b: p is aware of his/her level of the internal characteristic I
- Scope condition#1c: q can obtain more benefits from interacting with p, when p's level of the internal characteristic I is better

- Scope condition#1d: p has the expected benefit of r, which is the amount of benefit that p believes that p could obtain from interacting with q
- Scope condition#1e: q can choose whether to interact with p or not
- Scope condition#1f:p's expected signaling benefit r is independent from its level of I

Note that this set of scope conditions mostly corresponds with the basic assumptions of signaling theory. Similar assumptions are shared by both evolutionary biology (Searcy and Nowicki, 2005) and by Spence's (1973) job-market signaling model. I argue that situations that meet this set of conditions are surprisingly common. Many internal characteristics – such as loyalty, commitment, and intelligence – are not directly visible to others. At the same time, individuals who possess high level of these internal characteristics are often sought out for potential interactions. Therefore, being able to signal one's high level of such an internal quality is very beneficial in many situations. In addition, the model does not have a restrictive definition of what makes a cultural act. A cultural act is a behavior whose meaning is somewhat shared across its audience (assumption#3b) – the signaling model does not function at all if everyone has a different interpretation of the same behavior. An important element of culture is that it is being widely shared by the group (Chwe, 1998); it is this characteristic of culture which makes cultural acts the effective vehicle for signaling. Finally, assumption#1b and #1c suggest that a signal has its intended audience, but not all members of the audience can necessary recognize the signal. The likelihood of recognition is affected both by the visibility of the signal and the audience's knowledge of the signal. Tattoos on one's lower back are hidden in many social situations. One may read a highly sophisticated book in public, yet only those who are knowledgeable about the high literature can recognize the book – indeed,

many elaborate cultural signals (for example, highly sophisticated articles of fashion) are frequently invisible to those who do not possess the necessary knowledge to decode them. Therefore, a/A – the proportion of audience who can recognize the signal – can affect expected benefit from signaling. In a typical signaling model, this is a non-issue; evolutionary biology typically deals with highly visible mating signals, and education is always recognizable in any situations involving selection for employment.

Finally, we can note that the scope condition#1f is somewhat limiting – indeed, in many social situations, benefit that can be obtained from successful signaling may be positively correlated with the sender’s internal quality. For example, if someone benefits from signaling his/her preference of heavy metal, it is most likely because s/he does like heavy metal. Such correlations can diminish the relative benefit the sender can obtain from making untruthful signals that overstate his/her level of I . This will make the situation closer to a coordination game with communication than a signaling game; in coordination game, even cheap talk (no cost, non-binding communication) can facilitate coordination, as there is no incentive to exploit one’s partner (Farrell and Rabin, 1996). This distinction can have important implications for analysis of many social phenomena; the model in evolutionary biology actually incorporates this possibility (Graffen, 1990). Grose (2011) criticizes tendency of research utilizing signaling theory to emphasize association between cost and quality and ignore (potential) association between benefit and quality. Although I agree with the potential importance of benefit-quality association in many social situations, I argue that analysis on benefit-quality association falls outside of the aims of this research, as positive benefit-quality association can make the effect of intrinsic benefit redundant. In addition, there are two more scope conditions.

- Scope condition#2a: p obtains the intrinsic benefit of amount b from doing S
- Scope condition#2b: The higher p's level of I is, the larger the intrinsic benefit it provides

These scope conditions specifying that the amount of intrinsic benefit obtained from a cultural act increases with the increased level of an internal characteristic of the actor are the key element of the model. Evolutionary anthropologists on religion, such as Sosis and Bressler (2003), made a similar argument. I argue that many situations associated with cultural signaling meet this condition. For example, if I am a very devoted supporter of a presidential candidate, I would get more intrinsic reward from attending his rally. If someone has very sophisticated taste in literatures, s/he would get more enjoyment out of Finnegans Wake than otherwise. This association between an internal characteristic and an intrinsic benefit, combined with the cost of the cultural act, can create an inverse association between the internal characteristic and its signaling cost. Finally, we have to note that intrinsic benefit here excludes intrinsic benefit that is obtained by acts of signaling themselves – for example, feeling good about showing off.

The above assumptions and conditions lead to the following propositions about the sender of a signal.

- Proposition#1a: when $r * (a / A) = 0$, p will perform S when $b > C$
- Proposition#1b: when $r * (a / A) > 0$, p will perform S when $b + r * (a / A) > C$
- Proposition#1c: when $r * (a / A) > C$, p will perform S regardless of his/her level of I

Because a/A is the proportion of the audience who can recognize the signal, $r * (a / A)$ corresponds with the p's expected benefit from the interaction with q. This value can be

understood as the expected signaling benefit that p can obtain through signaling, which is the key idea behind the sender's side of the model presented here. Without an expected signaling benefit, the actor would perform a cultural act only when s/he can obtain enough intrinsic benefit to make up for its cost (proposition#1a). This can be seen as the equivalent of the separating equilibrium in Spence's job-market signaling model (1973). However, the expected signaling benefit will make actors perform a cultural act that is otherwise not beneficial enough for them (proposition#1b). This, in turn, will make it so that S will distinguish between the levels of I at a point that is lower than otherwise. For example, I may find attending a party to be an alright experience. It is not bad, but I don't think it is enough to make up for its cost in time and energy. However, if I also want to portray myself as sociable, then the combined benefit of the actual enjoyment from the party and appearing more sociable can make it worth its cost. If not for this expected signaling benefit, only the most sociable would attend the party; but with the signaling benefit, those who are only somewhat sociable would end up attending it. Therefore, attending the party no longer signals the high level of sociability; nevertheless, those who attend the party would tend to be more social than those who do not.

- Assumption#4: q has the naïve model of signaling, where p would perform S when the reward from S – both intrinsic and signaling – exceeds its cost C
- Proposition#2a: when q assumes that p expected no signaling benefit, q assumes that that p performed S because s/he could obtain more intrinsic benefit from S than its cost

- Proposition#2b: when q assumes that p expected the signaling benefit from S, q assumes that the level of intrinsic benefit p obtained from S may not be as high as its cost
- Proposition#2c: q estimates p's level of intrinsic quality I based on the amount of intrinsic reward that q perceives p to have obtained from S

These propositions describe the reaction by the receiver to a cultural act. It assumes that q possesses a naïve version of the model presented here, and uses it to evaluate an internal characteristic of p. If q assumes that p was expecting no benefit from signaling, then q would think that p performed for its intrinsic benefit alone (proposition#2a). In this case, q would assume that p's level of internal characteristic must have been sufficiently high to provide enough intrinsic benefit from S. However, if q has a reason to believe that p was motivated by a potential signaling benefit, q would take that into account (proposition#2b). In turn, q would assume that the amount of intrinsic benefit p receives from S might have been less than its cost, and p's level of internal characteristic I may not be as high as in the case where the signaling benefit expected by p is perceived to be zero (proposition#2c).

Note that this process can occur both consciously and subconsciously. When it is conscious, the whole process can be seen as a process of attribution, as seen in Jones and Davis (1965) and Kelly (1967). Here, q attempts to explain p's cultural act of S. If q perceives no extrinsic motivation for p's act of S, then q would assume p was motivated solely by intrinsic benefit, which in turn was caused by p's level of I. If q does suspect extrinsic motivations, q will discount the perceived intrinsic motivation of p. From this

reasoning, the strongest signal is a signal that is not perceived as a signal at all. If the expected signaling benefit is perceived to be very low – perhaps due to the low visibility of the signal, or because of the obscure, arcane nature of the signal – then q would assume that p was motivated solely by intrinsic benefit. On the other hand, highly obvious, obnoxious signals may fail to signal a high level of the internal quality; instead, the sender may end up being perceived as a poser. It is also possible for q to alter their perception of S from their past experiences. If I am constantly disappointed by the quality of the senders of a specific signal, I am going to take that into account the next time I encounter it. An important consequence of these propositions is that the actual signaling benefit may not be as high as what the sender of the signal has expected. Interestingly, this mechanism is what keeps the signal reliable in the long run. Thanks to it, the senders will have to eventually reduce their expected benefit, discouraging them to produce the signal that exaggerates their quality. In this model, it can be done through reducing r (others do not act as positively to the signaled level of I) or through reducing a/A (less people treat S as the signal). With this self-correction, the cultural act will correspond with its related internal quality once again. This is the equivalent of two-cycle in Spence's signaling game (Nöldeke & Samuelson, 1997); the system is not in equilibrium and actors with lower I will alternate their behaviors. In similar vein, in evolutionary biology, evolutionary stable strategy (ESS) model of signaling suggests that a signaling system will be stable only if the (total) benefit the receiver obtains from honest signalers is larger than the (total) cost the receiver suffers from dishonest signalers (Grafen & Johnstone, 1993); if this condition is not met, the receiver would stop accepting the signal due to the cost incurred by the dishonest signals.

Hypotheses

From the above model of cultural signaling, we can make the following hypotheses.

Hypothesis#1a: The actor who engages in a cultural act that is perceived as having a higher cost is perceived as having a higher level of the signal-relevant internal characteristic than the actor who engages in an equivalent cultural act that is perceived as having a lower cost

As seen in Proposition#2a, the perceiver of the cultural act would assume that the act was done primarily for its intrinsic benefit, if there is no information that suggests otherwise. Then, an act with high cost suggests that the actor obtains enough intrinsic benefit to make up for the cost. In turn, this shows that the actor has a high level of the internal characteristic that would lead to an intrinsic benefit. Therefore, the actor who engaged in a cultural act with a higher cost should be perceived as having a higher level of the internal quality that it signals than an actor who engaged in a cultural act with a lower cost.

Hypothesis#1b: The actor who is perceived to have an extrinsic motive for a cultural act is perceived as having a lower level of signal-relevant internal characteristic than the actor who is not perceived to have a motive based on signaling benefit and sending an otherwise comparable signal

Proposition#2b suggests that the perceiver of a cultural act is aware of the potential benefit that the actor can obtain from successful signaling. Therefore, the perceiver will use his/her perception of the potential benefit to discount the intrinsic benefit that is obtained by the actor. If the actor is perceived as seeking an extrinsic benefit from successful

signaling, the cultural act no longer serves as the signal of internal characteristic. If the cultural act is perceived as being motivated for its signaling benefit, its actor would be perceived as having a lower level of the signal-relevant internal characteristic than otherwise.

Hypothesis#1c: The receiver of the signal is more likely to initiate the interaction task with the sender of the signal when s/he is perceived to have a higher level of the internal characteristic and when the receiver's outcome from the interaction task is dependent on the internal characteristic of the sender

Signaling positive qualities can be beneficial for the sender because it can affect behaviors of the receiver. The receiver of the signal must actively employ the information obtained from the signal to decide whether to interact with the sender of the signal or not. Assuming that the internal characteristic is relevant to the interaction task and the receiver's outcome on the interaction task is dependent to the internal characteristic of the sender, the receiver should prefer interacting with those who are perceived to have a higher level of the internal characteristic than with those who are not. In turn, being selected for the interaction would be beneficial for the sender of the signal, creating incentive for signaling in the first place.

Hypothesis#2a: The actor with a higher level of a signal-relevant internal characteristic is willing to pay more for a cultural act than the actor with a lower level of a signal-relevant internal characteristic

This hypothesis is based on the core idea of intrinsic benefit model, that the same cultural act will provide more intrinsic reward for the actor with a higher level of the

signal-relevant internal characteristic than for the actor with a lower level of the signal-relevant internal characteristic. For example, individuals with higher level of religiosity can obtain more intrinsic reward from a pilgrimage than individuals with lower level of religiosity. Therefore, those with higher level of religiosity should be willing to pay more for the same pilgrimage than those with a low level of religiosity. Although this hypothesis has a common-sense quality to it, it nevertheless must be tested to ensure the validity of the signaling model.

Hypothesis#2b: When perceiving that it is desirable to signal the high level of the signal-relevant internal quality, the actor is willing to pay more for a cultural act with high visibility than for a cultural act with low visibility

Visibility could be seen as the most important factor for any signal, since a cultural act can only function as a signal when it can be perceived as such. Visibility also is the key variable for many studies of conspicuous consumption (Chao and Schor, 1998; Heffez, 2001; Griskevicious, 2010). However, in many models based on signaling theory, there is implied assumption that a signal is visible by definition; as a result, they would not place much attention on signal visibility. The general lack of interest in the role of visibility within the signaling theory can be explained by the nature of conventional models used in signaling theory. In typical models of signaling theory, there is a direct association between a cost of producing signals and a quality of internal characteristics being signaled. Because visibility of a signal does not affect the cost of producing the signal in these models, visibility has no direct relationship with the cost-quality association. Thus, visibility can be safely ignored in analysis based on typical models used in signaling theory. While

Ramaswami et al treated signal visibility as the key explanatory factor in their analysis of relationships between mentoring and career success (2010), their analysis was only loosely based upon the core theoretical characteristics of the signaling theory. Here, I suggest that visibility of the signal cannot be ignored in analyses based on the intrinsic benefit model (or, Henrich's model). In the intrinsic model, effectiveness of a signal hinges on the existence of the perceived association between an internal characteristic being signaled and an intrinsic benefit obtained from producing the signal. Because there is no direct association between cost of producing a signal and an internal characteristic being signaled, factors which may reduce or remove the perceived association would reduce the effectiveness of the signal. Thus, if visibility of a signal influences how the perceiver interprets intention behind the signal, it can in turn affect the signal's effect.

For the purpose of this model, visibility corresponds with a/A . A cultural act with high visibility would be able to access more members of the audience than a cultural act with low visibility. High visibility will increase the expected signaling benefit of the cultural act, thus the actor should be willing to pay more to engage in a cultural act with high visibility than in a cultural act with low visibility. Many successful mass philanthropic campaigns utilize highly visible tokens – such as pink ribbons in the breast cancer awareness campaign, or fake poppies on Remembrance Day in Canada – which again show the importance of visibility for signaling. On the other hand, Hypothesis#1b also suggests that high-visibility signals may have potential to have counter-productive effects; high visibility of signals can make the sender's motivation for signaling benefits to be more salient (as discussed in proposition #1b from the previous chapter), which may result in the perceiver to discount the amount of the intrinsic benefit contributing toward

production of the signal. For example, consumers engaged in conspicuous brand consumption are perceived to be motivated for ulterior motive (Ferraro, Kirmani, and Matherly, 2013). This, in effect, would place a strategic actor who wishes to maximize his/her benefit from signaling in a conundrum; s/he would strive to maximize visibility of his/her signal, while at the same time minimizing the impression that s/he is doing so for ulterior motive (obtaining benefits from successful signaling). A common observation is that many of those with high status prefer to use subtle signals to signal their high status, instead of using highly visible status display. For example, those who are “in the know” tend to prefer brands with more subtle cues (Berger and Ward, 2007). Such behaviors can be seen as a way to fine-tune characteristics of a signal so that it would be sufficiently visible to its audience without making its underlying ulterior motive salient. A signal based on subtle cues that are only recognizable to small numbers of people who are “in the know” (and are willing to pay attention) would be unlikely to arouse the suspicion of ulterior motive toward obtaining benefits from successful signaling. Assuming that those who are “in the know” overlap with the intended audience of the signal, the signal can be highly effective.

In addition, there is a potential interesting consequence of the need for fine-tuning visibility to minimize perceived signaling benefit. There are some potential signals which are prescribed to be highly visible. For example, if a Muslim woman wishes to signal her religiosity by following the religious restriction of wearing a hijab outside of home, her signal will automatically be highly visible. However, she can avoid potentially being perceived as driven by ulterior motive despite high visibility of her signal, as she could simply be following her religious obligation instead. Thus, signals which are normatively

prescribed to be highly visible can increase effectiveness, and in turn, may result in the widespread adaption of these signals.

Signaling of Pro-Social Qualities

The present study focuses on the signaling of pro-social/altruistic qualities. The image motive (or, signaling motive) – the desire to improve one’s social standing by improving how one is being regarded by others – is commonly recognized as one of the main motivations behind prosocial behaviors (Ariely, Bracha, & Meyer, 2009). The individuals are more likely to make charitable contributions when their contributions are made public, and they prefer their donation to be public if they have the option (Andreoni and Petrie, 2002). Reputational benefit is proposed to be one of the main solutions for the private provision of public goods (Willer, 2009). The signaling theory is a framework which is frequently used across disciplines to explain this phenomenon, as it provides an effective means to provide explanations of pro-social behaviors which is based on the rational-choice paradigm. Traditionally, researchers who utilize the rational choice paradigm tended to regard reciprocity as the driving factor behind prosocial behaviors (most famously, Axelrod (2009)). However, there are many examples of prosocial behaviors which are unconditional, which cannot be explained by reciprocity alone. The signaling theory can provide a different rational-choice based explanation for prosocial behaviors.

Some research – for example, Smith and Bird (2000) in evolutionally anthropology, and Glazer and Conrad (1996) in economics – treat pro-social behaviors as the signal of one’s resource or fitness. They tend to focus on conspicuous generosities – costly, highly

visible acts of generosity – and argue that such generosity could only be sustained by those with large amount of resource or fitness. Therefore, their arguments share the same structure as that of Zahavi's handicap model. However, I suggest that the models treating pro-social behaviors as the signal of one's resource/fitness is not satisfactory for sociological analysis of pro-social behaviors. Pro-social behaviors take many forms which occur under varying social contexts; not all pro-social behaviors are costly enough to be considered as conspicuous generosity, and it is unlikely that the signaling of resource (or fitness) can account for the large proportion of pro-social behaviors. Indeed, Grose (2012) criticized that the overemphasis on costly signaling (as with Zahavi's handicap model) has stunted the applications of the signaling theory.

In addition, the signaling of resource argument also fails to provide sufficient explanation for what makes pro-social behaviors special. We can note that the signaling of resource argument is in fact very similar to Veblen's classic theory of conspicuous consumption (1899) – in Veblen's argument, conspicuous consumption can signal one's resource because only those with large amount of resource to be wasted can engage in conspicuous consumption. Intuitively, it can be argued that conspicuous generosity must be theoretically distinctive from conspicuous consumption; a billionaire who donates the large amount of money to a charitable cause would be perceived differently from a billionaire who purchased a very expensive mansion. However, the cost of signaling alone cannot explain the difference between these two examples. Indeed, if we take the simple handicap model of signaling of resource, it can be argued that any wasteful behaviors would be able to function as the signal of one's resource. However, it is likely that many sociologists would find this claim to be absurd. One of the key points of Bourdieu's classic

argument on cultural capital was that conspicuous consumption is not sufficient as a signal by itself – instead, conspicuous consumption must be done in the “right” way to work as an effective signal (1984). The signaling of resource model would be unable to account for such observations.

Instead of treating pro-social behaviors as the signal of one’s resource, pro-social behaviors can be understood as the signal of underlying pro-social quality. This quality is called social preference within the rational choice paradigm. The concept of social preference was developed to account for the idea that the individuals are not strictly motivated by their material self-interests, but also are motivated by their concern for material outcomes of others (Fehr and Fischbacher, 2002). It can be understood as the blanket term encompassing one’s “selfless” tendencies, including concern for fairness and aversion to inequality. In many social interactions, especially the ones which are fraught with risk, it is very desirable to correctly determine social preference of others. Those with high level of social preference are more likely to treat you fairly, and less likely to exploit one’s trust (for example, by defecting in prisoner’s dilemma-type situations). In turn, this would make it highly desirable to signal one’s social preference. With this reason, Fehler and Przepirka (2013) have argued that the charitable giving can function as a signal of social preference, and having social preference can be evolutionary advantageous.

In addition, signaling of social preference would be highly relevant to the intrinsic benefit model because the concept of social preference is relevant to intrinsic benefit. The concept of social preference was developed to explain the behaviors which cannot be explained by the narrow self-interest. In practice, it assumes that those with high level of

social preference would experience utility gain from certain behaviors, such as reciprocating favors or avoiding inequality. Although the theory of social preference does not specify where this utility gain comes from, it would be reasonable to assume that the underlying process behind the effect of social preference is similar to the notion of the intrinsic benefit obtained from prosocial acts. In addition, the theory behind social preference assume that each individual possesses different levels of social preference (Fehr and Fischbacher, 2002), making it fit with the scope of the signaling theory.

Indeed, intrinsic benefit itself could be seen as the core component of the research of altruism. Within psychology, the intrinsic motivation is one of the key motivations behind pro-social behaviors, particularly the behaviors such as charitable giving and volunteering. Typically, research on intrinsic motivations tend to be concerned about their interactions with other motivations. For example, Baston et al (1978) suggested that the reliance on extrinsic motivation to solicit prosocial behaviors may result in the “crowding out” of the intrinsic motivation. In economics, the concept of “warm glow” is used in economics to describe how the motivation to feel good about giving can result in donation to public goods (Andreoni, 1990). Although warm glow is identified as being distinct from pure altruism, note that pure altruism here has the very narrow definition: altruism which does not care at all about “feeling good” from giving. Therefore, the arguments on warm glow does not contradict our argument on the association between social preference and intrinsic benefits being obtained from prosocial behaviors. The existence of warm glow giving is experimentally supported (Crumpler and Grossman, 2008); I argue that this adds additional supports for the importance of intrinsic benefit within research of altruism. With these reasons, the signaling of social preference is an appropriate topic for the

research on the intrinsic benefit model, as it has the potential to benefit both the research on the signaling theory and the research on altruism.

Conclusion

In this chapter I outlined the formal properties of the intrinsic benefit model. I first articulates the assumptions and scope conditions of the model, then I derived propositions based both on the signaling theory and on this study's emphasis on the (perceived) intrinsic benefit. Finally, I described the six hypotheses that I will test in this study.

In the next two chapters I present the results from three experiments. The first two experiments were conducted online using the vignette design; they tested how the purchase of electronic vehicle can function as the signal of environmental commitment, and the effects of the signal visibility in the process of signaling. The third experiment was conducted in the laboratory; it was concerned with whether the donation can function as the signal of trustworthiness, and how the signaling process is affected by the perceived intentionality of the signal.

Chapter 4. Experiments #1 and #2

The first two experiments focus on the commitment to the environment as the internal characteristic of the interest. Commitment to the environment – which can be defined as one’s willingness to engage in pro-environmental behaviors – has several desirable qualities as a test quality for this experiment. First, there is considerable interest in environmental problems both among researchers and among laypeople. Not only that it is a quality that is relevant to many within our society (including the potential subjects of the experiment), but also applying the signaling model to environmental commitment can be useful for many researchers studying solutions to the environmental issues. More importantly, commitment to the environment would be a highly desirable internal characteristic. As argued by Griskevicius (2010), commitment to the environment can be seen as a form of altruism: one person’s environmentally friendly behavior can potentially benefit everyone in the world. Altruism (or, cooperation), in turn, is theorized to be one of the most desirable qualities for signaling (Gintis, Smith, and Bowles, 2002). Therefore, commitment to the environment is a characteristic that is very suitable for the signaling model. Visibility is the key treatment variable for the experiments; the increased visibility can be represented by the larger value of a/A in the model, affecting the signaling benefit of a cultural act. The experiments focus on testing the propositions made by the model by manipulating visibility. In particular, the experiment will focus on the purchase of an electric vehicle (EV). Automobiles are strongly associated with status, signals, and conspicuous consumption (Heffetz, 2011), making them a natural choice of medium for signaling. EVs are receiving a large amount of media attention recently, and it should be relevant for many research subjects.

The two experiments were conducted on the subjects recruited through the popular crowdsourcing website (Amazon Mechanical Turk). The crowdsourcing service provides a more diverse population than it is available in traditional laboratory experiments. The participants were paid 30 cents for their participation, which is higher than the typical payment at the crowdsourcing website. Only users who currently reside in the United States were recruited. After recruited through the crowdsourcing website, the subjects are taken to the separate website where the vignette experiments were conducted. After the experiments, the subjects who submitted the answers that are clearly erroneous were replaced with the new subjects; in both experiments, the number of the replaced subjects was less than 3% of the total subjects.

Experiment #1

The first experiment tests hypotheses #1a and #1b, and it uses 2x2 design where the signaling cost (high/low) is crossed against the visibility of the signal (high/low). Note that the purpose of visibility here is to manipulate the perceived extrinsic motive; following Proposition #2b, cultural acts with high visibility are more likely to be perceived as being motivated by non-intrinsic motive. In turn, this will reduce the perceived level of the internal characteristics, as cultural acts are no longer seen as being motivated by intrinsic benefits alone. There were 60 subjects per treatment condition, for a total N of 240. The subjects were randomly assigned into one of four conditions. In the high visibility conditions, the focal actor in the vignette has bought an EV which has the distinctive appearance and is easily recognizable as an EV. In the low visibility conditions, the focal

actor has bought an EV which is visually indistinguishable from the popular model that is using a gasoline engine. In the high cost conditions, the focal actor has paid twice the cost of the comparable gasoline-engine car. In the low-cost conditions, the focal actor has only paid the same cost as the comparable gasoline-engine car.

After reading the vignette, the subjects were asked six questions evaluating the environmental commitment of the focal actor on the vignette. Each item in this compound scale has Cronbach's Alpha above 0.8, showing that it has at least moderate reliability. Finally, the subjects were asked a question about their evaluation of the motivation of the focal actor – whether they think the focal actor was motivated by the signaling benefit or not.

Table 1 shows the mean values of perceived environmental commitment and perceived signaling motive. The result from the analysis of covariance with perceived environmental commitment as the dependent variable is shown in the Table 2. Both the effect of visibility and the effect of perceived signaling motive are statistically significant at .01 level. Overall, the effects of the treatment variables are consistent with the predictions. Low visibility results in higher perceived environmental commitment, and high cost leads for higher perceived environmental commitment. Therefore, both hypothesis#1a and #1b are supported. The effect of cost and the interaction term are statistically significant only at .1 level. Table 3 shows the result from the analysis of covariance with perceived signaling motive as the dependent variable. Both cost (at 0.05 level) and visibility (at 0.001 level) are statistically significant, though the interaction term

is not. Again, the effects of the treatment variables are consistent with the intrinsic benefit model; both high visibility and high cost lead for the higher perceived signaling motive.

(Insert Tables 1, 2, 3 here)

Table 4 shows the result of path analysis, with perceived signaling motive as the mediating variable and perceived environmental commitment as the dependent variable; the interaction term of cost and visibility is omitted from the analysis. Both high cost and high visibility have positive, statistically significant effect on perceived signaling motive; the effect of visibility is far larger, suggesting that visibility impacts the perceived signaling motive more than cost does. Perceived signaling motive, in turn, negatively affects perceived environmental commitment, as predicted by the model. Therefore, the effect of visibility on perceived environmental commitment is at least partially mediated by perceived signaling motive, consistent with the intrinsic benefit model. The visibility also has the direct, negative effect on perceived environmental commitment, though this effect is statistically significant only at .1 level.

(Insert Table 4 here)

Finally, it can be noted that the direct and the indirect effect of cost somewhat cancel each other out. The direct effect of cost is positive, suggesting that high signaling cost is being perceived as the signal of increased environmental commitment. However, the indirect effect is negative; when the high signaling cost is perceived as the signal of the signaling motivation, the same high cost is interpreted as the signal of reduced environmental commitment. I argue that the contradicting effects of the direct effect and the indirect effect show the complexity inherent with the situations where effects of a

signal are contingent on perception of the receiver of the signal. With simple signaling models based solely on signaling costs, signals which are costly to produce usually are reliable indicators of underlying characteristics. Thus, a sender's decision on whether to produce a signal or not comes down to how easily the sender can afford the cost to produce the signal. However, when the effect of a signal is determined by the receiver's perception of the signal – the receiver's perception of the sender's intention for producing the signal, in particular – the decision would involve more complex strategic calculations to maximize the benefit of the signal while minimizing the potentially counterproductive impression the signal can make.

Overall, the result of experiment#1 is consistent with the intrinsic benefit model. The result supports the predictions based on intrinsic benefit model. In addition, it shows the importance of the receiver's perception in the effect of the signal.

Experiment #2

The second experiment tests hypotheses#2a and #2b. As with the experiment#1, experiment#2 is a vignette experiment conducted on the subjects from the crowdsourcing website. It includes one treatment variable and two non-manipulated independent variables. The visibility of the signal (high/low) is manipulated, while the subjects' attitudes toward the environment and the subjects' motivation for signaling are the non-manipulated independent variables. Again, the purpose of visibility here is to manipulate the signaling benefit; following proposition#1b, an act with higher visibility would result in more signaling benefit. On the other hand, the model does not specify the relationship

between the internal characteristic and the desire for signaling; therefore, the subjects' desire for signaling was measured separately. Note that because the experiment does not manipulate the subjects' motivation for signaling, the potential association between the signaling motivation and the characteristic being signaled can be only controlled statistically; this complicates the analysis of the results.

There are 141 subjects per treatment condition, for a total N of 282. The subjects were randomly assigned into one of two conditions. As the treatment, the subjects were asked to think as if they need to choose between two cars: a conventional gasoline car and an EV. In the high-visibility condition, they were told that although the EV has similar size and luxury as the gasoline car, it has a very distinctive look and has a different name from the gasoline car. In the low-visibility condition, they were told that the EV looks identical to its gasoline counterpart, and the two cars share the same name. The dependent variable is how much they are willing to pay for the EV, where the subjects have eight options (less than \$20000, \$20000, \$21000, \$25000, \$30000, \$40000, \$45000, \$50000 or more) to select from. After answering the question, the subjects were made to answer the bogus question. Afterward, the subjects were asked to answer the two set of questionnaires: six questions to measure their level of environmental commitment, and two questions to measure their desire for signaling. All questions use a 7-point Likert scale, and the actual texts from the questionnaire are available in the appendix. All items for the environmental commitment have Cronbach's Alpha above 0.8. Two items for the signaling motive have correlation of 0.56. For the dependent variable, no subjects selected \$45000 and only one subject selected \$50000 or more, therefore three highest categories are collapsed into a single category.

Ordinal logit regression was used to analyze the result of experiment#2. The cost the subjects were willing to pay for the hypothetical EV is the dependent variable, whereas visibility, the subjects' environmental commitment, and the subjects' signaling motive are independent variables. Visibility is the treatment variable, while environmental commitment and signaling motive are the covariates. The result of ordinal logit regression is reported on Table 5. The model includes three independent variables and the interaction term of visibility and signaling motive. The interaction term of visibility and environmental commitment was found to have no effect, and was omitted from the final model. Both visibility and environmental commitment are found to have the positive effect; the subjects in high visibility conditions are more likely to be in the categories of higher rank (therefore, willing to pay more for the hypothetical EV), and the subjects with higher environmental commitment are more likely to be in the categories of higher rank. These findings are consistent with the intrinsic benefit model, and hypothesis#2a is supported here. The main effect of signaling motive is not significant. However, the interaction effect of signaling motive and visibility is negative, and its significance level is close to .05; the subjects with higher signaling motive are less likely to be in the categories of higher rank when the visibility is high. Although this effect is not significant at .05 level, the result nevertheless does not support hypothesis#2b.

(Insert Table 5 here)

Although the interaction term does not support hypothesis#2b, this finding actually is not inconsistent with the intrinsic benefit model. As I have discussed in the earlier chapter, the intrinsic benefit model suggests that the signals with high visibility are less

likely to be perceived as the reliable signal of internal characteristic – the proposition is supported by the result of the experiment#1. Then, for the purpose of signaling, it would be more beneficial to use the signal which has low visibility. If we interpret high signaling motive as having more motivation for behaving more strategically (or, being more aware of the possible negative effects of high visibility signals), then it is plausible that those with high signaling motive are less willing to pay extra cost for high-visibility signals; in this case, they might be more concerned about being seen as “poser” due to their use of the high visibility signal. Such behaviors may not be uncommon; as I have mentioned in the previous chapter, individuals with more cultural capital in a particular domain tend to prefer products with more subtle cues (Berger & Ward, 2010). As it was explained in the previous chapter, the use of a subtle signal which is only recognizable to those “in the know” would be one of the ways to minimize the signaling motivation that is being signaled by the signal. Unfortunately, the design of the present experiment does not allow for validating potential explanations for the findings. In the later chapter, I will discuss about potential approaches for the sender to minimize the perceived signaling motivation, and potential research methodologies to explore them.

Overall, the finding generally supports the intrinsic benefit model. In addition, it suggests that individuals may be capable of manipulating their signaling behaviors to reduce the signaling motive that is perceived by others. Although the finding is not entirely consistent with the prediction, it appears that the finding is consistent with the reasoning behind the intrinsic benefit model.

Table 1: Experiment#1: Mean Perceived Environmental Commitment (PEC) and Perceived Signaling Motive (PSM)

	PEC	PSM
High Cost*High Visibility	5.23	5.51
Low Cost*High Visibility	5.24	5.02
High Cost*Low Visibility	5.78	4.31
Low Cost*Low Visibility	5.39	3.95

Table 2: Experiment#1: Analysis of Covariance on Perceived Environmental Commitment

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Cost (High)	1	2.30	2.30	2.81	0.095 .
Visibility (High)	1	7.16	7.16	8.74	0.003 **
PSM	1	6.16	6.16	7.52	0.007 **
Cost*Visibility	1	2.25	2.25	2.74	0.099 .
Residuals	235	192.46	0.82		

Table 3: Experiment#1: Analysis of Variance on Perceived Signaling Motive

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Cost (High)	1	9.8	9.85	5.08	0.025 *
Visibility (High)	1	77.3	77.31	39.86	<0.001 ***
Cost*Visibility	1	0.3	0.27	0.14	0.709
Residuals	236	457.8	1.94		

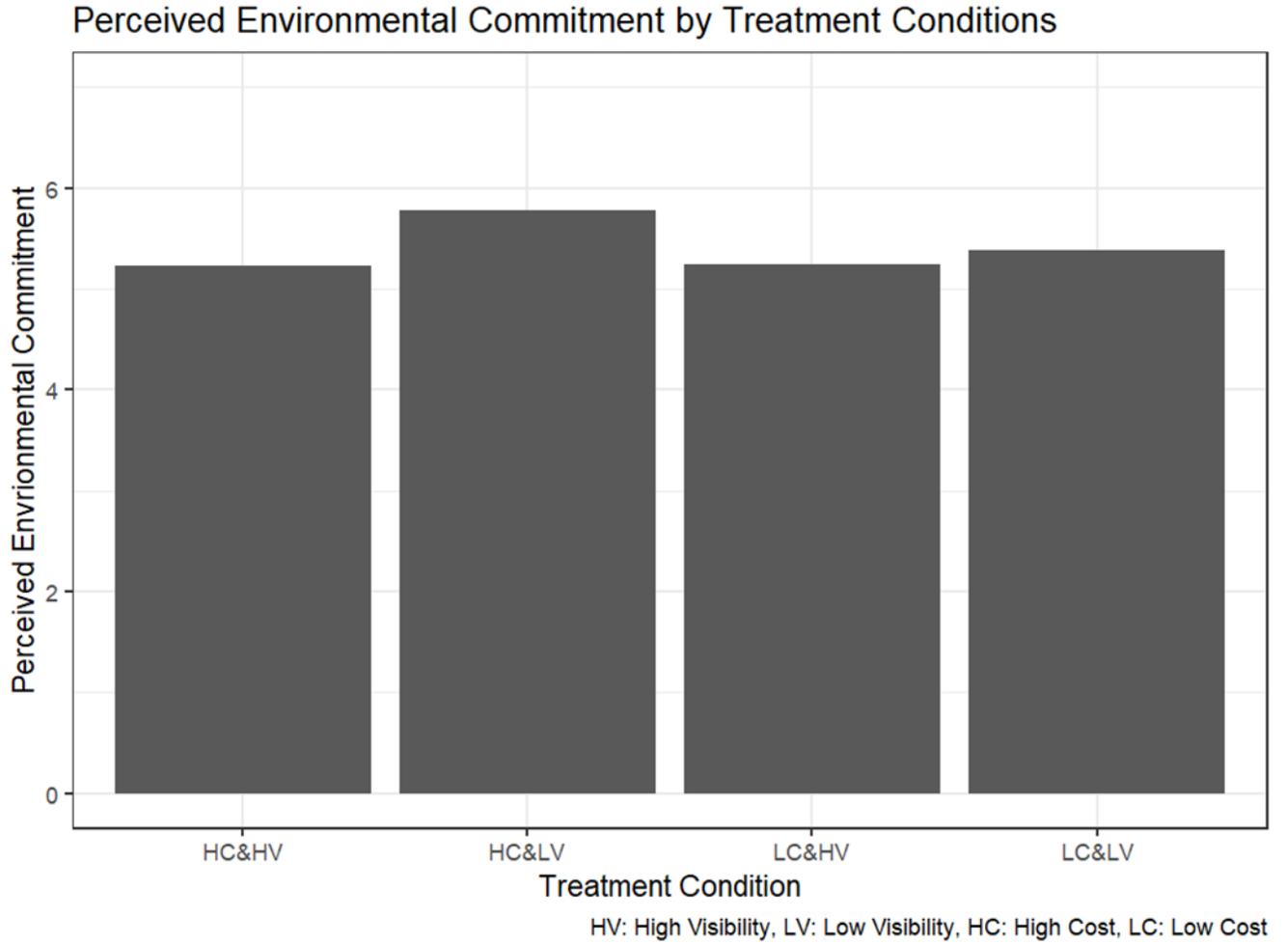
Table 4: Experiment#1: Path Analysis

	Estimate	SE	Z-value	P(> z)
PSM				
Cost (High)	0.424	0.178	2.377	0.017 *
Visibility (High)	1.135	0.178	6.365	<0.001 ***
PEC				
Cost (High)	0.239	0.118	2.034	0.042 *
Visibility (High)	-0.214	0.126	-1.701	0.089 .
PSM	-0.116	0.042	-2.755	0.006 **

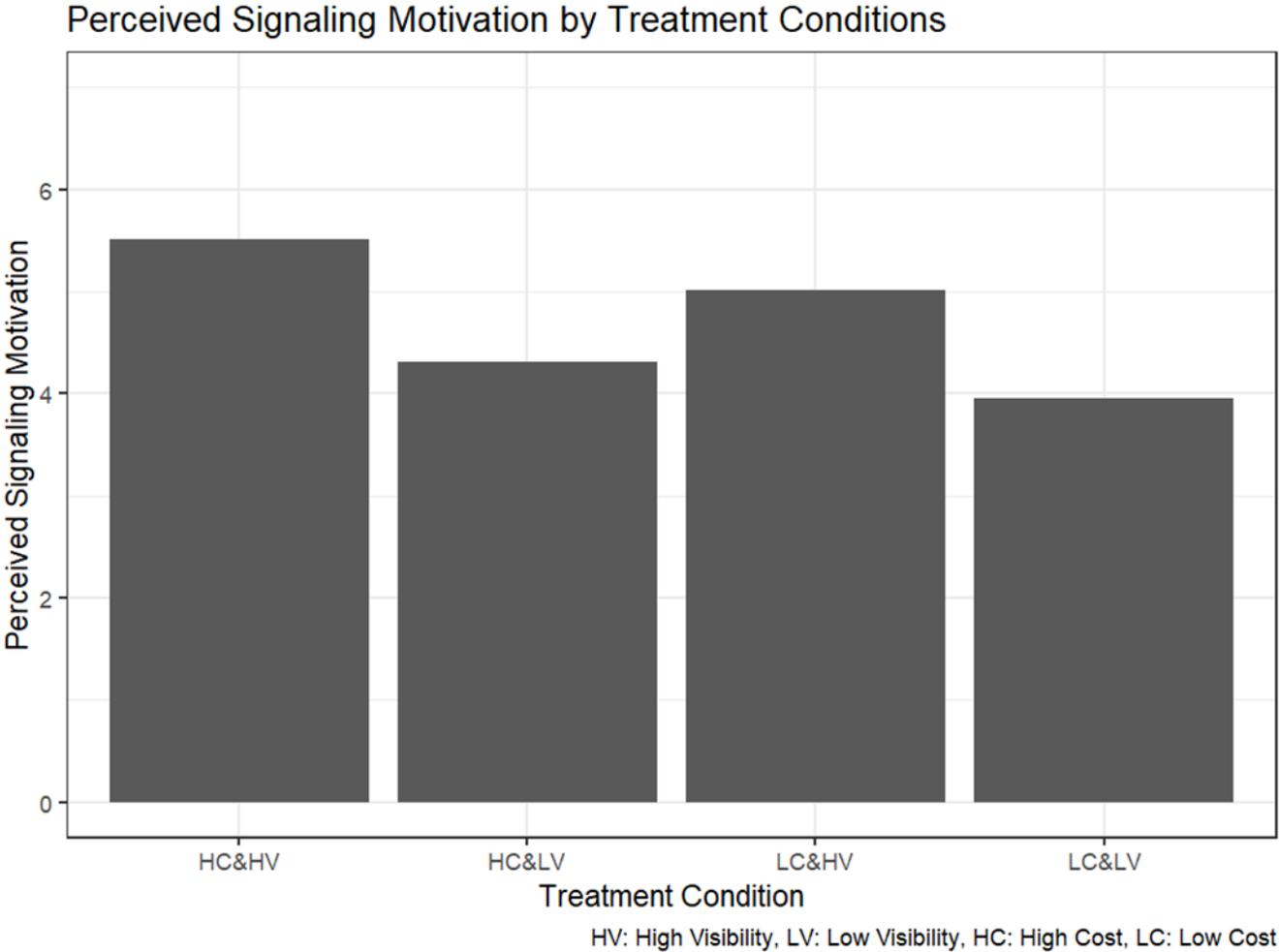
Table 5: Experiment#2: Ordinal Logistic Regression

	Estimate	SE	t-value	p value
Environmental Commitment	0.243	0.122	1.987	0.047 *
Visibility	1.619	0.758	2.136	0.033 *
Signaling Motive	0.189	0.116	1.621	0.105
Visibility*Signaling Motive	-0.308	0.160	-1.923	0.055 .
Less than \$20K \$20K	0.159	0.639	0.249	0.803
\$20K \$21K	1.202	0.634	1.897	0.058
\$21K \$25K	1.600	0.637	2.512	0.012
\$25K \$30K	3.755	0.672	5.588	<0.001
\$30K \$40K	5.188	0.711	7.291	<0.001
\$40K More than \$40K	6.328	0.801	7.899	<0.001

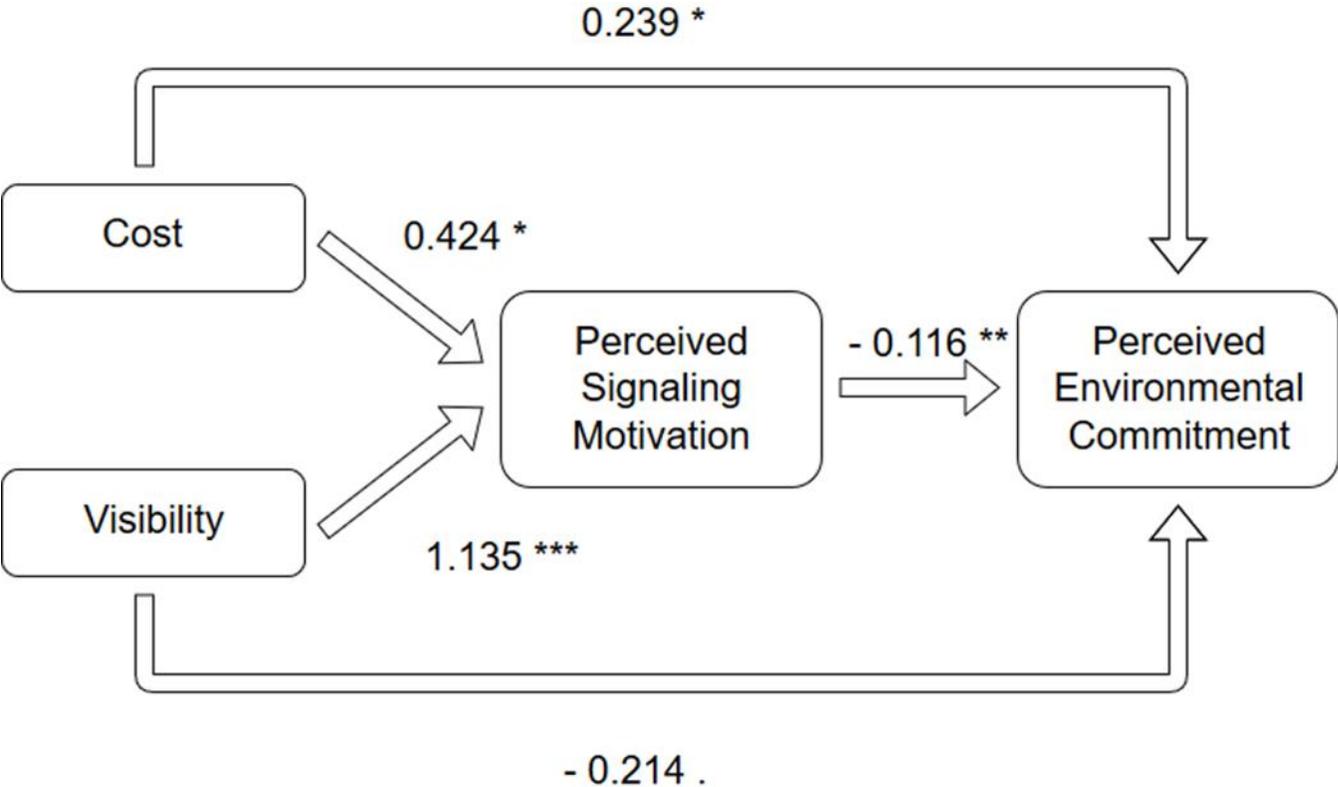
Graph#1: Perceived environmental commitment by treatment condition



Graph#2: Perceived signaling motive by treatment conditions



Graph#3: Path Analysis of the direct and indirect effects of cost and visibility on perceived environmental commitment



Chapter 5 Experiment#3

The third experiment has three goals. First, it will test the model with a different internal characteristic. Second, it aims to establish that a signal actually influences behaviors of the receivers (hypothesis#1c). Finally, the signal in this experiment will be actual behaviors, instead of hypothetical behaviors in vignettes. The central idea of signaling theory is that actions (with costs) speak louder than words (Henrich, 2009); therefore, an experiment based on actual behaviors would provide the stronger evidence for the model than the vignette experiments based on hypothetical behaviors alone. The experiment#3 applies the signaling model to trustworthiness, using behaviors in the trust-game to measure the perceived trustworthiness of an individual. In the trust game, there are two actors. The first actor decides whether to trust the second actor and transfer him/her with his/her own money. If the first actor transfers his/her credits, the researcher will increase this transfer by a set factor. Then, the second actor can decide how to split the resulting credit between him/her and the first actor. From the perspective of the first actor, s/he could get the best possible outcome when s/he gave the transfer to the second actor, and the second actor returns the favor by making (at least) a fair split. Therefore, detecting partner's trustworthiness is crucial for the first actor in the trust game. Tan and Vogel (2008) applied the signaling model of religion to the trust game, and found that religiosity can function as the signal of trustworthiness.

The experiment uses charitable giving as the signal of trustworthiness. Altruistic behaviors can serve as the signal of one's greater social preference – the psychological disposition to behave cooperatively (Feheler and Przepiorka, 2013). From the intrinsic

benefit argument, we can argue that those who have a higher level of social preference should obtain more intrinsic reward from making donations than those who are less trustworthy. It is well-established that cooperation in public good games can function as the signal of commitment to the group (for example, in Willer, 2009). In these situations, the receivers of the signal are directly benefiting from the public good. However, this is not the case with donations to charities, as typically donations are made to the benefit of third parties. Therefore, if someone donates to a charity, the potential benefit s/he can obtain from the donation must depend on the effect of signaling. The relationship between charitable giving and trustworthiness has been experimentally studied by Albert et al (2007) and Feheler and Przepiorka (2013). Both studies found that those who make charitable gifts tend to be more trustworthy, and are more likely to be trusted by others. Although Albert et al (2007) framed their finding in terms of the spillover across domains, the argument by Feheler and Przepiorka (2013) is based on the signaling theory. The experiment being proposed here extends these studies by adding the conditions of perceived signaling motivation, which are derived from the model being proposed here. This will be achieved through manipulating the prior knowledge of potential benefits from signaling (donation to a charity organization).

The experiment was conducted in the laboratory at the business school of a large public university in Southwestern United States. The subjects were the undergraduate students who registered as the members of the subject pool at the laboratory; typically, they are the students of business or related fields. The experiment was conducted through computer terminals connected as a network, and was programmed using Z-tree software (Fischbacher, 2007). Z-tree is originally created for conducting computerized experiments

in behavioral economics, and it is a popular software for social scientists who conduct computerized behavioral experiments on subjects connected on a local network.

The experiment#3 consists of two components (experiment#3a and experiment#3b) which shared the same goals – testing the hypotheses derived from the intrinsic benefit model with laboratory experiments – and the same treatment conditions. In the first half of the experiment#3, only experiment#3a was conducted. On the second half (which includes #3b) of the experiment#3, both experiment#3a and experiment#3b were conducted in a succession. Experiment#3b was conducted immediately after the subjects completed the interactions for experiment#3a, before the subjects know the result from the experiment#3a; thus, the outcomes of the experiment#3a did not affect the outcomes of the experiment#3b. With this reason, the results from the two halves of experiment#3a are combined in this analysis. A supplementary analysis conducted on the results from the first half of experiment#3a (the half which was conducted without experiment#3b) and the second half of experiment#3a (the half which was conducted together with experiment#3b) showed no difference compared to the analysis based on the combined data using the both halves of the experiment#3a. The random assignment of treatment conditions is done with the basis of batch. The majority of batches have mixed-gendered group of subjects numbering between 6 and 12, although several batches from the first half of the experiment#3 (where only the experiment#3a was conducted) contained more than 12 subjects. All batches from the second half of experiment contained the subjects numbering between 6 and 12.

At the beginning of the experiment, the subjects were asked to read and sign the informed consent form at the meeting room of the laboratory. Afterward, each subject was asked to move into the main laboratory room and sit in front of one of the computer terminals. In this experiment, the subjects completed their interaction through their computer terminals.

In all treatment conditions, the subjects were asked to decide on donating \$2 (from their payment) to one of the four charities after they were seated at their terminals. In the no-signaling condition and the no-signaling-motivation condition, the terminal displayed the prompt to the subjects to make the decision immediately after they seated in their terminals, and before they received the instruction for the main part of the experiment. The list of charities that the subjects could select from included Feeding America, American Cancer Society, UNICEF, and Smithsonian Institute. The prompt they saw on their terminals was the following.

Before the main phase of the experiment, we ask you to decide on the following.

You have \$2, separate from the money that will be used for the main part of the experiment -- you can either donate it to one of the three charities, or you can keep it yourself. If you decide to donate, I will donate \$2 to the charity of your choice under the name of "The experiment participants from the University of Arizona". If you decide to keep it yourself, \$2 will be added to your payout at the end of the experiment. What is your decision?

In the signaling-motivation condition, the subjects received the instruction sheet immediately after they were seated at their terminals. The sheet contains the detailed

description of the interactions they were going to engage in; although the sheet does not use the word “trust game” or “trust”, the subjects knew that it is in their benefit if their future partner decides to trust them. After they read the instruction sheet, the terminals displayed the prompt to the subjects to make the decision on donation. This prompt emphasized that their decision of their decision will be visible to their future interaction partners. The prompt they saw on their terminals was the following.

Before the main phase of the experiment, we ask you to decide on the following.

You have \$2 worth of experimental credits, separate from the money that will be used for the main part of the experiment -- you can either donate it to one of the three charities, or you can keep it yourself. If you decide to donate, I will donate \$2 to the charity of your choice under the name of "The experiment participants from the University of Arizona". If you decide to keep it yourself, \$2 will be added to your payout at the end of the experiment. Note that your partner will be able to see whether you made a donation or not, though your partner will not be able to see which charity you donated to. What is your decision?

In the no-signaling condition and the no-signaling-motivation condition, the experimenter distributed the instruction sheet after the subjects decided on donations. The instruction sheets are identical across the treatment conditions. A copy of the instruction sheet is available in the appendix.

The main part of the experiment#3a began immediately after the subjects had read the instruction and had decided on the donation. The program then randomly assigned the subjects into pairs, with each pair containing the first actor and the second actor. The first

actor was then asked to decide on their contribution; s/he could select from amounts between \$0 and \$5 in \$1 interval. In the no-signal condition, the prompt does not mention donation at all – the text from the prompt was the following.

Your partner is a fellow student at the University of Arizona, and was recruited for this experiment in the same way as you were.

Please decide how much you want to offer. You may offer up to \$5.

As you know, an offer of \$1 will result in the total pool of \$4 for your partner to split.

In the no-signaling-motivation condition, the prompt mentions whether the second actor (the partner of the first actor) made a donation or not. In addition, the prompt emphasized that the second actor did not know the information about his/her donation would be available for the experiment – the text from the prompt was the following. The text in bracket indicates the text for the partner who did not make donation, which replaces the sentence above it.

Your partner is a fellow student at the University of Arizona, and was recruited for this experiment in the same way as you were. You are randomly paired with your partner. In the beginning of the experiment, your partner was asked to make the same decision as you were; whether s/he would like to donate \$2 to one of the three charities. Your partner was unaware that this information would be available to his/her future interaction partner. Your partner decided to donate \$2 to one of the charities. (Your partner decided not to donate to any of the charities).

Please decide how much you want to offer. You may offer up to \$5.

As you know, an offer of \$1 will result in the total pool of \$4 for your partner to split.

In the signaling-motivation condition, the prompt mention whether the second actor made a donation or not. In addition, the prompt emphasized that the second actor was aware that the information about his/her donation would be available for the experiment – the text from the prompt was the following. The text in bracket indicates the text for the partner who did not make donation, which replaces the sentence above it.

Your partner is a fellow student at the University of Arizona, and was recruited for this experiment in the same way as you were. You are randomly paired with your partner. In the beginning of the experiment, your partner was asked to make the same decision as you were; whether s/he would like to donate \$2 to one of the three charities. Your partner was aware that this information would be available to the future interaction partner. Your partner decided to donate \$2 to one of the charities. (Your partner decided not to donate to any of the charities).

Please decide how much you want to offer. You may offer up to \$5, As you know, an offer of \$1 will result in the total pool of \$4 for your partner to split.

If the first actor decides to make a transfer, the researcher multiplies the amount by 4 to calculate the pool for the second actor to split. While multiplier of 3 is popular in the trust game experiments, the multiplier of 4 is used here to make successful signaling more appealing for those in the signaling motivation treatment. The first actor kept the remaining amount, which is \$5 subtracted by the amount of their contribution. If the first actor decided to make the contribution, the second actor is then asked to decide on how much they want to return from their pool (four times the amount contributed by the first

actor). After the second actor made the decision, the subjects were asked brief questions about their partner. The subjects were asked the following question.

While the system is processing the result, please answer how much you agree with the following statement.

"I think I can trust my partner"

The question used 5-points Likert scale. In the batches of experiments which only included the experiment#3a, the terminals then displayed the subjects' earnings, and the experiment ended there. In the batches which included the experiment#3b, the subjects did not learn about their earning yet, and the experiment instead proceeded to the second phase (experiment#3b).

The experiment#3b utilized the treatment conditions and donation information from the experiment#3a. Instead of the single-round experiment used in the experiment#3a, the experiment#3b utilized repeated trust-game interaction where the subjects were assigned a new partner and a new role randomly each period. To simplify the decision-making process for the subjects, both the first actor and the second actor were limited with binary choices. There were ten periods in the experiment#3b, and the subjects were paid for two periods which were randomly selected by the program. The repeated design allows us to increase the statistical power by ensuring that there is sufficient number of instances of trust game, which is impossible with the one-shot design of the experiment#3a. The experimenter handed out the instruction sheets at the beginning of the experiment#3b.

At the beginning of each period, the subjects who were assigned the role of the first actor were asked to decide whether to transfer \$2 to their partner or not. In the no-signal condition, the prompt does not mention donation at all – the text from the prompt was the following.

You are randomly paired with your partner for this round.

Please decide If you want to make a transfer of \$2 to your partner.

As you know, \$2 transfer will result in \$6 for your partner to split between you and your partner.

In the no-signaling-motivation condition, the prompt mentions whether the second actor (the partner of the first actor) made a donation or not. This study involved no deception, thus the subjects received accurate information on their partner's donation status. In addition, the prompt emphasized that the second actor did not know the information about his/her donation would be available for the experiment – the text from the prompt was the following. The text in brackets indicates the text for the partner who did not make donation, which replaces the sentence before it.

You are randomly paired with your partner for this round. In the beginning of the experiment, your partner was asked to make the same decision as you were; whether s/he would like to donate \$2 to one of the three charities. Your partner was unaware that this information would be available to his/her future interaction partner. Your partner decided to donate \$2 to one of the charities. (Your partner decided not to donate to any of the charities)

Please decide If you want to make a transfer of \$2 to your partner.

As you know, \$2 transfer will result in \$6 for your partner to split between you and your partner.

In the signaling-motivation condition, the prompt mention whether the second actor made a donation or not. In addition, the prompt emphasized that the second actor was aware that the information about his/her donation would be available for the experiment – the text from the prompt was the following. The text in bracket indicates the text for the partner who did not make donation, which replaces the sentence before it.

You are randomly paired with your partner for this round. In the beginning of the experiment, your partner was asked to make the same decision as you were; whether s/he would like to donate \$2 to one of the three charities. Your partner was aware that this information would be available to his/her future interaction partner. Your partner decided to donate \$2 to one of the charities. (Your partner decided not to donate to any of the charities)

Please decide If you want to make a transfer of \$2 to your partner.

As you know, \$2 transfer will result in \$6 for your partner to split between you and your partner.

If the first actor decided to make the transfer, the experimenter tripled this amount to \$6 to form the pool for the second actor to split. The multiplier of three was used in this part of the experiment because making successful signaling more appealing is not a relevant concern in this part of the experiment. The subjects who were assigned as the

second actor were asked to decide on whether they would split \$6 equally (both taking \$3), or to keep the entire \$6 by themselves. After the second actor had made this decision, the experiment proceeded to the next period, and the subjects were assigned new partners and new roles. Finally, the subjects learned their earning from the first part of the experiment (experiment#3a), which periods were selected for payments in the second part of the experiment (experiment#3b), their earning from the experiment#3b, and their total earning.

For the analysis of the experiment#3a, the results from the first half of the experiment (conducted without the experiment#3b) and the second half of the experiment (conducted with the experiment#3b) are combined. There were 164 subjects across three treatment conditions; 68 subjects were from the first half of the experiment, and 96 subjects were from the second half of the experiment. Of 164 subjects, 64 subjects made donation. Among the subjects who are assigned as the first actor, the majority of the subjects transferred some credits, with only 15 subjects not making any transfers. The average amount of transfer was \$2.8.

Table 6 shows the outcome of logistic regression where the treatment conditions are the independent variable, and donation is the dependent variable. Although there is a signaling-oriented incentive to donate in the signaling motivation condition, there was no statistically significant difference in the likelihood of making a donation made across the treatment conditions. This result contradicts the previous studies which found that people are more likely to make prosocial act (such as making a donation) in public than in private (Andreoni and Ragan, 2004; Soetevent, 2005; Ariely et al, 2009). However, I argue that this

result is consistent with the finding from the experiment#2, and show that the subjects were capable of behaving strategically in the manner consistent with the intrinsic benefit model. Because the subjects were aware that their donation could be interpreted as being motivated for soliciting transfer in the trust game, it is plausible that they thought that making a donation would not function as an effective signal of trustworthiness.

(Insert Table 6 here)

Table 7 shows the outcome of ordered logit regression, where the independent variable is the donation by the second actor, and the dependent variable is the transfer given by the first actor. This analysis was conducted separately for signaling-motivation condition and no signaling-motivation condition. Note that donation information is unavailable on the no-signaling condition, thus the result from the no-signaling condition is omitted. Contrary to the predictions, neither the no-signaling motivation condition nor the signaling motivation condition showed statistically significant difference in the amount of transfer made to the second-actors who made a donation and the second actors who did not make donations.

(Insert Table 7 here)

Table 8 shows the outcome of linear regression, where the independent variable is the donation by the second actor, and the dependent variable is the perception of trust of the second actor by the first actor. Consistent with the finding above, in both signaling motivation condition and no-signaling motivation condition, there is no statistically difference between the first-actors' perception of trust to donators and the first actors' trust to non-donator. This result contradicts the previous studies which found that people

are more likely to make prosocial act (such as making a donation) in public than in private (Andreoni and Ragan, 2004; Soetevent, 2005; Ariely et al, 2009). However, it is possible to argue that the finding shows that the subjects were capable of behaving strategically, which does not necessary contradict the underlying reasoning behind the intrinsic benefit model. Because the subjects were aware that their donation could be interpreted as being motivated for soliciting transfer in the trust game, it is plausible that they thought that making a donation would not function as an effective signal of trustworthiness. There is one previous research which gives support to this argument; Gambetta and Przepiorka (2014) found that the subjects were less likely to strategically cooperate in prisoner's dilemma when they were explicitly aware that their cooperation could be interpreted as being made for ulterior motive. Because the signaling motivation condition explicitly informed the subjects that their donation would be visible to their future partners, it is plausible to think that the subjects considered what they think their future partner would think. If they decided that their partner was likely to perceive the donation as being motivated for signaling, then it is likely that they would think that donation would not benefit them in the future trust game.

(Insert Table 8 here)

Table 9 shows the result of the ordered logit regression; the amount of transfer from the first or second (?) actor is the dependent variable, and the treatment conditions are the independent variables. There were no between-treatments difference in the amount of transfer from the first actor to the second actor. Similarly, there are no statistically

significant difference in perceived trustworthiness of the partner between treatment conditions (Table 10).

(Insert Tables 9 and 10 here)

Table 11 shows the effect of perceived trust on the amount of transfer from the first actor to the second actor. There is a strong association between trust and the amount of transfer, which is statistically significant at .001 level. This result indicates that the subjects based their decision on transfer based on perceived trustworthiness of their partners. However, the subjects did not use donation as the signal of trustworthiness.

(Insert Table 11 here)

Finally, an analysis on the second actors' behaviors on the experiment#3a was omitted; the sample size of the second actors who received transfer from the first actor is relatively low, especially considering that the model would include three treatment conditions in addition to one independent variable (the second actor's donation). In addition, varying amounts of transfers make it impossible to directly compare between the subjects who received different amount of transfers; thus, it is difficult to interpret results. The experiment#3b was designed to handle these shortcomings of experiment#3a.

There were 96 subjects in the experiment#3b, each playing 10 rounds of trust game; thus, there were 480 cases (or, pairs) for this analysis. 38 out of 96 subjects (39.6%) made a donation in the beginning of the experiment. The subjects made transfer in 276 out of 480 cases (57.5%).

The first analysis focuses on whether the subjects' donation status affect the likelihood of receiving transfer from their partner. For this analysis, separate analysis was applied to both the no signaling-motivation condition and the signaling motivation condition. A mixed-effect model was used to analyze the result, where a random effect was used to control for the baseline propensity of each subject to make transfer. Period effect was observed for the first period and the last period; the first period had noticeably higher rate of transfer, and the last period had the noticeably lower rate of transfer. The proportion of the subjects who made transfer each round is shown in Table 12. While the overall trend shows the gradual decline of the proportion of the subjects who made transfer, it fluctuated largely each period. Therefore, period effect was omitted from the subsequent analysis. A supplementary analysis was conducted using each period as a separate group, and the results were consistent with the analysis without the period effect. Another supplementary analysis was conducted with the period effect as a linear function of periods; one of the resulting models failed to converge.

(Insert Table 12 here)

Table 13 indicates the results of the analysis based on the mixed effect model for two treatment conditions (the no signaling condition is omitted); partner's donation is the independent variable. Both in the no signaling motivation condition and in the signaling motivation condition, partner's donation status does not have statistically significant effect on the amount of transfer. This result is contrary to the hypothesis, although it is consistent with the result of the experiment#3a. Again, this shows that the subjects do not appear to use charitable giving as a signal of trustworthiness.

(Insert Table 13 here)

Finally, Table 14 shows the comparison between three treatment conditions (with the no-signaling condition as the reference category) – partner’s donation was omitted from this analysis. The subjects in the signaling motivation condition had significantly lower likelihood of making transfer compared to the no-signaling condition. This finding is inconsistent with the finding from the experiment#3a, which showed no association between the treatment conditions and the amounts of endowment. However, the experiment#3b has larger sample size than the experiment#3a due to its repeated design; it may be the case that this leads to the difference in the results between two experiments. I argue that these findings suggest that while the subjects were not basing their decision on the donation status of their partner, they were being sensitive to the overall signaling environment itself. The environment which facilitates the strategic use of charitable giving as a signal decreased the level of trust among the subjects.

(Insert Table 14 here)

The second analysis focuses on trustworthiness of the subjects, which is operationalized as whether the subjects had returned credits to their partner after receiving transfer. For this analysis, 276 cases of second-actors who had received transfer from their partners were used for analysis. Out of 276 cases, the subjects returned the credit in 117 cases (42.4%). As we can see in Table15, the proportion of the subjects who returned the transfer tended to fluctuate, though it declined over time. Therefore, the period effect was omitted from the analysis as with the previous analysis; a supplementary analysis which conducted with the period effect did not alter the overall finding.

(Insert Table 15 here)

Table 16 shows that there was no statistically significant difference in the subjects' likelihood to return their transfers across the treatment conditions. The next table shows the result of analysis comparing donors and non-donors in each treatment condition. In the analysis which aggregated the subjects from three treatment conditions, donors are significantly more likely than non-donors to return their transfers, which confirms the prediction (Table 16). The next set of analysis is conducted on each treatment group separately. In both the no signaling condition and the no signaling motivation condition, there was no statistically significant differences in likelihood of returning transfer between donors and non-donors. In the signaling motivation condition, donors are significantly more likely than non-donors to return their transfers. Given that the donors are significantly more likely than non-donors to return their transfer in the subject pool which combined three treatment conditions, it is difficult to explain the lack of significant differences in the no-signaling condition and the no signaling motivation condition. Nevertheless, it is somewhat counterintuitive that donors in the signaling motivation condition were significantly more likely to return transfer, as the setup of the signaling motivation condition encouraged those without altruistic tendency to make donation for the potential gain in the trust game interaction.

(Insert Table 16 here)

The results of the experiment #3 were mixed. The donation was not perceived as the signal of trustworthiness by the subjects, and the incentive for successful signaling failed to induce the donation among the subjects. These two results are consistent with each other,

but they contradict the original predictions. On the other hand, the donation was found to be associated with trustworthiness, despite that the subjects did not perceive it to be so. Finally, it appears that the salience of signaling motive itself suppressed the overall level of trust among the subjects.

Table 6: Experiment#3a: donation by treatment conditions

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-0.49248	0.27061	-1.820	0.0688 .
Treatment(No Signaling Motivation)	-0.03815	0.39070	-0.098	0.9222
Treatment (Signaling Motivation)	0.26067	0.38884	0.670	0.5026

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 220.25 on 163 degrees of freedom

Residual deviance: 219.57 on 161 degrees of freedom

The no-signaling condition is the reference category.

Table 7: Experiment#3a: the effect of partner donation on transfer by treatment conditions

The no signaling condition is omitted from this analysis.

Treatment (no signaling motivation condition)

	Value	SE	t-value
Partner's donation	-0.2591	0.7076	-0.3662

Intercepts:

	Value	SE	t-value
0 1	-1.5838	0.5714	-2.7717
1 2	-0.7912	0.4902	-1.6141
2 3	-0.0159	0.4572	-0.0348
3 4	0.4416	0.4664	0.9467
4 5	1.1618	0.5240	2.2172

AIC: 107.3786

Treatment (signaling motivation condition)

	Value	SE	t value
Partner's donation	0.127	0.7417	0.1712

Intercepts:

	Value	SE	t value
0 1	-0.7704	0.4862	-1.5845
1 2	-0.5947	0.4775	-1.2455
2 3	-0.2675	0.4686	-0.5709
3 4	0.5133	0.4764	1.0775
4 5	1.7469	0.5979	2.9216

AIC: 96.6622

Table 8: Experiment#3a: the effect of partner donation on perceived trust, by treatment conditions

The no signaling condition was omitted from this analysis.

Treatment (no signaling motivation condition)

	Value	SE	t value
Partner's donation	-0.2591	0.7076	-0.3662

Treatment (signaling motivation condition)

	Value	SE	t value
Partner's donation	0.127	0.7417	0.1712

Table 9: Experiment#3a: the effect of treatment conditions on transfer

(No signaling condition is the reference category)

The effect of treatment conditions on endowment

	Value	SE	t value
No signaling motivation condition	-0.7219	0.4758	-1.517
Signaling motivation condition	-0.8261	0.4840	-1.707

Intercepts:

	Value	Std. Error	t value
0 1	-2.0229	0.4079	-4.9593
1 2	-1.5128	0.3759	-4.0245
2 3	-0.7900	0.3501	-2.2564
3 4	-0.1936	0.3444	-0.5620
4 5	0.6635	0.3566	1.8606

AIC: 297.2062

Table 10: Experiment#3a: the effect of treatment conditions on perceived trust

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.0690	0.2410	12.734	<2e-16 ***
No signaling motivation condition	-0.5875	0.3471	-1.693	0.0945 .
Signaling motivation condition	-0.5305	0.3505	-1.513	0.1341

Residual standard error: 1.298 on 79 degrees of freedom

R-squared: 0.04253

Table 11: Experiment#3a: the effect of perceived trust on transfer

	Estimate	SE	t value	Pr(> t)
(Intercept)	0.4351	0.3627	1.199	0.234
Perceived trust	0.8708	0.1207	7.212	<0.0001 ***

Residual standard error: 1.423 on 80 degrees of freedom

Multiple R-squared: 0.394

Table 12: Experiment#3b: numbers of subjects (the first actors) who made transfer, by period

All treatments (n=48)

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	37/48	32/48	31/48	24/48	26/48	27/48	25/48	26/48	28/48	20/48
Proportion	77%	67%	65%	50%	54%	56%	52%	54%	58%	42%

No signaling (n=15)

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	12/15	10/15	9/15	10/15	8/15	7/15	8/15	7/15	13/15	6/15
Proportion	80%	67%	60%	67%	53%	47%	53%	47%	87%	40%

No signaling motivation (n=16)

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	15/16	11/16	13/16	8/16	13/16	12/16	9/16	12/16	8/16	12/16
Proportion	94%	69%	81%	50%	81%	75%	56%	75%	50%	75%

Signaling motivation (n=17)

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	10/17	11/17	9/17	6/17	5/17	8/17	8/17	7/17	7/17	2/17
Proportion	59%	65%	53%	35%	29%	47%	47%	41%	41%	12%

Table 13: Experiment#3b: the effect of partner's donation on transfer

No signaling motivation condition

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	1.1146	0.3910	2.851	0.00436 **
Partner's donation	0.3943	0.4686	0.841	0.40015

Random effects:

Groups	Name	Variance	Std.Dev.
	SubjectID (Intercept)	2.044	1.43

Number of observations: 160, 32 subjects

Signaling motivation condition

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	- 0.5594	0.3499	-1.599	0.11
Partner's donation	0.6826	0.4267	1.600	0.11

Random effects:

Groups	Name	Variance	Std.Dev.
	SubjectID (Intercept)	1.817	1.348

Number of observations: 170, 34 subjects

Table 14: Experiment#3b: the effect of treatment conditions on transfer

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	0.7125	0.3634	1.961	0.0499 *
No signaling motivation	0.5759	0.5063	1.137	0.2554
Signaling motivation	-0.9934	0.4925	-2.017	0.0437 *

Random effects:

Groups	Name	Variance	Std.Dev.
SubjectID	(Intercept)	2.36	1.536

Number of observations: 480, 96 subjects

No-signaling condition is the reference category

Table 15: Experiment#3b: the number of subjects who returned transfer, by period

All treatments

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	18/37	17/32	14/31	10/24	12/26	7/27	10/25	12/26	11/28	6//20
Proportion	49%	53%	45%	42%	46%	26%	40%	46%	39%	30%

No signaling

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	6/12	6/10	3/9	2/10	3/8	1/7	2/8	5/7	4/13	0/6
Proportion	50%	60%	33%	20%	38%	14%	25%	71%	31%	0%

No signaling motivation

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	8/15	7/11	8/13	5/8	7/13	3/12	6/9	4/12	3/8	6/12
Proportion	53%	64%	62%	63%	54%	25%	67%	33%	38%	50%

Signaling motivation

Period	1	2	3	4	5	6	7	8	9	10
Yes/Total	4/10	4/11	3/9	3/6	2/5	3/8	2/8	3/7	4/7	0/2
Proportion	40%	36%	33%	50%	40%	38%	25%	43%	57%	0%

Table16: Experiment#3b: the likelihood of not returning transfer, by treatment conditions

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	0.9663	0.5332	1.812	0.0699 .
No signaling motivation	-0.9356	0.7088	-1.320	0.1868
Signaling motivation	-0.1765	0.7474	-0.236	0.8133

Random effects:

Groups	Name	Variance	Std.Dev.
subjectID	(Intercept)	4.322	2.079

Number of observations: 276, 89 subjects

No-signaling condition is the reference category

Table 17: Experiment#3b: the likelihood of not returning transfer, by donation

All subjects

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	1.2083	0.4018	3.007	0.00264 **
Donation	-1.5619	0.6057	-2.578	0.00992 **

Random effects:

Groups	Name	Variance	Std.Dev.
subjectID	(Intercept)	4.037	2.009

Number of observations: 276, 89 subjects

No-signaling condition

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	1.755	0.847	2.072	0.0383 *
Donation	-1.546	1.144	-1.351	0.1767

Random effects:

Groups	Name	Variance	Std.Dev.
subjectID	(Intercept)	4.743	2.178

Number of observations: 90, 28 subjects

No signaling motivation condition

Fixed effects:

	Estimate	SE	z value	Pr(> z)
(Intercept)	0.5202	0.5529	0.941	0.347
Donation	-1.5111	1.0060	-1.502	0.133

Random effects:

Groups	Name	Variance	Std.Dev.
subjectID	(Intercept)	3.843	1.96

Number of observations: 113, 32 subjects

Signaling motivation condition

Fixed effects:

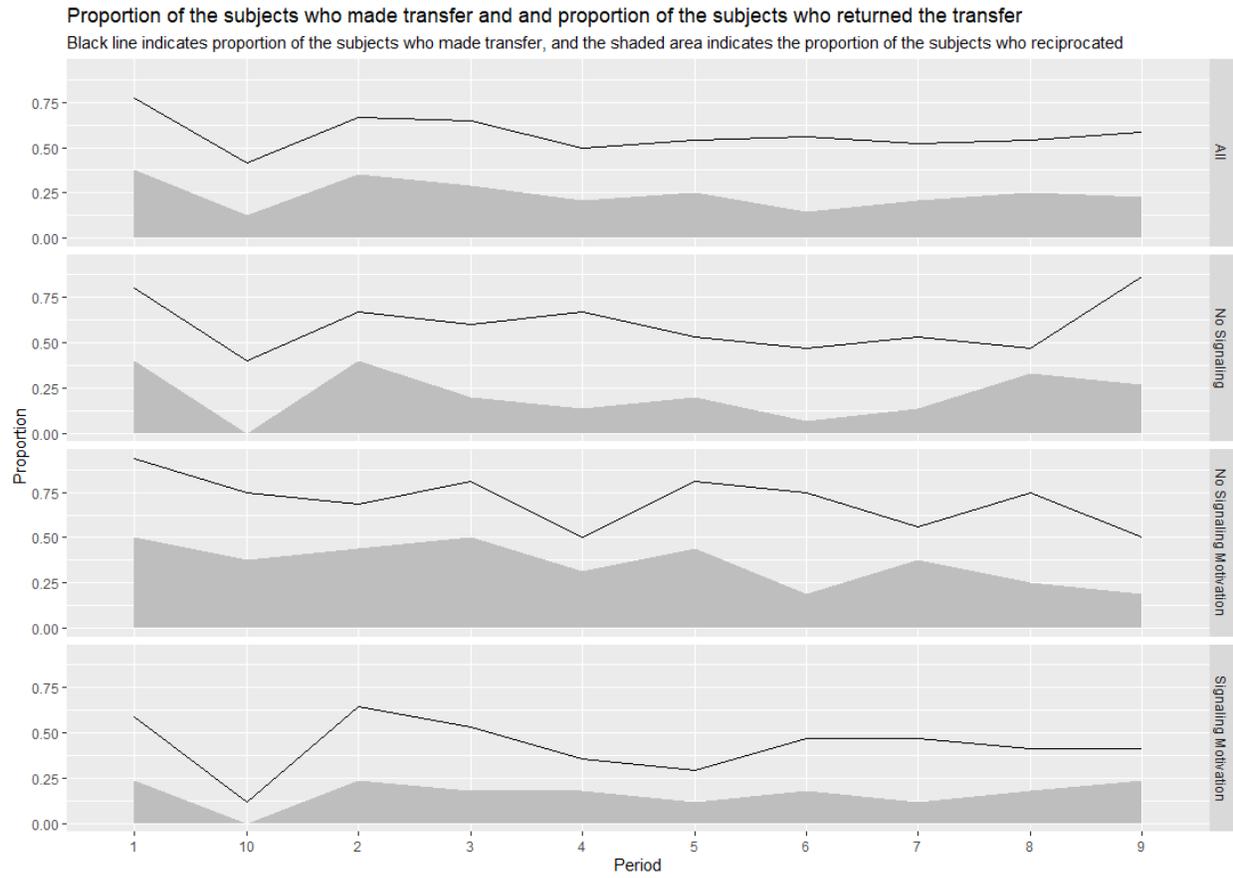
	Estimate	SE	z value	Pr(> z)
(Intercept)	1.6872	0.7338	2.299	0.0215 *
Donation	-2.0217	0.9941	-2.034	0.0420 *

Random effects:

Groups	Name	Variance	Std.Dev.
	subjectID (Intercept)	2.511	1.585

Number of observations: 73, 29 subjects

Graph#4: Outcomes from the repeated trust game over time, by treatment conditions



Chapter 6 Discussion and Conclusion

This chapter begins by summarizing our findings and relating them back to the intrinsic benefit model outlined in Chapters 2 and 3. Next, I will discuss about the several possible avenues of future research: further elaboration on the implications of “strategic” decision-making by the actors, considerations of the situations where the interests of the sender and the receiver converge, and the possible roles of type of the signal cost. Finally, I will describe the concluding thought of this study.

Summary of Results

The outcomes of the experiments show mixed support for the hypotheses proposed for this study. The results of the experiment#1 closely followed the predictions. As predicted, purchase of electronic vehicle (EV) was found to be used as the signal of environmental commitment. It was shown that the high signaling cost (for purchasing an EV) had the statistically significant positive relationship with the perceived environmental commitment, whereas the high visibility of the signal (the visual distinctiveness of the EV) had the statistically significant negative relationship with the perceived environmental commitment. In addition, the relationship between the visibility of the signal and the perceived environmental commitment was mediated by the perceived signaling motive. The result shows that the perceivers of the signal interpret high visibility signals as the sign of ulterior (non-intrinsic) motive behind the signal, which reduced the perceived

association between the signal and the characteristic being signaled – the purchase of EV and the environmental commitment, in this case.

The experiment#2 showed that those with higher levels of environmental commitment were willing to pay more for an EV, therefore indicating that (the willingness to make) purchase of EV is the reliable signal of environmental commitment. In addition, the subjects were willing to pay more for the EV with high visibility than for the EV with low visibility, indicating that the subjects thought the EV with high visibility had more potential for signaling – i.e., it should be worth more than the EV with low visibility, consistent with our predictions. Curiously, there was a weak, negative interaction effect between the environmental commitment and the signal visibility. It is difficult to interpret this result, as the association is not significant at .05 level, and is only significant at .10 level. However, it indicates the possibility that those with higher level of environmental commitment were less willing to pay more for the EV with high visibility. A potential explanation for this is that those with higher levels of environmental commitment were more aware of the potential drawback of using a high visibility signal, and they wished to avoid being perceived as a “poser” who buys the EV for the sake of signaling.

Compared to the first two experiments, the results of the experiment#3 diverged more from the predictions. Contrary to my hypothesis, the subjects did not use donation as a signal of trustworthiness at all, and the donation status of the partner did not affect the subjects’ transfer decision. On the other hand, consistent with the predictions, donors were more likely to return transfer than non-donors. While donation status did not affect the subjects’ transfer decisions, treatment conditions themselves affected the subjects’ transfer

decisions. The no signaling motivation condition facilitated the subjects to make transfers to their partners, whereas the signaling motivation condition reduced the amount of transfer subjects made. In addition, the signaling motivation condition reduced the subjects' perception of trust of their partners.

The results of the first two experiments largely support the intrinsic benefit model; the subjects mostly responded in ways consistent with the intrinsic benefit model. The experiment#2 illustrates the possibility that individuals are capable of behaving in the manner which is more strategic than initially predicted; nevertheless, this finding is not inconsistent with the logics underlying the intrinsic benefit model. Indeed, it can be argued that one of the main advantages of the intrinsic benefit model is that it has the room for the wide range of strategic behaviors. The functioning of the intrinsic benefit model depends on the perception of the perceivers; thus, many social factors come into strategic consideration of the actors, similar to what Goffman described in *Strategic Interaction* (1969). The present study focused on visibility, which is one of such social factors.

Although the third experiment failed to confirm the hypotheses – there was no association between donation and perceived trustworthiness among the subjects, and the subjects in the signaling motivation condition were no more likely to donate than the subjects in other conditions – it does not necessarily discredit the intrinsic benefit model. Indeed, given that the donation was not perceived as the signal of trustworthiness among the subjects, it can be argued that the subjects made the strategically sound decision to not donate for the purpose of appearing trustworthy.

The lack of association between donation and perceived trustworthiness among the subjects is more troubling. One potential explanation is that, from day-to-day interactions, the subjects had no opportunity to “learn” that charitable giving in small amounts do correspond with high levels of social preference/trustworthiness; thus, they did not know that donation is the signal of trustworthiness. This view is consistent with the assumption held by the typical models utilizing the signaling theory, as these models assume that individuals alter their strategies over time (through cognitive or evolutionary processes) based on previous outcomes. This process naturally requires the individuals to “update” their opinion on how well the specific signal represents the underlying characteristic. It is plausible to think that the subjects did not have any real opportunities to confirm or disconfirm whether donation signals trustworthiness or not. Thus, they had to on the different approaches; for example, by guessing what the other subjects were thinking, as discussed earlier.

In addition, the results of the third experiment provide additional support to the argument that the subjects could behave strategically within the signaling environment. In the signaling motivation condition, the donors were less likely to return their transfers than the non-donors; this finding suggests that the signaling motivation condition induced the less trustworthy subjects to make donations for their future gains. At the same time, the findings show that the subjects’ behaviors were affected by the treatment conditions. The signaling motivation condition reduced the rate of transfer among the subjects compared to the no signaling condition and the no signaling motivation condition; in turn, the no signaling motivation condition increased the rate of transfer among the subjects compared to the no signaling condition and the signaling motivation condition. In the first part of the

third experiment, the difference in the amount of transfer between the no signaling condition and the signaling motivation condition was mediated by the perceived trust. The results indicate that the subjects' awareness of their signaling environments – in particular, the environment which promotes strategic “altruism” for the sake of signaling – negatively affected their sense of trust in others. I argue that this finding suggests that the subjects understood the implications of charitable giving being used strategically for signaling; this further reinforces the view that the subjects themselves were capable of strategically utilizing signaling, both as the sender and as the receiver.

Finally, the findings from this study have a notable practical implication outside of making theoretical contribution to the signaling theory. As I have suggested earlier, many charity campaigns employ visible tokens to solicit donations. As indicated by the findings of the study, the intrinsic benefit model of signaling shows why such tactics are effective in soliciting donations. At the same time, the use of high visibility signals is a double-edged sword. Extrapolating from the finding of the third experiment, it is possible that the exposure to numerous high visibility signals of social preference can reduce one's perception of trust, as these signals crowd out the displays of “genuine” altruism. Traditionally, major donors of charitable organizations – who typically are members of the community elites – often preferred to contribute anonymously (Galaskiewicz, 1985). Such mechanisms may be effective in reducing the potential negative effects of strategic signaling of social preference. It may be the case that, in the world which is saturated by those who loudly claim to be virtuous, nobody would appear as virtuous. Although this argument falls outside of this study, it may warrant future research and theoretical developments.

Future Research

This study is planned to be expanded into several different directions. For further refinement of the intrinsic benefit model, the model currently lacks a formal analysis of its long-term consequences. Mathematical models of the signaling theory, by the definition, assume that the signaling interactions repeat over time. Indeed, the very notions of ESS (evolutionary stable strategy) in evolutionary biology (Grafen, 1990) and equilibriums in Spence's signaling theory (1976) assume that the signaling interactions, repeated over time, will eventually reach the stable outcomes. Because the actual structure of the model (sans the intrinsic benefit) does not deviate much from the existing models, it is likely that the preexisting findings would likely be applicable here. Nevertheless, it is also possible that innovations of this model, especially its emphasis on visibility, may lead to results with meaningful differences. More importantly, the model should be able to make long-term macro-level predictions that can be tested against empirical data. Such developments can greatly enhance the utility of the signaling theory within Sociology.

Signaling theory in economics also makes the distinction between signaling and screening. In a signaling game, the sender of the signal acts first. In a screening game, the receiver of the signal acts first (Riley, 2001). The difference between signaling and screening can lead to different outcomes, depending on specifications of the model. Some social phenomena, typically the ones that involve formal requirements, may correspond better with the screening model. Therefore, it may be beneficial to further elaborate the model so that it will be able to distinguish between signaling and screening.

1. Actors Behaving Strategically

In addition to refining the core model of the intrinsic benefit model, I suggest that there are several potential avenues for expanding upon the findings of this study. First, the findings from this study indicates that the subjects were capable of behaving strategically under the situations where the intrinsic benefit model is relevant. In particular, the intrinsic benefit model suggests that strategically manipulating the signal visibility to maximize the potential benefit from a signal and minimize the perceived signaling intention. Second, the intrinsic benefit model can be expanded to account for the situations where interest of the sender and the interest of the receiver converges. While the signaling theory in economics does not typically deal with such situations, the signaling theory in evolutionary biology is designed to account for these situations. Such situations can interact with the premise of the intrinsic benefit model to produce (theoretically and substantially) notable results. Finally, following the logic of the intrinsic benefit model, it is possible that the type of the signal costs – monetary and non-monetary, in particular – would affect what intention the signal is perceived to have. I will provide detailed discussions on the above points in the subsequent chapters.

2. The Interests of the Sender and Receiver Converge

The model can also be applied in situations where interests of the sender and the interests of the receiver converge. The intrinsic benefit model presented in this study assumes that the benefit the sender of a signal can obtain from successful signaling is

constant, regardless of the level the internal characteristic being signaled. This is the common assumption held in many models in the signaling theory. For example, it is reasonable to assume that being able to successfully find a mate through successful signaling is equally beneficial to animals, regardless of their fitness. Similarly, in Spence's classic model (1973), successful signaling would lead to an increased wage, which would be desirable to the sender of the signal regardless of the actual level of the characteristic being signaled. In regard to the experiments conducted in this study, it is reasonable to argue that this assumption is easily satisfied with the experiment#3, as it can be assumed that earning more money from successfully signaling high level of social preference is desirable regardless of one's actual level of social preference. Unfortunately, this assumption is somewhat more questionable for the experiment#1 and #2; environmentalism is somewhat of a partisan issue in the Unites States, and it may be the case that those who are not committed to the environment may not prefer to be perceived as being committed to the environment.

For the purpose of the formal analysis on the situations where the interests of the sender and the receiver converge, we assume the followings.

- Cost of producing signal for low-quality actor = CL
- Cost of producing signal for high-quality actor = CH
- Benefit obtained by low-quality actor from successful signaling = BL
- Benefit obtained by high-quality actor from successful signaling = BH

In the signaling model incorporating both benefits and costs, a signal can be reliable if the following condition is met.

$$(BH - CH) > (BL - CL)$$

Note that when $BH = BL$, the above condition is transformed into $CL > CH$, producing the familiar assumption of “simple” signaling models, including the main study of this dissertation.

Now, we have two general situations (Graph); $BL > BH$ (situations where low-quality individuals obtain more benefit from signaling than high-quality individuals do), and $BL < BH$ (situation where high-quality individuals obtain more benefit from signaling than low-quality individuals do). In such situations, dishonest signals would become more prevalent, as low-quality individuals would have more incentive to make successful signals. In evolutionary biology, threat display is a typical example of such situations; when individuals try to win competition for resources by displaying threat, the weakest individuals have the strongest incentive to make threatening displays, as doing so is the only way for them to secure resources, as they are incapable of winning any actual fights (Adams & Mesterton-Gibbons, 1995). It is likely that there are some sociologically relevant circumstances where the same conditions hold true. Indeed, “the smallest dogs bark the loudest” is a popular proverb (probably more so in Japan than in the US), suggesting that people tend to intuitively understand the argument underlying the signaling model of threat display. A possible “social” example of situations where there is negative correlation between benefit from successful signaling and internal characteristic being signaled is a firm with a product that is inferior to its competitors. Given that its product would not be

able to succeed based on its quality, the firm may instead invest heavily into signaling – by obtaining expensive celebrity endorsements, for example.

How would the situations where low-quality individuals obtain more benefit from signaling than high-quality individuals affect the intrinsic benefit model? Given the increased incentive for deception, it is possible that a receiver who is acting strategically would be more concerned about deceptive signals. In addition, I argue that it is likely for high visibility signals to function as signals of low quality. As shown in the experiment#1, a signal with high visibility is perceived to be associated with increased signaling motivation. Given that those with lower quality have more incentive for successful signaling in this condition, perception of increased signaling motivation, in turn, would signal lower quality. Thus, it is plausible to suggest that, under this condition, the potential negative effect of signal visibility to the receiver of the signal will be magnified. Under such conditions, I argue that fine-tuning of signal visibility will be particularly important; it may be the case that signals which are prescribed (either by norms, or by virtue of being de-facto standards – for example, it is listing one’s educational credentials when applying for a job would be seen as perfectly ordinary) tend to be most effective, as the mandatory nature of such signals can minimize the perceived signaling motivation.

While the situation where low-quality individuals obtain more benefit from signaling than high-quality individuals does not alter the core characteristics of a signaling model, the nature of signaling can be drastically altered in the opposite situations. Grose (2011) argued that this aspect of the signaling theory deserves far more attention than it currently receives. To understand it, we will go back to the inequation above.

$$(BH - CH) > (BL - CL)$$

First, we consider the special situation where the signaling cost is identical both for high quality individuals and low quality individuals. Under this condition, the above inequation will be transformed into a simple inequation.

$$BH > BL$$

This indicates that the difference (between high quality individuals and low quality individuals) in cost to produce a signal is not actually necessary for the signal to be reliable when the conditions above are met. Ohtsubo and Watanabe, in a rare application of this argument to human behaviors, suggested that an apology can function as the reliable signal of intention to restore relationships even when the cost of apology is identical for those who are sincere and those who are insincere, as long as only the sincere sender can benefit from the restored relationship (2009). It is also possible to interpret Ohtsubo and Watanabe as being similar to the intrinsic benefit model, except that the difference in benefit is concrete, instead of based on perception as with the intrinsic benefit model.

It is possible to further narrow down the condition added to the initial inequation. Assume that it is costless to produce the signal for both low quality individuals and high quality individuals ($CH = CL = 0$). In this case, the signal can be reliable if low quality individuals are incapable of benefiting from successful signaling, whereas high quality individuals can benefit from successful signaling ($BH > 0, BL = 0$), as there would be no incentive for low quality individuals to produce the signal in the first place. This can be seen as an equivalent of cheap talks (non-binding communication that is not costly to produce) in coordination (stag-hunt) games. In stag-hunt game, cheap talks are effective at

inducing cooperation (Charness, 2000), because both parties share the mutual interest and there is no incentive for producing a deceptive signal (that is, using a cheap talk to suggest cooperation, then betray the partner to defect instead).

What are the possible effects of the situations where high-quality individuals obtain more benefit from signaling than low-quality individuals on the intrinsic benefit model? I have argued that, in the situations where low-quality individuals have more to gain from successful signaling than high-quality individuals do, there is more incentive for producing deceptive signals. The opposite is the case for the situations where high-quality individuals have more to gain from successful signaling; low-quality individuals do not have much to gain from successfully signaling their (deceptive) “high quality”. This leads to a potentially very intriguing interaction with the signal visibility. If the receiver of the signal perceives the sender to have no incentive to make dishonest signals, the increased perception of signaling motivation is less likely to reduce the perceived level of the internal characteristic being signaled. Indeed, if the benefit from successful signaling can be only realized by those with high quality ($BH > 0$, $BL = 0$), the increased visibility in itself could function as the signal of high quality; if someone is willing to produce a high-visibility signal under such a situation, s/he must be high quality so that s/he would be able to benefit from it.

I suggest that one potential example of the above situation is signaling of memberships in categories which are commonly seen as undesirable. Instead of producing subtle signals, members of categories which are commonly perceived undesirable – for example, products which are associated with specific low-status niches, or the individuals who are the members of low-status subcultures – are often seen flaunting highly visible

cues of category memberships. On the surface, these behaviors contradict the predictions based on the intrinsic benefit model; in turn, the perceiver of a highly visible signal would assume that the sender of the signal must be motivated by potential benefits from signaling. If this perceived signaling motive reduces the perceived intrinsic benefit, as with the argument raised in this study, the signal would not be perceived as the signal of high quality.

However, as discussed above, the high visibility of a signal itself becomes the signal of high quality if it is perceived that low quality individuals have no incentive to signal. Thus, for example, perceivers would not question the reliability of highly visible, stereotyped signals of category memberships employed by low-status cultural products: such as mass-produced romance novels or albums of heavy metal music. A perceiver would assume that if a book has a title and a cover art which seem like stereotypical romance novel, the said book is a romance novel; there is no reason for the book of high literature to appear like a mass-produced romance novel, and the only reason why a publisher of the book wants its book to appear like a mass-produced romance novel is that the book is a romance novel. Thus, in cases such as the example above, the increased visibility signals the higher level of signal-relevant internal characteristic. As we can see here, through incorporating the association between individual characteristic and benefit obtained from signaling, the signaling theory can provide rational-choice explanation to the commonly observed phenomena that “high” culture tends to be associated with subtlety and “low” culture tends to be associated with distinctive cues. Note that it is also possible to explain the situation above by using the argument based on cost difference; it can be argued that actors belonging to the undesirable category would incur less cost from signaling its

category membership than actors not belonging to the undesirable category. However, this alternative explanation is not as straightforward – it assumes that actors who are not belonging to the undesirable category somehow wish to signal its membership to this category – and I suggest that it does not offer any theoretical advantage over the argument based on benefit difference.

Overall, it can be argued that the application of the signaling model based on benefit-quality association shares the main characteristic of the intrinsic benefit model. Both approaches allow the signaling theory to explain phenomena which could not be explained by the cost of signaling or the cost-quality association. Indeed, it may appear that the intrinsic benefit model covers the subset of situations with the benefit-quality association, where the benefit-quality association depends on the perceived intrinsic benefit. However, there is a key distinction between the intrinsic benefit model and a generic signaling model with the benefit-quality association; the intrinsic benefit is present regardless of the outcomes of the signaling, whereas the benefit in signaling models utilizing the benefit-quality association is contingent on successful signaling.

Otsubo and Watanabe's argument on how apologies are used as the signal of the intention to repair the relationship (2009)– the example of the research using the benefit-quality association, as identified by Grose (2011) – helps to illustrate the difference. O&W's model suggested that an apology is a reliable signal of the intention to repair the relationship because only those who wish to continue the relationship would benefit from repairing the relationship. On the other hand, with the intrinsic benefit model, an apology works as a signal of the sincerity of the apology because it can be assumed that the person

who is apologizing would obtain the intrinsic benefit from doing so depending on how sincerely they feel sorry; a sincere individual would be able to reduce their feeling of guilt from the apology. The difference has the practical implications for the application of the model; for example, O&W's model would not be applicable for apologies between strangers, as it is highly unlikely that the strangers would be interacting again in future, regardless of apologies. Therefore, I argue that the intrinsic benefit model can provide better explanations for apologies in such situations. Similarly, if the one who apologizing would have to interact with the receiver in future, O&W's model would not work; the benefit-quality association would no longer exist if everyone would obtain the same benefit from repairing the relationship. This hypothetical situation is similar to the experiment#3 in that everyone would obtain the same benefit from successful signaling; the intrinsic benefit model ought to be able to explain such phenomena, assuming that the cost of apology is high enough so that it is too costly for those without the sincere apology.

Suffering a public humiliation might be one of the possible options to make apologies costly, and apologies sometimes involve humiliation. For example, "dogeza" – prostration on the ground – is a common form of apology for serious transgressions in Japan, where doing dogeza is commonly seen as very humiliating. Similarly, there is a famous episode from the medieval history, where Henry IV, Holy Roman Emperor, waited outside for three days in the snow at the gate to the castle of Canossa for the purpose of begging for forgiveness from Pope Gregory VII after being excommunicated by the latter during the investiture controversy. A humiliating (and physically painful) apology such as this might be highly effective way to signal the sincerity of an apology. Of course, it is by no means that the apology is genuine, and it does not automatically ensure the signal would be

interpreted by the receiver as so. Much like the subjects in the experiment#3, the receiver may have the reason to doubt the intention behind the apology, when it is known that there is a very strong incentive to repair the relationship – this was the case in this particular example, therefore making O&W's model not applicable. We know that Henry IV's apology likely was not sincere; he resumed the conflict with the Papacy shortly after his excommunication was lifted (for this specific case, we may argue that inflicting the highly visible humiliation on Henry was desirable to Gregory so that it was worthwhile for Gregory to make Henry beg for forgiveness, regardless of what Henry did afterward). As we have seen here, it is plausible that the intrinsic benefit model often (if not always) provide a superior explanation than benefit-quality association in the case of apology.

3. Type of Signal Cost

It is also possible to integrate the type of signal cost into the intrinsic benefit model. In the conventional signaling theory, the type of signal cost should not matter; assuming the different types of signal cost can be somehow measured in the same metric, the cost of producing the signal is all that matters for the effectiveness of the signal. However, this position intuitively seems wrong; it strikes as implausible that donating \$10000 to a charitable cause and putting \$10000 worth of effort into the same cause (for example, by moving from a well-paid private sector position to a non-profit position which does not pay as well) ought to be perceived differently. Going back to Bourdieu's notion of cultural capital, the research on cultural capital informs the signaling theory that the cost of a signal itself is not sufficient. Instead, the signal must be produced in the "right" way – for example,

the discussion about the spending habits of “new money” and “old money” – and it is possible to argue that the “right” signal must be backed by the right type of the cost. Indeed, Bourdieu discussed how different groups tend to select signals with different costs, based on the availability of resources to each group. For example, educated, young, middle class people tended to pick the activities which were time-consuming, and which required some efforts but not much money – hiking, for example (Bourdieu, 1984).

The logic underlying the intrinsic benefit model suggests that, for a signal to be effective, it needs to be perceived as it was not motivated by the signaling benefit. Then, it is plausible that a signal whose cost is primarily monetary in nature may not make an effective signal. It can be argued that the perceived association between the monetary cost of a signal and the intrinsic benefit obtained from producing the signal is not very strong. Money, by its nature, tends to be perceived as impersonal and cold; indeed, a mere cue of money can induce individuals to act in more selfish (or, independent) ways (Vohs, Meade, and Goode, 2006). Additionally, wealthy individuals can easily absorb the cost of making donation – this, of course, is why wasting money can function as a signal of wealth. On the other hand, I argue that certain types of non-monetary costs – such as efforts or non-monetary sacrifice – tend to have a natural association with intrinsic benefit. One can put “heart and soul” into the effort (such as willingly enduring difficulties) of producing a signal, but one cannot do the same into monetary cost of producing a signal. It is natural to assume that the perceiver is more inclined to see the signal as being motivated by its intrinsic benefit when s/he sees the signal as the product of “heart and soul”. Many religions impose non-monetary costs to their believers, and those who are willing to go the extreme lengths to pay such costs were frequently revered. It seems highly unlikely that

those who purchased indulgences from the Catholic Church were regarded highly – instead, it is likely that the sales of indulgences harmed the reputation of the Church and contributed to the Reformation. For some characteristics, monetary costs are less likely to invoke the perception of intrinsic benefit; it is possible that the very nature of money can even make one's utilitarian motive salient.

Indeed, this might explain some of the (lack) of findings. All three experiments relied on the monetary costs of producing the signal. When the experiments were being designed, the monetary costs seemed like an appropriate choice, as the nature of money makes it possible to make precise comparison between different amounts. In the hindsight, perhaps I should have put more careful consideration of solely relying on the monetary cost, especially since this study is centered on the pro-social behavior. If money itself is perceived as the anathema of pro-social values, it may significantly weaken the effect of the signals relying on monetary costs. Indeed, in the experiment#1, the visibility – which is non-monetary cost – of an EV had much stronger effect than the cost of EV.

As far as the author is aware of, there are not many previous studies which aim to account for the effects of type of cost in signaling theory. This is not surprising, as the reasoning underlying two of the main disciplines associated with the signaling theory – economics and evolutionary biology – are not necessarily compatible with the type of signaling cost. Evolutionary biology does not presuppose sophisticated cognition, and the notion of utility in economics presuppose that any costs can be converted into utility. Nevertheless, there are studies which incorporated the type of cost into studies of the signaling-associated phenomena. For example, Cuypers, Koh and Wang studied how the

type of corporate giving affects how they are perceived (2015). They found that innovative giving (giving which is more than just donating money, and instead requires effort) is more likely to be perceived as more substantial than quantitative giving (generous giving), and this effect is more pronounced for the sin industry (such as alcohol or gaming). From the perspective of the intrinsic benefit model, a corporation whose organizational value is aligned with social responsibility would be perceived as being more likely to “benefit” from putting efforts into innovative giving. At the same time, the very nature of the sin industry would make the reputational/signaling motive more salient for perceivers, much like a sinner purchasing the indulgence. I suggest that similar approaches can be productive in analyzing various signaling-related phenomena, both at the individual-level and the organizational-level. However, it must be noted that straightforward experiments as used in this study may not be applicable, as different type of costs cannot be directly compared using the single metric. Identifying the methodology to sort out this conundrum may greatly benefit the research of signaling based on the intrinsic benefit.

Final thought

The main aim of this study is to expand the applications of the signaling theory by emphasizing the roles of the (perceived) intrinsic benefit from the signal production. I have discussed that the overall findings of this study are mixed. The experiment#1 was largely successful. The experiment#2 and the experiment#3 produced some results which appear to be consistent with the argument behind the intrinsic benefit model – in particular, the importance of the perceived intention of others – but failed to produce the outcome which

can be easily explained. I argue that the results are indicative of the complex nature of integrating perceived intention/motivation into the process of signaling. In the previous chapter, I suggested that the findings were consistent with the Goffman's notion of Strategic Interaction (1969). In this work, Goffman described the interactions between secret agents where both actors are trying to guess the true motive of the other, while at the same time being mindful that other is attempting the same. In a sense, Goffman's argument emphasizes the situations where individuals behave as if they are professional poker players. (Note that poker is the game with incomplete information – unlike, say, chess or go, which are games with complete information –incomplete information is the characteristic it shares with the situations described in the signaling theory.) A reasonable criticism of such an approach is that people usually do not act like professional poker players or agents in espionage. At the same time, I insist that high-stake social interactions with asymmetrical information are not unusual, and it is likely that many people would at least attempt to handle such situations to their benefits. Most people are not serious players of poker; however, there is nothing unusual about participating in job interviews (on either ends) or attempting to determine the character of their potential date – situations which would involve signaling where the perceived intrinsic benefit is relevant. Many high stake (and therefore, socially consequential) decisions, such as the ones pertaining to financial investments or coalition formations, most certainly involve some degrees of “strategic” signaling. I believe that it is reasonable to suggest that individuals behave strategically in such situations, especially when they are willing to consider what they think the other party would think.

It is important to note that there is a pitfall to this approach. Given that the effect of the signal depends on the perceived association between the intrinsic benefit and the characteristic being signaled, the same signal can hold different meanings depending on its audience. For example, let us go back to the Henry IV at the gate to Canossa. The most obvious – but somewhat unlikely – interpretation of the event is that it was a signal from Henry to Gregory, that he was sincerely regretting the past conflict and was willing to endure the extreme humiliation to atone for his error. Yet, it was not the only thing Henry signaled at the gate of Canossa. Perhaps Henry was signaling that he was willing to temporarily concede the defeat to sue for a truce by enduring the humiliation. If we consider his signal as the signal of penance, it would be considered as an unreliable signal, but this was not the case if he was merely displaying the willingness to concede. Perhaps Gregory (and Gregory’s supporters) was not even the intended audience of the signal; instead, he was signaling to the neutral parties in the conflict that his conflict with the Gregory was not based on his contempt of the Papacy. Or perhaps he was signaling to his political allies that he is willing to endure the humiliation to have his excommunication lifted for achieving the long-term victory in the investiture conflict if it was necessary, and therefore they could trust him to not betray his allies – thus they should keep supporting him. Indeed, any of the above interpretations – or, even the combinations of some of the above – might be possible, and it would be impossible to retrospectively identify the “correct” interpretation. Nevertheless, both Henry and Gregory were savvy political actors, it was likely that they were aware that they were playing the complex game of signaling. This, of course, is exactly the type of situations Goffman was discussing about in *Strategic Interaction*. Indeed, the type of analysis Goffman was known for – qualitative, in-depth analysis based on “thick

descriptions” – would make a suitable tool for analyzing individual situations with complex signaling interactions such as the one between Henry and Gregory. Yet, the question remains that how such analysis can scale to theoretical-based quantitative analysis within social science – if it is possible at all. The ability to provide an explanation to the specific situation is not particularly useful by itself.

Recall Bird and Smith (2005)’s argument that the main benefit of the signaling theory is its capability to integrate symbolic communication into rational-choice/strategic action framework. I argue that the main advantage of the rational-choice framework within social science, in turn, is how it can generate simple, falsifiable hypotheses based on the notion of the strategic actor behaving according to their environment. Therefore, any attempts to integrate symbolic perceptions into the rational-choice framework of the signaling theory can risk overcomplicating the analysis based on the signaling theory would require due caution so that it would not hurt the main strength of signaling theory. This study aimed to manage this by first identifying the specific characteristic of a signal which would have the specific implication under the intrinsic benefit model – the signal visibility – and test how it interacts with the processes of signaling. Although the results from this study are mixed, I am optimistic that it does not undermine the credibility of the underlying approach.

Finally, it is important to note that the signaling theory is not intended to explain everything; being able to explain a certain behavior through the use of the intrinsic benefit model does not mean there is no other explanation for the same behavior. Rather, the existence of the over explanation is the very reason the intrinsic benefit model of signaling

theory – the model requires that those with the high level of the certain characteristics tend to be willing to pay more cost to produce the specific signal, in general. For example, there are some altruistic behaviors which are solely motivated by the goodness of the heart. I am certain – and it is plausible to assume that most people would agree to it – that there are some altruistic acts that were made out of the goodness of their hearts. Just because some altruistic behaviors can be explained by the signaling theory, it does not mean all altruistic behaviors are produced for the sole purpose of signaling. In the third experiment, it seems to be the case that many subjects donated because of the intrinsic benefit they obtained from making the donation – even though they did not seem to have thought the others would actually do so. This creates the cost-quality association, making a signal reliable in the long run – the signaling of altruism is impossible if there is no altruism in the first place. The intrinsic benefit model, when applied to signaling by pro-social behaviors, implies that both signaling and “pure” altruism takes roles in production of altruistic behaviors. The intrinsic benefit model is not intended to supplant the other potential explanations for the behaviors which could be regarded as the signals, but rather, it is intended to provide the toolkit for the social scientific analysis of situations where both the intrinsic motivation and the signaling motivation potentially play the role in producing the behaviors which are regarded as the signals.

Appendix

Appendix A: Vignettes for Experiment#1

High Visibility, High Cost (corresponds to the *Visibility (High)/Cost (High)* treatment condition in the Experiment#1)

Alex is going to buy a new car soon, and has decided to buy an EV (electronic vehicle) instead of a gasoline car. The EV that Alex is going to buy has an attractive and distinctive appearance that stands out from other cars. People can easily recognize it as an electronic vehicle. Alex is very happy with the appearance of the new car. The EV costs about \$50000, whereas the gasoline cars with the same size and the same level of equipment will cost about \$25000. Please evaluate the following statements about Alex.

High Visibility, Low Cost (corresponds with the *Visibility (High)/Cost (Low)* treatment condition in the Experiment#1)

Alex is going to buy a new car soon, and decided to buy an EV (electronic vehicle) instead of a gasoline car. The EV that Alex is going to buy has an attractive and distinctive appearance that stands out from other cars. People can easily recognize it as an electronic vehicle. Alex is very happy with the appearance of the new car. The EV costs about \$25000, the same amount as a gasoline car with the same size and the same level of equipment. Please evaluate the following statements about Alex.

Low Visibility, High Cost (corresponds with the *Visibility (Low)/Cost (High)* treatment condition in the Experiment#1)

Alex is going to buy a new car soon, and has decided to buy an EV (electronic vehicle) instead of a gasoline car. The EV that Alex is going to buy is based on a popular gasoline car, but it operates on an electronic motor instead of gasoline engine. It has an attractive appearance, but it looks identical to the gasoline car it is based on. Alex is very happy with the appearance of the new car. The EV costs about \$50000, whereas the gasoline cars with the same size and the same level of equipment will cost about \$25000. Please evaluate the following statements about Alex.

Low Visibility, Low Cost (corresponds with the *Visibility (Low)/Cost (Low)* treatment condition in the Experiment#1)

Alex is going to buy a new car soon, and has decided to buy an EV (electronic vehicle) instead of a gasoline car. The EV that Alex is going to buy is based on a popular gasoline car, but it operates on an electronic motor instead of gasoline engine. It has an attractive appearance, but it looks identical to the gasoline car it is based on. Alex is very happy with the appearance of the new car. The EV costs about \$25000, the same amount as a gasoline car with the same size and the same level of equipment. Please evaluate the following statements about Alex.

Appendix B: Questions for the Experiment#1

[R] indicates reverse-coded items

Perceived Environmental Commitment (the mean of the six items corresponds with the variable *Perceived Environmental Commitment* (PEC) in the experiment#1)

Alex is more interested in convenience than conserving natural resources.

Alex is committed to protecting the environment.

Alex is concerned about global climate change.

Alex thinks that people worry too much about environmental problems. [R]

Alex thinks that people should be using more renewable energy sources. (solar, wind, and hydro, for example)

Alex is NOT concerned about using fossil fuels. [R]

Perceived Signaling Motivation ('yes' indicates high signaling motivation and 'no' low motivation and corresponds with the variable *Perceived Signaling Motive* (PSM) in the experiment#1)

Alex cares about what others think of him.

Appendix C; Vignettes for Experiment#2

High visibility (corresponds with the *Visibility (High)* treatment condition in the Experiment#2)

Imagine that you are considering purchasing a car. Skolden Verdant and Skolden Cerulean are both medium-sized cars being manufactured by the same automaker. Verdant was developed to operate on an electronic motor instead of a gasoline engine. Verdant has an attractive and distinctive appearance that stands out from other cars, and people can easily recognize it as an electronic vehicle. Cerulean is a popular car that uses a gasoline engine. It has an attractive appearance, although it does not stand out from other cars using a gasoline engine. Verdant and Cerulean are about the same size, and both cars are equally well-built and well-equipped. If the Cerulean is priced at \$20000, what is the highest amount you are willing to pay for the Verdant? Please choose from the following options: less than \$20000, \$20000, \$21000, \$25000, \$30000, \$40000, \$45000, \$50000 or more.

Low visibility (corresponds with the *Visibility (Low)* treatment condition in the Experiment#2)

Imagine that you are considering purchasing a car. Skolden Cerulean and Skolden Cerulean EV are both medium-sized cars being manufactured by the same automaker. Cerulean is a popular car using a gasoline engine. Cerulean EV is based on Cerulean, but operates on an electronic motor instead of gasoline engine. Cerulean has attractive appearance, but it does not stand out from other cars using gasoline engine. Cerulean and Cerulean EV looks identical from outside, and both

cars are equally well-built and well-equipped. If the Cerulean is priced at \$20000, what is the highest amount that you are willing to pay for the Cerulean EV? Please choose from the following options: less than \$20000, \$20000, \$21000, \$25000, \$30000, \$40000, \$45000, \$50000 or more.

Appendix D: Questions for the Experiment#2

[R] indicates reverse-coded items

The respondents' commitment to the environment (the mean of the six items corresponds with the variable *Environmental Commitment* in the experiment#2)

I don't care about saving natural resources (R)

I think we are too reliant on fossil fuels

I am willing to do whatever I can do to protect the environment

I think that the danger of global climate change is overly exaggerated (R)

I think we should be using more renewable energy sources (solar, wind, and hydro, for example)

I am sure that the environment will be fine regardless of human actions (R)

The respondents' signaling motive (the mean of the two items corresponds with the variable *Signaling Motive* in the experiment#2)

I do not care if others think that I am wasting natural resource (R)

I want others to know that I am committed to save the environment

Appendix E: The Instruction Sheet for the Experiment#3, Phase#1

It was distributed near the beginning of the experiment#3. In the *no signaling* condition and the *no signaling motivation* condition, it was distributed after the donation decision. In the *signaling motivation* condition, it was distributed before the donation decision.

You will be asked to make the following decisions. First, you will be randomly assigned into pairs; either as the first actor or the second actor.

If you are assigned as the first actor in a pair, here is what happens. From your \$5, you will be asked to decide on how much contribution you'll want to make to your partner. You can decide between \$0 (no contribution) and \$5, with \$1 intervals. You will keep the rest of your starting money.

You will make the decision by choosing one of the six options from the list, as shown on the picture to the right.



After the first actor makes the decision, the experimenter will add three times the amount of contribution made by the first actor to the pool. So, if the first actor makes \$3 contribution, the experimenter will add \$9, for the total pool of \$12.

Now, if you are assigned as the second actor, this is your turn to make a decision. You will decide on how to split the pair's pool of money. For example, you may decide to keep \$8 of the \$12 pool for yourself, returning \$4 to your partner (the first actor).

Here is an example screen for return decision. Note that you are typing in the amount you want to give back to your partner, not the amount you want to keep. You can type in any integral numbers between 0 and the total pool size.

Your partner made the contribution of 4 dollars.
 You have the total amount of 16 dollars to split between you and your partner.
 Please enter how much you want to give back to your partner.
 Once you enter the number, please press Confirm.

So, what will be the end result for the above example? The first actor will receive \$4 returned by the second actor, in addition to the amount s/he kept from the starting \$5. This makes the total of \$6 – \$2 from the starting amount (\$5 - \$3), and \$4 from the return. The second actor will take home \$8. Finally, all participants will receive additional \$6 from participating in this experiment, independently of the experiment's outcome.

Here is a table with more examples (A is the first actor, B is the second actor).

A contributes	A keeps	Total pool	B returns	B keeps	A earns	B earns
\$4	\$1	\$16	\$12	\$4	\$13	\$4
\$4	\$1	\$16	\$8	\$8	\$9	\$8
\$4	\$1	\$16	\$4	\$12	\$5	\$12
\$4	\$1	\$16	\$0	\$16	\$1	\$16
\$2	\$3	\$8	\$6	\$2	\$9	\$2
\$2	\$3	\$8	\$4	\$4	\$7	\$4
\$2	\$3	\$8	\$2	\$6	\$5	\$6
\$2	\$3	\$8	\$0	\$8	\$3	\$8
\$0	\$5	\$0	N/A	N/A	\$5	\$0

Appendix F: The Instruction Sheet for the Experiment#3, Phase#2

It was distributed before the beginning of the experiment#3b across all subjects who participated in the experiment#3b.

The second phase of the experiment is similar to the first phase, but there are some differences. The second phase is consist of 10 rounds of interactions between pairs (again, between the first actor and the second actor). For each round, the computer will generate new pairs; therefore, you will be randomly assigned a partner and a role for each round. Note that you don't know how many times you will be matched with a specific partner; you may be matched with the same person five times, or may never matched with a specific person – it is all up to the random number generator.

Although you will participate in 10 rounds of interactions, you will receive payments from outcomes of two rounds which are randomly selected by the computer. You will learn which two rounds are selected for payment at the end of the experiment. Finally, regardless of outcome from this phase, all participants will earn additional \$1 from this phase.

In each round, there will be the first actor and the second actor. If you are assigned as the first actor, you will decide on whether to transfer \$2 to your partner or not. The screen where you make this decision is shown on right.



If you decide to not transfer, you will keep \$2 and the round will end, and your partner will receive nothing from this round. If you decide to transfer, the experimenter will add \$4 (to the total of \$6), and the second actor will decide on how to split the resulting \$6. The second actor will have two options; split equally with his/her partner (both receive \$3), or take \$6 for himself/herself, while his/her partner (the first actor) receiving \$0. The screen where you make this decision is shown below.



The table below lists the possible outcomes for each round.

1 st actor does...	2 nd actor does...	1 st actor receives...	2 nd actor receives...
Not transfer \$2	N/A	\$2	\$0
Transfer \$2	Split equally	\$3	\$3
Transfer \$2	Take \$6	\$0	\$6

Note that when you are selected as the first actor and decide to make transfer, you will not learn the decision made by your partner until the end of the experiment. At the end of the experiment, you will learn about your outcome of two rounds which are selected for payment.

Again, thanks for your participation!

Appendix G: Structure of the Experiment#3

The First Phase: Experiment#3a Only

No Signaling Condition/No Signaling Motivation Condition

1. Donation Decision
2. Phase#1 Instruction
3. Transfer Decision by the Actor#1 (Behavioral Operationalization of Trust), Experiment#3a
4. Return Decision by the Actor#2 (Behavioral Operationalization of Trustworthiness), Experiment#3a
5. Question About the Perception of the Partner
"I think I can trust my partner"
6. The Outcome Information/Payout

Signaling Motivation Condition

1. Phase#1 Instruction
2. Donation Decision
3. Transfer Decision by the Actor#1 (Behavioral Operationalization of Trust), Experiment#3a
4. Return Decision by the Actor#2 (Behavioral Operationalization of Trustworthiness), Experiment#3a
5. Question About the Perception of the Partner
"I think I can trust my partner"
6. The Outcome Information/Payout

The Second Phase: Experiment#3a+#3b

No Signaling Condition/No Signaling Motivation Condition

1. Donation Decision
2. Phase#1 Instruction
3. Transfer Decision by the Actor#1 (Behavioral Operationalization of Trust), Experiment#3a
4. Return Decision by the Actor#2 (Behavioral Operationalization of Trustworthiness), Experiment#3a
5. Question About the Perception of the Partner
"I think I can trust my partner"
6. Phase#2 Instruction
7. Simplified Trust Game (Repeated Ten Times)
 - a. Random Assignment
 - b. Transfer Decision by the Actor#1 (Behavioral Operationalization of Trust), Experiment#3b
 - c. Return Decision by the Actor#2 (Behavioral Operationalization of Trustworthiness), Experiment#3b
8. The Outcome Information/Payout

Signaling Motivation Condition

1. Phase#1 Instruction
2. Donation Decision
3. Transfer Decision by the Actor#1 (Behavioral Operationalization of Trust), Experiment#3a
4. Return Decision by the Actor#2 (Behavioral Operationalization of Trustworthiness), Experiment#3a
5. Question About the Perception of the Partner
"I think I can trust my partner"
6. Phase#2 Instruction
7. Simplified Trust Game (Repeated Ten Times)
 - a. Random Assignment
 - b. Transfer Decision by the Actor#1 (Behavioral Operationalization of Trust), Experiment#3b
 - c. Return Decision by the Actor#2 (Behavioral Operationalization of Trustworthiness), Experiment#3b
8. The Outcome Information/Payout

References

- Ahlers, Gerrit K. C., Douglas Cumming, Christina Günther, and Denis Schweizer. 2015. "Signaling in Equity Crowdfunding." *Entrepreneurship Theory and Practice* 39(4):955–80.
- Albert, Max, Werner Güth, Erich Kirchler, and Boris Maciejovsky. 2007. "Are We Nice(r) to Nice(r) People?—An Experimental Analysis." *Experimental Economics* 10(1):53–69.
- Alsos, Gry Agnete, and Elisabet Ljunggren. 2017. "The Role of Gender in Entrepreneur–Investor Relationships: A Signaling Theory Approach." *Entrepreneurship Theory and Practice* 41(4):567–90.
- Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100(401):464–77.
- Andreoni, James, and Ragan Petrie. 2004. "Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising." *Journal of Public Economics* 88(7):1605–23.
- Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99(1):544–55.
- Austen-Smith, D., and R. G. Fryer. 2005. "An Economic Analysis of 'Acting White.'" *The Quarterly Journal of Economics* 120(2):551–83.
- Axelrod, Robert. 2009. *The Evolution of Cooperation: Revised Edition*. Basic Books.

- Barker, Jessica L., Eleanor A. Power, Stephen Heap, Mikael Puurtinen, and Richard Sosis. 2019. "Content, Cost, and Context: A Framework for Understanding Human Signaling Systems." *Evolutionary Anthropology: Issues, News, and Reviews* 28(2):86–99.
- Batson, C. Daniel, Jay S. Coke, M. L. Jasnoski, and Michael Hanson. 1978. "Buying Kindness: Effect of an Extrinsic Incentive for Helping on Perceived Altruism." *Personality and Social Psychology Bulletin* 4(1):86–91.
- Bell, R. Greg, Curt B. Moore, and Hussam A. Al-Shammari. 2008. "Country of Origin and Foreign IPO Legitimacy: Understanding the Role of Geographic Scope and Insider Ownership." *Entrepreneurship Theory and Practice* 32(1):185–202.
- Berger, Jonah, and Morgan Ward. 2010. "Subtle Signals of Inconspicuous Consumption." *Journal of Consumer Research* 37(4):555–69.
- Beverland, Michael. B. 2005. "Crafting Brand Authenticity: The Case of Luxury Wines*." *Journal of Management Studies* 42(5):1003–29.
- Bliege Bird, Rebecca, and Eric Alden Smith. 2005. "Signaling Theory, Strategic Interaction, and Symbolic Capital." *Current Anthropology* 46(2):221–48.
- Bird, Rebecca Bliege, Eric Smith, and Douglas W. Bird. 2016. "The Hunting Handicap: Costly Signaling in Human Foraging Strategies." *Behavioral Ecology and Sociobiology* 50(1):9–19.
- Boone, James L. 1998. "The Evolution of Magnanimity." *Human Nature* 9(1):1–21.

- Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgment of Taste*. Harvard University Press.
- Brandts, Jordi, and C. A. Holt. 1992. "An Experimental Test of Equilibrium Dominance in Signaling Games." *The American Economic Review* 1350–65.
- Camerer, Colin. 1988. "Gifts as Economic Signals and Social Symbols." *American Journal of Sociology* 94:S180–214.
- Carroll, Glenn. R., and A. Swaminathan. 2000. "Why the Microbrewery Movement? Organizational Dynamics of Resource Partitioning in the US Brewing Industry1." *American Journal of Sociology* 106(3):715–62.
- Chao, Angela, and J. B. Schor. 1998. "Empirical Tests of Status Consumption: Evidence from Women's Cosmetics." *Journal of Economic Psychology* 19(1):107–31.
- Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *The Quarterly Journal of Economics* 102(2):179–221.
- Chwe, Michael Suk-Young. 1998. "Culture, Circles, and Commercials Publicity, Common Knowledge, and Social Coordination." *Rationality and Society* 10(1):47–75.
- Correll, Shelly. J., and C. L. Ridgeway. 2006. "Expectation States Theory." *Handbook of Social Psychology* 29–51.
- Connelly, Brian L., S. Trevis Certo, R. Duane Ireland, and Christopher R. Reutzell. 2011. "Signaling Theory: A Review and Assessment." *Journal of Management* 37(1):39–67.

- Courtney, Christopher, Supradeep Dutta, and Yong Li. 2017. "Resolving Information Asymmetry: Signaling, Endorsement, and Crowdfunding Success." *Entrepreneurship Theory and Practice* 41(2):265–90.
- Crumpler, Heidi, and Philip J. Grossman. 2008. "An Experimental Test of Warm Glow Giving." *Journal of Public Economics* 92(5):1011–21.
- Cuyper, Ilya R. P., Ping-Sheng Koh, and Heli Wang. 2015. "Sincerity in Corporate Philanthropy, Stakeholder Perceptions and Firm Value." *Organization Science* 27(1):173–88.
- de Haan, Thomas, Theo Offerman, and Randolph Sloof. 2011. "Noisy Signaling: Theory and Experiment." *Games and Economic Behavior* 73(2):402–28.
- Densley, James A. 2012. "Street Gang Recruitment: Signaling, Screening, and Selection." *Social Problems* 59(3):301–21.
- Desjardine, Mark R., Emilio Marti, and Rodolphe Durand. 2020. "Why Activist Hedge Funds Target Socially Responsible Firms: The Reaction Costs of Signaling Corporate Social Responsibility." *Academy of Management Journal*.
- Eddleston, Kimberly A., Jamie J. Ladge, Cheryl Mitteness, and Lakshmi Balachandra. 2016. "Do You See What I See? Signaling Effects of Gender and Firm Characteristics on Financing Entrepreneurial Ventures." *Entrepreneurship Theory and Practice* 40(3):489–514.
- Emerson, Richard M. 1976. "Social Exchange Theory." *Annual Review of Sociology* 2:335–62.

- Farrell, Joseph, and Matthew Rabin. 1996. "Cheap Talk." *The Journal of Economic Perspectives* 10(3):103–18.
- Farthing, G. William. 2005. "Attitudes toward Heroic and Nonheroic Physical Risk Takers as Mates and as Friends." *Evolution and Human Behavior* 26(2):171–85.
- Fehrler, Sebastian, and Wojtek Przepiorka. 2013. "Charitable Giving as a Signal of Trustworthiness: Disentangling the Signaling Benefits of Altruistic Acts." *Evolution and Human Behavior* 34(2):139–45.
- Fehr, Ernst, and Urs Fischbacher. 2002. "Why Social Preferences Matter – the Impact of Non-Selfish Motives on Competition, Cooperation and Incentives." *The Economic Journal* 112(478):C1–33.
- Ferraro, Rosellina, Amna Kirmani, and Ted Matherly. 2013. "Look at Me! Look at Me! Conspicuous Brand Usage, Self-Brand Connection, and Dilution." *Journal of Marketing Research (JMR)* 50(4):477–88.
- Filatotchev, Igor, and Kate Bishop. 2002. "Board Composition, Share Ownership, and 'Underpricing' of U.K. IPO Firms." *Strategic Management Journal* 23(10):941–55.
- Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10(2):171–78.
- Galaskiewicz, Joseph. 1985. *Social Organization of an Urban Grants Economy: A Study of Business Philanthropy and Nonprofit Organizations*. Orlando, FL: Academic Press Inc.

- Gambetta, Diego. 2009. *Codes of the Underworld: How Criminals Communicate*. Princeton University Press.
- Gambetta, Diego. 2011. "Signaling." *The Oxford Handbook of Analytical Sociology*.
- Gambetta, Diego, and Wojtek Przepiorka. 2014. "Natural and Strategic Generosity as Signals of Trustworthiness." *PLOS ONE* 9(5):e97533.
- Gintis, Herbert, Eric Alden Smith, and Samuel Bowles. 2001. "Costly Signaling and Cooperation." *Journal of Theoretical Biology* 213(1):103–19.
- Glazer, Amihai, and Kai Konrad. 1996. "A Signaling Explanation for Charity." *American Economic Review* 86(4):1019–28.
- Goffman, Erving. 1970. *Strategic Interaction*. U. Pennsylvania Press.
- Grafen, Alan. 1990. "Biological Signals as Handicaps." *Journal of Theoretical Biology* 144(4):517–46.
- Grafen, Alan, and Rufus A. Johnstone. 1993. "Why We Need ESS Signalling Theory." *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 340(1292):245.
- Griskevicius, Vladas., J. M. Tybur, and B. Van den Bergh. 2010. "Going Green to Be Seen: Status, Reputation, and Conspicuous Conservation." *Journal of Personality and Social Psychology* 98(3):392.
- Grose, Jonathan. 2011. "Modelling and the Fall and Rise of the Handicap Principle." *Biology & Philosophy* 26(5):677–96.

- Heffetz, Ori. 2011. "A Test of Conspicuous Consumption: Visibility and Income Elasticities." *Review of Economics and Statistics* 93(4):1101–17.
- Henrich, Joseph. 2009. "The Evolution of Costly Displays, Cooperation and Religion: Credibility Enhancing Displays and Their Implications for Cultural Evolution." *Evolution and Human Behavior* 30(4):244–60.
- Hill, Geoffrey E. 1991. "Plumage Coloration Is a Sexually Selected Indicator of Male Quality." *Nature* 350(6316):337–39.
- Joel Hietanen and Joonas Rokka. 2015. "Market Practices in Countercultural Market Emergence." *European Journal of Marketing* 49(9/10):1563–88.
- Hopkins, Ed. 2012. "Job Market Signaling of Relative Position, Or Becker Married to Spence." *Journal of the European Economic Association* 10(2):290–322.
- Kim, Minjae. 2017. "A Man Is Known by His Cup: Signaling Commitment via Costly Conformity." *Academy of Management Proceedings* 2017(1):11239.
- Negro, Giacomo, Michael T. Hannan, and Magali Fassiotto. 2014. "Category Signaling and Reputation." *Organization Science* 26(2):584–600.
- Nöldeke, Georg, and Larry Samuelson. 1997. "A Dynamic Model of Equilibrium Selection in Signaling Markets." *Journal of Economic Theory* 73(1):118–56.
- Ohtsubo, Yohsuke, and Esuka Watanabe. 2009. "Do Sincere Apologies Need to Be Costly? Test of a Costly Signaling Model of Apology." *Evolution and Human Behavior* 30(2):114–23.

- Paik, Anthony, and Vernon Woodley. 2012. "Symbols and Investments as Signals: Courtship Behaviors in Adolescent Sexual Relationships." *Rationality and Society* 24(1):3–36.
- Podolny, Joel M. 2010. *Status Signals: A Sociological Study of Market Competition*. Princeton University Press.
- Ramaswami, Aarti, George F. Dreher, Robert Bretz, and Carolyn Wiethoff. 2010. "Gender, Mentoring, and Career Success: The Importance of Organizational Context." *Personnel Psychology* 63(2):385–405.
- Raub, Werner. 2016. "Hostage Posting as a Mechanism of Trust: Binding, Compensation, and Signaling." *Rationality and Society*.
- Riley, John G. 2001. "Silver Signals: Twenty-Five Years of Screening and Signaling." *Journal of Economic Literature* 39(2):432–78.
- Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. Norton.
- Searcy, William A., and Stephen Nowicki. 2010. *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems: Reliability and Deception in Signaling Systems*. Princeton University Press.
- Simmel, Georg. 1957. "Fashion." *American Journal of Sociology* 62(6):541–58.
- Smith, Eric Alden, and Rebecca L. Bliege Bird. 2000. "Turtle Hunting and Tombstone Opening: Public Generosity as Costly Signaling." *Evolution and Human Behavior* 21(4):245–61.

- Soetevent, Adriaan R. 2005. "Anonymity in Giving in a Natural Context—a Field Experiment in 30 Churches." *Journal of Public Economics* 89(11):2301–23.
- Sosis, Richard, and Candace Alcorta. 2003. "Signaling, Solidarity, and the Sacred: The Evolution of Religious Behavior." *Evolutionary Anthropology: Issues, News, and Reviews* 12(6):264–74.
- Sosis, Richard., and E. R. Bressler. 2003. "Cooperation and Commune Longevity: A Test of the Costly Signaling Theory of Religion." *Cross-Cultural Research* 37(2):211–39.
- Spence, Michael. 1973. "Job Market Signaling." *The Quarterly Journal of Economics* 87(3):355–74.
- Spence, Michael. 1976. "Competition in Salaries, Credentials, and Signaling Prerequisites for Jobs*." *The Quarterly Journal of Economics* 90(1):51–74.
- Spence, Michael. 2002. "Signaling in Retrospect and the Informational Structure of Markets." *American Economic Review* 434–59.
- Su, Weichieh, Mike W. Peng, Weiqiang Tan, and Yan-leung Cheung. 2016. "The Signaling Effect of Corporate Social Responsibility in Emerging Economies." *Journal of Business Ethics: JBE* 134(3):479–91.
- Tan, Jonathan. H. W., and C. Vogel. 2008. "Religion and Trust: An Experimental Study." *Journal of Economic Psychology* 29(6):832–48.
- Tyler, John H., Richard J. Murnane, and John B. Willett. 2000. "Estimating the Labor Market Signaling Value of the GED*." *The Quarterly Journal of Economics* 115(2):431–68.

- Van Vugt, Mark, and Wendy Iredale. 2013. "Men Behaving Nicely: Public Goods as Peacock Tails." *British Journal of Psychology* 104(1):3–13.
- Veblen, Thorstein. 1928. *The Theory of the Leisure Class: An Economic Study of Institutions*. United States: B. W. Huebsch.
- Vohs, Kathleen D., Nicole L. Mead, and Miranda R. Goode. 2006. "The Psychological Consequences of Money." *Science* 314(5802):1154–56.
- Wang, Yajin, and Vldas Griskevicius. 2014. "Conspicuous Consumption, Relationships, and Rivals: Women's Luxury Products as Signals to Other Women." *Journal of Consumer Research* 40(5):834–54.
- Weiss, A. 1995. "Human Capital vs. Signalling Explanations of Wages." *The Journal of Economic Perspectives* 9(4):133–54.
- Wilkins, John S. 2018. "Why Do Believers Believe Silly Things? Costly Signaling and the Function of Denialism." Pp. 109–29 in *New Developments in the Cognitive Science of Religion: The Rationality of Religious Belief, New Approaches to the Scientific Study of Religion*, edited by H. van Eyghen, R. Peels, and G. van den Brink. Cham: Springer International Publishing.
- Willer, Robb. 2009. "Groups Reward Individual Sacrifice: The Status Solution to the Collective Action Problem." *American Sociological Review* 74(1):23–43.
- Willer, Robb., K. Kuwabara, and M. W. Macy. 2009. "The False Enforcement of Unpopular Norms¹." *American Journal of Sociology* 115(2):451–90.

Willis, Paul E. 1977. *Learning to Labor: How Working Class Kids Get Working Class Jobs*.
Columbia University Press.

Xu, Xiaodong, Saixing Zeng, and Hongquan Chen. 2018. "Signaling Good by Doing Good:
How Does Environmental Corporate Social Responsibility Affect International
Expansion?" *Business Strategy and the Environment* 27(7):946–59.

Zahavi, Amotz. 1975. "Mate Selection—a Selection for a Handicap." *Journal of
Theoretical Biology* 53(1):205–14.