The Inescapability of Trust

Complex interactive systems and normal appearances

R.J. Anderson
Horizon Digital Research
University of Nottingham

W.W. Sharrock
Department of Sociology
University of Manchester

This is for inclusion in *The Complexity of Trust in Computing* edited by Richard Harper and to be published by Cambridge University Press 2013

March 2013

INTRODUCTION

As the contributions to the first and last sections of this volume indicate, trust is a problem for those who build internet services and those who are tasked with policing them.[1] If only they had good models and even better specifications of users, use and usage, or so they seem to say, they could build systems which would ensure and enhance the privacy, security and safety of on-line services. Understandably (but perhaps not wisely) they tend to be impatient with what appears to be overly precious concept mongering and theoretical hair splitting by those disciplines to which they look to provide these models and specifications. However, perhaps, an understanding of the provenance and distinctiveness of the range of models being offered might give those who wish to deploy them deeper insight into their domains of application as well as their limitations. Each is shaped by the presuppositions on which it is based and the conceptual and other choices made in its development. No one model, no individual summary of requirements can serve for all uses.

Awareness of this 'conceptual archaeology' is especially important when the model's presuppositions are orthogonal to those which are conventional in the field. In such cases, it is critical to understand why different starting points are taken and the benefits which are felt to be derived thereby. Difference is rarely an expression of simple contrariness but usually reflects deliberate choice made in the hope that, thereby, things might be brought to light which otherwise are left obscure.

There is a third point to be made here. Although a particular frame of reference or research mind set might seem to be the natural and obvious one to adopt and, indeed, be widely used, from the point of view of research method, no particular initial standpoint is mandated. We are not required to base our research investigations on any given set of presuppositions. The test of presuppositions is their fertility as a way of thinking about the research problem in hand not the fact that they happen to be popular in the research community.

All this is by way signalling our intention to adopt a frame of reference which is somewhat unusual in studies of trust and computational systems. We will start not in what might seem to be

the obvious place, namely with trust as a feature of the individual user's mental model (or whatever), but with trust as a constituent feature of a socio-technical system in use.[2] We start, then, in the same place as investigators like David Clark, Thomas Karagiannis (this volume) and Angela Sasse (this volume and Adams and Sasse 1999) with the system as a configuration of interacting users, services and other technical objects. However, we start there with a concern for users' experience as they are immersed within this socio-technical system and engaged in a flow of activities. From this departure point, we will eventually arrive at the issues that preoccupy Clark, Karagiannis, Sasse and others but will get there by a somewhat different route with, we hope, something distinctive to offer. Our intention is not to undermine, correct or criticise more conventional approaches. We see them as having much merit. However, in the collective conversation represented by this volume, we believe there should be a place for a distinctive voice and its alternative contribution.

To aid engagement with our proposals, our first task will be to summarise the key elements of the conventional approach so that we can contrast it with our way of re-conceptualising trust. We will illustrate just what this re=conceptualising makes available through an extended examination of trust in a technologically dense environment, namely air traffic control, as well as through a briefer description of some features of using internet based technology such as email. We finish with some recommendations for designers of internet services.

THE CHALLENGE OF TRUST IN CONTEXT OF THE INTERNET

Like many others who have a deep grasp of the internet and its technologies, in his stage-setting contribution to this collection David Clark bases his warning regarding the need to exercise care in their use on the distinction between actual and perceived risks. It is the gap between actuality and perception where the problem lies. There are risks everywhere and even the most technologically knowledgeable among us may not know or understand all of them. Certainly, those of us without their level of knowledge neither appreciate nor know how to mitigate them. The gap between

perceived risks and real risks is the challenge of trust in relation to the internet. Sara Bury and her colleagues (Bury et al 2010), for example, report that many (even most) people are aware of only some of the dangers that the internet poses and are particularly concerned about the security of personal information, images and photos, business information and so on. Though they don't put it in these terms, she says, they are preoccupied by the risks of phishing, spoofing, identity theft, malware and data vulnerability. They are, however, wholly unaware of other risks such as those identified by David Clark. Like Clark, Karagiannis and Sasse and Danezis, the general reaction to this is to advocate the development of a twin track response. By means of one-time passwords, public key encryption, smart cards, strong authentication, and other technical devices, technologists should develop ways of addressing the vulnerabilities of current technologies. This will make the technology more trustworthy. At the same time, major internet organisations and public agencies should undertake programmes of education and consciousness raising so that users are much more aware of exactly what the dangers are and how to reduce them. This will make the public less trusting.

This twin track approach is 'the conventional wisdom'. It has led to research which conceptualises and models trust in terms of levels of perceived vulnerability. Its starting point is an assertion such as this: once we understand how and why we think we are safe when using the internet, we should be able to design technologies and associated practices which could make us more secure. Examples of such studies have shown that users of Facebook (Fogel & Nehmad 2009) and other networking sites differ on demographic and other dimensions in their sensitivity to security issues and that these differences can be modelled by social contract theory. We know that the perception of trustworthiness (Fehr 2009) is driven by perception of likelihood of betrayal, social preferences and risk preferences and that has implications for the use of Behavioural Economics as a theory to drive design. We know that users have two kinds of trust beliefs based on their perceptions (Lankton & McKnight 2011); beliefs about technology and beliefs about other people. In relation to Facebook and other internet services, these beliefs are aligned. Users define interactive services both as a technology *and* as a quasi person. Thus they view such services in terms of the pairings: competence/functionality; integrity/reliability; and

benevolence/helpfulness. The result of all these analyses enables us to explain, or so it seems, why popular techniques like posting photographs of those delivering a service do not appear to make us more trusting of that service (Riegelsberger et al 2003).

There are several common features to these analyses. First, they are overwhelmingly cast in terms of what we will call 'bare psychology'. By this term we do not simply mean that they are motivated by reduced versions of explicit and formalised psychological theories. In addition to such simplification, trust is presumed to be obviously rooted in an individual's mental state. Trust is taken to be, first and foremost, an issue in relation to individuals and their interactions.[3] What David Clark, for example, refers to as the network level issues of trust are, on the conventional view, decomposed into the psychological properties of individuals. Parameterising these properties and controlling for those parameters is the *modus operandi* of the research.

Second, the studies are not really about trust at all. They are about *distrust* and how there is not enough of it. In order to talk about trust, investigators feel the need to frame their questions in terms of what people do or don't trust, should or shouldn't find suspicious, risky or dangerous. The concept of trust is determined by its supposed complement, distrust. It is as if we can't get a view of trust unless we go through distrust and risk.

Third, while everyone agrees the issues are interpersonal and essentially social, because trust has been taken to be a component of an individual psychological state, it is assumed the state of trusting must be the outcome of some entrusting action(s) we individually have taken or are taking which complements the ostensible actions we are engaged in. So, in using an ATM to draw out money, when posting photos on our Facebook site or ordering books from Amazon, we are doing both drawing money, posting photos, ordering books *and* trusting. But although event logs, keystroke capturing and screen scraping allow us to see what we do when we do the buying, posting etc., the trusting goes on 'in the background' as what Gilbert Ryle (1949) once referred to as an "occult process" in our heads (or somewhere) and is beyond direct (empirical) reach. It is for that reason trust has to be surfaced by asking about its absence.

TRUST AS A BARE PSYCHOLOGY

Faisal and Alsumait (2011) provide the most stripped down version of the psychology of trust. They see trust as a (mental) state where the actor presumes he or she will gain expected results from an encounter without suffering negative consequences. Al-Ani et al (2012) develop this into a pair of trust components: cognitive trust (that is, that the other has the knowledge and skills to do what is expected and reliably will do so); and affective trust (that is, that the other will satisfy the user's emotional and social expectations). Siau and Shen (2003) elaborate on this central idea and propose three core components to trust.

1. Trust occurs in dyadic relationships. There are two parties who rely on one another for mutual benefit.
2. Trust involves uncertainty and risk. There can be no guarantee that the other will prove to be trustworthy.
3. The trustor has faith in the honesty of the other who is being trusted.

The way the problem of trust is being conceived here might be summarised as follows. When two parties come together to engage in an exchange or extended interaction, trust must be built. To understand trust, we must model its build up (construction? accretion?) from a position where it does not exist to one where, through our interaction, we come to trust each other. Under this view, trust is a layering of assumptions or expectations we hold about competence, predictability and goodwill (Shau and Shen 2003) or competence and affectivity (Ani et al. 2012), reputation and identity (Clark) or, as in Sasse's case (Beautement et al. 2005), the compliance budget.[4]

The trouble which all commentators acknowledge is that the value of this analytic idealisation is undermined by users who seem to assume the basis of competence, capability, goodwill, affectivity, reputation and identity is already in place when the proposed processes for assuring that outcome have not been gone through. As a consequence, people trust the technology, or the other or both, too readily. In other words, people resolutely refuse to behave as the conventional conception of trust proposes they do (or should).

This refusal by users is the situation which the conventional approach to trust seeks to remedy. When faced with a disjuncture between what people do and what the theory says they should do, however, our response is to focus even more on what people actually do and seek an extensive description of the circumstances within which their actions are embedded. What is it about the context as they see it that allows them to be so trusting? Given we start from the point of view of the user, the central distinction of the conventional view between perceived risks and actual risks has little analytic edge for us. The only risks users orient to are those which they are aware of. The focus turns, then, to how from within the ordinary technologically enabled courses of action in which they are engaged, users take such contexts to be trustable. How is the ordinary, daily experience of using technology constructed so that trusting is the normal state of affairs it clearly is for most people most of the time.[5] We like to think of this kind of description as a 'third person phenomenology'.[6] In place of the contra-distinction between 'real' and 'perceived' risks we put a venerable sociological maxim, that of W.I. Thomas (Thomas & Thomas 1928): If men define situations as real, they are real in their consequences.[7] For us, the question becomes how trustworthiness comes to be defined as real in any given situation and so has the consequences it has.

Apart from the centrality of the 'perceived' and 'actual' risks dichotomy, the standard account of trust two further core elements.

1. Actors are defined as effectively *asocial*. Encounters, engagements and transactions happen between pairs of decontextualised and abstract 'agents' and not between members of social groups, communities or institutions. Obviously there is recognition that people do belong to groups, communities, and collectivities but these are viewed as the consequence of individual actions not as their frame. The emergence of technological and social networks is, for example, where David Clark's analysis ends.
2. From the above, it follows that the psychology invoked to motivate actors is a somewhat bare one. The social dimensions of actors' psychological states are absent.[8]

We ask what happens if one takes trust to be effectively (rather than consequentially) social in character and, therefore, to be rooted in our membership of groups, communities and collectivities. From that position, we ask exactly the same question that David Clark and most other commentators do. But in asking it, we attend to the social rather than the individual psychology of technologically mediated social interaction.[9] Such a social psychology treats trust as a routine, obvious, known and taken for granted feature of ordinary social technologically and non-technologically mediated interaction. For short, we will call this *commonsense social psychology*. The aim is to bring out what it is about social life that allows people act on the basis of a generalised assumption of trust and to assume that others are too. To put things slightly differently, we ask: What are the features of normal appearances in the ordinary (technological and non-technological) interactions of everyday social life which enable people to trust?[10]

Hopefully by now it is clear that contrast between our approach and the conventional view is not simply that simply it sets aside the 'real' and 'perceived risks' distinction and departs from a position that takes sociality as central. In addition, it demands that attention be focused first and foremost on descriptions of users' experience rather than the generalised explanations of their trust behaviour whether those be couched in terms of individual psychological traits, demographic characteristics or, as with Simpson (this volume), the structural properties of modern society.[11] Whilst we have a wealth of information about who is more trusting than whom and with regard to what, and we have a panoply of accounts of why modern industrialised societies people are (or are not) disposed to be more trusting, we know very little about what trusting looks like as a feature of commonsense psychology and especially commonsense psychology in regard to ordinary interaction with technology. Since this notion of a 'commonsense' social psychology is the pivot on which our analysis turns, we will now spend a little time explaining just what we take it to mean.

TRUST AND COMMONSENSE SOCIAL PSYCHOLOGY

Our view is one that sees trust is a routine collective *frame of mind* rooted in *experience*. It follows that sociality should be premised in an assumption of a shared *intersubjective* world. The social actors with whom each of us interacts are social beings much as ourselves. Our actions are based on the assumption they have the same order of understandings, feelings and attitudes as ourselves and act on the basis of them much as we do. They also share with us a common understanding of how social life is organised and of the rules to be followed. It is on shared understandings and shared intersubjectivity that the normal appearance of routine daily life is based.[12]

A number of further points are worthy of note.

1. Intersubjectivity is a condition of social life and our experience of it begins *in media res*, in the midst of the flux of daily life.
2. The unity of the social and material world[13] is pre-given for us. We are just as much "thrown" (to use Heidegger's expression) into the ongoing and pre-given social world as we are into the material one.
3. As fully competent participants in this social world, we all possess repertoires of skill for managing our actions and interactions. However, as with all skill, most of the time our skilfulness is rendered invisible. The fluency, deftness and elegance which we typically display when managing our social lives hides the skilful practices we use.
4. It is these social practices which are the routine joint deployment of these skills which co-produce the normal appearance of ordinary social life which we take for granted and rely on.

The research challenge is to find a way of bringing these skills into the open in order to describe how they are organised. To borrow a phrase of Harold Garfinkel's (2002) who was responsible for developing this approach, we have to "extract the animal from the foliage". If we

define successful social interaction as the outcome of skilful deployment of sets of shared practices, then, by examining the flow of the action itself, we can try to identify the practices and the skills by which they are deployed. This enables us to treat the smooth flow and seamless interconnection of actions in the social world as the jointly produced accomplishment of routinely used and methodical practice. Moreover, anticipation of the use of such skills becomes a key component of intersubjectivity and hence of the commonsense psychology which motivates the socially organised strategies whereby trust is accomplished.

We should enter an important caveat here. Nothing is being said about the goals, purposes or aspirations which participants might claim for themselves or attribute to others. These are only of interest to us as they surface in the flow of action and its outcomes, the achievements of those in the setting. Whatever the outcomes are, whatever the participants take the interaction to come to, our analysis treats the participants as "co-producing" those outcomes with just the normal appearances they take them to have. Since our topic is 'trust', this is analysed as an accomplishment, and participants are viewed as 'coproducing' trust. With this stance, we can look at how those achievements are organised and hence analyse what the mutual orientations, expectations and understandings which ground them are.

Some Working Concepts

The Phenomenological Epoché

The term epoché (or 'bracketing') is used to describe that part of the phenomenological method or attitude whereby the appearances or properties of the object under study are investigated. The analyst takes some object, say the tree outside the window and in order to reflect upon it, brackets off the-tree-in-our-field-of-consciousness from the tree about which we have experience, knowledge, understanding, presumptions, and theories. The object 'the tree-in-our-field-of-conscious' sets aside, for example, its species, the memories we have of sitting under it, what the use of its wood might be, its relationship to other trees around, the effect of the shade thrown on

the vegetable patch nearby and so on. The phenomenological epoché insists simply that we address this particular tree only as a datum of our experience here and now. Thus, the tree appears (or is 'appresented') as a locale of sensations and functions which are *egologically organised*. What this experience means, what it connotes, how it relates to other experiences we have, and the myriad of other no doubt important questions are not dismissed. They are simply set aside for now as not being relevant whilst our attention is turned to the tree as a phenomenon of immediate experience. The aim of this method is to move, step by step, through the levels of our experience, constituting each 'higher' level from those upon which it is premised.[14] In this way, what is taken for granted or assumed at one level can, in its turn, be subjected to scrutiny at another. In phenomenological analysis, each epoché is associated with a distinctive 'attitude' towards the world-under-view and a set of 'relevances' which shape the way the world is viewed. Some examples of 'attitudes' which have been examined in this way are philosophy (obviously), logic, mathematics, science, art, fantasy, theatre and religion. Each constitutes a different framework with a distinct attitude and set of relevances which bring different sets of 'objects' into view and puts others aside. As we will see, commonsense understanding in the ordinary wide awake world is an important attitude with its own distinct relevances.

The Epoché of the Natural Attitude

When turning to the examination of social life, phenomenological analysis noted that social actors adopt what is described as 'the epoché of the natural attitude' (Schutz 1982). This epoché involves suspending the possibility that the world is not as it appears. We assume the reality, facticity, veridicality, naturalness or obviousness of the world we experience in order to get on with ordinary social interaction.[15] This assumption is, of course, 'until further notice'. The epoché of the natural attitude is the commonsense basis for social life and the requirements of engagement with others in social life provide its relevances.

If we apply these ideas to the problem of trust in socio-technical systems, it soon becomes clear that it is the epoché of the natural attitude (i.e. what users 'bracket', take for granted, assume to

be the case) which causes the distinction between perceived and real risks to collapse. For someone engaged in a flow of activity using a system, perceived risks are the actual risks.

Asking how the normal appearances of trustability are constituted as the taken-for-granted facticities they are seen to be is an attempt to answer the same questions that David Clark asks at the end of his contribution. As we have just discussed, the commonsense properties of any social phenomenon (just what they are *for those in the setting*) can be treated as the jointly produced outcomes, the achievements, of those participating in that setting. With this as our frame of reference, analysis of the social character of trust becomes more tractable.

Three general analytic themes can be used to provide the description we seek. In the next section, we offer a little elaboration of them and how they might be used to analyse trust in relation to internet services. In the succeeding section, we will examine them in more detail in a very different domain, namely Air Traffic Control. The themes are:

1. *The natural metaphysics of a setting*: What 'objects' are oriented to and deployed in the setting and how are they organised, related and recognised? For example, what, from the flow of interaction, can we say such things as the internet, the user, a rendered web page, a broken link, a friend, malware, a site, a posted entry are? How are they constituted, recognised, oriented to, organised and related? How are they classified, grouped and arranged in this setting for the tasks that are underway? The metaphysics so described is, of course, a culturally given one.
2. *The situated reasoning of the setting*: What sets of relevances do social actors have and how do they recognise and account for similarities and differences, relationships and discontinuities between objects and classes of objects, projected and actual outcomes of actions and so on? How are causal and other sequences of actions produced so that a train of events becomes recognisably 'an attempted phishing', clearly 'a broken link', demonstrably ‘a friend's comments' and so on. Here two distinctive clusters of notions are important; "local historicity" and "natural accountability". The former refers to the precise course of actions and treatments

through which some particular phenomenon passes on any particular occasion. The latter is the set of practices whereby the sense or meaning which some phenomenon has, what it 'really' is, emerges in the context of particular events as the 'local historicity' of that object unfolds. In the ordinary flow of activity, what some object 'is' becomes the potentially revisable outcome of the interaction so far. The qualification is important since, if, as the interaction develops, things turn out to be different to what was anticipated, what the object 'is' undergoes revision.

3. *The specifics of the context*: How is the activity in hand produced so that it is obviously and recognisably *really that* activity and not a charade, mock up, spoof, scam or other construction? How do cultural objects display their "normality" so that they are immediately recognisable by any user. What are the "thises and thats" which demonstrate that this object is just what we take it to be.[16]

We will now provide a brief sketch of how these ideas might be applied to the description of very familiar aspects of internet interaction mediated by technology, namely managing email. This will act as a bridge to a broader description of trust in a dense technological environment, namely air traffic control.

TRUST, THE USER AND EMAIL

Imagine the following scenario. You have a set of documents to send to an organisation. They concern the estate of a relative or some such matter. You are not sure if the package is oversized and what the postal cost will be. You take the package to the nearby Post Office and ask the counter clerk. The counter clerk weighs the package, calculates the cost and puts stamps on the it. He then puts the package in a sack. You pay and walk out.

An episode, we hope, everyone recognises. If we look at it from the perspective we have been outlining, a number of facets surface.

1. Your actions and those of the counter clerk are premised in *sets of relevances* given by your different situations. You have documents to send. He has a job to do. This set of paired relevances is shaped by your and his reasons for acting.
2. You assume that these relevances provide *typical and dependable motives* for the counter clerk and others whom you expect to be engaged in the projected trajectory of action (various workers for the Post office, the members of the organisation who will process our documents, and so on). These motives complement yours. This assumption sets your *expectations* of what will happen.
3. These motives are associated with *types of social actors*; personal types such as the counter clerk and anonymous course of action types such as whoever processes the documents.
4. Given such expectations about typical relevances and motivations, you take for granted that these others, these typical actors, will do what you expect of them and so do not feel the need to ensure the clerk has calculated the cost correctly or has put the package in the right sack. You do not seek re-assurance that the Post Office employees will carry out their roles and take the package where it is supposed to go.
5. In addition, you anticipate they will hold expectations about yourself as a course of action type, someone who wants to post a package, and so you shape your behaviour to fit with their expectations. You write the address legibly and give the post code. You pay the price of the stamps, and so on. You *design* your actions to fit with your projection of their actions.
6. You do not feel the need to check that the documents have arrived or that they are being processed although of course you *can* do this, and sometimes you might. This time, though, the situation does not call for it. Your experience of the interaction allows you to take it for granted that the normal trajectory of actions will be carried out.
7. Finally, you expect to be able to do much the same thing any time you need. The typicality of the action just is that we can do it again.

So much so trivial. But it is in the trivial that we trust. The points we have just made are what 'trust' in the sending of a package comes to. In any course of action, it is the relevances we and the others engaged with us have which determine the patterns of typification we assume to be in place. The intersubjectivity of social life is the interlocking interpretations of these patterns of typicality. It is only when normal appearances fail with objects and relationships missing or used in ways we do not expect (the counter clerk throws the package in the air and calculates cost on the loudness of the 'thump' it makes when it lands) or projected courses of action do not materialise as we anticipate (the package never arrives) that we reach an 'until further notice' point and question what otherwise we have taken for granted. Commonsense rationality consists in following typical courses of action in expectable ways; in doing the same thing the same way, again and again. Being rational in ordinary life means, then, in trusting the counter clerk, post men, and officials in the processing organisation to do what we expect them to do.

We have chosen the posting of a package as our example simply because email systems attempt to capture much of the familiar processes of the postal system. When we send email, our relevances shape the patterns of typification we have. We assume 'someone' has designed the system to enable mail to be sent where we want it to go. We assume 'people' working for the network will keep the systems running to allow the mail to be delivered. We assume they have no interest in our mail *per se* and simply facilitate its passage through the system. We assume the designated recipient will receive the note and will read it. We assume the motivations of all these actors are complementary to our own and their relevances dovetail with ours. Because 'this time' it goes the way it always has, we take it that this is just another ordinary mail note being sent or received. Moreover, and this is crucial, the system has been designed on the presumption we will trust it. Our trust in the system is a pre-given in that the patterns of relevance we, as users, are assumed to have and the patterns of typification derived from them. The system works *because* we trust it.

A key term in the description we have just give is the *typicality of our experience*. This is experience is as we expect it to be. But how do we expect the experience of sending email to be?

Once we ask this question, what appears to be a fundamental and natural metaphysical distinction, that between the user considered as a bundle of social, cognitive and biological properties and the bundle of hardware and software properties designated as the email system can be brought into focus. As is shown by both Banks (this volume) and by Sasse (this volume), design of the user interface, be it command line, wimp, voice or whatever, is the explicit attempt to manage this deeply ingrained distinction. Putting it another way, those user interface designs which are successful (in whatever ways one wishes to measure such success), are so because of their capacity to solve the problems engendered by this separation. Somehow or other,[17] our experience consists of reaching across the divide to manipulate the system or to communicate through it in some way. The affordances they provide for closing the gap between the user and the system underpin the attractiveness of the most widely used metaphors for designing email and other interactive systems.

When we look at email use as a stream of experience, the natural distinction between the user and the system dissolves. Instead, what is foregrounded is the system-and-user as a cultural *object*. What, at any particular moment, is the system *for* the user and where the boundaries of *the-system-in-use* might be for those working with it can be brought into view. When composing an email we do not think about what the system is doing . Neither do we ask where the writing takes place. On the keyboard? in the mail window? Where? Just as writing with pen and paper is only rarely experienced as having a thought, moving the pen and expressing the thought, writing email is not pressing keys and watching bit maps. It is thoughts flowing through words. When scribbling a note to a friend, we are at one with the pen and paper. When typing email we too are at one with the system. In the experience of the task, there is no *perceived* difference between the user and the system. The commonsense reality of using email is that the system and the user are one. Because of their own distinctive relevances, our technical ways of describing the use of a system fail to capture this. Distinctions such as user and system are important in technical talk. But do they matter for the user in the midst of using the system? Does the practice of using email depend upon holding to a system-user dichotomy? Questions such as these indicate how we can make the

distinction between system and user move from being a given in the design to a topic for analysis in the service of design.

Bringing the ideas of relevances, typifications and commonsense realities to bear upon interactive systems such as email raises a number of possible lines of investigation. The first is that any description of the setting/working system becomes a description of the system-as-seen-and-produced-from-within flow of action. Second, and closely related, the constructs around which that system is organised (such as the unity of the user and the system) are resources for users in their sense making of the working system. The working of the 'working system' produces and re-produces the reality, validity and veracity of these constructs as the normal, ordinary and routine things they are. Immersion in the flow of action brackets questions about the dependability of the system or its trustworthiness.

Here lie the key insights we think our approach offers. First, it suggests that trust is the default mode of ordinary social life. We assume things are as they seem to be. Moreover, when we do call things into doubt, we do not suspend doubt *tout court*. It is not possible to doubt everything at once. To scrutinise some things, we still have to take other things for granted. Second, it follows that if we want to re-shape our encounters with technological systems such as email so that we are less trusting, we will have to design them in ways that do not, either overtly or tacitly, call into play the structures of typification we deploy in ordinary life. The metaphors and analogies which seem natural or intuitive and which allow for ease of use are precisely those which allow commonsense rationality to be deployed. If we don't want users to trust the system then we have to find ways of breaking the frame on which trust relies. In the next section, we will show just how hard this is to do.

AIR TRAFFIC CONTROL: A TECHNOLOGICALLY INTENSIVE ENVIRONMENT

Richard Harper's introduction to the volume summarises how the sheer complexity of modern technological (especially web-enabled) systems and services seems to have made the issue of trust

intractable. However, whilst the technologies might be different as technologies, as places where interaction in managed through technology they do not appear to be all that different to the environments and locales of work which we and others have studied in the past.[18] To demonstrate this similarity and further illustrate how the issue of trust in a technological system might be re-conceived, we will re-visit research carried out over 25 years ago at what was then the London Air Traffic Control Centre at West Drayton.[19] We will show the detailed social organisation of trust to be a constituent feature of ordinary computationally mediated work practice. Because the character and recognisability of ordinariness is embedded in the local physical and organisational detail, we will begin by summarising the setting as it existed then and how air traffic control was carried out (and, by and large, still is). For ease of reading, we will set our description in the historical present. We are not here concerned to mark how air traffic controlling has or has not changed. Nor are we concerned with its contemporary use of the internet as an infrastructure. Our interest is simply describing how, when we studied air traffic control, the routinisation of trust was evident all around.

The setting: London Air Traffic Control Centre

The organisation of air space above The British Isles is somewhat complicated. The simplest division is that between 'controlled' and 'non-controlled' air space. In the latter, aircraft are largely free to move at will. In the former, all aircraft must be controlled by an appropriate air traffic controller (ATC). Controlled air space takes three forms: en-route sectors where planes are at or approaching their cruising heights and speeds, Terminal Manoeuvring Areas (TMAs) where streams of planes seeking to land or take off are organised, and Aerodrome Approach where planes are taken into land.[20] The control of en route sectors and TMAs over England and Wales is located at London Air Traffic Control Centre (LATCC).

The suite at which a controller sits consists of buttons surrounding the circular luminescent green radar screen. Above these are screens displaying information of various different kinds. Away to each side are clacking line printers. Each controller is hooked into the suite by the trailing cable of a head-set and microphone. In the centre sits what is possibly the only anomalous feature,

a wooden tray holding printed strips with hand written notes scribbled all over them. These are the 'strips'.[21] To ATCs, the strips are the key to good controlling. As one controller said: "You have got to have a complete picture of what should be in your sector and what should be in your sector are on those strips." He went on to describe their use:

> It is a question of how you read those strips.....An aircraft has called and wants to descend. Now what the hell has he got in his way? And you've got ping, ping, ping, those three. Where are those three? There they are on the radar. Rather than looking at the radar. One of the aircraft on there has called. Now what has he got in his way? Well, there's aircraft going all over the place, now some of them may not be anything to do with you. It could be above them or below them. Your strips will show you whether the aircraft are above or below, or what aircraft are below you if you want to descend and aircraft, and which will become a confliction. .....You go to those strips. You pick out the ones which are going to be in conflict if you descend an aircraft, and you look for those on the radar, and you put them on headings of whatever. You find out whether those, what those two are...which conflict with you third one. It might be all sorts of conflicts all over the place on that radar. But only two of them are going to be a problem. And they should show up on my strips.

Flight data strips are about 1 inch wide and 8 inches long. They specify the flight path of an individual aircraft. This includes the aircraft name or "call sign" and type, its departure and destination point, its preferred route, height and speed. In addition, the estimated time of arrival at certain navigation points in the sector is printed at the side. Each sector has three of four key navigation points, a new strip being printed for each aircraft relative to each point. The strips are placed in racks or "bays" just above and behind the radar screens. Strips are printed 10 minutes or so before an aircraft is due at a point. Strips record the aircraft's passage through the sector. As each point is crossed, the relevant strip is discarded.

This record is what controllers attend to and use in their work. It is the material instrument and work site of controlling. However, strips do not determine the sequence of actions which controllers perform in the same way that whatever comes along the production line determines what the line worker has to do next. Rather, the controller has to organise the strips so they can become a resource to help organise the work of controlling. Strips are 'glanced at', 'searched for', 'taken heed of', 'ignored', 'revised', not just when they first arrive but continuously. This is the work of organising the 'doing' of the work. As a result, what the strips provide the controller at any moment is an 'at hand' and 'in hand' sequence of actions through which to create 'order in the skies'. Management of the strips is, then, a very large part of management of the traffic.

The controller's problem

At its simplest and most general, the controller's problem is a scheduling one. For any controller controlling any segment of air space, the traffic has to be taken as and when it arrives and threaded together into an orderly pattern before each individual plane is handed off to the next sector or controlling segment.[22] All of this scheduling and traffic management has to be achieved in and through making the traffic flow. Aeroplanes cannot be "parked" for a couple of minutes; nor can traffic jams be allowed to occur. Even when they are put into holding patterns of various sorts, aircraft are still on the move, part of the flow of traffic.

The controller utilises a number of different resources to solve the scheduling problem. Two are, in essence, technologically determined since they are related to or constrained by the hardware and associated software of the suites themselves. They are:

1. *Information resources*:
    a. the radar screen and its displayed data;
    b. the flight strips
    c. screens of weather conditions.etc
2. *Communication resources*:
    a. radio-telephone to aircraft
    b. telephone links to other controllers etc

c. face to face communication with the suite team

In addition, a further vital resource is the controller's working knowledge of the system itself. This accumulated, know-how, know-what of years of experience is brought to bear on the technologically provided resources to determine at any particular time what appropriate courses of action should be.

The point is an obvious and well known one (Reason 1986). The Air Traffic Control system comprises numerous complex sub-systems, instantiated in hardware, software, regulations for controlling, working practices and the like. In the face of the ordinary contingencies of practical working life, conflicts, inconsistencies and incompatabilities among these sub-systems are bound to arise. These constitute 'the normal, natural troubles' (Garfinkel 1967) of the controller's working life. In dealing with these troubles, the controller displays 'normally competent controlling'. Being able to recognise them as the 'normal, natural' phenomena they are is, to some extent anyway, what being a competent controller involves. Since these troubles can occur both with the traffic and with the technology, their solution is achieved by managing the traffic through managing the technology.

The skills required for this management (that is 'working with the technology') are multi-layered and interwoven. Moreover, they often seem to lack the sense of planfulness deliberation, cogitation, task-definition, specification, calculation and solution seeking that conventional approaches to system design presume. Rather, the process is an interpretive one but in a somewhat different sense than normal. The controller just *knows* what to do. What 'knowing' means here is ‘interpreting the conditions at this suite at this point in time against a background of what has gone so far, what time of day it is, where everything else is currently, what has yet to arrive, what is going on in neighbouring sectors, and everything else that the controller takes to be relevant’ to the task in hand. The whole is a *gestalt contexture* which provides the meaning of what is going on. The problem the controller faces is *this* problem here and now where *that* is obviously the appropriate course of action to take. The controller experiences problems and their solution, then, as part of a flow of work.

Air Traffic Control as a Division of Labour

At LATC, control of air space is managed within an extensive division of labour. This fact is known and used by all those who work there to explain and inter-relate the activities of controlling. Controllers, managers and others encounter LATCC as a working division of labour and depict the organisation of tasks accordingly. Consequently, the division of labour is a commonsense construct oriented to and used by controllers and others (Bittner 1965). Viewed from in the midst of its operation, the division of labour is encountered as a body of fragmented tasks and activities. Working within a division of labour, controllers encounter it not analytically as an integrated totality but practically as a stream of differentiated tasks to be performed. Tasks unfold as things to be done now; things that can be left until later; things that are tied to the completion of other tasks and so on. We can think of the division of labour as having almost a *transcendental presence*. Any individual task gets its sense from and hence contributes to achieving its overall structure. Seen from within the division of labour, the overall structure is not a unified and totalising rational scheme. Rather, it appears to be organised around a principle of *egological determination* in an environment which is information saturated.

The organisation of activities around any control suite consists in a number positions occupied by particular persons. Exactly how those positions are distributed is locally determined to best fit the management of the work in hand. From the point of view of any of these positions, the work to be done appears as a continuous, impersonalised stream. Within the bounds of training and regulation, it is immaterial who occupies what position. In other parts of the division of labour, of course, this might not be so. The differentiation of tasks and the related hierarchy of responsibility (as is depicted in the rational organisational scheme) is embedded in the flow of activity as an institutionalised structure of "decisions-I-can-make" and "actions-I-can-take" together with those that others deal with (hence our description of it as 'egological'). Processing the endless stream of

tasks means no more and no less than doing-what-I-do and passing things on so they can do what they do.

The egological principle is centred on the location of the individual within that structure of activities. The boundaries to spheres of operation vary from those that are permanently open and those under view, hence *near to hand,* and those which are *at remove* and so taken for granted. The distinction between tasks which are near to hand and tasks at remove is expressed in a number of ways.

1. There are tasks together with their associated rights and responsibilities whose performance are never a matter of enquiry. Other tasks, however, must be constantly appraised. Individual controllers need not concern themselves (in fact, cannot concern themselves) with the state of activities on a neighbouring sector, even one whose interaction with their own is intense. Similarly, it is of no concern why the upper limit of a holding stack may have been lifted from its normal setting at FL 130 (13000'). Surmises can be given from what "everybody knows" about controlling, but doing what has to be done does not require these issues to be even minimally investigated. The operating division of labour rests upon a presumed *symmetry of structure* and *reciprocity of location* without the need for an understanding of the precise detail. Controllers assume the organisation of work for others is much the same as it is for themselves.
2. Correlated with the presumed symmetry of structure and reciprocity of location is an *horizonal structure of relevances*. There is every reason for the inbound controller of the Heathrow northern TMA[23] to ensure that the 'squawks' (calls signs) displayed on the screen are correct by requesting identification from the aircraft when initial contact is made. Equally, it is crucial to ensure the statements of height transposed from the screen match those marked up on flight strips. There is much less need to know if the code for the airport from which the aircraft departed is correct. The correctness of this data is, literally, someone else's problem.

3. The egological principle provides a working solution to the problem of *task performance*. The division of labour specifies just which tasks one has to embed one's own activities within and which tasks are institutionally taken care of, so to speak. Competence is the achieving of such embedding or the calling into play of the institutional structures.
4. Coordination within the division of labour is both *ecological* as well as *egological*. Activities are performed in organisationally specified zones and niches. Some are technologically fixed; others are not. During busy periods, high work load sectors can only be really managed by further dividing the sector and allocating two controllers to the same screen. Similarly, the Sector Chief can only supervise the sector if he can see both screens and can manipulate the associated flight strips. Other activities are freed by technology. Because of telephone links, a geographically adjacent sector of sky with which a sector has a lot of interaction need not be a neighbouring suite. Indeed, fluid management of traffic requires that transactions between suites is independent of the layout of the Operations Room. At the same time, those who know the layout have a geography of institutionally specific locales where things get done. Thus, even the spatial layout of the Operations Room is saturated with information.

To summarise: from within, the division of labour is experienced as a fluid *gestalt contexture* of task performance. Tasks move back and forth, from foreground to background with an associated re-structuring of their relevances. The elements of this gestalt are tasks performed, those in process and those awaiting take up. All are thematised by the egological structures of relevance just described. This gestalt is evidenced in innumerable ways and is available at a glance to the competent controller and is what grounds the controller's trust of the system.

So far we have been describing how immersion in a division of labour is both a feature of and re-produces trust-in-the-system. Doing things normally, competently, routinely and appearing to do so is trust-in-the-system. To use the term we used earlier, the normal, natural attitude of competent controlling is premised in a bracketing of all sorts of matters which could, of course, be enquired into but are not. Trust in the normal, routine, expectable working of the division of

labour allows controller to embed tasks to be done now into the orderly flow of activity and thus to reproduce that orderliness. Being a competent controller means taking that trust for granted. Trust, then, is a given for the socio-technical system of air traffic control. Were that status, that bracketing of trust, to be withdrawn, normal, routine, competent controlling would become impossible.

We now want to turn away from general descriptions to look at two particular activities where this taken for granted character of trust can be exemplified. Both involve the collective or social use of technology to perform complex tasks.

Silent Handovers

For any particular controller, management of the 'blips-on-the-screen' and strips in the bay is the management of planes in the sky. The order on the screen and in the bay is proxy for the order in the sky. Controlling the progress of these objects is instructing the plane. This is the controller's task; a task which is completed with the transition of the object across the screen, down the bay and the plane out of sector. At such transitions, the task of controlling is passed on to someone else. One of the most striking things about routine controlling is the extent to which such transitions are managed with minimal or even non-existent exchanges between the controllers concerned. These transfers are called "silent handovers".

The accomplishment of a silent handover is evidence of the normal working of the system. Blips and strips appear on the relevant screens and in the relevant bays with the right codes and values, in the right order, and at the right time. 'Right', of course, means in correspondence with standard procedures and practices (or with whatever exceptions are in force). As the receiving controller, the routine controlling work done somewhere else has made your task unproblematic. We are not saying that all handovers are silent nor that explicit and extensive co-ordination work is never required. However, such explicit transitioning work is itself institutionalised, so that the 'repair work' required on the routine can be effortlessly undertaken. At times, such repair work is itself managed 'in the background', as when the Sector Chief accepts into sector a military crossover or an uncontrolled aircraft using visual flight rules (VSF).

The silent handover demonstrates that controllers can and do manage traffic flow secure in the knowledge that planes will be picked up and safely managed, whilst receiving controllers can be assured that the blip-on-the-screen *is* the relevant aircraft-in-the-sky and that it has arrived there in all the expectable ways it should. Silent handovers are testimony to the way normal appearances are used for the effective management of the system. Indeed, given traffic loads and complexities, silent handovers are a *necessary* way of managing the system. They are made possible because trust-in-the-system is built into the culture of air traffic control. Without silent handovers, managing transitions would impose more work on controllers and hence require more time to deal with individual aircraft. This in turn would make the system less efficient and less trustworthy than it is.

Stack Jumping

Stacks are located in the London Terminal Manoeuvring Area (LTMA), a sector of the airways roughly co-incident with South East England. For controlling purposes, LTMA is divided into a North and a South sector, each of which is further divided when required. The primary task of LTMA controllers is to separate outbound traffic leaving Heathrow and Gatwick and climbing to the levels stipulated by their Standard Instrument Departure patterns (SIDs) from aircraft inbound to the same airports. At busy periods, to ease congestion in-bounds are sent to one of several locations, or holding points, where they circle until space is available for a landing approach. These locations are the "stacks". Heathrow has four (over Lambourne, Biggin Hill, Ockam, and Bovington). Gatwick has two (north of Burgess Hill and over Mayfield). The number of aircraft in each stack and the number of stacks in use varies with how busy the sector is. As airspace fills, aircraft enter or are "stacked" at higher and higher levels. Each plane is separated from those above and below by 1000 feet. As planes leave the bottom of the stack, those above are directed down one level.

The purpose of stacking is, of course, to turn varied pulses of aircraft coming from all directions into a co-ordinated and predictable stream of planes which airports can handle. The controllers only have to direct aircraft to the top of the stack while Approach Control (situated at the airport) takes them out from the bottom. On the other hand, departures consist in a

continuous stream which has to be distributed across the various route ways. The LTMA Controllers receive traffic from the main airports and must direct it around the in-bounds before allowing it to turn away onto its designated routes. In practice, this involves threading planes around, through or over the stacks.

As one would expect there are sets of procedures for these tasks. Within LTMA sectors, the most important of these procedures relate to the standard profiles of aircraft inbound and departing from the airports and related stacks. These procedures are laid down in the operational manuals and take the form of

1. Standard Instrument Departures (SIDs) which detail the exact trajectory of outbound traffic and are designed to satisfy noise abatement requirements and ensure no flight conflicts with in-bounds.

2. Standard Arrival Procedures (STARs) designed to co-ordinate in-bound traffic. As with SIDs, STARs reflect the destination and route of the aircraft.

3. Agreements between LTMA and neighbouring sectors on which levels aircraft should be handed over.

From both the controller's and the aircraft's point of view, the standard procedural flight rules may not always be the most expeditious way of managing the plane or the traffic. Nor do they necessarily ensure safety. STARs and SIDs are complex and somewhat restrictive because they have been designed to weave traffic through but away from all other traffic. In addition, they do not take account of the differences in performance of aircraft. They are general purpose specifications which any aircraft can follow. As a consequence, following a SID can result in delaying the ascent of an aircraft to its optimal cruising level and speed, thus prolonging the flight, increasing cost and creating extra pilot work. In addition, it can create more work for controllers since planes on STARs and SIDs can be in a sector for much longer than they need and so use up airspace, RT time and require extended controller attention.

Not surprisingly, then, controllers have developed well known and shared procedures for dealing with the “troubles” which the conflicting demands of STARs, SIDs and stacks create. These procedures are an essential part of professional competence and controlling skill. They and the techniques associated with them avoid delay, reduce work-load and contribute to increased safety by reducing the time an aircraft remains in a busy sector. They are ways in which expert controllers apply their expertise by working within the system to manage the system. When faced with the possibility of, if not conflict then certainly inconsistency between sub-goals of the system — eg segregation of traffic and expeditiousness, controllers use the resources provided by the system to achieve working and workable solutions.

"Jumping" a plane through the stack can only be done because the controller trusts in and is at one with the system. Being at one with the system is a crucial element of trust in the system. Although the aircraft remains at a low speed to satisfy noise requirements, from the configuration of in-bounds, both in the stack and on their way to it, as well as those under the control of the Approach Controller, the controller senses there is enough "space", for the plane to jump through the stack. The value of "enough" here is "enough to satisfy the requirements of safety and competent handling". The former are defined by the Air Traffic Control Manual and the latter by the practices of ATC at LATCC. For example, two planes are circling in the Biggin Hill stack, one at 7000' and one at 8000'. An out-bound on its way to the Daventry sector would only have to climb to 9000' before or by Biggin Hill to be safely clear of the stack and so able to continue its climb out of the LTMA before it has reached the northern geographical boundary. Such a manoeuvre allows the out-bound to "jump" all the in-bounds under TMA control and will, almost certainly, allow it to continue its climb in the relatively empty sector above the TMA much sooner than allowed in the SID.

On the face of it, the practice of "stack jumping" looks to be a relatively straightforward tactic. Just what you might expect experts to do. The point, though, is not that controllers produce a "simple and easy" solution to a problem (which they do), but rather the work and skill which allows them to *see and feel* just how and where the system affords a solution to problems it

has itself created, and their trust of the working socio-technical system that enables them to employ it in the ways in which they do. Effortless though it appears, this work, this expertise, is by no means simple to describe nor easy to acquire.

To begin with, stack jumping requires a complex series of judgements about the changing structure of the traffic flow, the performance characteristics of particular aircraft and an awareness of everything else that might be relevant to the current state of the traffic flow. Controllers refer to this as their "picture". The need to hold this complex gestalt in mind when deciding whether to "jump" a plane is the reason stack jumping is rarely practised by novice controllers or those who have been off duty for some time. When jumping a stack, previous out-bounds have to be considered in case they are slow and thus likely to be in the way of subsequent, faster planes. Or there may be too many planes converging on the stack at its top level indicating that it could have to be raised before a possible "jumper" could get there. On the other hand, there may be the possibility of creating space at the top level by slowing down all the planes approaching the stack. Added to this is the fact that the speed of modern planes is such that often there are only moments to notice an opportunity and decide which out-bound to jump.

The advantages of stack jumping for both controllers and the system are obvious. It ensures quick exit of aircraft from the sector. It frees RT time, gives the plane to the En route controller earlier which can ease the handling problems of in-bounds. It enables planes to reach efficient operating height and speed quicker and, since it is a simpler trajectory, often increases passenger comfort. So keen are some pilots to jump that on their first contact with LTMA they "offer good climb rates" to controllers. On occasion this creates a situation where the issue of 'trust' (or lack of it) becomes explicitly visible and managed in the flow of controlling. Because of the efficiency and cost gains from stack jumping, pilots may offer climb rates which the controller knows or suspects are technically or practically infeasible. Given what depends on the rate being achieved, under certain circumstances the controller may decide not to trust the pilot and so refuse the climb rate offered and direct the aircraft on another routeing.[24] In addition, many aircraft do not have the capability to climb as fast as jumping requires. However, pilots know that informing a controller

early that such a climb rate is possible greatly increases the likelihood of being offered a chance of being put on that vector.

As is to be expected, stack jumping operates under limitations. Apart from the need for quick assessments of such situations, the most troublesome is the failure of the plane to reach its directed climb rate. There may be various reasons for this, but it has serious consequences. A controller may be depending on an outbound to climb in front of an in-bound, and if the out-bound does not climb fast enough a possible "confliction"[25] may occur. Other problems relate to the distribution of controlling responsibilities between the Approach and LTMA Controller. Occasionally, an aircraft might be directed a sector airspace without prior co-ordination. Rather than take a plane over the stack, the LTMA Controller, for example, may choose to route it around the middle of it and hence through the flight path of those emerging at the bottom of the stack (remember the aircraft in question is climbing all the time). On other occasions, a plane may be taken from mid-way up the stack. In cases such as these, one or both planes may have to be re-directed.

Stack jumping requires intimate knowledge of the routines of the sector and of the aircraft currently being controlled, traffic management requirements, an awareness of the amount of attention the controller must give to any one manoeuvre, and much, much more. Such knowledge is sector specific. This knowledge has to be applied and honed time and time again to allow the procedure to be effective, smooth and trouble free. It requires exact assessment of the progress of aircraft along their given vectors and where "in the sky" they are in relation to one another. These assessments are based on information 'seen at a glance' with the appropriate course of action being "executed' immediately and without deliberation.

Stack jumping not only depends upon trust in the system, it *is* trust in the system. For all these reasons, novice controllers generally shy away from it. The skill and the work by which it is brought off are made invisible by the very effortlessness of the achievement. That experienced controllers do not vacillate and ponder the possibilities; that they act smoothly and efficiently to produce the space for a jumper to jump through with no hiccoughs, finger crossing, wood

touching, drastic changes of mind or direction make it difficult to see the artful and professional handling of the system which makes it all possible. This is all the more so since such artfulness and skill are to be seen only ephemerally in the orderly progression of planes-on-the-screen, strips in the bay, inscriptions on the strips and exchanges with aircraft, controllers, and so on. This skill of competent controlling through stack jumping involves working the system to satisfy the procedural rules of air traffic control, where what counts as satisfying the rules is the production of smoothly flowing traffic and demonstrably competent controlling of whatever aircraft are in the sector at any moment.

CONCLUSION: WHAT TO DO ABOUT TRUST?

We have no doubt that the concerns that many commentators point to in relation to trust and computational systems are real and potentially dangerous. Like them, we think that technology providers, public agencies and those in relevant responsible positions have a duty to raise general awareness of the range of threats we face. However, we do not think that the approaches being advocated for the design of complementary technological solutions are likely to work. In fact they could even make things worse. This, we suggest, is the consequence of the limited construal of the notion of trust which is adopted and the 'bare' psychology which motivates it. We have suggested an alternative approach, one which is rooted in a particular social psychology in which the users of technologies are conceived *ab initio* as social beings sharing an intersubjective social world.

Adopting this point of view has allowed us to suggest that trust is not optional for normal social life. It is a given for and the basis of enacting the social relationships and functions with which we are all familiar. It is, we suggest, the default mode in all routine activity. Two things follow from this. If we want people to be less trusting in certain circumstances or in regard to specific particular types of technical object, then we will have to think very carefully about the use of what appear to be especially "natural" metaphors and analogies when designing systems for use in those settings. The resemblances we trade on in using such tropes predispose the user to take for granted

the relevances and typifications associated with them. This taking for granted is designed *into* the system and enables the systems to be used as they are. If we want to make the issue of trust more overt in the daily use of technologies then we will have no option but to question some of the core tenets on which usability as a design goal has been enshrined over the last half a century. That will be a severe challenge.

Second, even if we succeed in designing so as to reduce the level of trust we place in some technologies, we cannot design for its total suspension. In distrusting certain kinds of message, certain orders of instruction, certain types of location, certain types of technology, we will have to trust others. This will mean being very specific about when, where and with what and whom users should be careful, knowing full well that elsewhere and with others they will not need to be. Of course, this is just another way of saying that design will continue to be an endless game of catch up. As we engineer withdrawal of trust from some activities and technologies, so others will be at risk of being suborned.

Third, the panoply of activities that are underway in any complex operational socio-technical system constitute a multi-dimensional, tightly embedded mosaic. Designers would be wise to design for the whole mosaic rather than seeking to partition off of a particular enclave, since introducing a policy of distrust in some part of the complex will undoubtedly generate turbulence for the entire system. An organised and orchestrated management-endorsed-and-promoted policy of systematic distrust in an environment such as Air Traffic Control will generate far greater difficulties than the normal natural troubles which controllers and dealers are used to and comfortable with. If they cannot trust the system, they cannot make the system work. The same also goes for email, web browsing or internet banking. If we withdraw trust when using these systems, we cannot make them work either. Since we can't solve the 'problem' of trust once and for all, trying to do so will, in all likelihood, simply generate other, perhaps catastrophic, problems instead.

Finally, We have made great play with the suggestion that different departure points and different modes of reasoning bring out different features for examination. We recognise that, at

one level at least, this is trite. But at another level it is important. Too often in research, traditional ways of framing problems come to be unquestioned. As a consequence we are blinded to possibilities and insights they do not encourage or permit. We have suggested that we need not hold to the unquestioned assumption that our understanding of trust in relation to technical systems should be framed by contrasts between real and perceived risk, users and the system, psychological and social characteristics of the user, or investigated solely by the search for explanatory factors. Such assumptions may be the taken for granted of our research, but they don't have to be. But, as with trust, that does not mean we can proceed without any assumptions at all.

REFERENCES

Al-Ani, B., Trainer, E., & Redmiles, D.	2012	Trust and surprise in distributed teams. ICIC'12, March 21–23, 2012, Bengaluru, India. pp 97 - 106
Anderson, R. & Sharrock, W.	1993	Can organisations afford knowledge? *Journal of Computer Supported Collaborative Work* vol 1 no 1, pp 145 - 161.
Anderson, R., Hughes, J. & Sharrock, W.	1987	The Divsion of Labour in Conein, B., De Fornel, M, Quere, L. *Les Formes de la Conversation*. CNET Paris pp237 - 252
Baldamus, W.	1957	The relationship between wage and effort. *The Journal of Industrial Economics*, vol 5, no 3 pp 192 - 201.
Beautement, A. Sasse, A., & Wonham, M.	2008	The compliance budget. *NSPW '08*. March 22 -29, Olympic Valley, California, pp nn - nn
Beck, U.	1992	*The Risk Society*. Sage New York.
Bittner, E.	1965	The concept of 'Organisation'. ISocial research vol 32 no 3, pp239 - 55.
Bury, S, Ishmael, J, Race, N. & Smith, P.	2010	Designing for interaction with mundane technologies. *Personal Ubiquitous Computing* vol 14, pp 227 - 36.
Button, G. (ed)	1993	*Technology in Working Order*. Routledge. London
Cerf, V.	2010	Trust and the internet. *IEEE Internet Computing* September/October pp 95-6
Cook, K, & Gerbasi, A.	2011	Trust in P. Hedstrom & P Bearman (eds) The Oxford Handbook of *Analytic*

Sociology. OUP, Oxford pp 218 - 244.

Emory, F. & Trist, E. 1972 Toward a Social Ecology. Plenum Press, London

Ess, C 2010 Trust and new communications technologies. *Knowledge, technology and Policy*, vol 23, pp 287 - 305.

Faisal, M. & Alsumait, A. 2011 *Social network privacy and trust concerns*. *iiWAS2011*, 5-7 December, 2011, Ho Chi Minh City, Vietnam pp 416 - 419.

Fehr, E. 2009 On the Economics and Biology of trust. *Journal of the European Economic Association* April–May 2009 7(2–3):235–266

Fogel, J. & Nehmad, E. 2009 Internet social network communities. *Computers in Human Behavior*, vol 25, pp 153-160.

Garfinkel, H. 1967 *Studies in Ethnomethodology*. Prentice Hall, New Jersey.

Garfinkel, H. 2002 *Ethnomethodology's Program*. Rowman and Littlefield. Lanham.

Gibson, J. 1986 *The Ecological Approach to Visual Perception*. Lawrence Erlbaum. New Jersey

Harper, R. & Hughes, J. 1993 *'What a f*****g system!"* in Button, G. (ed) *Technology in Working Order*. Routledge. London pp 127 - 144.

Hughes, J., Shapiro, D., Sharrock, W. & Anderson, R. 1988 The automation of Air Traffic Control. *Final Report SERC/ESRC Grant No. GR/D/86157*. Economic and Social Sciences Research Council, Swindon.

Husserl, E. 1936 *The Cartesian Meditations*. Martinus Nijhoff. The Hague

Idhe, D.	2009	*Postphenomenology and Technoscience*. SUNY Press, New York
Lankton, N & McKnight, D.	2011	What does it mean to trust Facebook? *The DATA BASEfor Advances in Information Systems*, vol 42, number 2 pp 32 - 54
Largerspetz, O.	1996	*The Tacit Demand. Filosofiska Institutionen* Abo Akademi, Abo.
Mumford, E.	1996	*Effective Systems Design and Requirements Analysis*. MacMillan. London
Odlyzko, A.	2010	Providing security with insecure systems. *WiSec'10* March 22-24, Hoboken, New Jersey pp nn-nn
Pettit, P.	2008	Trust, reliance and the internet. *Information Technology and Moral Philosophy*. van den Hoven, J. & Veckert, J.(eds). Cambridge University press, London.
Pieters, W.	2010	Reve{a,i}ling the risks. *Techne* vol 14 no 3 pp 194 - 206
Reason, J.	1986	Recurrent errors in process environments. *Proceedings of the NATO Advanced Study Institute on Intelligent Decision Support in process environments*. Springer-Verlag New York pp 255 - 270
Riegelsberger, J, Sasse, A., McCarthy, D.	2003	Shiny happy people building trust? *CHI 2005 New Horizons* April 5–10, 2003, Ft. Lauderdale, Florida, USA volume 5, no 1 pp 121 - 128
Ryle, G	1949	*The Concept of Mind*. Hutchison, London.
Schutz, A.	1982	On multiple realities in *The Problem of Social Reality*, Collected Papers vol 1. Martinus Nijhoff. The Hague pp 340 - 345.
Schutz, A. & Luckmann,	1973	*The Structures of the Life World*. Heinemann. London.

T.

Siau, K & Shen, Z. 2003 Buiilding Customer Trust in Mobile commerce *COMMUNICATIONS OF THE ACM* April 2003/Vol. 46, No. 4 pp 91 - 94

Spafford, E. 2009 Cyber security. *ISIPS 2008, LNCS 5661, pp 20-33*.

Thomas, W. & Thomas, J. 1928 *The Child in America*. Knopf, New York

Turilli, M., Vaccaro, A., & Mariarosaria, T.

Zinn, J. (ed) 2010 The case of online trust. *Knowledge, Technology and Policy*, vol 23, pp 333-345.

2008 *Social Theories of Risk and Uncertainty*. Blackwell, London.

[1] Turilli et al (2010) provide a general introduction and summary of the issues. The general reactions range from those such as Cerf (2010) and Spafford (2009) who are very concerned to Odlyzko (2010) who takes a more low key view.

[2] Although we have deep respect for the studies of The Tavistock Institute (see Emory & Trist (1972) and Mumford (1996) for examples) we do not want to be heard as adopting their definition of this term. What we take it to denote will become clearer as we proceed.

[3] This much, then, we have in common with Largerspetz's (1996) original critique of the concept of trust as that is usually defined in Philosophy and the technological disciplines.

[4] As antediluvians, it is of some comfort to us to find the contemporary avant-garde re-inventing the sociological wheels of our youth. Beautement et al.'s compliance budget looks for all the world to be a reconstruction of the "effort bargain" first described by Baldamus more than 50 years ago (Baldamus 1957)

[5] None of which should be taken to imply that users are *always* at home with the system, *never* have moments of doubt about who they are dealing with etc etc. The point is that such 'breakdowns' are not the typical users typical experience. As we will see, it is on this typicality that normal appearances rest.

[6] This is a somewhat different approach to that of Pieters (2010) and Ess (2010 and this volume) in their phenomenological re-conceptualisations of risk and certainly to Don Idhe's 'postphenomenology' (Idhe 2009).

[7] It should be noted that this is a *methodological* (for want of a better word) point not a philosophical one.

[8] Let us be really clear what we are saying. We are *not* saying that the model which motivates conventional studies of trust should be expanded, extended or 'filled out' to include the social psychology we seek. We fully recognise that models are ways of reducing complexity with the reductive strategy being very much a matter of choice. We do not want to add anything to the bare psychology. We want to start from a less bare psychology, namely a social one.

[9] This is, then, a very similar conception to that of Watson (this volume).

[10] This could be seen, then, to be an attempt to provide the kind of psychology which Olli Lagerspetz (1996) asks for. Whether it is tainted by the 'romanticised Hobbesianism' he sees in other sociological accounts we will leave others to determine. Suffice it to say, we think not.

[11] Philip Pettit (2008) provides a philosophical interpretation of the implications of these changes. Ulrich Beck's (1992) well known volume develops a broad sociological position whilst a set of responses are contained in Zinn (2008). An attempt to define a possible 'middle range' approach is given by Cook & Gerbasi (2011)

[12] Again (we will stop this now), this is *not* to say that people always understand one another, always hold to the same shared expectations etc etc. All we are saying is that social life is grounded in a presumption that this is normally so; a presumption which holds until further notice. If it turns out the presumption needs to be revised for *any* occasion then, as normal social actors, we have ways of discovering what that requires. See Watson (this volume)

[13] Edmund Husserl's term for this world is the *lebenswelt*. As we will discuss later, Alfred Schutz's (Schutz & Luckmann 1973) interpretation of Husserl's analysis is one of the inspirations for our approach (Husserl 1960)

[14] The quotation marks here are important. The phenomenological method begins by withdrawing subscription to an ordering such as this. That we can point to hierarchies of experience reflects our capacity to organise and construct the facticities of the world, not any essential character the world must have.

[15] The important contrast here is with *the scientific or analytic attitude* which begins by accepting the possibility of doubting just these things in order to focus on how they are produced or what lies 'behind' appearances. Of course, in doing so, scientists bracket other things such as the dependability of their instruments, the stability of the properties of materials, and so on.

[16] In a previous discussion, we have indicated how the psychological notion of "affordance" (Gibson 1986) could be reformulated to provide the basis for this kind of description. (Anderson and

Sharrock 1993). Of course, those with malicious or mischievous intent trade on the recognisability of the specifics of the context.

[17] As Odom and Harper (this volume) demonstrate, because our knowledge of their success remains so vague, it has to be "somehow or other".

[18] See Button (1993) for an introductory overview collection

[19] The West Drayton Centre no longer exists. Its functions were transferred to a new facility at Swanwick near Southampton in 2007. For the original research report see Hughes et al (1998) and Harper & Hughes (1993).

[20] With the move to Swanwick, there was some change in nomenclature. TMAs became Terminal Control Areas. However the abbreviation TMA was retained.

[21] One critical change between ATC when we studied it and today has been the digitisation of paper strips. Nonetheless, that provision has been made for them to be used as a fall back in case of computer failure reinforces our view of the importance of the information they provide.

[22] We have discussed this in Anderson et al (1987)

[23] Terminal Manoeuvring Area i.e. the sector of air space that manages the approach to an airport.

[24] Pilots from particular countries and particular air lines are well known to be especially 'untrustworthy' in this respect.

[25] A confliction is the merging of the trajectories of two aircraft to violate the minimal requirement of two miles horizonal and 5000 feel vertical separation.