NET-PUBLIC SPHERE RESEARCH: BEYOND THE "FIRST PHASE"

LINCOLN DAHLBERG

Abstract

In recent years much has been said about the possibility of the Internet facilitating and extending the public sphere of informal rational-critical communication between private affairs and official decision making. However, the abundant speculation has not yet been matched by extensive empirical research. Ongoing theoretical debate about the validity and content of the public sphere conception makes empirical evaluation difficult, as does the Internet's constitution through a vast and dynamic array of human interactions, cultural contexts, and social institutions. Despite these difficulties, a number of pioneering studies have attempted to investigate the Net-public sphere relation. In this paper I offer a critical reflection upon some of these early studies, a reflection that explores three reoccurring methodological problems and their possible solutions, with the aim of providing strategies for developing more meaningful Net–public sphere research. These problems involve: developing a normative conception of the public sphere suitable for critical analysis; improving the transition between theory and empirical evaluation; and adequately explaining and extending findings. My suggested solutions include the specification of public sphere criteria from Habermas' theory of communicative rationality; the identification of a variety of case sensitive indicators that can facilitate the transition between criteria and practice; and the suggestion that explanation and extension of findings be undertaken by way of comparative case study research.

Lincoln Dahlberg is postdoctoral fellow at the School of Sociology, Social Policy, and Social Work, Massey University, e-mail: L.J.Dahlberg@xtra.co.nz.

Introduction

There has been much theoretical commentary and debate about the possibility of Internet communications facilitating public sphere(s) of rational-critical deliberation and public opinion formation (for some recent contributions see: Dahlgren 2001; Gandy 2002; Gimmler 2002; Papacharissi 2002; Poster 2001; Sparks 2001).[1] There has also been valuable institutional and political-economy analysis of systemic level opportunities and threats to the actualisation of the public sphere through the Internet. In particular, such research has highlighted incursions by economic and administrative power into the online world of communicative interaction, with resulting problems of discursive exclusions, social inequalities, and the colonisation of online space by instrumental-strategic rationality (for recent examples see, Bollier and Watts 2002; Dahlberg 2002; Fortier 2001; Gandy 2002; Levine 2002; McChesney 1999; Moore 1999; Schiller 1999). However, there have only been sketchy beginnings in the evaluation of the extent to which online communicative practices – whether interpersonal dialogue or the clash of discourses through media-publishing sites – actually constitute the public sphere. Much of the early work at this communicative level has been undertaken by graduate students (see, for example, the dissertations by Aikens 1997; Dahlberg 2000; Fang 1995; Graham 2002; Oblak 2001; Schneider 1997; Stromer-Galley, 2002; Thornton, 1996). This work can be seen as a first phase of research. Its importance lies not so much in providing definitive answers as in offering initial explorations of the Net-public sphere problematic, in terms of both highlighting some significant aspects of practice and developing methodological approaches to improve understanding. As these pioneering researchers recognise, their work is a beginning from which to develop (see Schneider 1997, 73; Wilhelm 2000, 41). In this paper I take up the challenge of building on this work, critically reflecting upon it with the aim of contributing to the methodological strategy of the next phase. Thus, this paper is very much situated within, and addresses the concerns of, a particular approach. It does not interrogate the political and epistemological assumptions behind the Net-public sphere question, but simply asks how we can refine and extend the methodology and concepts in order to improve on earlier research.[2]

In particular, I focus upon three of the important stages of Net-public sphere analysis. First, such research implies a critical analysis that requires a normative model of the public sphere for evaluation of the democratic value of practices. Second, a means of effectively undertaking this normative evaluation must be developed. Third, given the emancipatory goals of critical research, the results of evaluation need to be explained and ways to apply positive findings to other situations identified.

My examination of early Net–public sphere research highlights significant problems associated with each of these stages: inadequate specification of the public sphere conception; poor transition between public sphere definition and the process of evaluation; and underdeveloped explanation and extension of results. I consider each of these problems in turn, giving examples from early research. For each problem I explore and suggest solutions: the specification of public sphere criteria from Habermas' theory of communicative rationality; the identification of a variety of case sensitive indicators that can facilitate the transition between criteria and practice; and the suggestion that explanation and extension of findings be undertaken by way of comparative case study research.

Problem One: Delineating Public Sphere Criteria

Evaluation of the extent online practice constitutes democratic space requires a detailed elaboration of the public sphere conception. First phase research has put forward a diversity of public sphere definitions and criteria, drawing from a variety of democratic theories.[3] However, specification of these criteria are often inadequate. Some research draws criteria from dubious theory. For instance, several works draw criteria solely from Habermas' (1989, first published 1962) *Structural Transformation of The Public Sphere (STPS)*, even though Habermas (1992a, 1992b, 1996) himself has rejected the normative foundations he offered there. Other specifications suffer because they stem from rather sketchy and crude syntheses of overlapping elements of various public sphere theories (e.g., combining Habermas 1984, 1989; Barber 1984; Fishkin 1991; Guttmann and Thompson 1996), theories developed from different approaches and premises, and that do not completely gel without dismissing important elements. A number of studies have compounded these problems by drawing directly on the under-theorised conceptions offered by others. Finally, some strongly qualitative research simply does not provide specific normative criteria, offering instead general frameworks, lists of important "issues" for discourse analysis, and thick description.

To have confidence in the critical evaluation of online communicative practice we need a convincing normative conception of the public sphere. I suggest that a normatively and sociologically robust ideal can be drawn from Habermas' theory of communicative action, developing this ideal with particular consideration to his most recent developments of this theory (1996, 2001). There are a number of public sphere theories available, but Habermas provides the most systematically and rigorously developed basis for the public sphere. Moreover, many Net-public sphere researchers already draw inspiration from Habermas in their analysis, making it constructive to further utilise his conception here.

For Habermas, as with a number of other deliberative theorists, the public sphere is constituted by communicative rationality; action involving the public use of reason aimed at mutual understanding. Habermas' formal pragmatic analysis of communication reveals that every participant attempting to undertake communicative rationality must make reference to a number of idealising presuppositions. The elaboration of these idealisations provides a set of normative conditions of the public sphere. While public sphere commentators and researchers often invoke or refer in a general way to some requirements of communicative rationality, a detailed specification of the conditions is necessary for a comprehensive critical analysis of everyday practice.[4] Here I put forward a set of normative conditions as identified by my reading of Habermas' reconstruction of the presuppositions.[5] Simply stated, the criteria are as follows:

1. *Thematisation and reasoned critique of problematic validity claims.* Communicative rationality involves the reciprocal rational testing of validity claims in which the positions put forward and the subsequent questioning are backed by reasons.
2. *Reflexivity.* Participants critically examine their values, assumptions, and interests, as well as the larger social context.
3. *Ideal role taking.* Participants attempt to understand the argument from the other's perspective. This involves empathetic listening, which in turn means a

commitment to an ongoing dialogue with difference.

4. *Sincerity.* Deliberation is premised upon honesty or discursive openness in contrast to deception, including self-deception for which one must remain "critically alert." Further, rational judgment presupposes that participants make a sincere effort to make known all relevant information, including their intentions, interests, needs, and desires.

5. *Inclusion and discursive equality.* Debate is open to all those affected by the concerns under consideration, and each participant has an equal opportunity to introduce and question any assertion whatsoever and to express attitudes, desires, and needs.

6. *Autonomy from state and economic power.* Deliberation is driven by the concerns of publicly oriented citizens rather than by money or administrative power.

These criteria remain hypothetical; they invite critique and revision through further reflection upon, and empirical analysis of, the presuppositions of communicative action. This is an ongoing task and cannot be undertaken here. Instead, I will turn attention to how to utilise these criteria to effectively evaluate everyday practice. In other words, how are we to gauge the extent to which online communicative acts "fit" the criteria? This question constitutes problem two.

Problem Two: Deploying the Criteria in Evaluation

Early Net-public sphere research has been largely divided between quantitative and qualitative approaches. However, researchers utilising either approach have found significant problems in adequately deploying the public sphere normative conception for the critical evaluation of online practice.

Quantitative studies, mostly employing content analysis, have attempted to find reliable and valid measures for the public sphere conception. However, valid measures of the various public sphere criteria prove difficult to specify. Researchers can end up settling on different indictors for the same criteria. The affect upon findings of this use of different measurement instruments can be illustrated by comparing separate research projects that attempt to evaluate the same dimension of the public sphere. Schneider (1997) and Wilhelm (2000) both use content analysis to examine the deliberativeness of Usenet groups and define reciprocity or exchange of opinions as one of the central dimensions of the public sphere. However, they go on to obtain strikingly different results for the level of reciprocity. While Schneider (1997, 105-106) found high levels of reciprocity, Wilhelm concluded that "[f]ewer than one out of five messages represents a direct reply to a previous posting, which suggests the notion of an attenuated public sphere" (Wilhelm 2000, 98). Although we would expect variation in results from two studies of different online groups, the extent of disagreement here is highly likely a result of the difference in criteria used to measure reciprocity. Schneider (1997, 74, 91) considered a message to reciprocate another message if it appeared in the same thread as another post (within three or seven days)[6] or directly cited another message. Furthermore, he counted a message as reciprocated if other posts followed it in the same thread (within three or seven days) or if it was directly referred to by other posts. Wilhelm's measure of reciprocity was much stricter. He only counted messages as reciprocal where there was explicit reference to another message in the content of a post. Thus Schneider's more positive result is not surprising given his broader measure.

But the problem is not simply one of getting the operationalisation right by refining, agreeing upon, and expanding measures. The fundamental problem is that operationalisation requires researchers to focus upon those aspects of the public sphere for which narrowly defined and measurable indicators can be found, thus neglecting other aspects less amenable to quantification. The result is a serious loss of meaning. The best we can hope for is a limited and skewed (because some aspects are more measurable than others) understanding of the extent that the communicative practices under investigation fit the public sphere conception.

Take for example Schneider's content analysis of the Usenet group talk.abortion, which he uses to evaluate the Net–public sphere relation. From Habermas' *Structural Transformation of the Public Sphere* Schneider defines the public sphere via four dimensions – equality, quality, diversity, and reciprocity – and operationalises each with a specific measure. But the measures for the dimensions are so narrow that understanding is seriously jeopardised. This narrowness is clearly apparent in Schneider's measure of reciprocity discussed above, but is even more evident in his operationalisation of equality and quality. Schneider's operationalised measure of equality is simply "equal distribution of voices." This measures what is easiest to count, "the frequency of expression and average quantity of expression," but fails to adequately gauge the level of equality within the newsgroup concerned because frequency and quantity of expression are insufficient or even erroneous guides to equality. Other indicators of discursive equality need to be explored, such as the socio-cultural opportunities available for individuals and groups to take part in deliberations or the power relations involved that influence the agenda and style of communications. But, as mentioned above, the problem is not just one of improving the operationalisation by finding and including further indicators of equality, including for instance "opportunity" and "power." "Opportunity" and "power" remain largely resistant to being defined as indicators that can be measured. In order to avoid oversimplifying them, with associated loss of meaning, we need to leave them with certain roundness. Schneider's measure of quality also demonstrates the narrowing of meaning that takes place through operationalisation. Quality is "measured by the tendency of participants in the newsgroup to stay 'on-topic'" (1997, 2), which is "operationalised as the proportion of messages that are concerned with the abortion issue" (1997, 75). The quality of deliberation on talk.abortion was deemed to be lacking after the number of messages with "abortion content" was found to be disappointingly low compared to the number of messages without. In this measure, not only is the determination of what counts as "staying on-topic" a matter of debate, but also this indicator of deliberative quality is extremely narrow. Quality includes critical testing of claims, reasoning, reflexivity, sincerity, and respectful listening. But again these aspects of quality largely escape quantitative measurement and are ignored in the operationalisation.

These examples are of content analysis, but similar examples involving the operationalisation of the public use of reason can be found in survey research. For instance, some such research measures the quality of public reasoning simply by asking people about the quantity of time spent discussing certain topics and doing certain things (listening, talking) (e.g., Kim 1997; Muhlberger and Shane 2002).[7] Again, many aspects of the quality of deliberation that cannot be determined by counting are ignored, including the more subjective elements of communicative action (such as reflexivity and sincerity). Although invaluable for understanding those aspects of online communications that are readily quantifiable, measurement cannot by itself adequately account for the complex meanings involved in the communicative practices constituting the public sphere.

This critical reflection points to the need for qualitative research of the Net-public sphere relation in order to get at meanings not accessible through measurement. Determined to retain meaning by building, rather than predetermining, categories of understanding, first phase Net–public sphere qualitative research has avoided delineating indicators of the public sphere. At most, some initial definitions, general categories, or "sensitising concepts" are provided as a starting point. The results of such research often provide insightful and detailed descriptions of particular online practices. However, evaluations of the Net–public sphere relationship are vague or generalist at best. The critical aspect inherent in the public sphere conception is increasingly diluted as we move closer to naturalistic or thick descriptive approaches. Practices become accepted as they stand, with only general statements made as to the democratic ideals embodied in the public sphere conception. The lack of decisive judgement results from both avoiding the detailing of critical criteria (as noted in problem one above) and from not specifying how to go about gauging the extent that practices facilitate the public sphere. Thus evaluative weakness results from, in direct contrast to the quantitative research discussed, avoiding clear specification of indicators – those characteristics of practices that are significant in determining the existence and quality of online deliberation.

For example, in Tanner's (2001) "textual analysis" of an Internet forum's online news publication, "rational critical discourse or the public use of reason" is highlighted as one of the "four broad issues" of the [Habermasian] public sphere but is specified no further. While she notes that quality deliberation is important, what constitutes quality is not defined. The result is a lack of *evaluation*. Tanner's examination of the public use of reason simply wanders off into a *description* of what was being discussed in the forum. Similarly, Fang's (1995) "discourse analysis" of three e-mail discussion lists, which was meant to "provide an analytical reading of CMC and its potential as a public sphere" (1995, 147), simply describes or "illustrates" some of the content of the communications involved. Indeed, he makes some interesting observations about the types of exchanges on the lists in terms of Habermas' general notion of thematisation – the exchange and problematisation of various types of validity claims – but he does not provide a systematic evaluation of the public use of reason. While his study has a theoretically impressive front end, Fang backs off developing and deploying the normative criteria and associated indicators necessary for rigorous critique.

Hence, neither an approach that relies solely upon the operationalisation of discrete variables nor one that avoids all specification of critical criteria and indicators is adequate for enabling a comprehensive evaluation of the extent online communicative practices facilitate the public sphere. The complex practices that constitute the public sphere cannot be fully operationalised into discrete variables for measurement. At the same time, rejection of all criteria and indicators promotes research that is thick on description but thin on evaluation, resulting in vague generalities about the actual health of the public sphere.

Where do we go from here? I suggest dispensing with any methodological fundamentalism, and allowing the problem to drive the method rather than the other way around.[8] The problem here requires a critical evaluation of the Net-public sphere relation. We have a set of criteria already laid out specifying a set of ideal practices. Our task here is to gauge the extent to which these criteria are being met in online communications. Before we can do this, we need to think about what characteristics of practices will *indicate* the quantity and quality of online communicative action as determined by the criteria. Therefore we *do* need to consider indicators. But the drawing up of these indicators, and the way they are to be gauged, should not be determined by any particular approach (for instance, it is not a matter of choosing those things that can be readily measured or described). Indicators need to be chosen to maximise understanding, that is, to maximise the intersection between criteria and practice. To do this, we must stay open to the multiple and complex ways in which indication, and hence evaluation, can take place. The specific methods required will become clear as reflection on indicators proceeds.

Here I will provide a first sketch of possible indicators for each criterion. Likely indicators can be suggested by considering the public sphere criteria in the light of established understandings of online communications. My discussion will reflect on each criterion in turn, giving examples of possible indicators and corresponding methods for evaluation.

1. *Thematisation and reasoned critique of problematic validity claims.* For this criterion we are simply looking for the extent that positions are put forward and critically debated, with all positions being supported by reasons. This is reasonably straightforward to examine through content analysis, although some decision needs to be made as to what exactly counts as a reply (as already demonstrated in the Schneider-Wilhelm discussion above). Significantly, the quality of both claims and reasons put forward does not need to be determined. This is to be resolved in the process of rational-critical deliberation. We simply need to observe the existence of reciprocation and the extent to which claims and counter-claims are supported by reasons.

2. *Reflexivity.* This is relatively difficult to evaluate as it is largely an internalised process. Participants themselves are unlikely to, in any precise way, report levels of self-reflection and may not even know (or even want to admit) that their positions have undergone change. Moreover, written communications may only show traces of such a subjective process. But this does not mean we should simply ignore reflexivity as some deliberative researchers do. We can gather certain evidence and make interpretations even if we cannot clearly measure it. Participant observation and content or discourse analysis of online texts can identify some explicit instances of self-critique and change of position. However, we must largely rely upon participant interpretations – self-reporting through interview, survey, or focus group work – in order gain a reasonable appreciation of the extent interactions encourage self-critique and position alteration.

3. *Ideal role taking.* Here we must focus on the extent participants take into account and are sensitive to other participants and positions, not only those immediately present in the forum, but all affected by the problem considered. Online discussions contain some clear indicators of inconsiderate behaviours: lack of reciprocation, dogmatic ranting, abusive postings, and spam or prolific postings that dominate space and attention. We can also explore aspects of online communica-

tions that contribute positively to ideal role taking. The act of reciprocation, particularly if ongoing, indicates a basic respect for the other – the giving of one's time and effort to seriously engage in reasoned dialogue. We can detect higher levels of empathy and respectful listening where there are explicit attempts to understand the other, such as reflecting and summarising the other's position, requesting the clarification of statements, acknowledging and protecting the rights of all to be heard, and the putting forward of positions that are not one's own to enable broader reflection. Discourse and content analysis would be useful here, but self-reporting would also be helpful in strengthening understanding.[9] Observation and participant interviews are necessary to determine the time and effort people commit to dialogue.

4. *Sincerity.* We can clearly observe the provision of information through content analysis. But it is more difficult to judge the honesty (or the absence of deceit) involved. Chambers (1996, 208-209), drawing upon Gutmann and Thompson (1990), suggests that three "indicators that a speaker is acting sincerely are consistency in speech, consistency in speech and action, and coherence." Consistency in speech is when someone defends their position similarly in different situations or with different interlocutors. "Consistency in speech and action implies that speakers should act in ways consistent with their professed beliefs.... Coherence refers to a broader sense of consistency. We might question the sincerity of a speaker who, although consistent in speech and action, refuses to see the broader implications of her views for other issues or debates" (Chambers 1996, 208-209). To gauge these aspects of consistency requires intense observation of participants' ongoing interactions on and offline, which poses significant access and resource (particularly time) problems for the researcher. But as with evaluating reflexivity, we must not simply abandon attempts to understand sincerity due to the difficulty of the task – our research should not be driven by what is easiest to find. We need to look for clues to piece together understanding and develop the evaluation. We can detect aspects of sincerity through observation, such as instances of deceit or inconsistency.[10] However, such observation is not likely to be particularly comprehensive or rigorous. We need to turn to self-reporting to "thicken" understanding and sharpen evaluation, using further observation to help verify these reports.

5. *Inclusion and discursive equality.* Social resourcing problems (for example, limitations in Net access, time, or skills) can be direct indicators of both exclusions from deliberation and inequalities in communicative opportunities.[11] This requires social survey work, although data on such problems is increasingly available from public agencies and private research companies. Exclusion and inequalities are also indicated by formal restrictions; either state imposed legislation blocking sites and communications in certain jurisdictions or restrictions architectured into cyberspace such as membership passwords and site rules. But communicative exclusions and inequalities occur more subtly within communicative practices. For example, the domination of discursive space and attention by some participants, which may result in less opportunity for others to be heard. A seemingly obvious indicator for domination would be disparities in the amount (quantity and frequency) of postings in a particular forum. This is indeed tempting to use as an indicator as it would appear to clearly reveal discursive inequalities. In addition, it is straightforward to measure. However, as already seen in regards to Schneider's study, the amount of postings is a deceptive indicator of equality and must be used with care. Groups

and individuals may choose not to speak on an issue for any number of reasons, even when offered the opportunity to do so. Moreover, online communications are largely asynchronous, allowing participants to choose what to focus upon and what to ignore. Consequently the most prolific posters or positions do not necessarily command the most attention. Formal and informal (netiquette) rules (e.g., limits on time and frequency of speaking and restrictions on abusive postings) and the strict management of these rules, may be more reliable indicators of equality of opportunity to participate within online spaces. Yet we must also note that such rules do not necessarily signify equality of opportunity. Some groups or positions may dominate coercively even though they are not prolific or abusive. For instance, they may dominate the agenda and style of deliberation due to their status/authority, or because they follow dominant (e.g., masculinist tending) modes of online communication.[12] Hence, we need to address the complex and subtle relations of power involved. To do this, we can observe traces of such relations and subsequent inequalities by way of discourse analysis of online communications. However, possibly the best way to gain an in-depth understanding of the relations of power at work is simply to talk to participants about their experiences.

6. *Autonomy from state and economic power.* This criterion is easiest to evaluate negatively, as freedom from direct coercive controls (e.g., censorship and surveillance) and from strategic contents (e.g., advertising and publicity). It is quite straightforward to observe the levels of these aspects within online fora – by way of participant observations, discourse and content analysis, interviews with site organisers, and by inspecting formal rules and their management. But such "levels" only provide a reading of the explicit impact of money and administrative power. The influences of these instrumental media are subtle and their full impact upon deliberation cannot be determined by observation alone. Again, we need to turn to participant self-reporting to further gauge the extent of coercion and communicative distortions resulting from economic and administrative power. Participants may of course be unclear or even mistaken in their reports, but they do offer a first hand account of whose interests their arguments may serve or what coercions they experience.[13]

This outline gives a starting point for evaluation. It will be refined and developed as it is utilised in the critical analysis of online practices. Ongoing research will feed into – and consequently improve – the specification of indicators and corresponding methods. Thus, we cannot expect final answers. We must aim for persuasive evaluations in the light of current knowledge. The most convincing analysis of a particular situation in terms of the facilitation of the public sphere criteria will result from gathering a diversity of evidence through in-depth multimethod research.[14]

Problem Three: Explanation and Extension of Results

Given the expanse of the Internet and the limited resources available to researchers – coupled with the fact that online communication largely takes place within or across technologically defined "sites" – many researchers evaluating the public use of reason online have chosen to undertake case studies. Case study research provides useful understandings of the specific practices involved. It is particularly helpful for investigating the success of Internet-democracy initiatives. The previous discussion of public sphere criteria and indicators offers a general guide for undertaking such evaluations. However, critical research aims to go beyond case specific evaluation and to provide explanation and extension of findings: to identify the factors facilitating rational-critical communication and to determine which factors can be applied with success elsewhere. The question here is how we can undertake such explanation and extension.

So far extension of results beyond cases has been limited, if not flawed, and has largely focused on providing general evaluations rather than the much needed explanations. Some research has not attempted to generalise beyond the specific practices studied. Other research simply assumes that a case can be considered a microcosm (in some way typical or representative) of the larger population without this being clearly demonstrated. General conclusions about the Net-public sphere relationship end up being drawn from a small number (often just one) of case studies without adequate account of how such conclusions can be made.

Given his concerns about such generalisation, Wilhelm (2000) based his content analysis of online deliberation on a random sample of ten newsgroups from 57 Usenet and 14 America Online groups that he deemed the most "politically-oriented." He concluded from his case analysis that there is no evidence that online communications are cultivating the public sphere (2000, 102-104). But a small selection of cases from a select group (those deemed politically-oriented) of a particular form of online communication (Usenet and America Online) cannot be representative of the diverse and prolific practices of cyberspace. Wilhelm's sample is at best representative of the 71 groups deemed political, and not the wider newsgroup population, let alone the larger set of Internet communications. To be sure, Wilhelm's purposive sampling does lead to a useful study of one type of online communicative practice. Any attempt to randomly sample cases from cyberspace as a whole in order to apply statistical methods for the generalisation of results would be futile. Even if we could undertake a study of a large number of cases, random sampling requires us to assume a population distribution such that statistical generalisation would give meaningful results.[15] Such a distribution cannot be presupposed.

Statistical generalisation from randomly sampled cases can at best be used to provide some understanding of a particular type of online practice, that is, can be used to contribute to a larger case study. What about bypassing case study analysis and using other methods to undertake general evaluation and explanation of the Net-public sphere relation? Some researchers of the public use of reason online have employed survey and experimental methods. However, neither survey research nor experimental analysis *on their own* are adequate. While Net-public sphere survey research assists with understanding online political participation, it runs into the problems already discussed above in relation to the loss of meaning due to operationalisation. Experimental work offers a way to test particular aspects of deliberative theory, constructing controlled situations that isolate certain factors (variables) for analysis, but it is not adequate for evaluating and explaining real, complex practices or for determining how to extend the public sphere in such situations thick in context.[16] In addition, both survey and experimental research can be considered part of case study work. They necessarily focus on particular sets of online practices that are bound by geography, population, technology, site content, etc. In the end, we cannot escape case analysis.

How are we then to undertake explanation, and extend findings beyond cases? Explanation can take place to a certain extent within specific case analysis: causal factors affecting outcomes can be identified by observing patterns of variation between practices and outcomes. But this does not help us apply (or extend) successful factors (and findings) from a specific case to elsewhere – it does not tell us which factors successful in one case will be successful in a different situation. To determine this, I suggest we abandon representative sampling of cases and utilise instead a case comparative approach. Such an approach will not only help identify and explain factors impacting outcomes but will highlight those factors that can be applied to other situations with success.

Before exploring this approach I want to consider what type of causal factors we can expect to find. What sorts of things are likely to make up explanatory factors? To help clarify our thinking, we can differentiate between factors internal to deliberative forums, and more contextual factors. At the level of online communications, we can identify both formal or architectural, and informal or cyber-cultural factors. It is quite straightforward to identify and explore the formal factors that structure deliberations – the technological structures (software, forum configurations) and the formal management of online spaces (rules of participation, moderation systems). Informal factors are more difficult to identify prior to analysis. However, previous research points to various cyber-cultural practices that are likely to need attention, for instance, the domination of content and style of interactions by particular groups, or the development of netiquette. Architectural and cyber-cultural factors may overlap and mutually influence one another in complex ways. For instance, Internet architectures may encourage the development of new communicative forms, while online culture may generate new rules of interaction that in time become enshrined in software and management systems. This interlacing of factors is further complexified by the affects of the offline socio-cultural context.

Although underdeveloped in much Net-democracy research (apart from political economy analysis), the offline social context must be considered significant in influencing outcomes given that it frames the possibilities and limits of practice. To begin with, we need to consider the socio-cultural resources that enable or limit individual and group access and participation, including time, money, communicative competencies, computer access, and community support. Just as significant are the ideological frameworks that are brought to online practices by individuals and groups.[17] Online deliberations will be affected by the various political subjects that participate. This is demonstrated by the transformation of the early bourgeois public sphere as it extended to subjects less predisposed towards rational-critical deliberation (see, Habermas 1989). Behind these subjects are the agents of socialisation: the mass media, state institutions and agencies, civil society groups, and the formal political and legal system. Not only are these institutions involved in the constitution of political subjects but also in directly facilitating or impeding public deliberation.

Outcomes may be a result of a combination of factors at the two levels. For instance, respectful discussion may be a result of the architecture of a site as well as the background situation of those drawn to questions under consideration. Moreover, factors at the two levels may mutually influence one another. For example, the culture of a site is likely to both influence and be influenced by the identities of participants. To complicate things further, some explanatory factors may also operate as evaluative indicators. For instance, the domination of content and style may both indicate and cause exclusion, while rules of participation may both point to and promote inclusion.

The identification and analysis of the specific factors affecting the outcomes of a particular communicative practice can first be undertaken within a case study. The general idea is to examine variations in results as factors vary; for instance, observing the effects on deliberations of architecture changes. This examination can be carried out coterminous with the evaluation process described in the previous section of this paper and draw upon a similar array of methods.

Cases need to be selected on the basis of uniqueness rather than statistical representativeness. The aim is to choose cases that differ from others in such a way that significant lessons can be learned. Initially, we want to select exemplary cases, which means those that seem to be closer to achieving the public sphere attributes we desire. But how are we to decide which cases do this? To identify an exemplary case for analysis a background overview is needed of the public use of reason online. This overview will give some indication of the problems, limitations, possibilities, and successes encountered in practice, and of which characteristics of cases may be responsible for these, and thus what an exemplary case may look like. To provide this overview we have to rely on previous research, viewing it through the lens of the public sphere criteria so as to make sense of it for our purposes. Given that such an overview necessarily draws on a mixture of research that focuses upon different cyber-practices and that is undertaken with varying approaches, it cannot be used to provide a reliable evaluation or explanation of the general Net-public sphere situation. But it does help highlight what is going on and draws attention to the various factors at play, and as such provides a backdrop against which exemplary cases may be selected.

The background overview is the first step in explanatory comparison. The differences between overview and exemplary case, that assist with exemplary case selection, may also draw attention to factors crucial in influencing outcomes. For instance, we may be interested in studying a case partly due to noticing high levels of respectful interaction taking place in comparison to the overview, but this could also alert us to particular architectural or cyber-cultural or contextual factors facilitating civil communication. These factors may be explained to some extent within a case study, as noted above. However, to more fully specify causal factors, focused comparison with other cases needs to be undertaken. In fact, the significance of many factors is only likely to become apparent after comparison between different situations. The idea is to compare similarities and differences across cases to identify the significance of various factors for outcomes. Hence, to enable certain features to be focused upon, subsequent cases are to be selected not only in terms of exemplary features in relation to the general overview, but on the basis of similarities and difference with the initial case[s].

Once the various causal factors are identified and their significance determined, we can then reflect upon which type of situations we would expect to give which type of results – so instead of statistical generalisation of the situation, we can predict results by relating situations.[18] But more than this, we can start to think about which factors may be applied elsewhere with positive outcomes so that success can be attained in different situations. For instance, we may want to identify which of the factors found to be stimulating rational deliberation in Minnesota E-Democracy's forums can also be successfully applied to an online public space that draws on identities and issues in Wellington, New Zealand.[19] For such comparative analysis, both the "sending" and "receiving" cases (especially the social contexts) need to be known in detail. Even if a particular research project does not get to the point of comparison, it should at least offer enough detail of the distinctive features of the case in point to enable comparison at a latter date, or by other researchers.

Conclusion

My critical reflection upon pioneering research into the public use of reason online has highlighted three general methodological problems shared by many of the studies in the field. First, the public sphere as critical conception tends to be poorly theorised and specified. Second, the transition between public sphere definition and empirical evaluation is often poor. This is largely due to inadequate sensitivity to the development of evaluative indicators, which leads to sketchy results and significant loss of meaning. Third, explanation is poorly addressed, while extension of findings often takes place by erroneously assuming a statistical generalisability of results from a limited sample of cases. I have proposed solutions for each of these problems. First, for adequate critical analysis, it is essential that a robust normative conceptualisation be developed. I suggest a set of criteria drawn from Habermas' theory of communicative rationality. These criteria remain hypothetical and open to revision, but I believe they offer a strong starting point for critical analysis. Second, I argue that we need to focus on identifying how each criterion may be best evaluated by examining multiple possible indicators (and subsequently methods), with a sensitivity to maximising understanding. Third, in order to effectively evaluate and extend the public sphere through the Internet I argue that a case-comparative approach can be employed.

This "first draft" set of ideal solutions invites further development and refinement. The details of the process of evaluation and explanation need to be filled out. How exactly are we to carry out difficult tasks demanded by this model, like understanding reflexivity and sincerity, or identifying through comparative analysis the effects of contextual factors like political subjectivity? The answers to such questions can only be fully answered in the process of research. There are also a number of other questions regarding Net-public sphere research that I have not raised here but that need to be investigated. For instance, how are we to apply the normative criteria to the investigation of multi-media sites, and how might we explore public opinion formation online and its flow into and effects upon offline publics and policy-making. Decision also needs to be made as to how we are to interpret for practical purposes the findings of evaluations. For instance, can we say when a certain communicative practice has reached an acceptable "level" as against the public sphere idealisation or do we simply keep demanding more improvement? We also need to ask how the public sphere normative idealisation can be informed by findings. What is the feedback process between evaluative norm and practice? Furthermore, we have to accept that actual research will struggle to achieve the model solutions outlined in this paper due to resource limitations.

What this paper offers is a methodological reflection to assist the development of meaningful and comprehensive evaluation, explanation, and application. If nothing else, two general conclusions flow from the exploration here. First, individual studies need to be aware of and explicit about the limitations of their analysis and results. And second, comprehensive critical analysis of the Net-public sphere relationship demands international, trans-disciplinary, cooperative research, especially if comparison of different contexts is to be undertaken.

Notes:

1. I wish to acknowledge the New Zealand Foundation for Research, Science, and Technology for supporting this research.

2. Elsewhere I have addressed some of the important epistemological and political concerns raised by critics of this research, see Dahlberg (2000; 2001a).

3. I am not including in my discussion critical commentary and political economy analysis of the democratic possibilities of the Internet. Such work need only draw upon the general definition of the public sphere as a space for rational deliberation free of state and corporate power.

4. Net-public sphere researchers that have drawn upon Habermas' theory of communicative rationality in their evaluations include Fang (1995), Fung and Kedl (2000), and Wilhelm (2000). Graham (2002) is one of the few researchers that has attempted to comprehensively specify public sphere criteria from Habermas' and other theorists work on communicative rationality.

5. For the development of these criteria and discussion of critiques of the public sphere based on the theory of communicative action see my doctoral thesis (Dahlberg 2000). The central references I use in my formulation include Habermas (1984, 1-26; 1990, 43-115; 1993, 31-33, 56-76; 1996, 267-387; 2001).

6. It is unclear in Schneider's thesis whether a three or seven day period was used as the cut off period (see Schneider 1997, 74, 91).

7. Muhlberger and Shane (2002) carry out a mail survey gauging the quantity and quality of public deliberations in Pittsburgh community in general and not of Internet deliberations in particular. However, the survey is to be used as a base line to judge the success of a local online community project.

8. Graham's (2002) masters thesis, while exploratory, is one work that does undertake a reflection on the methodological requirements of the various aspects of the public sphere.

9. In terms of gauging levels of ideal role taking by way of self-reporting it may be more useful to ask participants if they feel they are treated with respect than to ask them if they treat others with respect.

10. In the idealised public sphere of rational-critical deliberation claims and reasons are challenged to the extent that deceit is revealed and consistency confirmed.

11. It is generally easiest to look for indicators of formal exclusion and inequalities in opportunity.

12. Refer Dahlberg (2001b) for discussion of how status and authoritative power differentials develop online and impact upon the equality of communications.

13. See Chambers (1996, 202-206) for further discussion on evaluating the coercive influence of money and administrative power on public deliberations.

14. It is necessary to emphasise that this process must not to be confused with "thick description" since we require a critical evaluation using normative criteria that gathers and judges evidence rather than a detailed description that builds theory.

15. Analysis of specific cases, no matter what the method used, necessarily involves abstraction from the inexhaustible supply of sensory detail available to the researcher, and hence always involves (often extensive) statistical and non-statistical sampling and generalisation. We may call this generalisation *within* cases. Why cannot we then undertake a similar type of generalisation *from* cases? The main reason is that a small sample is often all that can be taken in case research due to limitations of resources.

16. See Price and Capella (2002) for an example of experimental analysis of the public use of reason in a Web based environment.

17. By ideological frameworks I am not talking about the explicit political ideas adhered to, but the whole array of social meanings informing participants' actions.

18. As the number of cases analysed builds, adding to the rough understanding gleaned from the background overview of cyber-communications, a more general picture should develop of the Net–public sphere situation at large. This is not the same as the idea, discarded earlier, of statistical generalising of evaluations *from* cases. Findings of the comparative approach can only add up to give, at best, an approximate understanding of the situation at large, not simply because of the absence of statistical sampling but because there is always going to be variations in the research approaches undertaken in different case studies.

19. Minnesota E-Democracy can be found at http://www.e-democracy.org/ There is no similar project where I live in Wellington, New Zealand, despite the city being home to the parliament and the administrative power, and arguably being the hub of political activity in the country.

References:

Aikens, Scott G. 1997. American Democracy and Computer-Mediated Communication: A Case Study in Minnesota. Doctoral Dissertation. Cambridge: Cambridge University.

Barber, Benjamin. 1984. *Strong Democracy: Participatory Politics for a New Age.* Berkeley: University of California Press.

Bollier, David and Tim Watts. 2002. Saving the Information Commons: A Public Interest Agenda in Digital Media. New America Foundation and Public Knowledge (consulted June 2003). Available at: http://www.newamerica.net/Download_Docs/pdfs/Pub_ File_866_1.pdf.

Chambers, Simone. 1996. *Reasonable Democracy: Jürgen Habermas and the Politics of Discourse.* Ithaca: Cornell University Press.

Dahlberg, Lincoln. 2000. *The Internet and the Public Sphere: A Critical Analysis of the Possibility of Online Discourse Enhancing Deliberative Democracy*. Doctoral Dissertation. Palmerston North (NZ): Massey University.

Dahlberg, Lincoln. 2001a. The Habermasian Public Sphere Encounters Cyber-Reality. *Javnost—The Public* 8, 3, 83-96.

Dahlberg, Lincoln. 2001b. Computer-Mediated Communication and the Public Sphere: An Evaluation of the Prospects of Online Interaction Extending Rational-Critical Discourse. *Journal of Computer-Mediated Communication* 7, 1, http://www.ascusc.org/jcmc/vol7/issue1/dahlberg.html.

Dahlberg, Lincoln. 2002. Democratic Visions, Commercial Realities? The Corporate Domination of Cyberspace and the Prospects for Online Deliberation, *AntePodium,* April, http://www.vuw.ac.nz/atp/articles/Dahlberg_0204.html.

Dahlgren, Peter. 2001. The Public Sphere and the Net: Structure, Space, and Communication. In L. W. Bennett and R. M. Entman (eds.), *Mediated Politics: Communication in the Future of Democracy*, 33-55. Cambridge: Cambridge University Press.

Fang, Nien-Hsuan. 1995. *The Internet as a Public Sphere: A Habermasian Approach.* Doctoral Dissertation. Buffalo: State University of New York.

Fishkin, James S. 1991. *Democracy and Deliberation.* New Haven: Yale University Press.

Fortier, François. 2001. *Virtuality Check.* London: Verso.

Fung, Anthony Y., H. and Kent D. Kedl. 2000. Representative Publics, Political Discourses and the Internet: A Case Study of a Degenerated Public Sphere in a Chinese Online Community. *World Communication* 29, 4, 69-84.

Gandy, Oscar. 2002. The Real Digital Divide: Citizens versus Consumers. In L. A. Lievrouw and S. Livingstone (eds.), *Handbook of New Media: Social Shaping and Consequences of ICTs*, 448-460. London: Sage.

Gimmler, Antje. 2002. Deliberative Democracy, the Public Sphere and the Internet. *Philosophy & Social Criticism* 27, 4, 21-39.

Graham, Todd. 2002. Deliberating in Online Forums: New Hope for the Public Sphere? M. A.

Dissertation. Amsterdam School of Communication Research.

Gutmann, Amy and Dennis Thompson. 1990. Moral Conflict and Political Consensus. *Ethics* 101, October, 64-88.

Gutmann, Amy and Dennis Thompson. 1996. *Democracy and Disagreement.* Cambridge. Mass: Belknap Press.

Habermas, Jürgen. 1984. *The Theory of Communicative Action Vol. One: Reason and the Rationalization of Society.* Boston: Beacon Press.

Habermas, Jürgen. 1989. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society.* Cambridge, Mass: MIT Press.

Habermas, Jürgen. 1990. *Moral Consciousness and Communicative Action.* Cambridge, Mass: MIT Press.

Habermas, Jürgen. 1992a. Further Reflections on the Public Sphere. In C. Calhoun (ed.), *Habermas and the Public Sphere*, 421-461. Cambridge, Mass: MIT Press.

Habermas, Jürgen. 1992b. Concluding Remarks. In C. Calhoun (ed.), *Habermas and the Public Sphere*, 462-479. Cambridge, Mass: MIT Press.

Habermas, Jürgen. 1993. *Justification and Application.* Cambridge: Polity Press.

Habermas, Jürgen. 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy.* Cambridge: Polity Press.

Habermas, Jürgen. 2001. From Kant's "Ideas" of Pure Reason to the "Idealizing" Pressuppositions of Communicative Action: Reflections on the Detranscendentalized "Use of Reason." In W. Rheg and J. Bohman (eds.), *Pluralism and the Pragmatic Turn: The Transformation of Critical Theory*, 11-39. Minnesota: MIT Press.

Kim, Joohoan. 1997. *On the Interactions of News Media, Interpersonal Communication, Opinion Formation, and Participation: Deliberative Democracy and the Public Sphere.* Doctoral Dissertation. University of Pennsylvania.

Levine, Peter. 2002. Building the Electronic Commons: A Project of the Democracy Collaborative (consulted June 2003). Available at: http://www.democracycollaborative.org/ programs/ public/BuildingElectronicCommons.pdf.

McChesney, Robert W. 1999. *Rich Media, Poor Democracy: Communication Politics in Dubious Times.* Urbana: University of Illinois Press.

Moore, Richard K. 1999. Democracy and Cyberspace. In B. N. Hague and B. D. Loader (eds.), *Digital Democracy: Discourse and Decision Making in the Information Age,* 38-59. London: Routledge.

Muhlberger, Peter and Peter Shane. 2002. Prospects fro Electronic Democracy: A Survey Analysis, Version 1. Pittsburgh: Community Connections-InSITeS (consulted June 2003). Available at: http://communityconnections.heinz.cmu.edu/papers/MarkleFinal.doc.

Oblak, Tanja. 2001. Images of Electronic Democracy: Communications Technologies and Changes in Communication and Participation Processes. Doctoral Dissertation in Communication: University of Ljubljana.

Papacharissi, Zizi. 2002. The Virtual Sphere: The Internet as a Public Sphere. *New Media & Society* 4, 1, 9-27.

Poster, Mark. 2001. *What's the Matter with the Internet?* Minneapolis: University of Minnesota Press.

Price, Vincent and Joseph N. Cappella. 2002. Online Deliberation and its Influence: The Electronic Dialogue Project in Campign 2000. *IT & Society* 1, 1, 303-329.

Schiller, Dan. 1999. *Networking the Global Market System.* Cambridge, Mass: MIT Press.

Schneider, Steven M. 1997. *Expanding the Public Sphere through Computer-Mediated Communication: Political Discussion about Abortion in a Usenet Newsgroup.* Doctoral Dissertation in Political Science, Massachusetts: MIT.

Sparks, Colin. 2001. The Internet and the Global Public Sphere. In L. W. Bennett and R. M. Entman (eds.), *Mediated Politics: Communication in the Future of Democracy,* 75-95. Cambridge, Cambridge University Press.

Stromer-Galley, Jennifer. 2002. *New Voices in the Public Sphere: Political Conversations in the Internet Age.* Doctoral Dissertation in Communication. University of Pennsylvania.

Tanner, Eliza. 2001. Chilean Conversations: Internet Forum Participants Debate Augusto Pinochet's Detention. *Journal of Communication* 51, 2, 383-403.

Thornton, Alinta. 1996. Will Internet Revitalise Democracy in the Public Sphere. Masters Dissertation in Communications. Sydney: University of Technology.

Wilhelm, Anthony G. 2000. *Democracy in the Digital Age: Challenges to Political Life in Cyberspace.* London: Routledge.