

# Whistleblowing and Identity

---

## *Illustrating Identity as an Incentive to Blow the Whistle*

### *Abstract*

The incentives of whistleblowing have not been much in focus in economic literature. This is mainly due to conventional economics having a hard time explaining the incentives of a behavior that seemingly yields no benefits and carries heavy costs. The scarce economic research on whistleblowing incentives suggests that blowing the whistle is driven by monetary rewards and a wish to punish. These suggestions are however not supported by empirics. The importance of identity has been stressed by sociological and psychological research on whistleblowing. Therefore, the aim of this thesis is to propose an economic model on whistleblowing incentives which incorporates identity. By using the social perspective of identity economics, we abandon the individualistic and materialistic view proposed by current economic research. In order to exemplify our identity utility model, we apply the model to several whistleblowing cases.

*Keywords:* social perspective, social categories, identity, identity loss, prescribed behavior, whistleblowing, doubling

Spring 2011

### *Authors*

Elias Mettiji  
Louise Morsing

### *Supervisor*

Yoichi Sugita

## Table of Contents

1.	Introduction .....	2
1.1	Delimitation and Aim of Study .....	4
2.	Method.....	5
3.	Current State of Research .....	6
3.1	Criticism of Current Research.....	7
4.	The Social Perspective .....	8
4.1	Identity Economics .....	9
4.1.1	The Identity Utility Function .....	10
4.1.2	Identity Loss.....	12
4.1.3	Dealing with Identity Loss .....	13
5.	Is Identity a Valid Incentive for Whistleblowing?.....	14
5.1	Evidence of Identity .....	15
5.2	Identity Utility in the Whistleblowing Context .....	16
5.3	General Utility Formulation.....	17
5.4	Whistleblowing Incentive Discussion.....	18
5.4.1	Discussing the Costs of Whistleblowing .....	19
5.4.2	The Sensitivity of Deviating from Prescribed Behavior .....	20
6.	Blowing the Whistle.....	22
6.1	Blowing the Whistle in Ahlmark's Case.....	23
6.2	Blowing the Whistle in Bratt's Case .....	25
6.3	Blowing the Whistle in Magnusson's Case .....	27
6.4	Summarizing Analysis .....	28
7.	Conclusion .....	30
8.	Discussion and Implications .....	31
	References .....	33

## 1. Introduction

What have I gotten myself into? The train is nearing Stockholm. I am on my way there to reveal the illegal business of the Swedish weapons manufacturer Bofors—about how Swedish weapons have been illegally sold to certain countries. For a year I have been the anonymous source in the so-called Bofors-affair and now I shall come forth. I have much to think about. Is it possible to overcome a massive Swedish company, whose very own directors have broken the law, and not just any company, but Sweden's biggest weapons industry, a corner-stone in the Swedish defense industry?/.../What has driven me to go against the company that I have belonged to for so many years? (Bratt 1988: 9)

In May 1985 Ingvar Bratt, an engineer at the Swedish weapons manufacturer Bofors, revealed the occurrence of bribery and the illegal sales of weapons. Bratt's revelations came to be known as the Bofors-affair, which is one of the most media-covered stories in Sweden (Haglunds 2009). The affair shed light on surprisingly corrupt and immoral behavior of a well-known Swedish company, as well as governmental involvement. Just as Bratt had feared, by blowing the whistle on Bofors, he came to suffer severe consequences. Apart from being threatened anonymously and by fellow peers, he eventually felt compelled to leave his job.

The Bofors-affair is one of many international whistleblowing cases, and these types of cases are seen even today. In recent years, inhumane circumstances at a caretaking facility for the elderly were disclosed by the assistant nurse Sarah Wågner, which even resulted in new public legislation.<sup>1</sup> Some other examples are the widespread bribery in the municipal administration of Gothenburg that was exposed thanks to a whistleblower in 2010, and the assistant nurse Peter Magnusson who blew the whistle on allegedly racist behavior at Södertälje hospital.

Typically, the whistleblower experiences heavy costs such as unemployment and physical or emotional retaliation. Simultaneously, they receive little tangible reward. Therefore it is hard to understand the incentives behind this type of behavior from an economic perspective. This is the primary reason for why economic research on the incentives of whistleblowing has been scarce. The latest of this economic research, as presented by Villena and Villena (2010), assumes that these incentives are primarily driven by pecuniary rewards. Moreover, Heyes and Kapur (2008)

---

<sup>1</sup> The commonly used name for this legislation is *Lex Sarah*. In short, it dictates that anyone who is a witness to grave misconduct in the care of the elderly or handicapped individuals has a duty to immediately report this to the authorities.

emphasize that the main incentive to whistleblowing consists of the urge to punish misbehavior. This approach is however not supported by empirics. Empirical studies of whistleblowing have shown that the social dimension and non-monetary incentives are more prominent driving forces. Therefore, our intention is to explore these non-monetary incentives in a social-economic framework—namely the framework of identity economics.

Whistleblowing yields positive externalities to society as it is sometimes the only way to monitor and report misbehavior which in the long run is regarded as being economically inefficient. Because of this, national governments are trying to promote whistleblowing by employing mechanisms such as legal protection, monetary rewards and independent reporting offices. The current European debate on whistleblowing is focusing on the protection of whistleblowers. In 2010, the EU Council drafted Resolution 1729 which specifies the demands on whistleblowing protection. This resolution led to the formation of Recommendation 1916 which urges EU member states to look over their national legislation and its implementation with regards to whistleblowing. In the USA, apart from legal protection, creating incentives for whistleblowers through monetary rewards has been part of the policy debate.

Whilst the above policy process is understandable, what is less clear is on what grounds these policies are made. Both monetary rewards and legal protection have something in common. They both assume that the consequences of whistleblowing are irretrievably negative for the whistleblower. In other words, the policies are made to compensate for the damages and risks that the whistleblower experiences. But is the assumption of the irretrievably negative consequences necessarily accurate? By gaining a better understanding of the incentives and drivers of whistleblowing, the policies would probably better stimulate whistleblowing. Analogously, a policy that aims to alleviate the health risks during an icy wintertime would do best by salting the streets rather than employing more doctors. In essence, we believe that the above policies are more reactive than proactive. In order to become more proactive, we need to understand the core incentives that initiate whistleblowing behavior.

Nobel laureate George Akerlof and Rachel Kranton introduced the theory of *identity economics* (Akerlof and Kranton 2000). This theory incorporates a social perspective into an individual's utility function. More specifically, the social construction of *identity* is regarded as being a primary drive in human behavior. Bringing identity into economic analysis can allow for a new view of several economic problems, ranging from the economics of smoking to game theoretic

frameworks. This paper will thus introduce identity economics in order to extend the concept of a utility function for whistleblowers as proposed by Villena and Villena (2010).

We will demonstrate that this new formulation of the utility function helps to understand empirics better. We believe identity economics to be a powerful toolbox for explaining whistleblowing incentives since empirics gives evidence of identity playing a key role in whistleblowing. Also, several psychological and sociological studies propose identity as an antecedent to whistleblowing. ***Thus the aim of this thesis is to illustrate identity as an underlying incentive to whistleblowing and to capture this incentive in an economic whistleblowing model.***

The structure of this thesis is as follows. Next, we present the aim and delimitation for this thesis. In section two, we present the method used to achieve the aim. Section three contains a brief summary and our criticism of the current state of economic research on whistleblowing. In section four, we introduce the social perspective of whistleblowing in terms of identity economics. Section five verifies the validity of identity as an incentive to whistleblowing. Here we propose a whistleblowing utility function. In section six, three prominent whistleblowing cases are used to illustrate and exemplify our proposed utility function. Section seven concludes this thesis. Finally, in section eight, we discuss the implications of our whistleblowing utility function.

## 1.1 Delimitation and Aim of Study

We will utilize Glazer and Glazer's (1989) definition of a whistleblower since this is the most frequently used definition in academic literature. This definition was also used in the latest economic research on whistleblowing incentives. Glazer and Glazer (1989) define a whistleblower as an individual who (i) acts to prevent harm to others, and not to oneself, while (ii) possessing evidence that would convince a reasonable person. In addition to Glazer and Glazer's (1989) definition, we find it imperative to add that the individual is not reporting anonymously, and thereby the whistleblower is to a varying degree aware of the risks of blowing the whistle.

Many academic papers attempt to give a full picture on the incentives of whistleblowing behavior (see Miceli et al. 1991, and Trevino and Victor 1992). In effect, there are several different theories on the incentives of whistleblowing. To give a full picture is however not our ambition. We will

instead add a new perspective that hopefully will enrich the current economic discourse on whistleblowing. This new perspective might be valuable for policy work as it might promote accuracy. However, policy implementation is not the focus of this study.

Typically, whistleblowing behavior implies heavy costs and seemingly no utility for the whistleblower. This has made it hard for classical economics to explain the incentives behind whistleblowing. The focus of this study is consequently to ***illustrate identity as an underlying incentive of whistleblowing and to capture this incentive in an economic model.***

## 2. Method

In order to illustrate Akerlof and Kranton's (2000) concept of identity as an incentive of whistleblowing, we gathered and utilized (i) empirical case studies (ii) sociological and psychological research and (iii) our own case studies and interviews with a few well-known whistleblowers.

In order to capture this identity-driven incentive in a conceptual model, our starting platform was the latest economic research from Heyes and Kapur (2008) and Villena and Villena (2010) on whistleblowing incentives. Thereon, we proceeded to suggest an alternative to Villena and Villena's (2010) concept of a whistleblowing utility function by utilizing Akerlof and Kranton's (2000; 2005) identity model. The aforementioned case studies and sociological and psychological research was finally used to illustrate and exemplify our model.

We conducted interviews with two prominent whistleblowers in Sweden. The first interviewee was assistant nurse Peter Magnusson who blew the whistle in 2009 on racist and discriminatory behavior at Södertälje Hospital. After being ostracized by several colleagues and superiors at work, he was eventually demoted and sent to work in a different hospital ward.

The second interviewee was cartographer Anders Ahlmark, who is perhaps the most famous whistleblower in Sweden. He blew the whistle in 1977 after his superiors had manipulated nautical charts in order to avoid blame for the grounding of Russian freight ship *Tsesis*. As a result of his whistleblowing, he came to suffer ostracism from his superiors and some of his colleagues. He was eventually forbidden to work as a cartographer by his employer, the Swedish Maritime Administration.

In addition to this, we utilized the written account by weapons engineer Ingvar Bratt who was employed by weapons manufacturer Bofors in the small city of Karlskoga. He describes how he blew the whistle on Bofors's illegal arms trade in 1985, resulting in a public scandal for the company. Karlskoga's primary employment opportunity was at Bofors, which is why Ingvar Bratt faced much resistance and attention in blowing the whistle. His story was of particular importance as it provides an extensive and detailed description of his incentives.

### 3. Current State of Research

Whistleblowing behavior has not been much in focus in economic theory. This is because whistleblowing is not regarded as rational behavior in accordance to *Homo Economicus* as it is seemingly everything but utility-maximizing for the whistleblower. The cost of unemployment and retaliation are too high to yield positive utility (Heyes and Kapur 2008; Villena and Villena 2010). As a result of this, current policy processes focus on the alleviation and compensation of this seemingly unavoidable utility loss.

Heyes and Kapur (2008) propose an economic framework for whistleblowing behavior, aiming to create a more systematic platform in the policy-formation process. They emphasize the importance of understanding the core incentives of whistleblowing. Their paper states that standard economic theory is insufficient to model whistleblowing, and therefore it combines economic theory with evidence from sociology and psychology. From this, the paper proposes three schools of thought, each presenting a distinct motivation for blowing the whistle.

According to Heyes and Kapur (2008), an individual blows the whistle to either cleanse her conscience, to punish others or to act in the interest of society. As a consequence, these different motivators should be taken into account when creating policy on whistleblowing. Heyes and Kapur (2008) propose two tools for stimulating these motivations, namely to impose a punishment fee on misbehavior and to increase the probability for the whistleblower to be heard. For example, if assuming that most potential whistleblowers are punishment-motivated, policy should aim to impose a heavy punishment fee on those who misbehave. On the other hand, if the whistleblower is assumed to be driven by his moral conscience, there need be no punishment fee.

Villena and Villena (2010) reiterate the difficulties in modeling whistleblowing behavior in economics, as the consequences are too grave and the rewards are virtually inexistent. According to Villena and Villena (2010), this explains why little attention has been given to whistleblowing in economic literature. Thus, they incorporate the concept of social norms into the whistleblowing motivational function. Villena and Villena (2010) identify three social norms: one is that of an honest bureaucrat, another of a corrupt bureaucrat, and the third of an honest whistleblower. Each social norm has its own utility function.

The honest whistleblower's utility function is determined by wage  $w$ , a policy-based reward for whistleblowing  $\sigma$  with respect to the probability of detecting corrupt behavior  $\vartheta$  and the population share of corrupt bureaucrats  $p_2$ . The disutility is two-fold, where the first is a transaction cost  $\tau$  with respect to the probability of detecting corrupt behavior  $\vartheta$  and the population share of corrupt bureaucrats  $p_2$ . The transaction cost can be seen as the cost of reporting misbehavior (e.g. ostracism or retaliation by the corrupt, unemployment, or even physical violence). The second disutility is a monitoring cost  $m$  which is a function of the share of corrupt bureaucrats  $p_2$ . The monitoring cost is the cost associated with detecting misbehavior. Villena and Villena (2010) thus propose the following utility function for a whistleblower:

$$\pi_3 = w + p_2\vartheta(\sigma - \tau) - m(p_2)$$

If there is no policy-based reward  $\sigma$  or if  $\sigma < \tau$ , there would be no incentive to be a whistleblower. The critical implication of Villena and Villena's (2010) model is that if there is no intervention by policy-makers such as through a reward for whistleblowing, then there would be little incentive to be a whistleblower. Thus even potential whistleblowers would remain corrupt in a stable equilibrium.

### 3.1 Criticism of Current Research

Heyes and Kapur (2008) bring a social dimension and social categories into the analysis of whistleblowing, thereby acknowledging non-material drivers of blowing the whistle. Their paper is to our knowledge the first to portray whistleblowing in such a way into economic analysis. This is in accordance with many sociological, psychological and anthropological studies. For example, Bjørkelo et al. (2010) conclude that certain personality traits, such as extrovertism and low



agreeableness drive whistleblowing behavior. According to Trevino and Victor (1992) group norms play an important role in inhibiting whistleblowing, and consequently the probability to blow the whistle depends on the strength and the overall acceptance of the norms. Furthermore, there is a wide array of empirical evidence that supports the notion of non-material incentives to whistleblowing, which will be discussed later on.

What is unclear however, is how Heyes and Kapur (2008) came to identify the three aforementioned social categories within whistleblowing. The categories seem to be arbitrarily chosen—why only three categories and why these in particular? Another limitation to the model is that the proposed incentive instruments—the probability to be heard by an enforcement agency and the punishment fee imposed on the corrupt—also seem arbitrarily chosen. Contrary to these suggestions, sociological studies show that punishment is not a primary incentive in whistleblowing. The paper's purpose is to affect policy, but their propositions seem minor with respect to empirics and sociological theory.

The social dimension that is allegedly employed by Villena and Villena (2010) is not apparent. In their model individuals are still primarily driven by material incentives and would easily shift between social categories and norms. For example, an honest whistleblower would instantly turn into a corrupt bureaucrat if this would yield a higher utility. This view is, however, not supported by empirical evidence. For example, in the cases that we have studied, none of the whistleblowers could imagine themselves becoming corrupt, regardless of any possible gains. In other words, although Villena and Villena (2010) claim to employ a social perspective, it is questionable to what extent the model is social. Thus, with respect to empirics, the limitation of their model is that the impact of the social context is overlooked as a material reward is seen as the primary incentive to blow the whistle.

#### **4. The Social Perspective**

In accordance with current sociological and psychological research on whistleblowing, we will apply a social perspective. Sociologist Haglunds (2009) states that:

[Whistleblowing] points to a social dilemma connected with having strong social bonds to different and conflicting social networks or organizations. (Haglunds 2009: 135)

What do we then mean by a social perspective? According to Becker's (2000) theories on social economics, an individual's utility of a certain action is directly dependent on the action of others. For example, an individual would be more likely to vote Democratic if that individual's friends vote that way. Also, the utility of keeping a tidy garden is usually larger in a wealthy neighborhood than in a ghetto environment. This is because a tidy garden is more appreciated by others in the wealthy neighborhood. Therefore, the social perspective consists of an interdependent utility concept—an individual's utility depends both on her own actions *and* the actions of others'. Below, we will demonstrate Akerlof and Kranton's (2000) identity model which assumes this interdependency.

## 4.1 Identity Economics

In 2000, George Akerlof and Rachel Kranton introduced identity economics, which stems from Becker's perspective of social economics. According to Akerlof and Kranton (2000) the social context is imperative in understanding the mechanisms of identities and norms. Their theory brings identity and norms into economic analysis, which gives a new view of many economic problems. Akerlof and Kranton (2000) were the first to incorporate identity into an economic utility function. They were, however, not the first to bring identity into economic analysis. For example, Amartya Sen had already in 1985 stressed the importance of identity as an influence on economic behavior (Sen 1985).

Because of its explanatory power, psychology, sociology and anthropology have adopted identity as a central concept to explain behavior. Consequently, identity economics blend elements of these fields into economics. Akerlof and Kranton (2000) define identity as an individual's *sense of self*. This concept originates from Tajfel and Turner's (1979) social identity theory, in which an individual's sense of self is dependent on her perceived membership of a social group. Also, the individual's attitudes and behavior are strongly affected by the attitudes and behavior of other group members, as in accordance with Becker's social economic perspective.

The identity economic model of behavior begins with social difference. These differences translate into different *social categories*, such as being a man or a woman. To these categories belong a set of *prescribed behaviors*. For example, prescribed behavior for a man would be not to wear a dress, whilst this is strongly associated with prescribed behavior for being a woman. A violation of prescribed behavior evokes anxiety and discomfort in oneself and in others. In many

cases, a man wearing a dress would experience internal discomfort by not feeling true to oneself. Akerlof and Kranton (2000) refer to this discomfort as an *identity loss*.<sup>2</sup>

Each social category has a large set of prescribed behaviors. Thus, the prescribed behaviors for the social category of being a man stem from gender norms such as working in masculine jobs (a male nurse and a male secretary break these norms), and behaving heterosexually (for example marrying a woman) amongst many others.

There is no universal conception about prescribed behavior and social categories and thus there is much dispute about who dictates them in the first place (Akerlof and Kranton 2000). Social categories and prescribed behaviors are to a large extent dependent on context and culture.

Akerlof and Kranton (2010) finally use the term identity as shorthand to bundle together the terms identity, prescribed behaviors and social categories.

#### 4.1.1 The Identity Utility Function

The economic identity utility function, as proposed by Akerlof and Kranton (2000), incorporates identity which is based on social categories. There is a set of social categories  $\mathbf{C}$  such as “man” and “woman”. Any individual  $j$  has a conception of her own belonging to a certain category and all others belonging to the same category. For example, in most cases, a biological male has a conception of himself belonging to the social category “man”, and also has a conception of all others who belong to the same category. Each person  $j$  belongs to an almost infinite number of social categories, such as “man”, “son”, “athlete”, “Swede”, “Stockholmer”, “doctor” and so on. A single category can become more salient in a certain context. For instance, in the context of the Olympic Games, the social category of being “Swedish” will be more important for Swedish citizens than categories such as “Stockholmer”.

Each social category is assigned a set of prescribed behaviors  $\mathbf{P}$ , which is expected in a certain situation. To exemplify, an ideal “janitor” is expected to clean at the workplace, but is not expected to clean when on holiday, whereas an ideal “doctor” is expected to save lives regardless

---

<sup>2</sup> This anxiety can be related to the concept of *cognitive dissonance* which was coined by Festinger in the 1950's. Cognitive dissonance occurs when an individual hold two or more ideas which are related but inconsistent, which causes discomfort (Festinger 1957).

of where he is. A general definition of conforming to prescribed behavior is to *do the right thing* with respect to the social category (Akerlof and Kranton 2000).

Akerlof and Kranton (2000) propose the following utility function:

$$U_j = U_j(\mathbf{a}_j, \mathbf{a}_{-j}, I_j)$$

An individual  $j$ 's utility  $U_j$  is thus dependent on her own actions  $\mathbf{a}_j$ , others' actions  $\mathbf{a}_{-j}$  and  $j$ 's identity or self-image  $I_j$ . Since  $\mathbf{a}_j$  and  $\mathbf{a}_{-j}$  also capture an individual's consumption of goods and services, utility is derived from two sources: material (which is in accordance with standard economic theory) and identity.

The following representation of identity  $I_j$  is proposed:

$$I_j = I_j(\mathbf{a}_j, \mathbf{a}_{-j} | \mathbf{c}_j, \varepsilon_j, \mathbf{P})$$

As shown by the equation, an individual's identity  $I_j$  depends first of all on the assigned social category  $\mathbf{c}_j$ . Identity is also dependent on the degree to which the individual  $j$ 's own characteristics  $\varepsilon_j$  correspond to the prescribed behavior  $\mathbf{P}$  given  $j$ 's assigned category. Thus,  $\varepsilon_j$  is not a shared social construction, but reflects the individual's unique personality. Also, identity is dependent on the extent to which  $j$ 's own actions  $\mathbf{a}_j$  and others' actions  $\mathbf{a}_{-j}$  match the actions dictated by the prescribed behavior  $\mathbf{P}$ . In other words, it is  $j$ 's own actions and the actions of other people (belonging to the same category) that affect  $j$ 's identity, given the category  $\mathbf{c}_j$  that  $j$  belongs to, and ultimately that category's prescribed behavior  $\mathbf{P}$ .

The utility that an individual derives from *belonging* to a certain social category is denoted by the constant  $I_c$ . Increases and decreases in identity utility are called *identity gains* and *identity losses*, respectively. It might seem less problematic to understand that an individual's own actions will affect her identity. For example, a doctor failing to save a patient's life might suffer identity loss as a doctor. What might be less intuitive is how others' action can affect an individual's identity. An example could be that an unprofessional act by a fellow doctor could harm the reputation of all doctors, and consequently also the sense of self for other doctors.

#### 4.1.2 Identity Loss

Akerlof and Kranton (2005) developed their identity model in a principal-agent setting where they formalize the mechanics of how identity loss affects utility.

Recall the general identity function:

$$I_j = I_j(\mathbf{a}_j, \mathbf{a}_{-j} | \mathbf{c}_j, \varepsilon_j, \mathbf{P})$$

Akerlof and Kranton (2005) show how the first parameter  $\mathbf{a}_j$ , i.e. the individual's own actions, affects identity and thus utility. Utility is lost by diverging from a certain prescribed behavior belonging to a social category. This is denoted by the following:

$$t_c |e^*(c) - e|$$

where  $t_c$  represents the importance of conforming to the ideal behavior. It can also be interpreted as the *sensitivity of deviating from prescribed behavior*. Recall that each social category has a set of different prescribed behaviors, and thus to every single prescribed behavior there exists a  $t_{c_x}$ . In other words, each behavior has a varying sensitivity or importance of conformity.  $e^*(c)$  is denoted as the prescribed behavior that is related to a social category.  $e$  denotes a certain action or behavior in a certain setting. The identity parts of the utility function can thus be represented as follows:

$$I_c - t_c |e^*(c) - e|$$

As can be seen, an individual will experience an identity loss if her behavior  $e$  deviates from the ideal behavior  $e^*(c)$  as dictated by her social category. To illustrate, the social category of being a military professional is distinct from that of being a civilian. Within this military category, there exist several prescribed behaviors which, within sociology, are described as being war oriented, masculine and differentiated in structure and culture from civilian society (Moskos et al. 2000). These prescribed behaviors are enforced by strong rituals and traditions. To break from these prescribed behaviors  $e^*(c)$  by acting cowardly or crying  $e$  during an operation will cause an identity loss with magnitude  $t_c$ . Why would then that individual indeed act cowardly or cry? Let us recall that each individual belongs to a wide array of social categories and thus several identities which may be more salient in certain situations. The situation of a soldier killing a child

soldier resulting in him to start crying may signal the salience of another more dominant identity such as that of a caring father. Another way of interpreting this is that the military identity is affected by the father identity in a way that the magnitude of the military  $t_c$  changes.

Up to this point we have, by using Akerlof and Kranton (2000; 2005), shown that an individual gains utility by belonging to a certain social category  $I_c$ , which is equivalent to belonging to an identity. Also, an individual's utility is affected by conforming to or breaking from the prescribed behaviors of a certain identity. As previously stated, an individual's identity utility is not only affected by her own actions but also by the actions of others  $\mathbf{a}_{-j}$ . We have above illustrated Akerlof and Kranton's (2000) mechanics of identity loss from the individual's own actions. This identity loss can be interpreted as a reduction of identity utility  $I_c$  of the size  $I_s$  where  $s$  denotes "self". This mechanism was introduced above, and can be summarized below as:

$$I_{self} = I_s(\mathbf{a}_j | \mathbf{c}_j, \mathbf{P}) = t_c |e^*(c) - e_s|$$

In a similar fashion, the mechanics of identity loss from the actions of others can be written as:

$$I_{others} = I_o(\mathbf{a}_{-j} | \mathbf{c}_j, \mathbf{P}) = t_c |e^*(c) - e_o|$$

#### 4.1.3 Dealing with Identity Loss

The reason why an identity loss is experienced is that the violation of prescribed behavior of a group member will cause an inner anxiety for himself and other group members (Akerlof and Kranton 2000). This is due to the fact that these individuals have internalized these prescribed behaviors. When an individual internalizes certain rules of behavior, she genuinely believes in the rules, as opposed to being forced or enforced to follow these rules. Thus, in accordance to psychoanalytic theory, the anxiety that arises stems from repressed emotions and wishes that are aroused when internalized rules are violated by oneself and others (Thomas 1996). The process of internalizing and having internalized these prescribed behaviors is what Akerlof and Kranton (2000) denote as *identification*, which is the identification of oneself to a social category, its rules and ultimately an identity.

An illustrative example of how identity loss  $I_o$  occurs in reality is described in an anthropological study by Coyne and Mathers (2009) who examine puzzling phenomena associated with female

genital mutilation. Elderly circumcised women who have themselves experienced the physical pain of mutilation often function as promoters for this tradition. This phenomenon is counterintuitive, as they promote something so painful and lethal. Coyne and Mathers (2009) explain this to be due to the identity loss  $I_o$  the old women would suffer if other women were to deviate from this internalized prescribed behavior of circumcision.

Akerlof and Kranton (2000) propose that it is possible to restore or minimize any identity loss, be it from the actions of others or your own actions. In the case of  $I_s$ , the individual can restore any self-inflicted identity loss by conforming to prescribed behaviors as  $e^*(c) - e_s = 0$  and thus  $I_s = 0$ . When an individual suffers identity loss  $I_o$  due to others' deviation from prescribed behavior, the individual can *react* and restore this identity loss. This type of reaction can be seen in a variety of different situations according to Akerlof and Kranton (2000). For example, children in school playgrounds who behave differently and not in accordance to certain norms are taunted or mocked for this. This mocking ostracizes the individual and restores the identity loss  $I_o$  for the others as they demonstrate their distancing from the deviating individual.

This reaction comes at a cost, however, wherein the costs could be physical, emotional or even material. For example, the individual can be sued for discrimination, counter-ostracized and even punished physically as a retaliatory response. This is what Akerlof and Kranton (2000) propose as a trade-off an individual considers for restoring any identity loss inflicted by others. Thus, if the costs of restoring the identity loss are too high, the individual may have to refrain and accept the anxiety of identity loss  $I_o$ .

This interdependence on others and their actions is what signifies the social framework of the identity economic model. In a similar fashion, we believe that the mechanics of whistleblowing can be illustrated from this social perspective.

## 5. Is Identity a Valid Incentive for Whistleblowing?

Now that we have introduced Akerlof and Kranton's (2000) theory on identity economics, we will explore whether this concept is valid in explaining whistleblowing. In order to do this, we have studied whistleblowing cases as well as sociological and psychological literature on whistleblowing. In addition to this, we have gathered our own empirical material by interviewing two prominent whistleblowers.

## 5.1 Evidence of Identity

To begin with, sociology tends to include identity as a valid concept in explaining the incentives of blowing the whistle. For example, Schwartz et al. (1969) describe whistleblowing as a result of an individual's need to affiliate to a certain reference group. More specifically, when having witnessed misbehavior that goes against the norms of that reference group, an individual is driven to blow the whistle. This need for affiliation and to react against misbehavior can be translated into Akerlof and Kranton's (2000) theory on identity and prescribed behavior.

Moreover, Haglunds (2009) describes whistleblowing as a *social identity transformation process*. This process implies an identity change for the whistleblower from being a colleague to being considered a threat to the community.

According to Haglunds (2009), an individual who blows the whistle on the employer does not do so because of hate or revenge. This is supported by Near and Miceli (1996) who state that whistleblowers are generally not driven by a desire to harm others. Instead, Haglunds (2009) explains that the reason for whistleblowing is due to the fact that an individual belongs to several social groups and organizations. By belonging to several social groups, an individual is bound by several different and often conflicting social norms and rules. Thus, when an employer misbehaves, an employee may feel the urge to set things right with reference to a social group outside the misbehaving firm (Haglunds 2009).

Glazer (1983) exemplifies this urge to set things right in the case study of Dr. Grace Pierce, a physician at Ortho Pharmaceutical Corporation. Dr. Pierce explained her whistleblowing as caused by "my responsibility to my profession first and to my corporation second" (Glazer 1983: 35). Her employer Ortho Pharmaceutical had insisted on withholding significant health risks of a new drug product, but as a physician Dr. Pierce could not obey and finally blew the whistle.

In identity terms, Dr. Pierce's strong identification with the social group of a physician was decisive for her to blow the whistle. By not conforming to her employer's request to approve a risky drug, she conformed to the prescribed behaviors of the social category of a physician. In terms of Akerlof and Kranton's (2000) identity model, by conforming to the prescribed behavior of a physician, Dr. Grace Pierce can be said to have avoided identity loss  $I_s = t_c |e^*(c) - e_s|$  as  $e^*(c) = e_s \rightarrow I_s = 0$ .



Another example of conflicting social groups is when weapons engineer Ingvar Bratt found himself in the moral crossroads of the Swedish Environmental Party (*Miljöpartiet de Gröna*) and his employer Bofors:

I was still very much a Bofors man. But what about my commitment to the Third World and the Environmental Party's ideas? This tormented me/.../I tell them nothing about Bofors selling to forbidden countries, but within me, solidarity towards a new group of people grows stronger [namely, the Environmental Party]. (Bratt 1988: 59)

Bratt was tormented by anxiety caused by participating in the illegal arms trade (Bratt 1988). This was because he knew that this was against the prescribed behaviors not only for the category of all engineers but for the category of all Swedes (Bratt 1988). As opposed to the case of Dr. Pierce, Bratt felt the angst of identity loss  $I_s = t_c |e^*(c) - e_s|$  for quite some time until he finally chose not to conform any longer. He for example deliberately avoided any work involving the illegal arms trade and reported himself ill in order to avoid any further anxiety or identity loss  $I_s$ . In other words, he conformed to prescribed behavior  $e^*(c)$ .

Psychologist Alford (2001) further supports the notion of several social memberships causing conflicts and dilemmas as whistleblowers "...inhabit worlds in which different moral languages are spoken" (Alford 2001: 72).

As we have shown in this section, identities and prescribed behaviors do not seem foreign to the whistleblowing context. Sociology and psychology have since long regarded identity as a valid instrument for understanding the incentives to blow the whistle. Empirics also seem to support this notion.

## 5.2 Identity Utility in the Whistleblowing Context

By employing Akerlof and Kranton's (2000) theories on identity economics, we propose that a potential whistleblower gains identity utility  $I_c$  by belonging to a social category—be it the category of an engineer or a cartographer and so on. This utility can be affected by the potential whistleblower's own actions and the actions of others at the workplace. Thus, she will experience utility loss  $I_s$  if she deviates from prescribed behavior and utility loss  $I_o$  when her colleagues deviate.

As was introduced and will be further seen in empirics, whistleblowers avoid  $I_s$  by conforming to prescribed behavior. This supports our proposition of the existence of the constant utility  $I_c$  in a whistleblowing context. This is because there is an incentive to preserve or maximize this identity utility. We propose that whistleblowing is a response to regain the identity loss  $I_o$  that the whistleblower suffers from her colleagues deviating from prescribed behavior. This whistleblowing response is derived from Akerlof and Kranton's (2000) proposition to react to restore identity loss.

Below we propose a general utility function of a whistleblower by employing Akerlof and Kranton's (2000) identity parameters.

### 5.3 General Utility Formulation

We propose the following whistleblower utility function:

$$U = I_c - I_s - I_o + W$$

As can be seen, the parameters  $I_c, I_s, I_o$  are those that were developed by Akerlof and Kranton (2000). In addition to these parameters, we propose  $W$  that denotes whistleblowing as a possibility to restore the identity loss  $I_o$  caused by others' misbehavior. More specifically,  $W$  is  $I_o - K$ , where  $K$  denotes the cost of blowing the whistle.  $W$  is activated whenever  $I_o > 0$  and when  $I_o > K$ .

Recall that  $I_s = t_c |e^*(c) - e_s|$  and  $I_o = t_c |e^*(c) - e_o|$  and  $W = (I_o - K | I_o > 0, I_o > K)$ . Thus the utility function can be rewritten as follows:

$$U = I_c - t_c |e^*(c) - e_s| - t_c |e^*(c) - e_o| + (I_o - K | I_o > 0, I_o > K)$$

Remember that  $W$  gives an opportunity to restore lost identity utility caused by others' actions ( $I_o = t_c |e^*(c) - e_o|$ ), though at a cost  $K$ .

In relation to Villena and Villena's (2010) concept of the *transaction cost*  $\tau$ , we have chosen to integrate this cost into  $K$ . In our model  $K = \tau + \varphi$ , where  $\tau$  represents costs such as being

ostracized, fired or physically harmed after blowing the whistle. A *presentation cost*  $\varphi$ , on the other hand, is our own conception of a cost associated with presenting or transferring the information of misbehavior. In other words, it is the whistleblower's effort needed to tell her side of the story. We became aware of this type of cost after having reviewed whistleblowing cases. This type of cost will be discussed later. Also, we have chosen to eliminate the monitoring cost  $m(p_2)$  which was proposed in Villena and Villena's (2010) model. This will also be discussed later.

## 5.4 Whistleblowing Incentive Discussion

As can be seen in our proposed whistleblowing function, an incentive for blowing the whistle will occur whenever the costs of blowing the whistle  $K$  are less than the restored identity loss  $I_o$ . Formally, this can be written as when  $I_o > K$ . The whistleblowers we have studied must thus have perceived  $I_o > K$ , if they are to follow our model. But what makes these individuals unique? In other words, why did no one else blow the whistle if they belonged to the same social category and shared the same prescribed behaviors?

In the case of assistant nurse Magnusson, his colleagues knew that racist and discriminatory behavior deviated from the prescribed behavior of hospital staff. But according to Magnusson, the misbehavior did not cause much concern or anxiety for everyone. For instance, a nurse colleague admitted that her racist and discriminatory actions were wrong, but at the same time claimed that "she had no problem with being a racist" (interview Magnusson).

In the case of cartographer Ahlmark, several of his colleagues were unhappy with the Maritime Administration manipulating the nautical charts: "Many of us knew that it was wrong, but people were afraid to openly complain—they didn't want to risk their jobs" (interview Ahlmark).

The patterns of others not reacting against misbehavior described by Magnusson and Ahlmark were also seen in several other cases. We propose that there are two reasons that prevent individuals who suffer identity loss  $I_o$  from blowing the whistle. The first is that the costs of blowing the whistle  $K$  is perceived as being too high in relation to the potential restoration of identity loss  $I_o$ . The second reason is that individuals are not equally sensitive to the deviation from prescribed behaviors implying a varying magnitude of identity loss  $I_o$ .

In the following, we will take a closer look at these obstacles to whistleblowing.

#### 5.4.1 Discussing the Costs of Whistleblowing

Recall that  $K = \tau + \varphi$ . While the presentation cost  $\varphi$  is well defined as the whistleblowers' effort needed to tell their side of the story, the rest of  $K$ , namely the transaction cost  $\tau$  consists of a broader set of costs. So far, we have proposed these costs as being associated with losing the job, ostracism and physical and emotional retaliation. In general,  $\tau$  consists of any costs that are related to the response by others to the whistleblowing action. This is why we choose to distinguish the presentation cost  $\varphi$  from the transaction cost  $\tau$ .

We have chosen to split  $\varphi$  into a *rhetorical constraint* and a *medium constraint*. We define the rhetorical constraint as the “the frustration and difficulty in explaining the apparently obvious misbehavior to the public” (interview Ahlmark). In other words, as most whistleblowers are not media trained, substantial effort is required to inform about misbehavior in an understandable manner. We define the medium constraint as the emotional stress caused by “the watchful eyes of the media” (interview Ahlmark). In other words, it is the discomfort about being exposed and scrutinized in public. These costs will be clear in the below cases.

As previously mentioned, we have chosen to eliminate the cost of detecting misbehavior, the monitoring cost  $m(p_2)$  which was proposed by Villena and Villena (2010). The reason is that this cost does not seem to be supported by empirics, as the misbehavior seemed to be visible for the majority at the workplace. This visibility differentiates whistleblowing situations from wrongdoings discovered by financial auditors. The latter puts in considerable time and effort to detect wrongdoings, and consequently carries heavy monitoring costs. Take note that our definition of a whistleblower does not include the auditor's work role.

##### 5.4.1.1 Perception of the Costs of Whistleblowing

Are the costs of whistleblowing  $K$  fully known to the potential whistleblower? One reason for why most people choose not to blow the whistle is that they perceive the costs as being too high (Haglunds 2009). This perception may build from earlier whistleblowing cases that have been portrayed in media showing the intense hardships and obstacles that whistleblowers have to face (Haglunds 2009). Thus, even though the cost  $K$  is seldom fully known before blowing the whistle, there is often a close to accurate conception of its magnitude (interview Magnusson, interview Ahlmark).

Evidence shows that knowledge about the magnitude of whistleblowing costs in a specific case develops over time, and that this evolution of cost is one important reason for potential whistleblowers to back out (interview Magnusson, Haglunds 2009). For example, assistant nurse Magnusson pointed out that initially he was part of a small group who reacted to the misconduct. After some time, however, Magnusson stood alone, and although those few others still supported him, they could not proceed to be a part of the whistleblowing process. Magnusson explained that as the costs (ostracism and risk of unemployment) became more noticeable, his colleagues were not able to bear them (interview Magnusson).

Conclusively, in terms of our model, this inability to whistleblow can be captured by the interaction of  $I_o$  and  $K$ . In other words, the perception of  $K$  had grown too large to regain the identity loss  $I_o$ .

As we have proposed, the other obstacle of whistleblowing in our model is due to the varying sensitivity of deviation from prescribed behavior. This will be discussed next.

#### 5.4.2 The Sensitivity of Deviating from Prescribed Behavior

Although there is consensus of what the prescribed behaviors dictate, some people seem not to care when these prescriptions are violated. This sensitivity, as proposed by Akerlof and Kranton (2005) is captured by  $t_c$ . Recall that identity loss from others' actions is denoted by  $I_o = t_c |e^*(c) - e_o|$ . Thus, if  $t_c$  is low or zero, little to no identity loss  $I_o$  is felt.

Generally, this was seen in all whistleblowing cases as the whistleblowers' colleagues seemed to care about the misbehavior to a varying degree. One prominent example of a low sensitivity was when Ahlmark's colleague "with ease sacrificed forty years of professional integrity by cheating and covering up the misbehavior at the Maritime Administration" (interview Ahlmark). Another example is that of Magnusson's nurse colleague who "had no problem with being a racist" (interview Magnusson). Obviously, both of these individuals seemed to have suffered little anxiety from deviating from prescribed behavior which translates into them having a low  $I_o$ . According to our model, their magnitude of  $t_c$  can be said to be low. In our model, the magnitude of  $t_c$  determines the magnitude of identity loss of both  $I_o$  and  $I_s$ . Consequently, an individual with a low magnitude of  $t_c$  will be less sensitive to both own deviations and others' deviations from prescribed behavior.

Ultimately, an individual may misbehave because the deviation of prescriptions would yield a low identity loss  $I_s$ . An individual who gravely misbehaves is translated into a large magnitude of  $|e^*(c) - e_o|$ . If this individual's  $I_s$  is low, then this implies a low magnitude of  $t_c$ . This in turn means that the individual is insensitive to others' deviations and would suffer little identity loss  $I_o$ . This low  $I_o$  finally implies that the individual will have little incentive to react or blow the whistle.

The concept of a low  $t_c$  suggesting some individuals of being less sensitive to deviations from norms and prescriptions finds support in the whistleblowing literature. This will be reviewed next.

#### 5.4.2.1 *Sensitivity of Deviation and the Concept of Doubling*

What explains the differences in sensitivity of deviations from prescriptions,  $t_c$ ? Alford (2001) suggests that some individuals have the ability to keep different parts of their life—work, family, leisure—mentally apart. He refers to this ability as *doubling* and suggests that whistleblowers are deprived of this ability. The concept of doubling was originally introduced by psychologist Lifton (1986). He discovered this ability when exploring the psychological mechanisms of Nazi doctors in Germany who performed cruel experiments on Jews at work, whilst living a peaceful family life in their spare-time.

In our whistleblowing model, the ability to double can thus be captured by  $t_c$ . We define a high ability to double as being equal to having a low magnitude of  $t_c$ . An individual with a very low  $t_c$  will suffer little identity loss from deviating from prescribed behavior. In the extreme case, when  $t_c = 0$ , the individual will suffer no identity loss when deviating, and is thus highly capable of doubling. On the other hand, an individual with a very high value of  $t_c$  will experience a large identity loss when deviating from the prescribed behavior, as  $I_s = t_c |e^*(c) - e_s|$ . Consequently, we propose a whistleblower to be an individual with a low capacity to double and thus a high value of  $t_c$ .

#### 5.4.2.2 *Multiple Identities and Doubling*

We have proposed that the ability to double is captured by  $t_c$ . But is  $t_c$ , the sensitivity of deviation from prescribed behavior, fixed or can it be changed?

Recall that Akerlof and Kranton's (2000) model assumes an individual belonging to an almost infinite number of identities which we exemplified as being a man, Swede, Stockholmer, doctor etc. Myers et al. (1991) describe how different identities are interrelated in a complex manner. They explain that the norms of one identity can interfere and alter the norms and values of another. This interdependency of identities has also been problematized by Roccas and Brewer (2002) who developed the concept of social identity complexity, which describes how an individual's different social identities affect each other.

By utilizing the ideas of Myers et al. (1991) and Roccas and Brewer (2002), we thus propose that  $t_c$  of one identity can be affected by another identity and its prescribed behavior. In other words,  $t_c$  is a function of the set of social categories  $\mathbf{C}$  that an individual belongs to. For example, Ingvar Bratt's most salient identity was initially that of an engineer at Bofors. When first discovering that Bofors illegally sold weapons to certain countries, he acknowledged that this was not right, i.e. that it was a deviation from prescribed behavior. However, he did not react as strongly as he did later on, after having internalized the values of the Environmental Party (Bratt 1988). This can be interpreted as if the magnitude of Bratt's  $t_c$  as a Bofors engineer had been affected by the Environmental Party identity. In other words, the magnitude of  $t_c$  had increased and consequently, the deviation from prescribed behavior yielded a higher identity loss. It is important to note that it is the sensitivity of deviating from prescribed behavior  $t_c$  that is affected, and not the prescribed behaviors themselves.

Another example of how different identities affect each other is that of Magnusson's case. He thought that he might have been extra sensitive to the racist behavior at his workplace as he had two adopted sisters from India (interview Magnusson). Thus, his identity as an assistant nurse can be said to have been influenced by his identity as a protective elder brother.

## 6. Blowing the Whistle

Now that we have elaborated a whistleblowing utility function and discussed its parameters, we intend to illustrate our model. This will be done by applying the model to three prominent whistleblowing cases, namely those of Anders Ahlmark, Ingvar Bratt and Peter Magnusson.

## 6.1 Blowing the Whistle in Ahlmark's case

When Ahlmark was faced with his colleagues and superiors manipulating charts, he suffered an identity loss  $I_o$ , since manipulating and cheating  $e$  was contrary to the right thing to do  $e^*(c)$ . The reaction in Ahlmark's case was to blow the whistle so as to restore  $I_o$  at a cost  $K = \tau + \varphi$ . In Ahlmark's case,  $\tau$  consisted of being ostracized by other employees and by his superiors, of being forbidden to work as a cartographer and ultimately of being demoted from his work.

Ahlmark stressed that he experienced great difficulty in presenting his version of the story without appearing as if he was acting in pure self-interest or punishing his organization. More specifically, it was painful to be questioned by many and to present a clear case to the media. Detailed technical knowledge was required to fully understand the misbehavior and therefore, Ahlmark had to put a lot of effort and time in making the case understandable to people both outside and inside of the organization. In other words, Ahlmark experienced a clear rhetorical constraint as part of the presentation cost  $\varphi$ . Ahlmark expressed that the enormous media attention was greatly unpleasant since he "wasn't media trained" (interview Ahlmark). Thus, Ahlmark also suffered a medium constraint.

Ahlmark explained that he witnessed several cases of grave misbehavior at the Maritime Administration. However, after his first and only whistleblowing he refrained from reporting as he was unwilling to be perceived as a "public trouble-maker" (interview Ahlmark). This can be interpreted as if his presentation cost  $\varphi$  had become infinitely large after his first whistleblowing. Thus, substantial media exposure will cause the presentation cost  $\varphi$  to increase and prevent the whistleblower from ever blowing the whistle again.

How did then Ahlmark view the tradeoffs between restoring his  $I_o$  and his heavy cost of whistleblowing  $K$ ? Clearly, Ahlmark's identity loss  $I_o$  was larger than his whistleblowing cost  $K$ . His large magnitude of  $I_o$  stemmed from his large  $t_c$  as a cartographer. His ability to double the identities of a cartographer and a loyal employee was thus very low. He explained his inability of doubling by saying that he "couldn't live with something that is pretty much criminal!" (interview Ahlmark). Unlike the Nazi doctors (Lifton 1986), Ahlmark could not at the end of the day go home and live with what had happened at work.



Although Ahlmark admitted that his costs  $K$  were considerable, he stated that his professional commitment as a cartographer and as a specialist in his field was too valuable for him to sacrifice. An example of how he valued his cartographer identity was apparent when a cartographer colleague of his decided to manipulate the nautical charts. Ahlmark perceived his colleague's behavior to be perplexing and that his colleague had "sold his forty-year old professional honor" (interview Ahlmark).

In the end, Ahlmark was clearly aware of the costs  $K$  but his identity loss  $I_o$  as a cartographer would have been too high by not reacting to the misbehavior. He confirmed that by not blowing the whistle would have caused him an anxiety that he "would not be able to live with" (interview Ahlmark).

Let us go through Ahlmark's entire whistleblowing process by using the model:

$$U = I_c - I_s - I_o + W$$

which could be rewritten as:

$$U = I_c - t_c |e^*(c) - e_s| - t_c |e^*(c) - e_o| + (I_o - K | I_o > 0, I_o > K )$$

Ahlmark gained utility from his identity as a cartographer  $I_c$  which was more salient than his identity as a loyal employee. He avoided identity loss  $I_s = t_c |e^*(c) - e_s|$ , by refusing to manipulate the charts  $e_s$  which followed prescribed behavior  $e^*(c)$ . Hence  $e^*(c) = e_s$  which implies that  $I_s = 0$ . As his ability to double was low, his  $t_c$  was high. Thus, when others manipulated the charts, Ahlmark suffered substantial identity loss  $I_o$ . As  $I_o > 0$ , Ahlmark had an incentive to blow the whistle  $W$ . And finally, as Ahlmark explained, the anxiety or identity loss  $I_o > K$ , and thus he activated  $W$  and blew the whistle. In the end, his utility function can be illustrated as follows:

$$U = I_c - K$$

Ultimately, it is critical to note that Ahlmark's utility from blowing the whistle was larger than the utility obtained from conforming and remaining quiet:

$$I_c - K > I_c - I_s - I_o$$

## 6.2 Blowing the Whistle in Bratt's Case

Before becoming involved with the Environmental Party, Bratt suffered little identity loss  $I_o$ , since his  $t_c$  was still low:

It seemed okay to break the rules and I turned a blind eye. The company's attitude meant more to me and my actions than following the law to the letter. (Bratt 1988: 50)

However, after having undergone identity therapy and internalized the values of the Environmental Party (Bratt 1988), his  $t_c$  as a Bofors engineer increased in magnitude.

This supports the notion of a complex interdependence of norms and different identities which was described by Myers et al. (1991) and Roccas and Brewer (2002). Bratt explained that as he became more knowledgeable of the Environmental Party's perspective on illegal arms trade, he came to feel an increasing affiliation with the Party (Bratt 1988). It can be discussed to what extent this affiliation stemmed from the strong opinions within his family against arms trade. In any case, his view on the gravity of illegally selling weapons became stronger and thus ultimately increased his magnitude of  $t_c$ . This increase greatly amplified his identity loss  $I_o$  to the point that Bratt felt that he had to react. An example of Bratt's inability to double is seen in the below quote:

I realized that it was unbearable to work for Swedish weapons export during the day and during the evenings work against this as an Environmental Partyist. (Bratt 1988: 58)

What can be said about Bratt's  $K$ ? Bratt devoted considerable time on evaluating the magnitude and nature of the cost that he risked to face if he blew the whistle on his employer. One possible part of  $K$  was the risk of getting the population of Karlskoga against him. He was aware that: "the whole city would fall if Bofors fell" (Bratt 1988: 122) and thus would drive many to think badly of him to say the least. He thus experienced a loyalty conflict between his "conscience and the interests of the city of Karlskoga" (Bratt 1988: 122). Other costs that he reflected on before blowing the whistle were:

/.../how badly Bofors would strike back, if my prospective job as a teacher would be compromised, how badly my family would be affected and if I would be able to deal with the social pressure/.../(Bratt 1988: 81)

These are clearly a part of the transaction cost  $\tau$  that he evaluated. Concerning the other part of  $K$ , namely the presentation cost  $\varphi$ , Bratt was initially very reluctant to go public: “TV’s impact is so vast and the reactions to me can be strong” (Bratt 1988: 97). In other words, he experienced a considerable medium constraint, i.e. he did not want to appear as a public figure. Initially, this medium constraint prevented him from blowing the whistle  $\mathcal{W}$ . Bratt mentioned that: “the mere thought of sitting in front of the TV cameras and trying to explain my actions in front of the whole Swedish population makes me weak in the knees” (Bratt 1988: 97). Thus, he simultaneously experienced a rhetorical constraint.

Eventually, however, Bratt made the decision to whistleblow  $\mathcal{W}$ . In accordance to our whistleblowing model, we propose two main reasons for his final decision to come forth. Firstly, his perception of  $K$  decreased in magnitude as he began to regard the threatening costs as whistleblowing as: “a challenge and an opportunity to face my deep fear for conflict and confrontations/.../to come forth is not just my obligation but an opportunity to grow” (Bratt 1988: 97). Secondly, his identity loss  $I_o$  stemming from the illegal arms trade at Bofors, continued to increase with his increasing identification with the Environmental Party.

In a similar fashion to Ahlmark’s case, Bratt’s whistleblowing process can be illustrated by the model:

$$U = I_c - t_c|e^*(c) - e_s| - t_c|e^*(c) - e_o| + (I_o - K|I_o > 0, I_o > K)$$

As a Bofors engineer, Bratt gained identity utility  $I_c$ . Firstly, he felt no identity loss  $I_s$  by refraining from work activities related to the illegal sales. His initially small magnitude of  $I_o$  was due to his initially low  $t_c$ . When his identity as an engineer became influenced by the identity and values of the Environmental Party, his  $t_c$  for the illegal arms trade increased. He thus eventually experienced a large identity loss  $I_o$ . This activated his incentive to whistleblow  $\mathcal{W}$ , and even though the obstacles of  $K$  were high, they were not higher than the restored identity loss  $I_o$ . As with Ahlmark, Bratt gained more utility from whistleblowing than living with the identity loss:

$$I_c - K > I_c - I_s - I_o$$

### 6.3 Blowing the Whistle in Magnusson's Case

Magnusson stated that he felt very uncomfortable about the discriminatory behavior towards patients directed by some of his colleagues. This can be interpreted as him feeling an identity loss  $I_o$ . Magnusson expressed this discomfort or identity loss  $I_o$  by saying:

The intensive-care unit—it's not a game, people are dying there! It's unacceptable that *intensive-care hospital staff* behave in a racist and prejudiced manner. Maybe this behavior wouldn't have been as bad in a car repair shop or at McDonald's. (interview Magnusson)

As can be seen, the racist behavior was a deviation from prescribed behavior  $e^*(c)$  for the social category of hospital staff. Magnusson strongly refrained from taking part in this, thereby avoiding an identity loss  $I_s$ . One important reason for why Magnusson so strongly opposed racism and discrimination was that he had two adopted foreign-born sisters. Magnusson said that he had witnessed his sisters being subject to racism, and this led him to assume a defensive stance against racism and discrimination. In identity terms, Magnusson's identity as a protective elder brother affected his identity as an assistant nurse. This resulted in his sensitivity for discrimination  $t_c$  to be high. Magnusson's case thus provides yet another illustrative example of the interdependence of identities as was proposed by Myers et al. (1991) and Roccas and Brewer (2002).

Magnusson's high  $t_c$  implied an incentive to avoid identity loss  $I_s$  and  $I_o$ . As described above, Magnusson avoided  $I_s$  by refraining from misbehaving himself, as  $I_s = t_c |e^*(c) - e_s|$  and  $e_s = e^*(c)$ . After having failed to affect change by discussing the discrimination problems at the intensive care unit, Magnusson began to evaluate the alternative to turn to the media. In other words, he assessed the alternative to blow the whistle  $W$  and its costs  $K$ .

Initially, Magnusson did not think that he risked losing his job, or that there would be much resistance. By persisting to affect change at his workplace, Magnusson got his first taste of the types of costs. For example, certain colleagues ceased to greet him, and soon thereafter he realized that even his job was at risk. Since Magnusson owned a private parachuting company, he felt that he had an economic safety-net in the case of losing his job at the hospital. However, he

feared being ostracized and being put in an extremely uncomfortable social situation (interview Magnusson). Thus, he perceived the transaction cost  $\tau$  to be considerably high.

When it comes to the presentation cost  $\varphi$  he explained that he had no problem in expressing himself as he considered himself a verbally talented individual. This implies that he had an insignificant rhetorical constraint. When it comes to the medium constraint, however, Magnusson felt a heavy anxiety about appearing in public television: “In the taxi on my way to the television studio I intensively thought about how I could bail” (interview Magnusson). Also, Magnusson mentioned that “it may have seemed that I was highly keen to appear in the media, but it was the complete opposite!” (interview Magnusson). Consequently, Magnusson perceived his medium constraint as high.

In the end, his fears of  $K$  were not enough to prevent him from blowing the whistle to restore the identity loss  $I_o$ : “I would have done it again even after having actually endured the costs” (interview Magnusson). His utility function can in a similar fashion be depicted as in the previous two cases:

$$I_c - K > I_c - I_s - I_o$$

## 6.4 Summarizing Analysis

By analyzing empirical cases and whistleblowing literature, we have claimed that identity as a concept is a key parameter in the incentives of whistleblowing. Thus, by adding identity to the utility function as proposed by Akerlof and Kranton (2000), we have created a whistleblower utility function. Each identity has a set of prescribed behaviors and others deviating from these behaviors yield an identity loss  $I_o$  to the potential whistleblower. By blowing the whistle  $W$ , she can restore this lost utility at a cost  $K$ .

Individuals have several identities that affect each other in terms of the sensitivity of deviating from prescribed behaviors. Through doubling, individuals can handle conflicting norms of these different identities. This ability to double is equivalent to a low value of  $t_c$ . A critical part of our whistleblowing model is thus that a whistleblower has a low ability to double, i.e. a high value of  $t_c$ . A high  $t_c$  yields a high identity loss  $I_o$  which in turn is the primary incentive to react or to blow the whistle.

The secondary step in this process consists of the trade-off evaluation between the restoration of identity loss  $I_o$  and the whistleblowing costs  $K$ . Conclusively, what differentiates someone who blows the whistle from those who do not is that the whistleblower has a low ability to double and a perceived low whistleblowing cost  $K$  in relation to her lost identity  $I_o$ .

## 7. Conclusion

The aim of this thesis was to illustrate identity as an underlying incentive of whistleblowing from a social perspective by using Akerlof and Kranton's (2000) theory of identity economics. The incentives of whistleblowing have not been much in focus in the economic literature. This is mainly due to that conventional economics has a hard time explaining the incentives of a behavior that seemingly yields no benefits and carries heavy costs. However, the attempts made by Heyes and Kapur (2008) and Villena and Villena (2010) suggest motivations to whistleblowing, such as monetary rewards and a wish to punish those who misbehave. Contrary to this current research, we have abandoned the individualistic focus and pecuniary incentives of whistleblowing in favor of a non-pecuniary and social framework.

Research and theories on whistleblowing in other academic fields such as in sociology and psychology as well as from empirical case studies support the notion of a social approach in investigating a whistleblower's incentives. By exploring several whistleblowing cases, described in literature and by interviewing whistleblowers, we have been able to illustrate identity as a driving force to blow the whistle. These studies were aimed to complement theories from sociology and psychology in order to come one step closer to the minds and incentives of the whistleblowers.

By using psychological and sociological theories with our own case studies and applying these findings to Akerlof and Kranton's (2000) socially calibrated model, we have suggested a whistleblowing utility model. By blending elements of psychology and sociology with traditional economic analysis, what we have achieved is an awareness of a different type of incentive to whistleblowing. We propose that these incentives are strongly identity-driven.

## 8. Discussion and Implications

Our model assumes that the prescribed behaviors for a certain identity are more or less fixed and universally agreed upon. This suggests that all individuals of a social category would acknowledge misbehavior in the whistleblowing context as being wrong. This assumption can, however, be discussed. For example, conflicts between members of the same social category can often arise as a result of disagreement on what is the right thing to do. This occurrence can be observed in areas such as politics, religion, and so on. Thus, even individuals who are strongly linked by a common social category have different opinions on prescribed behavior. One circumstance that supports our assumption, however, is that the colleagues and superiors of the whistleblower put effort in hiding or covering up their misbehavior. In other words, they acknowledged that misbehavior was a deviation from prescribed behavior.

Further, our model assumes that the sensitivity for deviations from prescribed behavior  $t_c$  can be affected by an individual's multiple identities. Thus  $t_c$  is a function of the set of social categories  $\mathbf{C}$  that an individual belongs to. We do not, however, suggest that  $\mathbf{C}$  is the only parameter affecting  $t_c$ . We suggest that  $t_c$  could also be a function of a certain context or a function of Akerlof and Kranton's (2000) personality parameter  $\varepsilon_j$ . An example of when context affects the importance to keep to the prescribed behavior is when one group or social category interacts with another. For instance, at a workplace, more emphasis is often put on behaving and dressing correctly when in contact with external visitors. This can be related to Tajfel's (1974) theories on intergroup behavior. The function of  $t_c$  can thus be subject to further research.

Our whistleblowing function does not focus on monetary parameters. This is because the aim of our thesis was to illustrate identity as an incentive to whistleblowing. We do not, however, suggest that monetary incentives are completely insignificant. The parameter  $K$  includes a possible monetary cost, especially of losing your job. Thus we do not fully overlook the possible incentive mechanism of money in whistleblowing. On the other hand, as in accordance with Titmuss's (1970) theory on crowding-out, there is reason to assume that monetary rewards might prevent rather than promote whistleblowing.

In the introduction of this thesis we discussed the importance of understanding whistleblowing incentives in order to create effective policies. We also stated that economics does not contribute much to this policy work. The implication for policies of our proposed whistleblowing model is



to add an identity perspective to these incentives. More specifically, in order to promote whistleblowing, we suggest policy to take non-monetary rewards into account.

Lundquist (1998) suggests symbolic rewards to promote whistleblowing, such as medals and official recognitions. His suggestion supports the proposition of the importance of the social context and of non-monetary incentives. By using symbolic rewards and positive recognitions, the view of whistleblowing as yielding solely negative costs  $K$  could be altered. By including a symbolic reward  $R$  into our model, the magnitude of  $K$  becomes smaller in relation to the utility loss  $I_o$ . This type of symbolic reward  $R$  can be incorporated into  $W$ , such that  $W = I_o - K + R$ . Thus, the utility by whistleblowing would be higher than by remaining silent and conforming if:

$$I_c - K + R > I_c - I_s - I_o$$

## References

Ahlmark, Anders. Interview. 2011-04-27.

Akerlof, George and Kranton, Rachel. (2000), "Economics and Identity". *The Quarterly Journal of Economics* 115(3), 715-753.

Akerlof, George and Kranton, Rachel. (2005), "Identity and the Economics of Organizations". *The Quarterly Journal of Economic Perspectives* 19(1), 9-32.

Akerlof, George and Kranton, Rachel. (2010), *Identity Economics*. Princeton: Princeton University Press.

Alford, C. Fred. (2001), *Whistleblowers—Broken Lives and Organizational Power*. Ithaca: Cornell University Press.

Becker, Gary S. and Murphy, Kevin M. (2000), *Social Economics: Market Behavior in a Social Environment*. Harvard University Press.

Bjørkelo, Brita, Einarsen, Ståle and Matthiesen, Stig B. (2010), "Predicting Proactive Behavior at Work: Exploring the Role of Personality as an Antecedent of Whistleblowing Behavior". *Journal of Occupational and Organizational Psychology* 83, 371-394.

Bratt, Ingvar. (1988), *Mot rädslan*. Malmö: Carlssons Bokförlag.

Coyne, Christopher J. and Mathers, Rachel L. (2009), "The Identity Economics of Female Genital Mutilation". Working Paper. George Mason University, Fairfax.

Festinger, Leon. (1957), *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.

Glazer, Myron P. (1983), "Ten Whistleblowers and How They Fared". *The Hastings Center Report* 13(6), 33-41.

Glazer, Myron P., and Glazer, Penina. (1989), *The Whistleblowers*. New York, N.Y.: Basic Books.

Haglunds, Magnus. (2009), *Enemies of the People*. Stockholm: Acta Universitatis Stockholmiensis.

Heyes, Anthony and Kapur, Sandeep. 2008, "An Economic Model of Whistle-Blower Policy". *The Journal of Law, Economics & Organization* 25(1), 157-182.

Lifton, Robert J. (1986), *The Nazi Doctors: Medical Killing and the Psychology of Genocide*. New York: Basic Books Inc.

Lundquist, Lennart. (1998), *Demokratins Våktare*. Lund: Studentlitteratur AB.

Magnusson, Peter. Interview. 2011-04-18.

Miceli, Marcia P., Near, Janet P. and Schwenk, Charles R. (1991), "Who Blows the Whistle and Why?". *Industrial and Labor Relations Review* 45(1), 113-130.

Moskos, Charles C., Williams, John Allen and Segal, David R. (2000), *Armed Forces After the Cold War*. Oxford: Oxford University Press.

Myers, Linda J., Speight, S. L., Highlen, P. S., Cox, C. I., Reynolds, A. L., Adams, E. M. and Hanley, C. P. (1991), "Identity Development and Worldview: Toward an Optimal Conceptualization". *Journal of Counseling & Development* 70, 54-63.

Near, Janet P. and Miceli, Marcia P. (1996), "Whistle-blowing: Myth and Reality". *Journal of Management* 22, 507-526.

Roccas, S. and Brewer, M. (2002), "Social identity complexity". *Personality and Social Psychology Review* 6, 88-106.

Schwartz, Shalom H., Feldman, Kenneth A., Brown, Michael E. and Heingartner, Alex. (1969), "Some Personality Correlates of Conduct in Two Situations of Moral Conflict". *Journal of Personality* 37, 41-57.

Tajfel, Henri. (1974), "Social Identity and Intergroup Behavior". *Social Science Information* 13, 65-93.

Tajfel, Henri and John Turner. (1979), "An Integrative Theory of Intergroup Conflict", in Worchel, Stephen and Austin, William (eds.), *The Social Psychology of Intergroup Relations*. Monterey, CA: Brooks/Cole.

Thomas, Kerry. (1996), "The Defensive Self: A Psychodynamic Perspective", in Stevens, Richard (ed.) *Understanding the Self*. Thousand Oaks, CA: SAGE Publications.

Titmuss, Richard M. (1970), *The Gift Relationship: From Human Blood to Social Policy*. London: Allen and Unwin.

Trevino, Linda K. and Victor, Bart. (1992), "Peer Reporting of Unethical Behavior: A Social Context Perspective". *The Academy of Management Journal* 35(1), 38-64.

Sen, Amartya K. (1985), "Goals, Commitment, and Identity". *Journal of Law, Economics and Organization* 1(2), 341-355.

Villena, Mauricio G. and Villena, Marcelo J. (2010), "On the Economics of Whistle-Blowing Behavior. The Role of Incentives". Paper presented at the annual meeting of the BALAS Annual Conference, ESADE, Barcelona, Spain, Mar 24, 2010.