From Universalizing Transparency to the Interplay of Transparency Matrices: Critical insights from the emerging social credit system in China

Hans Krause Hansen
Copenhagen Business School, Denmark

Richard Weiskopf
University of Innsbruck, Austria

Abstract
Building on research on organizational transparency and surveillance, mediated visibility and Foucauldian dispositional analytics, we develop the concept of transparency matrices for studying the interplay of mediating technologies and normative arrangements in the formation of transparency as a heterogeneous regime of visibility. Using the emerging and controversial social credit system in China as a critical case, we make two contributions. First, we enrich the conceptual vocabulary for the study of transparency as a dynamic multiplicity with varying power effects in different contexts, challenging in this way universalizing and otherwise reductive notions of transparency. Second, by emphasizing the operation of several coexisting matrices we provide a novel way of approaching the ‘post panoptic’ condition which can account not only for the power effects of different modalities of transparency, but also for the instabilities and potential openings that are generated by their interplay. Transparency is heterogeneous, irreducible to a single reality and open to future becoming.

Keywords
dispositional analytics, mediating technologies, social credit systems, surveillance capitalism, transparency matrices

Introduction
On 26 October 2015, BBC News published an article entitled China ‘social credit’: Beijing sets up huge system (Hatton, 2015). It describes how the Chinese government is building an ‘omnipotent
“social credit” system that is meant to rate each citizen’s trustworthiness, based on ‘a vast national database that compiles fiscal and government information, including minor traffic violations, and distills it into a single number ranking each citizen.’ Warnings about the advent of ‘digital dictatorship’ and phrases like ‘Big Data meets Big Brother’ (Botsman, 2017) have proliferated in research and Western public media ever since, and they reflect a rapidly growing focus on the contemporary global process whereby power and control become entwined with digitalization and result in new and often concerning forms of transparency.

But have we, as some critics fear (e.g. Han, 2015), finally reached the age of universal transparency and surveillance? Is it really the case that information systems can now ‘provide a computer age version of universal transparency with a degree of illumination that would exceed even Bentham’s most outlandish fantasies’ (Zuboff, 1988, p. 321)? Do we face a kind of ‘absolute transparency’ (Zuboff, 2018, p. 397), driven by digitally based cybernetic control that leads to the ‘effective closure of formerly open systems’ (Diab, 2017, p. 1)? This paper suggests that the answer is ‘no’ and develops a conceptual framework that indicates why. Our central purpose is to demonstrate the complexities and intricacies of transparency in contemporary processes and practices of organizing and, by implication, the unlikeliness of universal transparency. We show how transparency comes about in multiple forms with varying effects of power and control and alongside the production of new invisibilities. We argue that critical knowledge of these processes can nuance our understandings of the workings of transparency in contemporary organizational settings and help question deterministic interpretations of the matter.

Contemporary scholarship regards transparency as an important mode of governing (e.g. Etzioni, 2010; Hood, 2006), managing (Bernstein, 2017; Roberts, 2017; Schnackenberg & Tomlinson, 2016) and organizing (e.g. Christensen & Cheney, 2015; Hansen & Flyverbom, 2015; Ringel, 2019). Transparency is part and parcel of an emerging ‘culture of surveillance’ (Harding, 2018). Corporations and governments monitor and profile citizens and users, who knowingly or unknowingly become subjects to and of surveillance (Haggerty & Ericson, 2000). Underexplored in this research is the systematic analysis of transparency’s normative embedding and forms of technological mediation, as well as the effects it has on what is purportedly made visible. Here, we understand transparency as a product of observation, which is always shaped by technologies and normative arrangements. None of these are homogeneous or uniform, however, but take different contours and vary relative to contextual features and logics, all of which results in the production of multiple modalities of transparency, if not transparencies. To examine such variations dynamically we develop the concept of transparency matrices.

According to the Oxford English Dictionary the term matrix refers to ‘the cultural, social or political environment in which something develops’, a ‘mould in which something is cast or shaped’, ‘a substance between cells’ (in biology), or an ‘array of quantities or expressions in rows and columns’ (in mathematics). Etymologically the term is related to the Late Latin word for ‘womb’ or ‘source of origin’. For us, normative arrangements and technologies constitute the matrix’s main elements which, when linked and operating together, can be conducive to specific forms of transparency. By putting the term matrix in the plural we suggest not only that there are several matrices from which transparency can emerge, but also that these matrices can operate simultaneously. They can support each other, but they may create instabilities, contradictions and new openings with transformative potential.

In developing the concept of transparency matrices we build on literatures on ‘mediated visibility’ (Thompson, 2005) and Foucauldian studies of ‘dispositives’ and ‘dispositional arrangements’ (Garland, 2014; Raffnœ, Gudmand-Høyer, & Thaning, 2016). Mediated visibility refers to transparency’s reliance on ‘mediating technologies’, which range from Taylor’s (1911) scientific observation tools to today’s computerized forms of data capture (Bernstein, 2017; Kitchin, 2017). The
concept of dispositive suggests how normative arrangements shape observation and thereby what we are inclined to see and not to see. By integrating these insights in the concept of transparency matrices we extend the conventional view on transparency and enrich the conceptual repertoire for critical reflections on the matter.

The first section situates our approach in the context of contemporary transparency studies. We mainly contribute to the ‘constitutive paradigm’, which provides an alternative to the ‘informational paradigm’ that dominates the field. Next, we detail the concept of transparency matrices. In section three we scrutinize several transparency matrices using the emerging Chinese social credit system (SCS) as an illustration. The fourth and fifth sections discuss their interplay and power effects. We show how attending to the interplay of coexisting matrices reveals a much more complex picture than suggested in assumptions about universal transparency (e.g. Diab, 2017; Han, 2015; Zuboff, 2018). How subjects are addressed and power effects generated varies considerably, making the effective, final closure of systems such as the SCS unlikely. By way of conclusion, we suggest the wider relevance of our framework for studying the intersections of transparency and power in modern organizations.

The SCS is our main empirical reference point, yet we do not pretend to present a full-blown empirical analysis. Our aim is more modest. The SCS is widely considered as ‘one of the most ambitious and alarmingly dangerous big data initiatives under construction’ (Jiang & Fu, 2018, p. 378). It is an ‘extreme case’ (Flyvbjerg, 2006, p. 229) and as such a rich resource for conceptual development. It is also a ‘critical case’ by having strategic importance regarding the widespread assumption that universal transparency is on the rise. Instead of taking that assumption for granted, we open up an analytical space for examining the production of transparency as a dynamic multiplicity with varying power effects in different contexts, in this way challenging dominant understandings of the matter.

The ‘Informational’ and the ‘Constitutional’ Paradigm

Notions of transparency go back to ancient times (Hood, 2006). With the rise of secularism and modern scientific ideals in the West, transparency begins to be understood as visibility contingent upon observation (Bernstein, 2017; Brighenti, 2007). As Comaroff and Comaroff (2003, p. 291) observe, ‘lives and deaths [come] to be read less as a sign of cosmic metaphysical forces than as the sum of mundane biophysical processes, knowable primarily through the modest art of observation’.

Most contemporary research on transparency embodies this legacy and is anchored in the ‘informational paradigm’. The collection, processing and provision of information is understood as (ideally) representing a pre-given reality. Observation – if done in a methodologically accurate way – literally makes certain aspects of reality ‘shine through’ (from the Latin trans-parere). Transparency aspires to transcend ‘bias’ and ‘idiosyncrasies’ in the production of knowledge and becomes ‘key to good governance’ (Garsten & Lindh de Montoya, 2008; Hood, 2006). In studies of principal–agent relations, for example, ‘transparency corresponds to the ability of the principal to observe what the agent does’ (Prat, 2006, p. 91). Information asymmetry to the advantage of the agent could be conducive to abuse of delegated power and market failure. On the other end of the spectrum, the free flow of information within organizations and societies more generally appears as a condition for democracy. In this vein, Fung (2013, p. 183) has developed a broader programme of a future ‘Infotopia’, which, by realizing how much humans depend on information, can ‘unleash the democratic power of transparency’.

While the informational paradigm has fuelled speculations about total oversight and control by the sovereign (principal) (Bowles, Hamilton, & Levy, 2014), it has also served as an argument for the redistribution of centralized power. In contemporary models of ‘open government’, for
example, information disclosure is the basis and measure of a democratic state. Fenster (2017) paraphrases the normative ideal of openness in this way: ‘Government doors should never be closed, and all government information should be available to the public as it is created or collected.’ In the ‘Age of transparency’, Sifry (2011) claims, ‘[s]ecrecy and the hoarding of information are ending; openness and the sharing of information are coming’ (p. 17). ‘Disclosure programmes’ offered by corporations promise to meet the information needs of other actors (Heimstädt & Dobusch, 2018). Examples include public meetings, financial disclosure statements, CSR and ‘transparency’ reports (Power, 1999; Flyverbom, Christensen, & Hansen, 2015; Fung, Graham, & Weil, 2007).

The informational paradigm dominates social science and business studies and provides many important insights (Bernstein, 2017). It has also spurred critical reflections (Ananny & Crawford, 2018; Flyverbom et al., 2015; Hansen & Flyverbom, 2015). It underestimates the complexity of information gathering and processing, the role of strategic selection, ambiguity and (self) presentation of information (Stohl, Stohl, & Leonardi, 2016). It downplays the place of legitimate forms of invisibility, including anonymity, confidentiality and privacy arrangements, business secrets and intelligence (Birchall, 2011; 2015; Costas & Grey, 2016). The frequent assumption that information provision itself will produce an informed and engaged public, holding accountable people in positions of power, does not convince in times of ‘information overload’ (Mol, 2010). Finally, the technological mediation necessarily involved in the collection, processing and provision of information often appears as a neutral process of transmission, obscuring the power that is involved in the sorting of information and overlooking the distinctive properties of media technologies themselves.

A different stream of literature has recognized the important role of information in the production of transparency but also emphasized the need to explore the pursuit of transparency itself (Albu & Flyverbom, 2016; Christensen & Cheney, 2015; Flyverbom et al., 2015; Garsten & Lindh de Montoya, 2008; Hansen, 2012; Hansen & Flyverbom, 2015; Roberts, 2009). In the ‘constitutive paradigm’, observation is never a passive or neutral process of registering of what is out there, but rather an active process of ‘making up’ (Hacking, 2007) social reality, a ‘wider-ranging principle of organizing that dramatically influences social behavior, while potentially eliciting new forms of closure, manipulation, control and surveillance’ (Christensen & Cheney, 2015, p. 85). The transparency pursuit is always performative as it creates objects of knowledge that can ‘react’ or ‘loop back’ on social reality (Espeland & Sauder, 2007; Hacking, 2007). It is ‘organizing’, understood here as a reflexive process of ordering people, practices and processes. Moreover, transparency is always and necessarily selective, with ‘invisibility’ as its faithful companion.

The constitutive paradigm is informed by Foucault’s analysis of Bentham’s panopticon, which underscores the constitutive role of the centralized, hierarchical gaze. Scholars have suggested significant modifications to the panoptic diagram given the proliferating ‘liquid’ contexts of late modernity. Here ‘the world of fixity dissolved into flows’ (Bauman & Lyon, 2013, p. 55), and ‘new surveillance practices, based on information processing . . . permit a new transparency in which . . . [we] are constantly checked, monitored, tested, assessed, valued and judged’ (Lyon, 2014, p. 12). New media also allow for the public scrutiny of leaders in ways that were not possible until recently (Mathiesen, 1997; Thompson, 2005). All these insights are important as they suggest the necessarily heterogeneous character of contemporary transparency. Gazes come from multiple angles, they are selective, never absolute nor complete, and they can therefore never result in universal transparency. At the same time, we argue, such ‘post-panoptic’ perspectives are in danger of overestimating the ‘fluid’ nature of the ‘new transparency’, potentially overlooking the persistence of the older forms which emerge in new
clothes (Anteby & Chan, 2018; Caluya, 2010). Given that our central purpose in this paper is to demonstrate the complexities and intricacies of transparency in contemporary processes and practices of organizing, it is paramount to consider the coexistence of different modalities of transparency as well as to think of the relation between these in terms of layering and entanglement rather than historical succession. Necessarily heterogeneous and irreducible to a single reality, transparency always rests on observation guided by varying purposes and shaped by a plethora of norms and mediating technologies. In order to develop a conceptual vocabulary that allows us to analyse these intricate dynamics in more detail we now outline the concept of transparency matrices and spell out its implications for empirical studies.

**Transparency Matrices**

The concept of transparency matrices integrates important insights from literatures that are not commonly brought together. Its central building blocks revolve around the concepts of ‘mediated visibility’, ‘dispositive’ and ‘dispositional arrangements’.

**Transparency as mediated visibility**

As Thompson (2005, p. 35) notes, observation ‘is never a matter of simply opening one’s eyes and grasping an object or event’, it is never ‘pure vision’. While there was a time before the advent of modern media where vision was severely constrained by the here and now, today vision is generally shaped ‘by the distinctive properties of communication media, by a range of social and technical considerations. . . and by the new types of interaction that these media make possible’. Such properties or ‘affordances’ (Lievrouw, 2014) invite a specific use of mediating technologies and help to produce often varying kinds of transparency (Flyverbom, 2016, p. 112; Hansen & Flyverbom, 2015, p. 876). The term mediating technologies can refer to words and narratives, numbers and images, as well as to machines and devices like video cameras, mobile phones and remote sensing, mass and social media, digital platforms and so on. These are not simply registering what is out there, but shape how people and things in the world become visible and knowable to the observer.

In the context of digitalization, ‘datafication’ provides the basis for the production of particular forms of visibility. Datafication transforms human experience into machine-readable ‘big data’ (Mayer-Schönberger & Cukier, 2013, pp. 73–97), which can be algorithmically processed and analysed. Gillespie (2014, p. 184) notes that we have come to make ourselves ‘algorithmically recognizable in all sorts of ways . . . for all of us, social networks and the web offer some analogous kind of “mediated visibility”’. Human activities leave ‘digital traces’, which coded algorithms can search for patterns and organize into ‘digital doubles’ or ‘profiles’, through algorithmic selecting, filtering and framing of relevant information (Flyverbom & Murray, 2018). The process does not deliver proof of causality or any conclusive reasoning, but mathematical correlations are sometimes taken as indicative of expected behaviour (Hildebrandt, 2013). Algorithms help produce visibility, albeit selective and with problematic side effects. Brayne (2017, p. 27) notes that ‘Institutional actors making decisions based on big data may assume that data doubles are more accurate, or unbiased, representations of a person’s profile than are those gleaned from “small” data, such as personal observations.’ Algorithms can also work contrary to their initial design as they develop exponentially. Neither their designers nor users always fully understand or control them (Ananny & Crawford, 2018; Gillespie, 2014).

All this raises complicated questions about the character of transparency in the digital context. A major insight is perhaps the recognition that the production of transparency is more
multidirectional, fluid if not more uncontrollable than portrayed in classical studies. Foucault’s treatment of Bentham’s ‘panopticon’ has often been taken as an example of how the ‘few (in power) can observe the many’, leaving aside the question as to ‘how the many can see the few (in power)’ (Haggerty & Ericson, 2000; Mathiesen, 1997; Thompson, 2005). With computerization and algorithms the significance of the question as to who observes whom, what is being observed based on what kind of data, becomes much more complex (see e.g. Ananny & Crawford, 2018; Anteby & Chan, 2018; Bauman & Lyon, 2013; Caluya, 2010; Uldam, 2016). These insights aside, there is obviously much more to the production of transparency than the technological mediation of observation – whatever its directions and the kinds of transparency produced – than we have discussed so far. As Thompson (2005, p. 36) notes, ‘seeing is always shaped by a broader set of cultural assumptions and frameworks’. Social norms and relations of power play a central role in the production of transparency and hence for our understanding of what transparency means and how it comes about in specific contexts and situations. We now turn to Foucault’s concept of the dispositive (Foucault, 2007) to help clarify such normative shaping of the gaze and its implications for transparency and relations of power.

**Transparency as normatively shaped visibility**

While the panopticon illustrates how mediating technologies such as prison architecture allow for central observation of inmates, it is better understood to be epitomizing wider shifts in social norms, mechanisms of power and visibility in modernity; the ‘disciplinary society’ where conscious and permanent visibility ‘assures the automatic functioning of power’ (Foucault 1977/1991, p. 201). But the term ‘disciplinary society’ and the ideas of ‘permanent visibility’ and power’s ‘automatic functioning’ are too generalizing. By engaging with the type of ‘dispositional analysis’ to be explained below we extend original insights derived from Foucault’s study of the panopticon to include a much wider set of conceptualizations that better help us understand the multiple facets of transparency and power at work in modern organizations.

A dispositive (sometimes also called dispositif and related to concepts such as apparatus or assemblage; see Walters (2012) and Raffnøe et al. (2016) for a detailed exegesis of Foucault’s use of the term), is an arrangement of heterogeneous elements which normatively shape experiences and organizational practices (Garland, 2014, p. 378; Raffnøe et al., 2016). The term dispositive does not refer to an entity, but rather to a ‘net of relations’ that can be established between these elements. A dispositional analysis does not take objects and subjects as given but is mainly concerned with how these emerge as elements that become assembled through the operation of different dispositional logics. In our context, central questions include: How do dispositional arrangements and their logics affect observation, and so condition what becomes visible and invisible? How do varying dispositional arrangements produce different modalities of transparency? What is the interplay between them? And what are their power effects?

Foucault (2007) analysed three particularly important prototypical dispositives – law, discipline and security – characterized by specific dispositional logics (Raffnøe et al., 2016, p. 285). These imply ‘particular ways of managing and organizing a multiplicity’ (Foucault, 2007, p. 9) and, as we will show, involve modalities and exemplary practices of transparency. However, we also take inspiration from Foucault’s study of pastoral power4 (Foucault, 1983; 2007, pp. 125–130) and ethics in Greek antiquity (Foucault, 1986; O’Leary, 2002), and extend the three prototypical dispositives to include a fourth, the dispositive of care.5 This makes it possible to highlight important modalities and exemplary practices of transparency that would otherwise not be considered in detail, including the confession and parrhesia (Folgers, 2016). Our dispositional
analysis enables us to go well beyond Foucault’s influential work on the panopticon and provide a more nuanced and multifaceted picture.

We now bring our analysis of mediated visibility (especially the concept of mediating technologies) and normatively shaped visibility (especially the concepts of dispositional arrangement and logics) together in one framework termed transparency matrices. Importantly, we posit that these transparency matrices should not be seen as a historical sequence, where one follows or replaces the other. Rather, they can coexist, presuppose or even contradict each other.

**Four transparency matrices**

First, within the dispositional arrangement of law, transparency emerges from observations guided by a distinction between allowed and forbidden. ‘Misconduct’ and ‘unlawful behaviour’ is made visible and linked to sanctions. The logic is prohibitive. The revelation of behaviour (potentially labelled as ‘criminal’ post factum) followed by conviction and determination of guilt in court is one example, the public pillorying of malefactors another. The two examples reflect a modality of transparency that can be traced back to classical sovereignty, including the spectacles of public punishment surrounding the exercise of sovereign power (Foucault, 1977/1991). In contemporary contexts this transparency modality rests typically on mediating technologies such as oral, written and visualized evidence and intelligence, produced through institutionalized practices such as crime investigations and processes of public blacklisting on websites (Hu, 2016). Blacklisting classifies the activity of individuals or organizations as ‘wrong-doing’ on the basis of pre-established legal criteria, exposes it to the public and applies sanctions. It signals the inclusion into or exclusion from a community and helps to maintain order by preventing what is forbidden (Foucault, 2007, p. 46).

Second, within the dispositional arrangement of discipline, the distinction between allowed and forbidden is supplemented by a distinction between normal and abnormal, which shapes observation. The logic is prescriptive. In contrast to law and its focus on compliance, discipline concerns what one is in relation to a model of what one should be (Foucault, 2007, p. 57). Ideally, this dispositional arrangement produces behaviour that conforms to the norm by rewarding desired behaviour and sanctioning deviant or less well-performing behaviour. The transparency modality here is therefore the comparison, which serves to make such differences visible. The paradigmatic institutional practice is the examination, which ‘establishes over individuals a visibility through which one differentiates them and judges them’ (Foucault, 1977/1991, p. 184). Exemplary mediating technologies include performance rankings, ratings and benchmarking systems. These classify and measure (individual or group) performance typically in numerical form, making differences or ‘gaps’ between the listed entities visible. Lists highlighting acceptable or celebratory achievements are also examples (Espeland & Sauder, 2007; Hansen, 2015; Kornberger & Carter, 2010; Townley, 1998).

Third, the dispositional arrangement of security supplements law and discipline to organize and manage the circulation of humans, information and materials. The distinction between allowed and forbidden, and between normal and abnormal, is supplemented here by a distinction between good and bad circulation, which allows ‘maximising the positive elements and minimizing what is risky and inconvenient’ (Foucault, 2007, p. 19). The security logic is anticipatory and is about acting ‘in a space in which a series of uncertain elements unfolds’ (p. 21), which entails organizing and managing instability and dangers. The transparency modality here is the forecast or prognosis, i.e. the analytical work of making visible unknown futures and uncertain developments, which (although they have not yet materialized) shape action and decision-making in the present (Amoore, 2013). Exemplary mediating technologies include the elaboration of population metrics or statistical correlations (e.g. rates of crime, unemployment, birth, illness, etc.), displayed in charts, tables and
numbers; real-time screening, remote tracking and biometrics; as well as predictive forms of algorithmic profiling (Hildebrandt, 2013; Rouvroy, 2013).

Fourth, the dispositional arrangement of care has as its objective the care for self and others and implies forms of managing and organizing people (‘multiplicities’) that can help establish beneficial relations within a moral community. The transparency modality here is the disclosure or exposure of a subjective truth, anchored in paradigmatic practices of ‘truth-telling’, such as the confession and parrhesia (Folkers, 2016). Especially in the digital context many confessional forms of truth-telling are associated with the promise of positive rewards or even with the pleasure of self-presentation. According to Ball (2009, p. 641), ‘the pleasures of performative display (often) override the scrutinies that come hand-in-hand with self-revelation’. Parrhesia – often translated as ‘fearless speech’ – was a highly valued form of truth-telling in Greek antiquity (Foucault, 2001). Here, the subject courageously speaks an uncomfortable truth to an authority or to the public, making visible what is subjectively conceived as ‘intolerable’. Whistleblowing is a contemporary form of parrhesiastic truth-telling (e.g. Kenny, 2019; Vandekerckhove & Langenberg, 2012; Weiskopf & Tobias-Miersch, 2016).

Today, confession and parrhesia are rarely purely verbal acts of speaking, but are relying on a broader panoply of mediating technologies, including images and sound, for expressing and exposing the (situated, perspectival) ‘truth’ and making it heard in a wider audience. So-called ‘portable electronic confessionals’ (Bauman & Lyon, 2013, p. 30) like Facebook profiles accessible via mobile phones can assist networked subjects reveal their inner secrets and make themselves visible to others (Ellerbrok, 2010). The same technologies can also be deployed to disrupt established normative orders through ‘parrhesiastic exposures’ which bring uncomfortable realities to light (Walters, 2014, p. 287; see also Munro, 2017; Nayar, 2010). By introducing a rupture in a given normative order, parrhesiastic exposures can open up a space for reflection, debate and potential transformation. Table 1 summarizes the major components of the four transparency matrices.

Our conceptualization of transparency matrices makes it clear that the production of transparency comes in different modalities. These relate to the operation of dispositional logics and are facilitated by mediating technologies. Recalling that our aim is to demonstrate the intricacies of transparency in contemporary processes of organizing, we will now consider our critical case – the Chinese social credit system (SCS) – and provide a more detailed analysis of the interplay of transparency matrices, their dispositional logics and mediating technologies, and power effects.

**The Chinese SCS**

‘Since imperial times’, notes Zeng (2016, p. 1450), ‘China’s central government has been struggling to obtain real information at ground level. For thousands of years, the central regime has relied on the regional and local governments to collect and report information for decision-making.’ Under Mao Zedong, the performance of individual citizens was recorded in the dang’an (file). This document contained personal information and performance appraisals by superiors and peers; educational and employment records; criminal convictions and administrative penalties; club memberships; and political history, including Chinese Communist Party (CCP) membership. In contemporary China, with 772 million internet users and 1.36 billion mobile phone users, the paper dang’an and neighbourhood informers are gradually being replaced by cameras, big data algorithms and cloud storage of data stemming from geo-location to online payments (Creemers, 2017, p. 88; Jiang & Fu, 2018).

At the beginning of the 21st century the Chinese government began to make systematic use of new mediating technologies to measure the financial creditworthiness of individuals and
businesses. Establishing online financial credit systems was seen as key to better the integration of China in the global market economy. The first credit-scoring companies were established in the 1990s, and later the People’s Bank of China (PBC) established an enterprise and personal credit database. While initially focused on financial aspects of credits (like loan lending activities), the system was extended to the measurement of personal and corporate trustworthiness in relation to contract fulfilment and legal commitments. Public discontent with corruption, public institutions and fraudulent company practices further encouraged the Chinese government to launch a plan in 2014 for a full-blown SCS as an ‘important component part of the Socialist market economy system and the Social governance system’ (SCC, 2014, emphasis added; Liang, Vishnupriya, Kostyuk, & Hussain, 2018). The Planning Outline for the Construction of a Social Credit System (2014–2020) focuses on restoring trust, improving ‘sincerity’ in governmental, commercial and societal affairs, on raising judicial credibility and the ‘honesty mentality and credit levels of the entire society’, with international competitiveness and a ‘harmonious Socialist society’ as ultimate goals (SCC, 2014). The SCS is run by the state. But its development and functioning rests on close cooperation with big Chinese internet companies like Alibaba, Baidu, Tencent and SenseTime. These provide technical infrastructures, and mine and process huge amounts of data harvested from countless sources and data-points, including geo-location and online payments. The government is constructing several national data platforms for collecting, storing, sharing and mining population data and focuses on ‘integrating these separated platforms into centralized data infrastructures’ (Liang et al., 2018, p. 426).

How do the different dispositional arrangements and their logics described in the previous section and Table 1 come about with respect to the SCS? The prohibitive logic of the dispositional arrangement of law is apparent in the set-up of the SCS, which aims to make violations of legal rules and procedures visible through ‘perfect(ing) systems’ that ‘record and expose information concerning trust-breaking’ (SCC, 2014). Violations are exposed to the public on blacklists, and a

<table>
<thead>
<tr>
<th>Transparency matrices</th>
<th>Law</th>
<th>Discipline</th>
<th>Security</th>
<th>Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logic</td>
<td>Prohibitive</td>
<td>Prescriptive</td>
<td>Anticipatory</td>
<td>Supportive</td>
</tr>
<tr>
<td>Objective</td>
<td>Avoiding misconduct</td>
<td>Normalizing conduct</td>
<td>Managing circulation</td>
<td>Care for self and others</td>
</tr>
<tr>
<td>Transparency</td>
<td>Revealing and publicly punishing misconduct</td>
<td>Comparing what (one) is in relation to what (one) should be; making differences visible</td>
<td>Forecasting; making visible unknown futures and uncertain developments</td>
<td>Disclosing/exposing a subjective truth</td>
</tr>
<tr>
<td>modality</td>
<td>The criminal</td>
<td>The examination</td>
<td>The statistical correlation</td>
<td>Confidential and parrhesiastic truth-telling</td>
</tr>
<tr>
<td>Paradigmatic practices</td>
<td>investigation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exemplary</td>
<td>Oral, written and visualized evidence and intelligence, the exposure on public blacklists</td>
<td>Performance scores, league tables, ratings and rankings, white- and red-lists</td>
<td>Social statistics, big data, spotting and visualizing trends and patterns, (algorithmic) profiling</td>
<td>‘Portable electronic confessionals’; mass and social media</td>
</tr>
<tr>
<td>mediating technologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telos</td>
<td>Preventing disorder through law</td>
<td>Creating order through rewards and punishment</td>
<td>Optimizing populations</td>
<td>Guiding people towards a specific end; transformation</td>
</tr>
</tbody>
</table>

Table 1. The four transparency matrices.
line between lawful and unlawful behaviour is drawn. So-called ‘discredited people’ are sanctioned. Botsman (2017) quotes Meng Xian, head of the executive department of the Supreme Court: ‘We have signed a memorandum (with over) 44 government departments in order to limit “discredited people on multiple levels”.’ Yang (2017) reports that the Chinese Supreme Court has operated a blacklist of long-term debtors since 2013:

All 6.73m people on the list have been banned from boarding flights and high-speed trains. The ban works by blocking personal ID card numbers, which are required for buying and checking in for flights and trains as well as staying in hotels. Anyone can search the blacklist via the Supreme Court’s website by entering a full name and ID card number.

The prohibitive logic behind these measures is also well expressed in a statement by Li Ming, a credit-counselling expert at the Beijing Institute of Big Data Research: ‘to create a disciplinary mechanism in which all of us will unite to pressure people who behave badly or commit crimes to come back to the right track’ (quoted in Vanderklippe, 2018). Chen and Cheung (2017) describe the public disclosure by the Chinese Supreme Court and local governments of ‘serious discrediting behaviour’ and note that especially the ‘public naming of judgement defaulters is acknowledged to be necessary for inhibiting the prevalent circumvention of obligations imposed by effective judicial rulings’ (p. 25).

The prescriptive logic of discipline becomes apparent in the ‘models of sincerity’ which are not only promoted through ‘propaganda and the guiding roles of television, radio, newspapers’ (SCC, 2014) but also inscribed in the SCS’s algorithms. These serve to rank and compare citizens and include so-called ‘red lists’, which ‘record particularly conspicuous merit’ (Chorzempa, Triolo, & Sacks, 2018, p. 9; Creemers, 2018) to ensure ‘that those keeping trust are incentivized and rewarded, and those breaking trust are restrained and sanctioned’ (SCC, 2014). An otherwise compliant Chinese citizen may be notified that he or she engages socially or is somehow connected with an individual who is low rated or even blacklisted. ‘Those who wear virtue on their sleeves’ as Greenfield puts it (2018, p. 4), perhaps by taking public transit consistently instead of driving to work, taking out the recycling regularly, or even denouncing a misbehaving neighbor – might enjoy new benefits, like being able to rent a flat with no deposit, or earning the right to send their children to exclusive schools.

The ‘joint reward and punishment mechanism’ (Liang et al., 2018, p. 432) at the heart of the dispositional arrangement of discipline links profiles generated from multiple sources of behavioural data to the structure of incentives. It not only objectifies individuals, but also works as a subjectifying force, inducing people to reflect on their own ‘trust-keeping’ or ‘trust-breaking’ behaviour, including all the dimensions that influence the score.

The SCS might seem to produce ‘responsible subjects’ capable of participating in a market economy, to introduce freedom in the form of choice, and to provide incentives to use this particular freedom in ‘trustworthy’ ways. But such subjectivation, typically associated with Western (neoliberal) governmentality (e.g. Miller & Rose, 1990), clearly takes a modified form here: the creation of market subjects and their taming takes place within the moral framework of ‘sincerity’ and ‘honesty’ as defined by the Chinese Communist Party, and in addition self-discipline might be linked to Confucian traditions (Kipnis, 2008). Moreover, the use of credit scoring and the ranking of individuals in the SCS differ from their use in other contemporary contexts. While credit scoring has become a pervasive practice in the West and its mechanisms are often black-boxed (Fourcade & Healy, 2016; Pasquale, 2015), it has commonly been
restricted to finance and commerce, and some level of privacy protection can be claimed to exist. In China, such protections are lacking. Authorities can draw on all sorts of data for rating citizens and enterprises, including credit data, traffic data and criminal records, data on shopping and online habits, data from social networks, and so on (Chen & Cheung, 2017). According to Jiang and Fu (2018, p. 384) there is ‘a growing state–market symbiosis where the Chinese government and China’s Internet giants have developed a largely mutually beneficial partnership to maximize their relative economic and political gains.’

Next, the SCS relates to the dispositional arrangement of security by providing knowledge for the optimization and control of the circulation of individuals, goods and information. Credit-score points, for example, facilitate the regulation of access to employment or education, and international education for technical and economic elites is seen as absolutely crucial for China’s emerging economy. Especially important is the prognostic ‘anticipatory transparency’ generated by the big data produced to provide ‘early warning about the occurrence of potential unrest and incidents’ (Creemers, 2017, p. 98). In certain areas of the province of Xinjiang, for example, the information generated by video cameras and the compulsory installation of spyware apps on mobile phones help to rank people’s ‘trustworthiness’ (The Economist, 2018, pp. 6–7). Creemers reports that in urban areas, ‘“grid management” techniques integrate information and communication technologies with street-level policing, social services and both coercive and cooperative forms of management. The objective is responding to social unrest, and preventing the materialization of potentially destabilizing risks’ (Creemers, 2018, pp. 1–2). This system, which facilitates the visualization and filtering out of ‘discredited persons’, and thus minimizes what is perceived as risky and dangerous, is related to China’s vast programme of Internet governance and censorship, which blocks and filters information and communication flows on multiple levels.

Finally, the SCS relates to the dispositional arrangement of care. On the one hand, it implies forms of managing and organizing that seek to establish a moral community of sorts. Chinese authorities’ references to a ‘harmonious socialist society’ are echoing not only CCP ideology but also ‘a long Confucian tradition of attempts to bolster public rectitude’ (Greenfield, 2018, p. 4; Chen & Cheung, 2017, p. 5, Kipnis, 2008). Here, transparency comes about when subjects are encouraged or obliged to speak the truth about themselves and others. In general, people are officially encouraged to report corruption, use blogs to blow the whistle and urge corrupt officials to confess their crimes (Zhu, Huang, & Zhang, 2017). For example, Ding (2015) quotes Xinhua, which is characterized as ‘a mouthpiece of the Chinese government’, saying that ‘recent exposures of official wrongdoings have been helped by central government . . . China’s fight with corruption, with enthusiastic participation of netizens, has entered a new stage . . . ’ (Ding, 2015, p. 11). Yuen (2014) reports that in order to reinforce CCP discipline ‘local Party committees are launching self-criticism sessions among officials as part of a mass line education campaign’. Here, people can ‘confess their mistakes and criticise one another’ (Yuen, 2014, p. 45).

Another government initiative is the mobilization of ‘good citizens’ as informers. Page and Dou (2017) report the launch of the ‘Safe Zhejiang’ app that ‘enables users to notify authorities of problems ranging from leaky drains and domestic disputes to traffic violations and illegal publications, in text or photographic form, as long as the informants reveal their location and identity’ (Page & Dou, 2017, p. 1). Such apps are not infrequently met with suspicion. Some citizens see them as tools to snitch on neighbours or to disclose information about the reporting citizen’s identity herself. Informing can be rewarded with ‘discounts at upmarket coffee shops and coupons for taxi hailing and music streaming services, as well as for the Alipay online payment system, run by the financial affiliate of local tech giant Alibaba Group Holding Ltd’ (Page & Dou, 2017, p. 1). But ‘truth-telling’ refers also to critical bloggers and countermovements enacting forms of parrhesia. They are censored and controlled, but nevertheless emerge again and again (Liu, 2015). While such
parrhesiastic exposures pose challenges to the government, they are fundamentally ambivalent. As Qin, Strömberg and Wu (2017) note,

social media users (in China) have an incentive to speak out about local problems as they expect the central government to address these problems. In the corruption posts, we find frequent explicit appeals for central government action. The posts about protests and strikes may help organize the events, but, importantly, they also make these events visible to the central government. (p. 137)

The system seems to extend through the social fabric and is widely believed to strengthen authoritarian rule and state intrusion into the private sphere (Chen & Cheung, 2017). Nonetheless, it is not a fully integrated and unified system, which is often the way it is portrayed in the West. Subject to considerable controversy (Horsley, 2018) is the idea of establishing a single national score to rate companies’ and individuals’ financial creditworthiness and social trustworthiness by awarding positive points for desired behaviour and minus points for undesired behaviour. In practice, the system is highly fragmented and technically complex, consisting of a variety of social credit initiatives and projects at the local and regional levels that will probably never be fully integrated at the national level (Creemers, 2018). Moreover, there are examples of popular resistance, with locals reportedly having rejected a digital tool designed to let them inform on neighbours, ‘on the grounds that it reminds them of the Cultural Revolution–style denunciations that remain within the living memory of elders’ (Greenfield, 2018, p. 4). Eventually, parrhesiastic exposures might help create a space for public debate and transformation.

To summarize, in its emerging form the SCS displays the central dynamics laid out in our conceptualization of transparency matrices. The operation of dispositional logics implies that transparency comes about in different modalities. Most of these are produced mainly by the new mediating technologies embodied in the system as such, but they are clearly complemented by or responded to through mediating technologies that are not necessarily part of the SCS. Examples include blogs and social media and other forms of communication that enable transparency modalities such as parrhesiastic truth-telling.

Effects of Power

Our analysis makes clear that multiple forms of transparency and dispositional logics are at play, and these address subjects and generate power effects in different ways, thus making the role of power and social norms highly complex and inherently unstable. For example, central elements of the SCS embody the logic of the law. Typically, subjects are addressed as loyal subjects or citizens expected to comply with the rules of law. Public blacklists are deployed as exemplary mediating technologies making misconduct transparent and public shaming possible. While blacklisting has a long tradition in China, the SCS is ‘meant to give blacklists teeth’ (Ahmed, 2019) by sharing and pooling of information between different government bureaus and also private companies. This ensures that blacklisted individuals cannot avoid punishment and makes it an effective tool in reinforcing the law. Power effects are generated by deterrence or by inducing fear of punishment or embarrassment. Here it is the fear of losing reputation or of suffering from material disadvantages that influences behaviour.

On the other hand, we have also identified elements that embody the disciplinary logic. The main transparency modality is the comparison of what one is in relation what one should be, as elicited in mediating technologies such as performance scores and rankings. Here individuals are addressed as subjects who contribute as ‘productive’ and ‘trustworthy citizens’ to economic and social wellbeing and development. The normalizing effect generated by rewarding ‘good’ and punishing ‘bad’ behaviour is fundamental for producing a specific type of behaviour that is
constructed as ‘trustworthy’. Disciplinary practices typically constitute self-disciplining forms of subjectivity. Transparency here ‘comes to constitute, either directly or through the comparisons it makes possible, an ideal of what the self should be’ (Roberts, 2009, p. 965). It is the anxious awareness of being observed, evaluated and judged (as untrustworthy) together with the desire for recognition (as a ‘trustworthy person’) that generates power effects of subjectification (Fleming & Spicer, 2014, pp. 267–275). By mobilizing technologies that incite subjects to turn the gaze upon themselves the SCS shapes the conditions that influence how people experience themselves as people. A number of studies suggest that the system is widely accepted by Chinese citizens, in particular by better educated and wealthier citizens in urban areas (e.g. Kostka, 2019), which speaks to their subjectification by governmental techniques.

Other elements, like the management of uncertainties and risks associated with the flow of people, goods, money and information, are more directly related to the logic of security. Making uncertainties and risks visible is the main transparency modality and facilitated by mediating technologies like big data and algorithmic profiling. The security logic addresses subjects indirectly and has effects even by ‘bypassing consciousness and reflexivity’ (Rouvroy, 2013, p. 153). Based on pattern recognition from large data-sets, algorithmic profiling generates images of the ‘potential person’ and visualizes future developments. Power effects are generated through anticipatory action and pre-emptive decision-making. As Rouvroy (2013, p. 153) observes, ‘the focus on anticipation and pre-emption shifts the target of “power” from actuality, and from the present wilderness of facts, to potentiality, to the risks and opportunities’. Koopman (2019, pp. 155–164) calls this phenomenon ‘info power’. Information ‘fastens’ subjects in the double sense of ‘tying subjects down’ and of ‘speeding subjects up’. In algorithmic profiling, subjects (the ‘information person’) are pinned down to predefined formats, categories and conceptions of the ‘trustworthy’ person. This operation accelerates (social) recognition and interaction and so optimizes flows and the ‘fluidification’ of economic and social life.

SenseTime, one of China’s ‘national champions’ in artificial intelligence (AI), illustrates this. The company provides the face recognition software and algorithms used by the government to identify and profile members of the ethnic minority, the Uighurs, for the purpose of law enforcement and ideological indoctrination (Mozur, 2019). This example demonstrates the use of AI for conscious racial profiling and the entanglement of the logic of care. According to SenseTime, the systems on offer ‘are concerned about the wellbeing and safety of individual citizens’ (quoted in Mozur, 2019). There are other elements which can be related to the logic of care. Persons are addressed as moral subjects and asked to ‘give an account’ of themselves. Power effects are generated through the promotion of ‘self-disciplining identities’ (Fleming & Spicer, 2014, p. 269) and ‘subjectification’ (pp. 267–275). When subjects make themselves visible and evaluate themselves in categories provided by authorities their sense of self may be (re)shaped and tied to specific identities, such as the ‘trustworthy person’. We have identified confessional forms of truth-telling, which work as exemplary technologies here. Yet, the very technologies that mediate confessions can also empower subjects and generate, at least potentially, parrhesiastic exposures, which problematize established power relations and open a space for potential transformation.

It has been argued that the SCS embodies a ‘big dream of total awareness and perfect certainty mediated by algorithms that filter a perpetual flood of data flows from private and public suppliers’ (Zuboff, 2018, p. 392), but understanding it in terms of the ‘panopticon becoming total’ (Han, 2015, p. 49), the ‘closure of formerly open systems’ (Diab, 2017, p. 1) or ‘absolute transparency’ (Zuboff, 2018, p. 397) would be simplistic and misleading. The complex power effects of the SCS as well as the dynamics and intricacies of the transparency pursuit more generally need to be unpacked with conceptual tools that offer more nuances and facets.
Concluding Reflections

It was exactly to this end that we have proposed transparency matrices as a conceptual tool for exploring the dispositional arrangements and mediating technologies that enable observations and hence the production of multiple forms of transparency, shaping human actors’ relations to self and others within historically specific contexts. Transparency is not linked to one specific telos only, but implies a variety of normatively shaped forms of observation and visibility, mediated by different technologies. Here, ‘various kinds of dispositional logics come into play and concur to create a multi-layered field of normativity’ (Raffnsøe et al., 2016, p. 285).

Our analysis provides important nuances to those recent bodies of literature in the constitutive paradigm that emphasize the fluidity of contemporary data-driven forms of transparency. Often implied is a generalized move from panoptic to ‘post-panoptic’ or ‘synoptic’ forms of surveillance and transparency (e.g. Albu & Flyverbom, 2016; Bauman & Lyon, 2013; Brivot & Gendron, 2011; Hansen & Flyverbom, 2015). Our analysis instead emphasizes the operation of several coexisting dispositional logics and avoids the trap of periodization, which often assumes that one form is replaced by another. It also helps address the varying conditions for observation, their logics and hence the different transparency modalities at stake, including their ensuing power effects.

This, we argue, provides a novel way of approaching the phenomenon of transparency. Given that transparency matrices can operate simultaneously, they can support each other, for instance when they coincide, at least partially, in shining light on the subject. Importantly however, they can also work to create instabilities, contradictions or new openings. We have provided some illustrations of this. The dispositive of care, for example, involves subjective forms of truth-telling. Here the transparency modality of parrhesiastic exposures can question and destabilize established practices and hierarchies even if the SCS incites or rewards the confessional form that stabilizes them. This not only problematizes simplified notions of panoptic transparency but also concurs with Lyon’s (2018, p. 159) call for more research that addresses ‘the experience of being transparent to others, and how people may contest but also be content, comply or even covet transparency’. Another example relates to the algorithmically generated blacklists. While these can work as effective tools for law enforcement, they can also impose sanctions with less than proper investigation, examination and trial. This undermines legal principles like ‘due process’ (Rouvroy, 2013). Finally, it is also worthwhile to note that since the algorithms used to create individual scores remain largely unknown (Ahmed, 2019) the possibility of creating subjectifying effects may be undermined and ‘counter-subjectifying tactics’ (Fleming & Spicer, 2014, p. 275) evolve. As already noted, the ways in which algorithms can take part in the production of transparency is not a straightforward matter that is always fully understood and under control (Ananny & Crawford, 2018; Brayne, 2017; Gillespie, 2014). Overall, we suspect the transparency pursuit will rarely result in a stable regime of visibility. Rather, it constitutes a dynamic ‘play of light and shadow . . . true and false, hidden and manifest, visible and invisible’ (Foucault, 2016, p. 17), which affects people in many different ways. Even totalizing aspirations can probably never fully determine the course of events, since the different modalities of transparency will rarely add up to a unified and closed system. The most sophisticated forms of predictive analytics can never fully anticipate a future becoming.

We have developed and illustrated our concept of transparency matrices with reference to a large-scale example of managing and organizing a population. Future research might explore the ways in which the tensions between different logics and their ensuing modalities of transparency are played out and enacted in other organizational contexts and in relation to specific managerial versions of the transparency pursuit (Anteby & Chan, 2018; Bernstein, 2017). ‘People Analytics’ (PA), for example, is a proliferating approach to managing the workforce, pioneered by researchers
Hansen and Weiskopf

at the MIT Human Dynamics Lab (Pentland, 2014; Waber, 2013). It builds on the idea of a ‘data driven society’ and combines advanced methods of data collection, processing and mining for making ‘transparent’ and thereby shaping human behaviour in the workplace. However, not only is it to be expected that different modalities of transparency at play in the practice of PA, but also that different dispositional logics, ranging from discipline and law to security and care, are shaping the managerial gaze and the subjectification of employees. Exploring these logics and their intersections with transparency modalities in more detail might produce new insights into the dynamics of power and control in modern organizations.

In conclusion, using the concept of transparency matrices suggests that we speak of ‘transparency’ in the plural, rather than in the singular, and, by implication, that we should study multiple modalities, practices and technologies of transparency as well as their specific power effects. An analysis of this kind deconstructs universalizing notions of transparency and opens up a space for practical and theoretical engagement with its multiple facets and for potential transformation.

Acknowledgements

For constructive feedback on earlier versions of this article the authors especially thank Lars Thøger Christensen, Mikkel Flyverbom, Leopold Ringel, Leonhard Dobusch and the Organization Studies Group in Innsbruck. The thorough engagement of the three anonymous reviewers and Senior Editor Peter Fleming was invaluable. The usual disclaimer applies.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Other articles in Western media have referred to the social credit system as ‘a long-tentacled regime of social control’ with ‘digital surveillance’, a ‘system of reward and punishment’, and, not least, with ‘we-know-best-paternalism’ at its core (The Economist, 2016).
3. ‘Transparency’ and ‘surveillance’ are terms that are often used interchangeably in these literatures to denote such organizing processes. For instance, what Bernstein (2017, p. 222) calls ‘control-focused, purposeful observation’ in the context of transparency can hardly be distinguished from Lyon’s classical definition of surveillance as ‘systematic attention to personal details, with a view to managing or influencing the persons and groups concerned’ (Lyon 2003, p. 16).
4. Foucault theorized ‘pastoral power’ (1983, 2007) as a ‘power of care’ (2007, p. 127). It is a ‘beneficent’ (2007, p. 126) and ‘salvation oriented’ (1982, p. 214) form of power that seeks to guide or lead a multiplicity of individuals towards a specific end. Pastoral power implies specific forms of surveillance: first, the shepherd is ‘someone who keeps watch’ (Foucault, 2007, p. 127) and second, it is a form of power which cannot be exercised without knowing the inside of people’s minds, without exploring their souls, without making them reveal their innermost secrets’ (Foucault 1983, p. 214).
5. It could be argued that the dispositive of care forms part of the dispositive of security, since the latter embraces notions from Foucault’s work on biopolitics and governmentality in which the issue of care has a prominent place (Garland, 2014; Raffnsøe et al., 2016, pp. 281–282).
6. Hatton (2015) quotes the Chinese blogger Wen Quan for saying ‘a credit system puts people’s past history on the record. It’ll build a better and fairer society.’

References


Prat, A. (2006). The more closely we are watched, the better we behave? In C. Hood & D. Heald (Eds.), *Transparency: The key to better governance*? (pp. 91–103). British Academy.


**Author biographies**

**Hans Krause Hansen** is Professor of Governance and Culture Studies at Copenhagen Business School. His research focuses on public and private governance, corruption and anti-corruption, transparency and surveillance. He has published widely in the fields of international studies, organization and communication studies.

**Richard Weiskopf** is Professor of Organization Theory at the University of Innsbruck. His research focuses on ethical and political dimension of organizational practices. He is particularly interested in transparency and surveillance, whistleblowing and truth(telling) in organizations.