Evaluating the impact of EHR on clinical reasoning performance: A TTF perspective

Matthew J. Wills  
*Dakota State University*

Omar F. El-Gayar  
*Dakota State University*

Amit Deokar  
*Dakota State University*

Follow this and additional works at: [https://scholar.dsu.edu/bispapers](https://scholar.dsu.edu/bispapers)

**Recommended Citation**
Evaluating the impact of EHR on clinical reasoning performance: A TTF perspective

Matthew Wills
Informatics, Indiana University East, Richmond, IN, United States., willsm@iue.edu

Omar El-Gayar
Information Systems, Dakota State University, Madison, SD, United States., omar.el-gayar@dsu.edu

Amit Deokar
College of Business and Information Systems, Dakota State University, Madison, SD, United States., Amit.Deokar@dsu.edu

Follow this and additional works at: http://aisel.aisnet.org/amcis2012

Recommended Citation
Wills, Matthew; El-Gayar, Omar; and Deokar, Amit, "Evaluating the impact of EHR on clinical reasoning performance: A TTF perspective" (2012). AMCIS 2012 Proceedings. 27.
http://aisel.aisnet.org/amcis2012/proceedings/ISHealthcare/27

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2012 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Evaluating the Impact of EHR on Clinical Reasoning Performance: A Task-Technology Fit Perspective

Matthew J. Willis  
Indiana University East  
willsm@iue.edu

Omar F. El-Gayar  
Dakota State University  
omar.el-gayar@dsu.edu

Amit V. Deokar  
Dakota State University  
amit.deokar@dsu.edu

ABSTRACT

The proliferation of Electronic Health Records (EHR) as a backbone for healthcare delivery necessitates adequate tools for evaluating and improving the efficacy of such investments particularly as it relates to clinical reasoning performance. However, a thorough review of the literature indicates lack of a tested, validated instrument for evaluating and predicting the impact of EHR use on clinical reasoning performance. Despite successful application to a variety of other industries, TTF has not been adequately adapted to healthcare, EHR technology or the clinical reasoning task. Accordingly, the objectives of this research are to: 1) produce a valid instrument with diagnostic and predictive capabilities for evaluation of clinical reasoning performance with electronic health records, and 2) extend and validate the TTF model to the clinical domain with an emphasis on specification of the clinical reasoning task and EHR technology characteristics.

Keywords

Electronic Health Records, Task-Technology Fit, Clinical reasoning, Performance

INTRODUCTION

In the U.S., electronic health records (EHR) have emerged as the foundation of health information technology. Although fewer than 20 percent of physician practices have adopted the technology (DesRoches, 2008), recent directives and incentives from the U.S. federal government call for significant expansion of EHR adoption. Historically the field of medicine has been slow to embrace information technology, relying more on complex diagnostic tools and devices, as well as advances in treatment and surgical techniques (e.g. surgical robotics). So many interesting questions arise as human beings, smart medical devices and information systems become intertwined in support of effective, efficient and safe patient care.

For more than three decades information systems research has explored how and why people accept and use technology. Information systems researchers have also considered how technology impacts individual (Goodhue and Thompson, 1995) and group (Zigurs and Buckland, 1998) decision-making performance. Practitioners who implement and manage EHR’s would benefit from a method of identifying factors that either inhibit, or enhance user performance.

However, a thorough review of the literature indicates lack of a tested, validated instrument for evaluating and predicting the impact of EHR use on clinical reasoning performance. While a variety of instruments exist for evaluating a number of important research questions pertaining to EHR, it does not appear as those instruments deal specifically with the key issue of clinical reasoning performance. The study discussed in this article aims to address this research gap.

This research uses task-technology fit (TTF) theory as the foundation for development of an evaluation instrument. The premise of TTF is that individual performance will be enhanced when the functionality of the technology meets the user’s needs, i.e., fits the task at hand. The original TTF instrument was developed for the evaluation of multiple information systems and focused on managerial decision-making in the transportation and insurance industries (Goodhue, 1995b). Goodhue (1995b) developed the TTF construct and demonstrated that the characteristics of both the task and the technology impact “fit”, which in turn impacts perceived performance. Despite successful application to a variety of other industries, TTF has not been adequately adapted to the healthcare domain.
Accordingly, the objectives of this research study are to: 1) produce a valid instrument with diagnostic and predictive capabilities for evaluation of clinical reasoning performance with electronic health records, 2) extend and validate the TTF model to the clinical domain with an emphasis on specification of the clinical reasoning task and EHR technology characteristics, and 3) assess the on-going viability of TTF theory to explain the factors influencing individual performance.

The remainder of the paper is organized as follows: The next section provides a brief background and review of pertinent literature followed by a description of the theoretical model. The Research Method section describes the underlying research methodology followed by a presentation and discussion of results. The last section concludes the paper with a highlight of theoretical and practical implications as well as limitations and directions for future research.

BACKGROUND

Electronic Health Records and Clinical Reasoning

The electronic health record is an aggregate electronic record of health-related information on an individual that is created and gathered cumulatively across more than one health care organization. It is managed and consulted by licensed clinicians and staff involved in the individual’s health and care. The EHR is not one specific technology; rather it is often understood as a composite of technologies including computerized provider order entry, clinical decision support plus administrative, laboratory and imaging systems.

Although research on clinical reasoning has a tradition spanning decades, there exists no unified theory or explanation for how clinicians reason (Croskerry, 2005). Despite the theoretical variation of existing decision models, common themes and strategies have emerged from the cognitive literature. For example, present-day models generally agree that clinical reasoning can be understood as being either informal/intuitive or formal/analytical in nature, or some combination of both (Croskerry, 2002; Edwards et al. 2004; Elstein, 2002; Elstein, 1978; Norman, 2005).

Health IT Evaluation

The use of health information technology offers a number of opportunities to improve health care. From reduction of clinical errors to improving efficiency and quality of care, there is mounting evidence that information technology plays a critical role in the future of health care (Cantrill, 2010). At the same time, health information technology is expensive, and the failure of such systems could have negative effects on patients, staff and organizations. Given what is at stake, evaluation of health information technology is a valuable and necessary activity.

There are a number of challenges to evaluating health information technology. Chief among them is the complexity of the information technology itself, the complexity of the evaluation project, and the motivation for the evaluation (Ammenwerth 2003). Information systems are defined not only by their hardware and software components, but also by the social and behavioral processes of system use. This socio-technical complexity makes evaluation of information technology difficult on a number of different levels.

The specific challenges addressed in this research include evaluating individual performance with an EHR from a TTF perspective, as well as developing the task and technology characteristics constructs. TTF is itself a multi-dimensional construct whose components have been designed from the perspective of managerial decision-making. It will be critical to explore the relevance of Goodhue’s original eight TTF constructs to the clinical decision-making domain, from both theoretical and qualitative perspectives.

Information Systems Utilization and Performance Research

With respect to the behavioral determinants of use, the Technology Acceptance Model (TAM) represents the first theory specifically established for the information systems (IS) context (Davis, 1989). Other variations followed, including Combined Technology Acceptance Model –Theory of Planned Behavior (TAM-TPB) (Taylor, 1995), Technology Acceptance Model 2 (TAM2) (Venkatesh, 2000), the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, 2003) and Technology Acceptance Model 3 (TAM3) (Venkatesh, 2008).

Contrasted with models that predict acceptance and use, TTF attempts to explain user performance with information systems. The theory measures task-technology fit along multiple dimensions. Goodhue also demonstrated the validity of an instrument for information systems user evaluation based on TTF (Goodhue, 1998). TTF has been examined in group performance situations (Shirani, 1999; Zigurs, 1998), as intended with the focus on managerial decision-making (Ferratt, 1998), and has been further examined with an emphasis on ease-of-use (Mathieson 1998). TTF has also been extended with the technology acceptance model (Dishaw, 1999; Klopping, 2004; Pagani, 2006).
The application of TTF in the healthcare domain has been limited to date. With the exception of Kilmon et al. (Kilmon, 2008) and Wills (Wills, 2009; 2012) and El-Gayar (El-Gayar, 2009, 2010), there are few studies employing TTF in user evaluation of EHR systems. Kilmon et al. (Kilmon, 2008) utilize the TTF instrument presented in Goodhue (1995b) as a diagnostic tool to evaluate a first-phase implementation of an EHR at a university hospital, while Wills and El-Gayar (2010) used a reduced TTF model to evaluate performance.

Whereas past efforts (Kilmon, 2008; Wills, 2009; 2012; El-Gayar, 2009, 2010) in the health care domain have focused on testing reduced TTF models, this study follows Goodhue (1995b) and is adapted to the clinical environment with respect to the task and technology characteristics and performance constructs. The TTF dimensions employed in this study include data quality (QUAL), data locatability (LOCAT), correct level of authorization (AUTH), data compatibility (COMP), IS production timeliness (PROD), systems reliability (RELY), ease-of-use and training (EOU), and IS department’s relationship to users (RELAT). Task characteristics include the dimensions of task complexity (TSKCMP), task uncertainty (TSKUNC), and task significance (TSKSIG). Technology characteristics are defined by the dimensions of information (INFO), knowledge (KNOW) and inferencing support (INFER). The task-technology fit characteristics are posited to directly influence performance (PERF). Figure 1 illustrates the theoretical model.

Research Hypotheses

According to the TTF model, the strength of the link between information systems and performance is a function of the extent to which system functionality corresponds to task needs. This research hypothesizes the following:

H1: User evaluation of task-technology fit will have explanatory power in predicting perceived performance impact. This can be further divided among the 8 TTF dimensions as follows:

H1a: Data quality will positively influence user performance.
H1b: The locatability of the data will positively influence user performance.
H1c: Data authorization will positively influence user performance.
H1d: The compatibility of data from other systems will positively influence user performance.
H1e: Ease of use and training will positively influence user performance.
H1f: Production timeliness will positively influence user performance.
H1g: Systems reliability will positively influence user performance.
H1h: The IS departments’ relationship with users will positively influence user performance.

H2: The characteristics of the clinical reasoning task, complexity, uncertainty and significance, will have a significant positive influence on the TTF construct.

H2a: Task complexity will positively influence the fit between EHR technology and the clinical reasoning task.
H2b: Task uncertainty will positively influence the fit between EHR technology and the clinical reasoning task.
**H2c:** Task significance will positively influence the fit between EHR technology and the clinical reasoning task.

**H3:** The characteristics of EHR technology will have a significant positive influence on the TTF construct.

- **H3a:** Information capability will positively influence TTF.
- **H3b:** Knowledge capability will positively influence TTF.
- **H3c:** Inferencing capability will positively influence TTF.

**H4:** Task-technology fit will positively influence utilization.

**H5:** Utilization (Use) will positively influence performance.

### RESEARCH METHODS

#### Setting, Subjects, and Data Collection

The setting for this study was the main campus of a regional medical center in the U.S. Midwest, and also included affiliated primary care clinics in the region. Subjects included physicians (M.D., D.O.), certified nurse practitioners (CNP) and physician assistants (PA-C) who currently use an EHR. Data was collected in a variety of ways including semi-structured qualitative interviews, instrument pre-testing and the survey.

#### Survey Instrument

The survey instrument included demographic (age range, gender, EHR experience) questions, as well as 67 questions relating to the constructs evaluated in the study. Questions take the form of declarative statements and were measured on a 7-point Likert scale. The survey instrument was based on Goodhue (1995b) and adapted to the clinical reasoning and EHR domains.

#### Data analysis

**Assessing instrument validity**

We initially used qualitative data analysis techniques to refine the instrument. After the quantitative survey, exploratory factor analysis was also performed to validate the instrument. Finally, during the subsequent analysis using partial least squares, construct validity was further ensured. Following Goodhue (1998), this research uses the framework for measurement validity of new instruments as suggested by Bagozzi (1979; 1980). The framework identifies six components of construct validity that address each stage of the proposed research, including initial definition and operationalization of the theoretical constructs, instrument development and testing.

**Theoretical Meaningfulness:** The theoretical foundation upon which this research rests has been derived from previous research on task-technology fit and clinical reasoning, resulting in constructs that are consistent with prior theories. A literature review was conducted and TTF theory was selected as the underlying model for the study. TTF constructs were adapted from existing constructs for TTF based on Goodhue (1995). Other constructs such as task and technology characteristics were similarly developed, using the TTF and clinical reasoning literature as a basis for construct explication and operationalization.

**Observational Meaningfulness:** Observational meaningfulness, the second of Bagozzi’s validity concerns, refers to the extent to which a measure represents all facets of a given construct. In developing the questionnaire, the challenge was to devise questions that cover the target domain and are closely linked to the defined theoretical constructs. For each of the constructs (TTF, clinical decision task characteristics, EHR characteristics, utilization and performance), two to three questions have been developed for each indicator.

**Internal Consistency Reliability:** Internal consistency reliability is a measure of how well a scale addresses different constructs and delivers reliable scores. The survey instrument for this research project was designed with at least two parallel questions for each construct. Internal validity will be measured using Cronbach’s alpha.

**Discriminant Validity:** Discriminant validity addresses the possibility that supposedly distinct constructs may be indistinguishable from each other. Discriminant validity is evaluated through partial least squares by comparing item loadings to variable correlations and by examining the square root of the AVE of each variable to the correlations of this construct to all other variables (Gefen and Straub, 2005).
Convergent Validity: Convergent validity is used to evaluate the degree to which two or more measures that theoretically should be related to each other are, in fact, observed to be related. Convergent validity of the variables is assessed by examining the t-values of the outer model loadings (Gefen, 2005).

Nomological Validity: This form of construct validity is the degree to which a construct behaves as expected in a system of related constructs (nomological net). This version of validity is of particular importance in this research project, particularly relative to proposed task and technology characteristics constructs, as these have been adapted to the clinical decision task and EHR technology and are therefore slightly different from previous versions of these measures. Overall, nomological validity will be confirmed by the extent to which theoretically grounded predictions are realized.

Structural model testing

To assess the structural model, partial least squares (PLS) was employed. For this study, PLS was chosen for two reasons: 1. As an SEM technique, PLS is designed to explain the significance of the relationships in the model, as is the case in linear regression, and for this reason PLS is better suited to predictive models than covariance-based SEM approaches which focus on overall model fit, and 2. in contrast to covariance-based SEM, the estimation of significance in PLS does not require parametric assumptions, thus allowing analysis of comparatively small data sets, such as in this study (Gefen, 2000).

RESULTS AND DISCUSSION

161 of the 258 eligible participants successfully completed the survey resulting in a 62.4% response rate. During data exploration, 17 observations were omitted resulting in 144 observations or a 55.8% completion rate. Subjects were asked to respond to questions using a 7-point Likert scale, which ranged from 1=strongly disagree to 7=strongly agree. Subjects included 106 physicians, 18 advance practice nurses and 20 physician assistants who currently use an EHR system in clinical practice. Forty-four subjects were between the ages of 55-64, fifty-one subjects were aged 45-54, thirty-eight subjects were aged 35-44 and eleven noted their age in the 25-34 year range. Of the 106 physicians, 69% (73) were male and 31% (33) were female. For nurse practitioners, 95% (17) were female and 5% (1) were male. Of physician assistants 75% (15) were male, 25% (5) female.

Measurement Validity

In this section, the results are presented in the context of the six components of construct validity (Bagozzi, 1979; 1980; Straub, 1989).

Theoretical and observational meaningfulness

The first two criteria of validity, theoretical and observational meaningfulness, involve semantic issues as opposed to statistical tests. They refer to the internal consistency of the language used to represent a construct and the conceptual relationship between a theoretical construct and its operationalization (Papoutsakis, 2008). Input from interviews with clinicians was used to revise the language of some of the items and ultimately confirmed the theoretical and observational validity of the individual items developed through a comprehensive review of pertinent literature.

Internal consistency

Reliability is an empirical validity designed to assess the degree of internal consistency. The requirements include more than one observational indicators or variables for each theoretical construct. For attitudinal measurements, Cronbach’s alphas above 0.6 are generally considered acceptable. Results are summarized in Table 1, where all alphas are above the 0.6 cutoff.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Quality (QUAL)</td>
<td>0.89</td>
</tr>
<tr>
<td>Data Locatability (LOCAT)</td>
<td>0.77</td>
</tr>
<tr>
<td>Data Authorization (AUTH)</td>
<td>0.77</td>
</tr>
<tr>
<td>Data Compatibility (COMP)</td>
<td>0.92</td>
</tr>
<tr>
<td>Production Timeliness (PROD)</td>
<td>0.92</td>
</tr>
<tr>
<td>Systems Reliability (RELY)</td>
<td>0.75</td>
</tr>
<tr>
<td>EOU/Training (EOU)</td>
<td>0.61</td>
</tr>
<tr>
<td>Relationship with Users (RELAT)</td>
<td>0.94</td>
</tr>
<tr>
<td>Task Complexity (TSKCMP)</td>
<td>0.89</td>
</tr>
<tr>
<td>Task Uncertainty (TSKUNC)</td>
<td>0.71</td>
</tr>
</tbody>
</table>
Discriminant validity

Discriminant validity refers to the degree to which one theoretical construct differs from another. With the outer model loadings confirmed higher than any other loadings, discriminant validity is confirmed according to Gefen & Straub (2005) as the diagonal elements (representing the square root of AVE) are significantly higher than the off-diagonal values (Chin, 1998b) as illustrated in Table 2.

Convergent validity

The criterion for convergent validity is that the correlation between measures of the theoretical construct should be different from zero and significantly large to support additional investigation. Two criteria are used to determine convergent validity. In the first approach, convergent validity is assessed by comparing the t-values of the outer model loadings. t-values are estimated using a nonparametric bootstrapping technique of 100 samples. The t-values of the outer model loadings exceed 1.96 (p < 0.05) verifying the convergent validity of the instrument (Gefen and Straub, 2005).

Nomological validity

The last component of Bagozzi’s construct validity is nomological validity. Nomological validity is interpreted by the consistency with which one’s theory can be confirmed within with a wider body of theory and whether it contributes to that theory. Assessment to nomological validity takes place with reference to related research. The theoretical framework for this research is firmly planted in the domain of task-technology fit and clinical reasoning theory.

Structural model testing results

A number of relationships in figure 1 are statistically significant. Quality to Performance is significant (β = 0.216, p < 0.0005), as is locatability with a negative path sign (β = -0.199, p < 0.0028). Authorization is significant (β = 0.218, p < 0.003), compatibility (β = 0.216, p < 0.0017), system reliability (β = -0.166, p < 0.017), EOU/TRN (β = 0.158, p < 0.022), and clinical informatics/IT relationship with users (β = 0.562, p < 0.0001) are significant. Production timeliness was the only endogenous construct that did not have a significant relationship with performance.

Looking toward the left hand side of the model, table 3 includes the path coefficients for the exogenous constructs in table form. Statistically significant relationships in table 3 are denoted with an asterisk indicating p < 0.05. Five of the 26 significant relationships in table 3 are both significant and negative.
Table 3 and figure 2 illustrate the overall model testing results. Due to the complexity and number of relationships involved, the exogenous relationships from task and technology characteristics to TTF were not included in figure 2. Table 3 describes these exogenous relationships while figure 2 focuses on the paths from the TTF constructs to performance.

**Table 3: Exogenous construct path coefficients**

<table>
<thead>
<tr>
<th>Construct</th>
<th>TSKCMP</th>
<th>TSKUNC</th>
<th>TSKSIG</th>
<th>INFO</th>
<th>KNOW</th>
<th>INFER</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUAL</td>
<td>0.126</td>
<td>-0.016</td>
<td>0.261</td>
<td>0.227*</td>
<td>0.245*</td>
<td>0.130</td>
<td>0.216*</td>
</tr>
<tr>
<td>LOCAT</td>
<td>0.134</td>
<td>0.327*</td>
<td>0.129*</td>
<td>-0.310*</td>
<td>0.209*</td>
<td>0.059</td>
<td>0.218*</td>
</tr>
<tr>
<td>AUTH</td>
<td>0.305*</td>
<td>0.358*</td>
<td>0.182*</td>
<td>0.451*</td>
<td>-0.290*</td>
<td>0.093</td>
<td>0.216*</td>
</tr>
<tr>
<td>COMP</td>
<td>-0.031</td>
<td>0.174*</td>
<td>0.267*</td>
<td>0.309*</td>
<td>0.089</td>
<td>0.218*</td>
<td>-0.199*</td>
</tr>
<tr>
<td>PROD</td>
<td>0.127</td>
<td>-0.060</td>
<td>0.154</td>
<td>0.220</td>
<td>0.435</td>
<td>0.131</td>
<td>-0.171</td>
</tr>
<tr>
<td>RELY</td>
<td>0.408*</td>
<td>0.115</td>
<td>-0.215</td>
<td>0.310*</td>
<td>0.130</td>
<td>0.179*</td>
<td>-0.166*</td>
</tr>
<tr>
<td>EOU</td>
<td>0.135</td>
<td>0.269*</td>
<td>-0.005*</td>
<td>-0.206*</td>
<td>0.072</td>
<td>0.439*</td>
<td>0.158*</td>
</tr>
<tr>
<td>RELAT</td>
<td>0.109</td>
<td>-0.141*</td>
<td>0.030*</td>
<td>0.083</td>
<td>0.353*</td>
<td>0.232*</td>
<td>0.562*</td>
</tr>
</tbody>
</table>

**Figure 2. Model Testing Results (TTF to Performance)**

**CONCLUSION**

This study identified several positive, significant relationships from TTF to performance. Similarly, both task and technology characteristics demonstrated several significant positive links to TTF constructs. Utilization did not significantly impact performance, nor did TTF show a significant relationship to utilization. As a result, utilization was not included in the final model.

The main contributions from this research include: 1) a validated instrument with diagnostic and predictive capabilities for the evaluation of clinical reasoning performance with an electronic health record, 2) an extension and validation of the TTF
model to the clinical domain and an emphasis on clinical reasoning task characteristics and EHR technology characteristics, and 3) a reaffirmation of the on-going viability of TTF theory to explain the factors influencing individual performance.

One implication of this study is the development of a validated model and instrument that can be the basis for further study in the clinical reasoning/EHR fit domain. Because of the challenges inherent in measuring performance with respect to clinical decision-making, and due to the lack of a validated instrument, the results from this study may open doors to other researchers who wish to evaluate this phenomenon. Another implication for research concerns theory adaptation. This approach adapted TTF to a new domain and the approach used here could serve as a template for further adaptation of TTF. It may be possible to adapt this model to other performance-related questions in the clinical domain.

One of the most important limitations of this study is the TTF constructs. This project examined the final eight TTF constructs in Goodhue’s (1995b) original study without deviation. That is not to say the study was not open to new ideas. This was the central point in the interview process; the goal was to familiarize clinicians with the construct and the theory behind it, and then explore other avenues. No other TTF concepts arose during the process.

Another concern relative to the TTF construct is the lack of support for production timeliness, system reliability and locatability. Although production timeliness was not significant, system reliability and locatability were, however as discussed the direction of the relationships are reversed from the hypothesized orientations. Future work should re-examine these constructs and attempt to determine if the constructs are simply not valid in the context of EHR, or if the language and question content should be changed and retested.

One must also consider the generalizability of these findings. The study was conducted in a medium-size health system in the Midwest, and focused on one brand of EHR in the context of one of a myriad of possibilities with respect to implementation status. The major limitations of PLS include a higher risk of overlooking true correlations as well as sensitivity to the relative scaling of the descriptor variables. Finally, sample size and potential issues concerning common method bias should also be considered limitations to this study.

Task-technology fit research is experiencing resurgent interest with the theory being applied across a wide variety of domains and problems. In the health care domain, TTF is being applied to problems of utilization and perceived benefits (Lepanto, 2010). Other domains include the use of TTF theory for knowledge management systems (Kuo, 2011), online learning systems (Yu, 2010) and hotel information systems (Taegoo, 2010).

Directions for Future Research

Other areas within the TTF construct worth exploring include data locatability, system reliability and production timeliness. Production timeliness is concerned with the turnaround times for clinical informatics and IT with respect to reports and information requests, and in this study it did not have a significant relationship to performance. This is a curious result when one considers how vital such service may be to clinicians. Also curious were the negative, inverse relationships from locatability and systems reliability to performance. Future research should address these constructs in the context of other implementations of EHR and attempt to determine their overall contribution to the model.

REFERENCES


