The language of information need: Differentiating conscious and formalized information needs

Ian Ruthven

Department of Computer and Information Sciences, University of Strathclyde, Glasgow, United Kingdom

A B S T R A C T

Information need is a fundamental concept within Information Science. Robert Taylor's seminal contribution in 1968 was to propose a division of information needs into four levels: the visceral, conscious, formalized and compromised levels of information need. Taylor's contribution has provided much inspiration to Information Science research but this has largely remained at the discursive and conceptual level. In this paper, we present a novel empirical investigation of Taylor's information need classification. We analyse the linguistic differences between conscious and formalized needs using several hundred postings to four major Internet discussion groups. We show that descriptions of conscious needs are more emotional in tone, involve more sensory perception and contain different temporal dimensions than descriptions of formalized needs. We show that it is possible to differentiate levels of information need based on linguistic patterns and that the language used to express information needs can reflect an individual's understanding of their information problem. This has implications for the theory of information needs and practical implications for supporting moderators of online news groups in responding to information needs and for developing automated support for classifying information needs.

1. Introduction

The concept of information need has been fundamental to many models and studies within Information Science and has been the motivation for much research on information seeking, information use and interactive systems design. In his seminal work on information needs, Taylor proposed that information needs exist across four levels: the visceral, conscious, formalized, and compromised levels (Taylor, 1968). The visceral need is the unexpressed need that may only reflect a '*vague feeling of dissatisfaction*' one that is '*probably inexpressible in linguistic terms*' and, as Cole later put it, '*unspecifiable even to the user herself*' (Cole, 2015) The conscious need is '*a conscious mental description of an ill-defined area of indecision*' that results from the conscious recognition of a problem that requires attention. The formalized need is '*a qualified and rational statement of his question*', i.e. a clear expression of an information need. Finally, the compromised need is the question that is posed to the (human or technical) information system.

Although Taylor seemed concerned that what he was stating what was obvious, at least to library professionals, his conceptualisation of information need levels has been immensely influential and his paper has become one of the most cited works in Information Science (Tyckoson, 2015). It has influenced research within Information Science, including work on system design, e.g. (Hoenkamp, 2015), and information behaviour studies, e.g. (Pálsdóttir, 2017), and outside of Information Science, e.g. the work of Jansen who used Taylor's four levels to model the process of selling online (Jansen & Hoppenbrouwers, 2017).

Chang's recent citation analysis of Taylor's paper demonstrated that the four levels of information need has been the most cited contribution from this work (Chang, 2013). This includes a recent and very considered article by Cole, who proposed a re-

interpretation of Taylor's four levels, not as stages in information need development, but as different levels of understanding of the same need (Cole, 2011). The idea that different levels of need have different characteristics has also been supported by Lundin (Lundin, 2014) who sees the compromised need level as one way of framing desired goals whilst Agarwal (Agarwal, 2015) sees the visceral level being the area where serendipity is most likely to occur.

In spite of decades of work on information need, it is still a concept that is poorly understood (Ormandy, 2011; Savolainen, 2017; Timmins, 2006) with Savolainen noting in a recent review '*even though information need is probably the most widely used construct explaining why people engage in information seeking, this concept is still vague*' (Savolainen, 2017). Lundh posits that one reason for this lack of specificity around what we mean by information need is that information needs, as internal cognitive states, are difficult to investigate directly and it is easier to investigate the information behaviours and activities arising from information needs instead (Lundh, 2010).

Taylor's division of information needs is described typically in conceptual or anecdotal terms with the division taken as a useful axiom about information needs but rarely analysed itself. In preparing for the work in this paper, we analysed the 300 odd hundred papers that cited Taylor's work in the five preceding years, (January 2013 to April 2018) and none provided, or cited, any empirical way to differentiate between the levels. That is, we lack empirical investigations on the difference between levels of need that go beyond Taylor's original conceptual work and that can be used in practical settings. In this paper, we conduct a linguistic analysis of written information need statements that demonstrates we can differentiate between different levels of information needs based on the language used to describe the needs.

Our proposal is that people reveal their inner states when posting needs to online forums and by analysing these posts we can identify what level of information need they are experiencing. As well as providing a new understanding of how information needs differ, and therefore how the response to such needs may differ, this also opens up the opportunity to detect automatically what level of need is expressed in requests for help made online and therefore how moderators of discussion groups should respond to posts.

We consider first the related literature to motivate four hypotheses on differences between information need statements that correspond to Taylor's conscious and formalized information needs. Following this, we present a series of empirical investigations followed by a discussion of our findings and their implications for future research.

2. Literature review

In this section, we consider four areas of relevant literature on information needs. The literature on information needs is vast and so here we restrict our analyses to contributions that were useful to develop our hypotheses. We particularly focus on the differences between Taylor's conscious and formalized information needs as the two need types that are most amenable to linguistic analyses (see Section 3.1 for more on this point).

2.1. Information needs and problematic situations

Even though information needs are a core concept in information seeking, it is not a given that information needs are what we should be investigating when we study information seeking. Other authors have proposed that the situations that require information are a better focus of study. Taylor himself noted that '*inquirers frequently cannot define what they want, but they can discuss why they need it*', i.e. people often cannot say what is their information need but can talk about the situations that have given rise to information needs (Taylor, 1968).

Others have also commented on the issue that people do not think in terms of formal statements of need but rather on problematic situations that have to be turned into expressions of information need in order to obtain information. For example, Belkin et al. in their famous contribution on Anomalous States of Knowledge (ASK) stated '*information need is in fact not a need in itself, but rather a means toward satisfying some more basic need, typically, in the situations with which information science is concerned, the resolution of a problem*' (Belkin, Oddy, & Brooks, 1982). They also proposed that this situation-based understanding of need should affect how we design information systems '*for our representation of ASKs are narrative statements by the users of the IR system, of the problems which brought them to the system*' and therefore are better ways for users to present needs to systems than queries.

These authors placed an emphasis on the *situation* where information might be useful rather than on the information *need* itself. This situational view is still a popular approach to understanding information behaviour, see for example the discussions in (Dervin, 1998; Ormandy, 2011; Savolainen, 2012).

People in problematic situations have the challenge of either turning their situation into a need statement (translating their conscious need into a formalized one) or presenting their situation to someone who can help this transformation. Problematic situations are closest to Taylor's conscious level of need, which he felt would need dialogue with someone else to clarify. In his '*ambiguous and rambling*' depiction of conscious information needs, Taylor noted that conscious needs might lack focus due to their emerging nature. Similarly, Belkin et al.'s description of '*narrative statements…of the problems that brought them to the system*' suggests that early stage information needs are more likely to be descriptive accounts of a problem situation rather than a precise expression of a need. Formalized needs though, from Taylor, are '*qualified and rational statement of his question*' – the use of the word 'question' suggesting that the situation has resolved into a need statement.

This division between narrative descriptions of problematic situations (conscious need level) and focused descriptions of need statements (formalized need level), leads us to hypothesize that textual descriptions of conscious information needs will be longer than those of formalized needs due both to the need to describe a situation and the lack of an ability to express a precise (formalized) need.