Workload and Performance Factors Associated with Multimedia Job Aids for Community Health Workers

PhD Dissertation

Jose F Flórez-Arango, MD MS

The University of Texas Health Sciences Center at Houston

School of Health Information Sciences

Advisory committee:

From SHIS:

M Sriram Iyengar, PhD

Kim Dunn, MD, PhD

Jiajie Zhang, PhD

Cynthia Phelps PhD

From NASA JSC:

Kathy Johnson MD PhD

July 2009

Workload and Performance Factors Associated with Multimedia Job Aids for Community Health Workers

A

DISSERTATION

Presented to the Faculty of

The University of Texas

School of Health Information Sciences at Houston

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

by

Jose Fernando Florez-Arango, MD MS

Committee Members:

M. Sriram Iyengar, PhD[1]

Jiajie Zhang, PhD[1]

Kim Dunn, MD MPH[1]

Cynthia Phelps, PhD[1]

Kathy Johnson, PhD[2]

[1] School of Health Information Sciences, The University of Texas Health Science Center at Houston

[2] NASA Johnson Space Center

PhD Dissertation

Workload and Performance Factors Associated with Multimedia Job Aids for Community Health Workers

By

Jose F Florez-Arango, MD MS

July 24th 2009

APPROVED:

____

M. Sriram Iyengar, PhD

____

Jiajie Zhang, PhD

____

Kim Dunn, MD MPH

____

Cynthia Phelps, PhD

____

Kathy Johnson, PhD

To Any, Juanis and Susy

Table of Contents

Table of Tables

Table of Figures

Table of Equations

Introduction

This dissertation focuses on factors of multimedia job aids that modify workload, protocol adherence and clinical errors in community health workers. Literature shows that community health workers performance is not acceptable even with support of paper job aids. There are cognitive theories that try to explain reasons why the performance of community health workers is poor regardless of the access to paper based-job aid. Based on cognitive science and multimedia design theories an intervention was designed to compare alternative representations for the information contained on paper job aids and the capability of this new designed job aids to enhance community health workers performance.

The dissertation is divided in 5 main parts: 1. identification and description of the problem, 2. a methodological approach to create and evaluate an intervention, 3. Presentation of results of the intervention evaluation, 4. Discussion of findings and 5. Conclusions.

Problem statement

Sub-optimal performance is common in health care. Consequences of health care providers' poor performance are reflected in increased patients morbidity and mortality (Morrato, Dillon, & Ziegler, 2008). Literature shows that lack of knowledge or skills is one source of sub-optimal performance in health care (Knebel, 2000). One strategy to reduce suboptimal performance of health workers by filling gaps in their knowledge and skills, is the implementation of tools named *Job Aids* (Folley & Munger, 1961; Knebel, 2000). A formal definition for job aid is an **external device or cognitive artifact** that provides **just-in time knowledge and information** to help **individuals with tasks** by directing, guiding and enhancing **performance** (Elsenheimer, 1998; Knebel, 2000; Rossett & Gautier-Downes, 1991). **However** literature demonstrates that even well motivated and trained health workers that use manual job aids can present sub-optimal performance expressed as **errors in treatment and diagnosis**, as well as low **protocol compliance** (S. Y. Rowe et al., 2007).

Controlled studies that deal with health workers performance and job aids are few in number. Further, the few studies that have been conducted are focused on manual job aidsand do not provide an explanation about sources of failure or success of the job aids supporting particular tasks (Knebel, 2000). Possible explanation for **suboptimal performance** displayed by health providers using job aids could **be related to the complexity and the format** used for presentation **of the job aids, situation that increases the user's workload** (V. L. Patel, Arocha, Diermeier, Greenes, & Shortliffe, 2001; S. Y. Rowe et al., 2007). Frequently job aids are poorly designed. Main design problems are terms are not fully explained nor is there complete information, information is presented in a higher reading level that the one that corresponds to the user, and workflow is not explicit stated (Knebel, 2000; V. L. Patel et al., 2001; Quaglini & Ciccarese, 2006).

It is important to note that **most of the job aids** currently in use by community health workers **are paper-based**. Job aids presented in static format as paper are passive in the decision making process, and have limited interactivity (Quaglini & Ciccarese, 2006). The static format has additional limitations in the way that information is presented, e.g. to represent common elements in health care as sounds, processes, procedures and activities.

Basic elements of the problem space has been presented and stated how job aids play an important role in human performance, and how they are related with additional human factors. Figure 1 shows different elements that are part of the problem space and the way that they are interrelated.

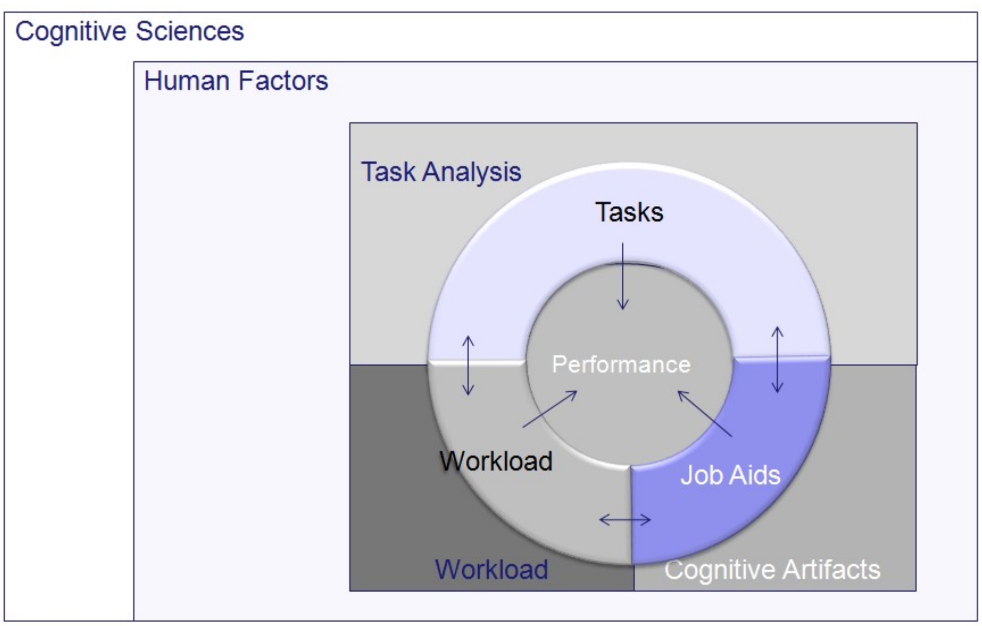


Figure 1. Problem space diagram

Cognitive science and learning theory support that multimedia is a rich format to transmit information (Clark, Nguyen, & Sweller, 2006; Mayer, 2001). Multimedia format in the context of learning can help to illustrate terms, provide supporting information, reduce the complexity and reduce the cognitive load of the learner (Mayer, 2001). In addition, persuasion theory understood as 'attempt to change attitudes or behaviors or both without using coercion or deception' (Fogg, 2003) gives elements to use multimedia as a possible platform to enhance job aids. The main persuasion elements that differentiates structured multimedia job aids from paper job aid are suggestion (express urgency and severity situations), tunneling (provide an structure that matches workflow), reduction (reduce cognitive workload), and tailoring (adaptable to users) . However, the benefits of multimedia in the context of job aids to improve health care performance, while potentially intriguing, have not been formally investigated in health care (Clark et al., 2006; Knebel, 2000). This gap in knowledge on the possible benefits of multimedia job aids is the study of this dissertation.

So far the problem elements include **poor performance** of **health workers**, even with support of **job aids**, with a possibility to enhance their performance by **designing better** job aids. Following is a detailed description of these three elements of the problem. Community health workers (CHWs) were selected as the instance of the health care providers to be studied. Reasons and justification for this decision are presented in the next section. Subsequent sections elaborate previous knowledge about job aids and CHWs as well as an exposition of the multimedia job-aid approach that creates the basis for this dissertation.

1 Community health workers

Community health workers (CHWs) are typically the sole providers of medical care to millions, especially in developing countries and rural areas, where physician density is very low and accessibility to health resources is scarce. CHWs have multiple titles as community health advisors, patient navigators, public health auxiliaries, and *promotores/as,* barefoot doctors among others. For the purpose of this work, a CHW will be considered as any individual that provides health services to members of a specific community. This individual is required to have specific training for the health service offered, but without a formal post-secondary degree or certification in the clinical field (Dower, Knox, Linder, & O'Neil, 2006).

Most CHWs are the primary point of contact with health services (World Health Organization, 2006) in certain communities; in fact, CHWs are very often the only opportunity that sick people have to receive health care in those communities. Today, CHWs are very important in developed countries given the cost of the health care, and the difficulties to access medical services for a big portion of the population (Dower et al., 2006). CHWs are filling a variety of roles. Following are seven basic roles attributable to CHWs: (Dower et al., 2006):

1. Cultural mediation between communities and health and human services systems,
2. Informal counseling and support,
3. Providing culturally appropriate health education,
4. Advocating for individual and community needs,
5. Assuring people get services they need,
6. Building individual and community capacity, and
7. Providing direct services.

An important characteristic of CHWs to be taken into account is that usually they have not completed high school and only received a very basic level of clinical training. The average CHW's reading level is equivalent to 7th grade (S. Y. Rowe et al., 2007). The low literacy level of the CHWs demands that the information made available is at the proper reading level.

Additional to limitations in the basic knowledge of CHWs, It is a challenging to keep CHWs up-to-date in terms of continuing education and training. CHWs' training is an expensive process, especially when accounting for the social cost of removing the only source of health care for a community for the length of the training. Alternatives such as online education have been tested, but this is not a complete solution, given the lack of computers or connectivity in their working sites (S. Y. Rowe et al., 2007).

CHWs are very important elements in the health care system, but who require special attention given their education, roles and function. They are prone to error and, in general, their performance is below expectations in the clinical domain. Results of low performance measurement are called performance gaps. Experts have identified three main sources of performance gaps, they are: Lack of knowledge/skills for the task, flawed work environment or processes; and lack of motivation for the job. For each of the sources of performance gaps, several strategies have been proposed. Use of job aids is one of the proposed interventions as a complement for training and coaching, when possible gaps are originated in subjects' lack of knowledge or skills. (Knebel, 2000). This dissertation is centered in understanding the role of job aids to fill the performance gap derived from lack of knowledge and skills of CHWs.

2 Job Aids

Performance aids in the context of cognitive sciences (D. A. Norman, 1993) and arteology (science of products and professions), (Routtio, 2004) are items of auxiliary equipment or document provided to enhance on-the-job **performance.** In other words, they are products or devices designed with the specification to enhance human performance by extending physical, cognitive or social capabilities (Folley & Munger, 1961).

Further, Cognitive artifacts are subclasses of performance aids. Cognitive artifacts are defined as external representations of human knowledge in artificial devices (not necessary physical devices) designed to maintain, display, or operate upon information (D. A. Norman, 1991). Therefore, cognitive artifacts are representations that help to extend working memory and additional cognitive process as attention, perception, action, language and thought (D. A. Norman, 1993). Examples of cognitive artifacts are language, books, and memory cards, among others. Job aids are considered specialized cognitive artifacts. This topic has been studied broadly by the military, but it has not been studied in the context of CHWs domain (Knebel, 2000).

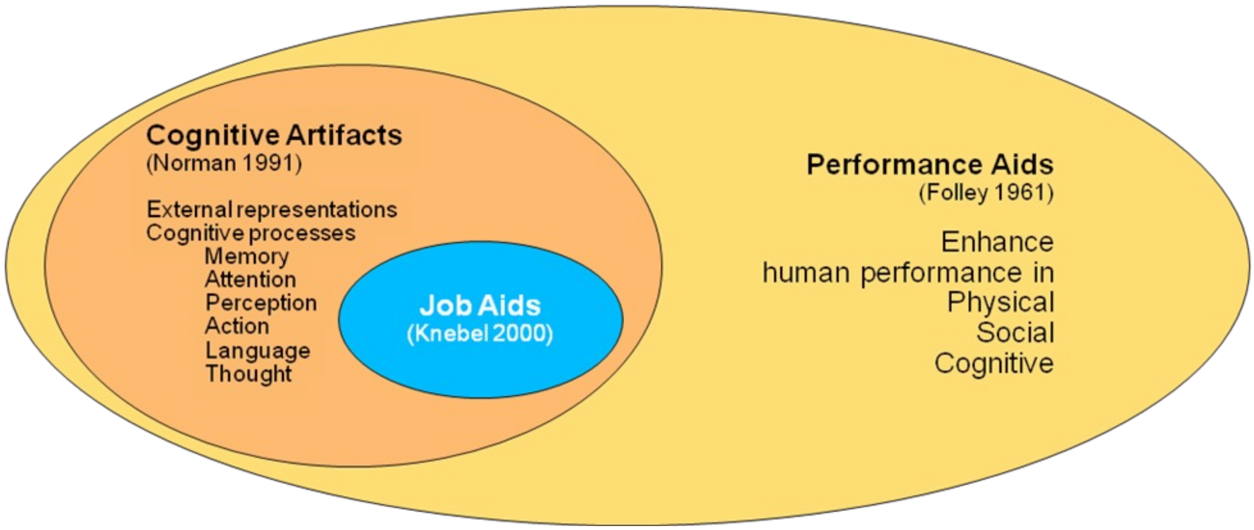


Figure 2. General Concept of Job Aids

Job aids are an additional subclass of cognitive artifacts and performance aids. Job aids are those devices that accomplish both functions: to externally represent information and to operate with the aim of reducing errors and user workload, supporting the CHWs job. A formal definition of job aid is an **external device** that provides **just-in-time knowledge and information** to help **individuals with work and activities tasks** by directing, guiding and enhancing their **performance.** Concrete examples of job aids include but are not limited to clinical guidelines, flow charts, flipcharts, clinical algorithms, drug manuals, reminders and calculators (Knebel, 2000).

Figure 2 summarizes the relationships and the main concept to understand the definition of job aid used for this research. Each oval represents a class where main properties are expressed. Then it can be said that job aids are cognitive artifacts that support the cognitive performance of humans.

The following sections present elements to understand the concept of job aid and the relevance in CHWs work. The sections describe types of job aids, components of job aids and the relationship of job aids and CHWs.

2.1 Types of Job aids

Job aids are in essence **interactive products** according to arteology (science of products and professions). An interactive product is a device that performs some operations steered by the user, and under the user's demand (Routtio, 2004). But not all interactive products are the same. Some of them interact with the user in a **passive** way, while others are more **active**. Therefore, **active job aids** are those where the device is able to receive input and changes its state according to user inputs, while**passive job aids** are not able to receive inputs (Routtio, 2004).

An example will illustrate the job aid classification. A clinical guideline can be considered to be a job aid since it helps with the classification of a disease processby a set of predefined observable variables (external representation) with the aim of standardizing care to a certain level (performance aim). The guideline can be **passive** or **active** depending on the way it is encoded and presented. When the guideline is presented as a text document or even a flow chart it is passive since the user needs to understand the structure of the content and create a strategy to navigate the information in order to obtain some result. But when the same guideline is presented such that the guideline requires interaction or intervention from the user in order to present a relevant portion of content, it is an active job aid. This study compares **passive and active** job aids.

2.2 Components of job aids

Job aids inherit cognitive artifacts properties. Then, job aids have two main components: **content and representation**. Any **change in** some of the **job aids components** can be reflected in the **usability of the job aid** and as consequence in the user **workload and performance** (Knebel, 2000; V. L. Patel, Arocha, Diermeier, Greenes, & Shorliffe, 2001; Quaglini & Ciccarese, 2006).

Possible explanation for **suboptimal health providers performance** using job aids with **proper content** could be related to the complexity of the content and the format used for presentation **(representation**) of the job aids that increases the user's workload (V. L. Patel et al., 2001; S. Y. Rowe et al., 2007). Literature shows that it is common that job aids are poorly designed, terms are not fully explained, there is incomplete information, are presented in a higher reading level thanthe one that corresponds to the user, or workflow is not explicit stated (Knebel, 2000; V. L. Patel et al., 2001; Quaglini & Ciccarese, 2006). Therefore there is a mismatch in the representation, and the content.

Most of the job aids currently in use **are static and paper-based**. Job aids presented in static formats are passive in the decision making process and have limited interactivity (Quaglini & Ciccarese, 2006). This format has additional limitations in the way that common pieces of information in health care are presented. For example paper is not strong media for job aid that requires representation of sounds, movements or activities, as frequently is required in health care. This dissertation presents how different representation of the basic content modifies CHWs providers.

2.3 Job aids and CHW

Most of the studies done on job aids and health care have focused on physicians and nurses. But even in this context many questions have not been addressed. As a matter of fact, when projects involving CHWs are designed, job aids are taken into account. A problem of described studies is that usability is barely studied after project's implementation. There are few studies where the utility of the job aids has been established. The most frequent evaluation of utility has been performed within one month of the project's implementation (Knebel, 2000). The main questions regarding job aids for CHWs include strategies for adoption, design methods, training and usability (Knebel, 2000). Most of the recent studies are dedicated to studying manual job aids, those that are passive and lack of interactivity (Y. M. Kim et al., 2005).

A well know type of job aid is the clinical guideline. A clinical guideline is a class of job aid that supports a health care provider in the acquisition of information, classification of the patient, and intervention in a standardized way. Clinical guidelines for CHWs have been developed in an effort to keep health services standardized and to impact public health outcomes. Usually clinical guidelines are developed by panels of experts. Delivering guidelines can be done in multiple modes such as flow-charts, clinical algorithms, and textbooks, among others (Knebel, 2000).

The study of the use of clinical guidelines by CHWs is limited (Knebel, 2000). One well studied clinical guideline for CHWs is part of the integrated management of childhood illness (IMCI) initiative. IMCI was developed by the World Health Organization (WHO) to support CHWs at first-level health care facilities to assess, classify, and treat sick children, as well as to counsel sick children's mothers for highly prevalent diseases (Gove, 1997).

IMCI has been in existence for more than 10 years, is one of the most important strategies for reducing mortality and morbidity. Studies using IMCI demonstrated that training CHWs in the use of the guidelines improves their performance (Simoes et al., 1997), yet, due to the cost of training, this type of program is not reaching those who need it most (Victora et al., 2006). In the cases where the program has been implemented, long-term adherence and usage of the guidelines is not optimal. As demonstrated by Rowe, diagnosis and treatment error rate of CHWs can be around 61% (Figure 3) (S. Y. Rowe et al., 2007). Other reports have shown that the clinical competence of health workers in developing countries is between 71% and 79%, where clinical competence is measured as **history-taking, physical examination, diagnosis and treatment** (Ashwell & Freeman, 1995). In any case, this level of performance is considered mediocre, since the error rate can lead to increased morbidity and mortality (S. Y. Rowe et al., 2007).

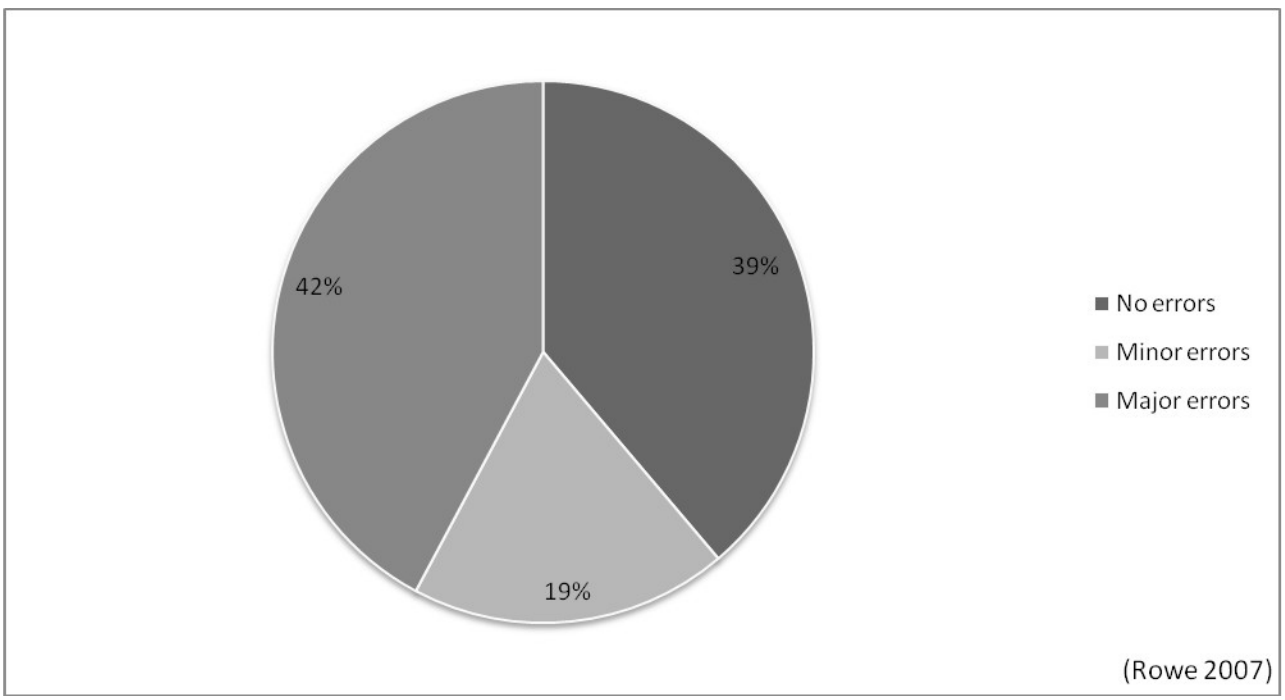


Figure 3. Community Health Workers Performance

In one of the IMCI guidelines adherence studies errors were classified as minor errors or major errors. A minor error was defined as a case where the CHW did not choose all recommended treatments but the treatment provided was potentially a life-saving intervention; a major error is the case where the treatment provided was considered inadequate to save the child's life (S. Y. Rowe et al., 2007). Overall, CHWs had minor errors in 19% of the cases and major errors in 42% of cases. (S. Y. Rowe et al., 2007) The consequences of such mistakes can be profound, causing avoidable high mortality and morbidity. The patients that required some critical intervention given their conditiononly received all the life saving interventions required 59% of the cases. Previous finding means that by omission of possible interventions by CHWs, the patients' lives are put into jeopardy 41% of the time. Authors try to explain the results based on that the process of classification and treatment for the group of patients is complex (S. Y. Rowe et al., 2007).

There is conflicting literature about the type of interventions that improves the adherence to job aids. Studies showed that the use of job aids have good results (Knebel, 2000), negative impact (Grol et al., 1998), and no difference compared with using just the training and the knowledge of the CHW (S. Y. Rowe et al., 2007). One interesting finding is that use of flip-charts does not modify the adherence to the guideline. Authors attributes this mainly to the poor design of the flipcharts for the guidelines (Ashwell & Freeman, 1995; S. Y. Rowe et al., 2007). This dissertation adds quality data to the body of knowledge regarding the use of two types of job aids and the impact in performance gaps of CHWs.

3 Active Interactive Persuasive multimedia job aids - An experimental approach

Based on the need to improve job aids a multimedia job aid is proposed as alternative to paper based job aids. This selection intended a better match in representation of tasks. Multimedia job aids support information that is not easily represented in paper, as sounds and movement. Concept of multimedia used in this dissertation considers it as a type of knowledge representation. In this context, multimedia requires that two or more sensory channels work together to aid the human cognitive processes (Mayer, 2001). Most studies on multimedia have focused on instructional messages, only one of the properties needed to make a job aid easy to use (Knebel, 2000). The multimedia instructional message can help with both information acquisition and knowledge construction (Mayer, 2001).

Multimedia has been studied from two points of views by instructional designers: as information acquisition tool and as a knowledge constructor. Multimedia as an information acquisition tool is a passive vehicle that adds information into memory. Multimedia as a knowledge constructor is more a communicator and cognitive guide to build coherent mental structures, making the learner an active actor (Mayer, 2001). In the proposed research, multimedia behaves more like an information acquisition tool given the tunneling action that has in the problem solving. There is no intention to create long-term knowledge retention after use of multimedia job aids. Proposed multimedia job aid is designed to provide tunneling (from captology) of tasks. Repetition of actions in a tunneled way makes the CHW a passive learner.

There are guidelines to create good multimedia content; the guidelines are based in the learner-centered approach for multimedia. The learner-centered approach tries to understand how the human mind works and searches for ways of adapting multimedia to enhance human learning (Mayer, 2001). Following good practice guidelines, multimedia instructional design has shown that multimedia content improves knowledge retention; reduces learning time and testing answering time; and increases recall, recognition and generalization of application (Clark et al., 2006).

The literature suggests, but there is no complete evidence, that implementing similar practices for multimedia instructional design into the design of multimedia job aids will provide benefits to the user of job aids (Clark et al., 2006). Although the use of multimedia as a format for learning tools has been shown to improve learning and promises just-in-time performance improvement, it is also known that a increased workload imposed by an activity or a device can hinder its adoption. This dissertation intends to reduce the knowledge gap about the effect of multimedia job aids in performance, learning and workload.

4 Conceptual Model

This research is framed in the **cognitive sciences**, centered particularly in **human factors and cognitive artifacts**. The intention of the dissertation is to understand the **relationship** between **clinical tasks, user workload, and user performance and multimedia job aids (a cognitive artifact) (Figure 4)**. In this research, concepts from human performance theory, learning theory, work-centered research, captology (study of computers as persuasive technologies) and cognitive ergonomics, are interrelated. In a concrete perspective the **performance of a subject reflects the relation between task factors**, the **associated workload** and **possible supporting job aids**. This research studied the relationships between such elements (Figure 1).

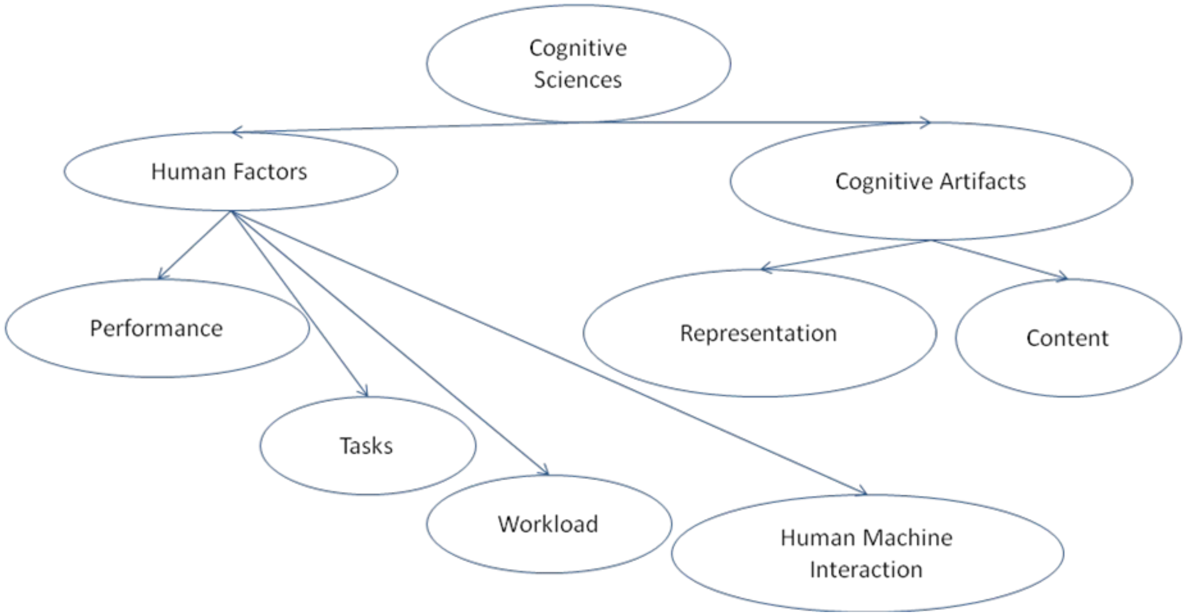


Figure 4.Conceptual Framework

Based on the framework and the problem statement, this research is an instance that **focuses in clinical tasks**, where **subjects** of study are **community health workers**, the **performance measurements** are **protocol compliance and errors**, and **workload** is an associated factor measured in a **multidimensional** way, while the **job aid** under evaluation has a multimedia format.

5 Summary of Problem Statement

To summarize, performance in health care and in special CHWs' performance is sub-optimal, influenced by lack of knowledge and skills. CHWs are in need of tools to reduce the knowledge and skills gap. Reduction of such gap could help to improve their performance by enhancing their protocol adherence and by reducing error rates. A possible solution for the problem is to improve the content representation of job aids with multimedia. This work provides elements to understand the existing relationships between CHWs performance using job aids, the task factors and the workload.

6 Hypothesis

Based on the facts presented in the problem statement that:

- community health workers have poor performance,
- there are not differences of community health workers performance when using traditional paper based job-aid
- and there is an evident mismatch representation between tasks and traditional job-aids;

And following concepts from :

- multimedia theory,
- cognitive load principles (Clark et al., 2006)
- captology principles (Fogg, 2003).

This dissertation was driven by the hypothesis that use of multimedia job-aids can enhance CHWs performance and modify workload reducing cognitive workload without increasing others factors. To test this hypothesis an experimental approach was used. Detailed explanation of methods is presented in the following section.

Methodological Framework

The methodological framework used is a mix of existing and newly created methods. This dissertation evaluated Colombian CHWs performance and workload in a controlled and standardized environment, the human subjects section describes inclusion criteria and the procedures followed to protect them. The experimental design section explains how a crossover design was selected and implemented. There is a section dedicated to present variables in detail. Variables section includes newly developed methods to create scenarios, create interactive rich media job aids and performance measurements based on clinical tasks. The sample size calculation, data collection, data quality assurance and the data analysis are exposed in separated sections as well.

7 Experimental Design

In order to test the hypothesis, a crossover study with two treatments and two periods was conducted. Crossover studies permit to expose all subject to all treatments under evaluation. In the case of a crossover studies with two treatments – two periods half of the subjects receive treatments in one order or sequence, while the remaining subjects receive treatments in reverse order. One advantage of crossover studies is that smaller sample sizes are required than in cohort studies. Crossover studies permit a reduction in the variability between subjects, since each subject serves as its own control. The crossover design also allows accounting for effects created by the order intervention. Order of treatments needs to be accounted because it can induce changes not related with the treatments (Dallal, 2007). Crossover studies can be affected by a residual effect when switching treatments. This means that in the second period is a residual effect of the treatment in the first period. The possible residual effect in this dissertation is provided by procedural learning, since subject will be exposed to multiple cases with common tasks.

In this dissertation design study factors will be:

- TREATMENT: refers to the interventions to be evaluated in the study. This study has two types of interventions, which will be named “Use GV” and “No Used GV”. Use GV refers to the experimental method under evaluation, the multimedia job aid, more details regarding treatments are in Sections 7.1 and 7.2.

- PERIOD: referrers to the moment of the study where actions are performed. Since this study has two treatments, then is it necessary to have two periods. Periods in the experiment are called as '1' and '2'. Subject where exposed to 15 randomly selected clinical cases in each period. Period 1 contains cases 1 to 15, while period 2 contains cases 16 to 30. More details regarding cases are in Section 10.3

	Pediatrics	**Trauma**	**No Trauma**
Multimedia job aids	5 cases	5 cases	5 cases
Current job aids	5 cases	5 cases	5 cases

Table 1. Distribution of cases in periods

- GROUP: is a way to distribute subjects and is determined by the order or sequence in which treatment are offered over period to subjects. This study will have two groups with levels 'A=Used GV then No Used GV' and 'B=No Used GV' then Used GV'.

	Period 1	Period 2
Group A	Used GV	No Used GV
Group B	No Used GV	Used GV

Table 2. Crossover study design.

This study used two treatments, one experimental (Use GV) and one control (No used GV). The experimental treatment is the use of multimedia job aid and the control treatment is the use of traditional methods and job aids. Following is a detailed description of both treatments.

7.1 Control treatment paper-based job aids

Control treatment consisted of paper-based job aids made available in a binder for every research subject. Paper-based job aids were selected according with cases selected for evaluation (see section10.3). Selection criteria for this job aids included, to be evidence based, peer reviewed, published and available freely for CHWs in Colombia. For this reason two major sources of guidelines were used. The first source was the AIEPI (Atención Integral para las enfermedades prevalentes de la Infancia)/IMCI (Integral Management of Childhood Illneses) technical manual published by the World Health Organization. The second was the Pre-Hospital care guidelines written by the Colombia Association of Prehospital Care and published by the Colombian Ministry of Social Protection (Health and social security). Both sources are used in training of CHWs in Colombia and expected to be used in a daily basis in the work.

Guidelines were grouped in three sections in the binder, pediatrics, adult trauma and adult clinical conditions. Pediatrics sections contained two guidelines, IMCI for 1 week to 2 months old children and IMCI for 2 months to 5 years old children. The trauma section included, initial assessment, management of thorax wound, general management, burning management, immobilization guidelines. Finally No trauma section included, stroke, myocardial infarction, job aids. At the beginning of each period subjects received a briefing about location and use of each guideline.

7.2 Experimental treatment: An active interactive multimedia communication-enabled job aid

The experimental treatment in this study consisted of a series of job-aids presented in a multimedia job aid delivery system named GuideView (GV). GuideView is owned by the University of Texas Health Science Center at Houston and licensed exclusively to Advanced Guidance Systems LLC. It is a free distribution tool for academic purposes. GuideView was conceived to create and deliver clinical guidelines in an easy-to-use, step-by-step multi-modal format for the benefit of non-physician health care providers. At each step, clinical advice is presented simultaneously in text, voice, images, and video/animations. These modes synergistically create a rich information delivery environment in which guidelines can be easily understood and complied with even by poorly trained persons.

The two main components of GuideView are GuideView Author and GuideView Viewer. The Viewer delivers the guidelines and GuideView Author enables the development of GuideView compatible guidelines (called guideviews) using a GUI authoring environment. Typically, a team consisting of clinicians and subject-matter experts develops guideviews taking into account the training of final users. GuideView architecture strictly separates presentation from content, making it easy to accommodate multiple languages.

GuideView has been successfully integrated with PDA-based electronic medical records and can acquire data from medical sensors with subsequent real time execution of branching logic. GuideView was tested at NASA Johnson Space Center. A small usability study was conducted using 10 biomedical engineers that work in the surgeon console at Johnson Space Center. In the NASA study subjects showed a preference for voice over text to receive instructions, especially when performing hands-on work, this was an encouraging finding since is a feature not implementable in paper and serves as basis for the purpose of the proposed study (See Figure 5) (M. S. Iyengar & Florez-Arango, 2007; M. S. Iyengar et al., 2005).

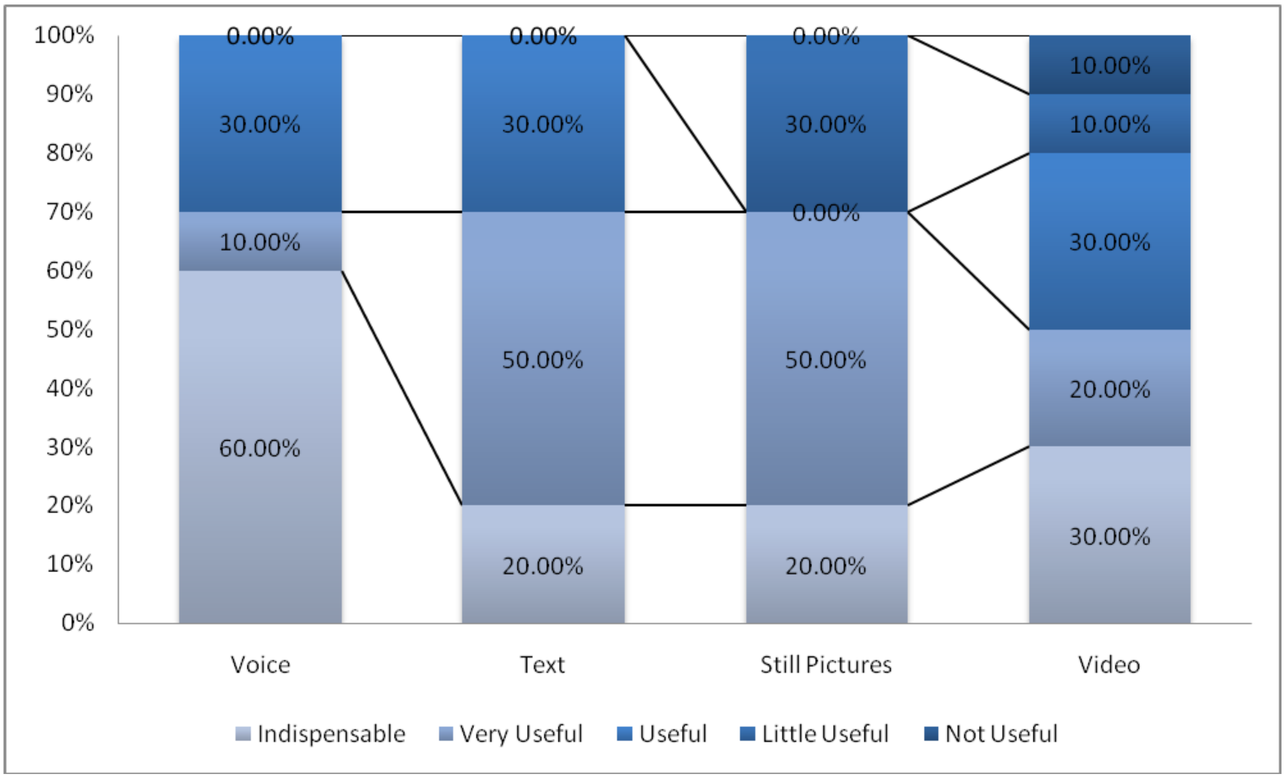


Figure 5. Usefulness of output modes of GuideView

The GuideView Viewer executes on multiple platforms including web, Windows PCs, and Windows Mobile for PDAs. Since the cell phone is rapidly becoming a cheap and ubiquitous computing and communications device in many areas including Central / South America, India, and Africa, GuideView Viewer was recently ported to Windows mobile (smart) cell phones. GuideView running in cell phones is called CGV, this system was tested on a Cingular 3125, and a HTC TYTNII cell phone. Quality of voice, image display, and video playback were found to be comparable from that on PDAs or lap/desktop machines (Florez-Arango & Iyengar, 2007).

Guideviews for CGV are stored on an SD-card inside the cell phone or PDA and can be updated by synchronization with a host server. The phone numbers of a remote physician or hospital can be encoded into specific steps in the guideview. When execution arrives at such a step CGV automatically dials that number (Florez-Arango & Iyengar, 2007). The user can then transmit images, data, and audio to the remote expert for further advice. Since the guideview was created previously by medical experts, CGV can enforce the following of predefined paths in protocols. CGV documents patient encounters in a small local database that can be synchronized later to a server.

In contrast to other forms of telehealth that typically require the physician to participate in a voice or video-conferencing session, CGV can initiate such synchronous communication only when needed. This feature enables better utilization of all care providers, ensuring that CHW handles routine care in an appropriate and standardized form locally. Expert assistance from physicians and specialists is called upon only when necessary.

In selecting a proper device to deliver multimedia job aids, a literature review showed that the worldwide adoption of cellular phones is without precedent in the history of technology. While the number of cellular phone users in developed countries in Western Europe is estimated to be close to 380 million, the most stunning finding is the growth in cell phone usage in developing countries in Latin America, Asia, and Africa. The annual growth in the cellular phone market in these regions is around 17% (Netsize Group, 2006). In India alone, it is estimated that there are more than 125 million cellular phone users and new subscribers sign up at the rate of 3 to 5 million every month (Migration Dialogue, 2006). Media reports have documented the innovative ways in which rural and economically disadvantaged populations utilize cellular phones to improve their daily lives and reduce the technological gap (Kelly, 2006; Rice & Katz, 2003). In Colombia, the experimental location, people with income levels below the country's mean represent 72% of those with cellular phones (ASOCEL, 2006). The number of cellular phone subscribers worldwide is estimated at 2 billion (32.2% of world population). These subscribers give access to mobile telephony to more than 5 billion people (80.6% of world population) (Netsize Group, 2006).

Competition among cell phone manufacturers and service providers results in constant introduction of new cell phone models with increasingly powerful processors, higher data storage capacities, and higher display resolutions. Cell phones have gone from being purely communications devices, to becoming sophisticated, powerful, mobile communications-computing devices.

There are previous experiences with the use of cellular phones in health care. Cellular phones have been used to monitor cardiac function for occurrences of arrhythmias that are then transmitted to a doctor's office (Schickendantz, Pillekamp, Emel, Sreeram, & Brockmeier, 2006); short message service for data exchange between nurses and patients (H. Kim, Kim, & Abn, 2006); and, traditional on-demand direct voice connection between caregivers and patients (Farmer et al., 2005). Additionally, the researcher have shown how cellular phones can be job aids, that can become a *de facto* complete synchronous telehealth tool (Florez-Arango & Iyengar, 2007)

This study was strongly motivated by the idea of exploring the use of cell phones for health care in developing countries, where health care providers are mobile, traveling from village to village with limited resources and poor phone landlines, connectivity and transportation infrastructure. Thus, portability and availability of medical information at the point of care is highly desirable. It is for these reasons that cellular phones can potentially contribute to telehealth and health care delivery in general. And the justification for selecting a technology that is compatible with cellphones delivery as the experimental treatment in this dissertation.

7.2.1 **<u>Construction of GuideViews</u>**

All guideviews used in this study where created *de novo.* In order to create a guideview an iterative multi-stage procedure was followed. The process was divided in three main steps: logic representation, media population and testing. Following is the explanation of each of the three steps.

7.2.1.1 Logic representation

The logical representation is the basis for any guideview, this computational structure allows GuideView viewers to navigate through the gudieview content. A logical representation is an XML file that contains information regarding the guidelines, and a directed graph with nodes and links. Every node will contain the multimedia elements that will be presented to the subject in each step and links represent buttons to control navigation in the user interface.

Experimental guideviews were built using a Delphi approach. Delphi method consists in iterative anonymous revision of content, until saturation (no new changes) or agreement is achieved (Clayton, 1997). The anonymous participation in Delphi method allows reducing the hierarchical power relationships among individual that can affect its participation or opinion driven by the authority. At least 3 physicians participated in developing each guieview, and at least 3 iterations of revisions where allowed.

The initial version of the logic representation was created in GuideView Author by a team of 2 physicians using as initial guidance tasks defined in the Ontology of Clinical Tasks (OCT) (See Section10.2.1) and already published guidelines for the cases of interest (See Section7.1). Physicians that participated in the creation of the original logic layout are the same that created the OCT. Both physicians have knowledge of GuideView authoring system and the guidelines to be used.

In the process of creation of guideviews each task in a guideline, and defined in OCT, was mapped to a node in the guideview logic representation. A heuristic of priorities was applied to link nodes in the guideview. If a logical workflow is already stated, as in CPR, that workflow was respected. All critical and time constrained tasks were located at the beginning of the workflow segment, providing a priority within them according with the time constraint. *Decisions* and *Information Acquisition* tasks permitted to segment the guideview and nest proper tasks according with answer (only steps related with fever where followed, once fever was established as present). Resulting workflow can be as divergent – convergent graph.

Once the initial logic was created additional revisions were conducted by at least 2 physicians that did not participated in creation of the original logic representation. Iterations were submitted anonymously for cross-revision and afterwards consolidation, following Delphi method. Once agreement and saturation of changes were established, logical representation was locked and then each node populates according with a set of predefined heuristics.

7.2.1.2 Heuristic for content population of guideviews

Each node in guieview has 4 slots to allocate content and create the multimedia environment. Heuristics were extrapolated from previous knowledge from multimedia design for learning (Mayer, 2001). The four slots are *Instruction, Explanation, Audio, Image. Instruction* and *Explanation* slots only can contain verbal information as text. While Audio and Image slots can contain verbal and no verbal information. *Image* slot receives both to static (pictures) and in-motion (video and animations) visual information. Following is an explanation how to populate content into slots to keep a proper representation of the task. First there is an explanation for the verbal content, followed by the visual information.

Regarding verbal content it is required that each node in guideview have the *Instruction* slot completed. Content of the *Instruction* is determined by the type of task of the node according with OCT. Characteristics of *Instruction* slots content includes to be a short single sentence and to be written in easy to understand language. There are three types of task (*Actions, Decisions and Plans*) , recommendations how to populate verbal content (Instruction, Explanation and Audio) are summarized in Table 3. For the task with type *Action* it is recommended to use a short command as “do x...” in the *Instruction* slot and an explanation of terms or steps in the *Explanation* slot. For tasks with type *Decision* the heuristic is to provide a short question (“What size is …”) in the *Instruction* slot and the explanation to the optional answers (Large,Medium,Small) in the *Explanation* (Large means more than X …) . For *Plan* tasks *Instruction* will contain the name or title of the plan, while the *Explanation* will contain the aims and summary of the plan.

Type of Task	Instruction	Explanation	Audio
Action	Short command	Description of steps to be follow	Command
Decision	Short question	Explanation of the options	Question
Plan	Title of the plan	Aims of the plan	Description of the plan

Table 3. Heuristics of verbal content based on type of task

Visual information associated with a node in a guideview follows a heuristic that is informed as well by the OCT, allowing some relationships between factors associated with the task and the type of content to enhance representation. It is important to notice that visual information can be both verbal and non-verbal see in Table 4. One example is that in cases where a task requires a cognitive process that involves movement (an intervention procedure) it is better represented by a video, but when the important thing in the task is to recall something or compare a visual finding with an standard it is better supported by a still image .

Factors	Video	Still Image
Cognitive Process	Perception action with movement	Memory
Action in tasks	Operative actions	Informative actions and information transmission
Skills	Device	Manual

Temporal constrains	*Limited	
Complexity	Complex	Simple

Table 4. Heuristic to support non-verbal content by factor in task

Temporal constrains limits the use of video to a minimum, since wating for a video to play in an emergent or urgent matter could be life-threatening.

7.2.1.3 Testing

Final versions of each GuideView were tested systematically by 2 physicians looking for inconsistencies, missing content and proper workflow. Systematic approach consisted in visiting each node in a guideview creating standardized pathways. Evaluation was emphasized in necessary pathways to solve selected cases. In cases were changes were required, a new Delphi iteration was used.

7.2.2 **Experimental guideviews**

For the purpose of the investigation guidelines of interventions for all cases were grouped in three guideviews, two for pediatrics and one for adults. Pediatrics guideviews required was split into two following the structure of the original guideline, which provides two guidelines based on the age of the patient, . Since most of adult protocol have a common group of steps then a single guideview was possible to implement. Pediatrics guideviews were IMCI for 1 week to 2 months old children (Figure 6), and IMCI for 2 months to 5 years old children(Figure 7). Adult guideview reviewed adults basic life support in trauma and no trauma conditions (Figure 8). Adult basic life support included life-saving procedures allowed to be performed by CHWs in Colombia.

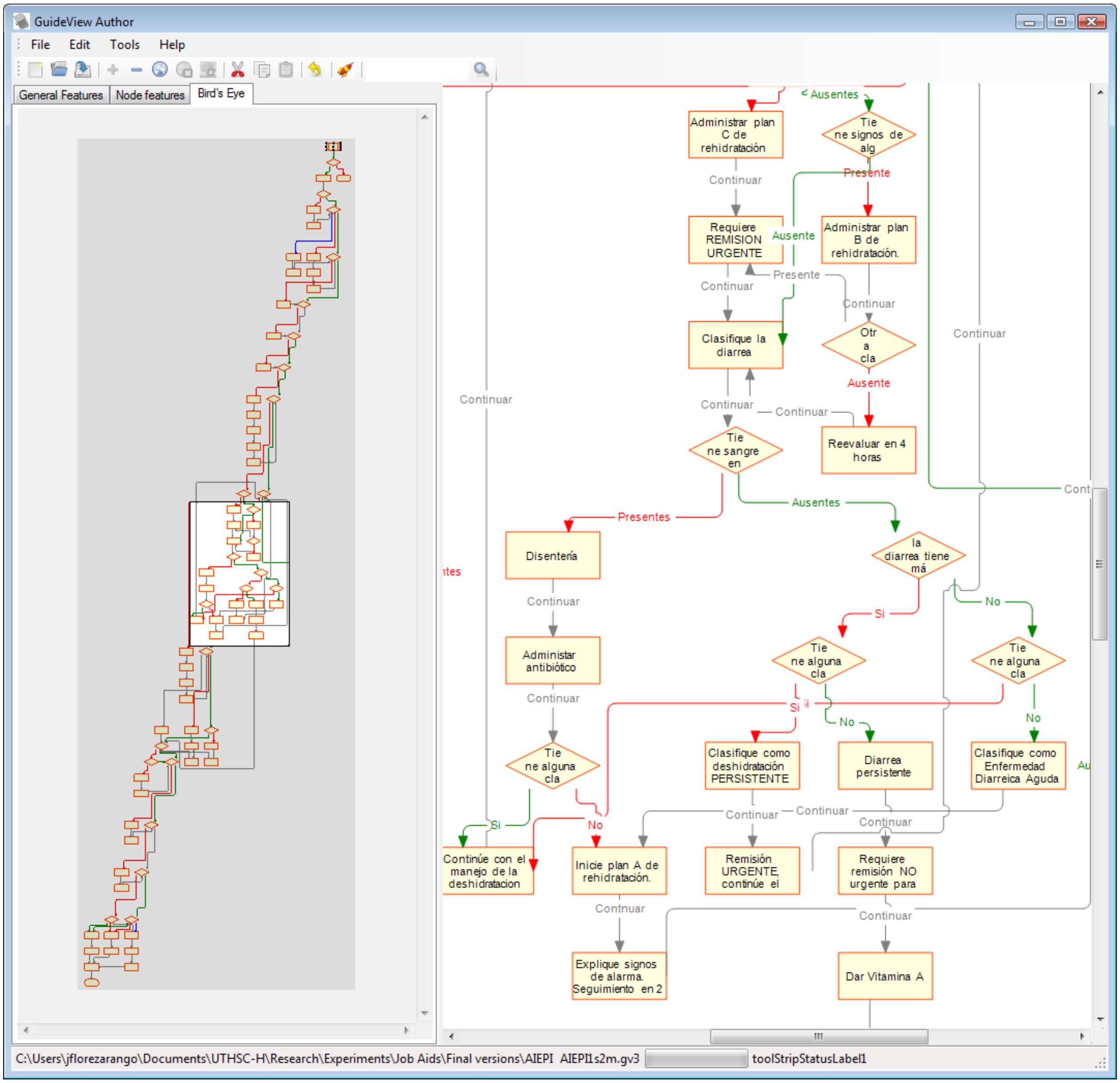


Figure 6. Detail of experimental guideview IMCI 1 week to 2 months in GuideView Author

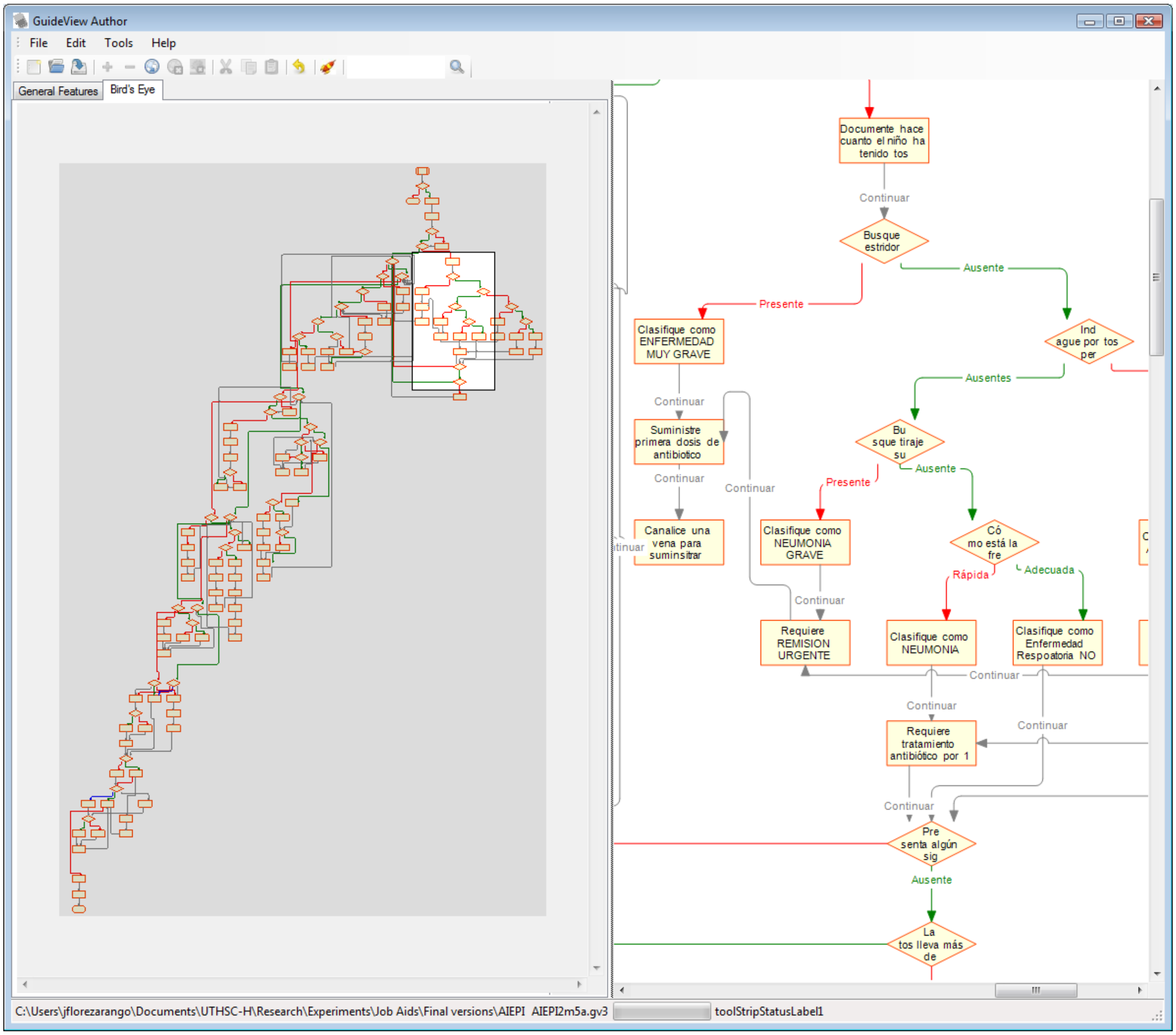


Figure 7. Detail of experimental guideview 2 months to 5 years old in GuideView Author

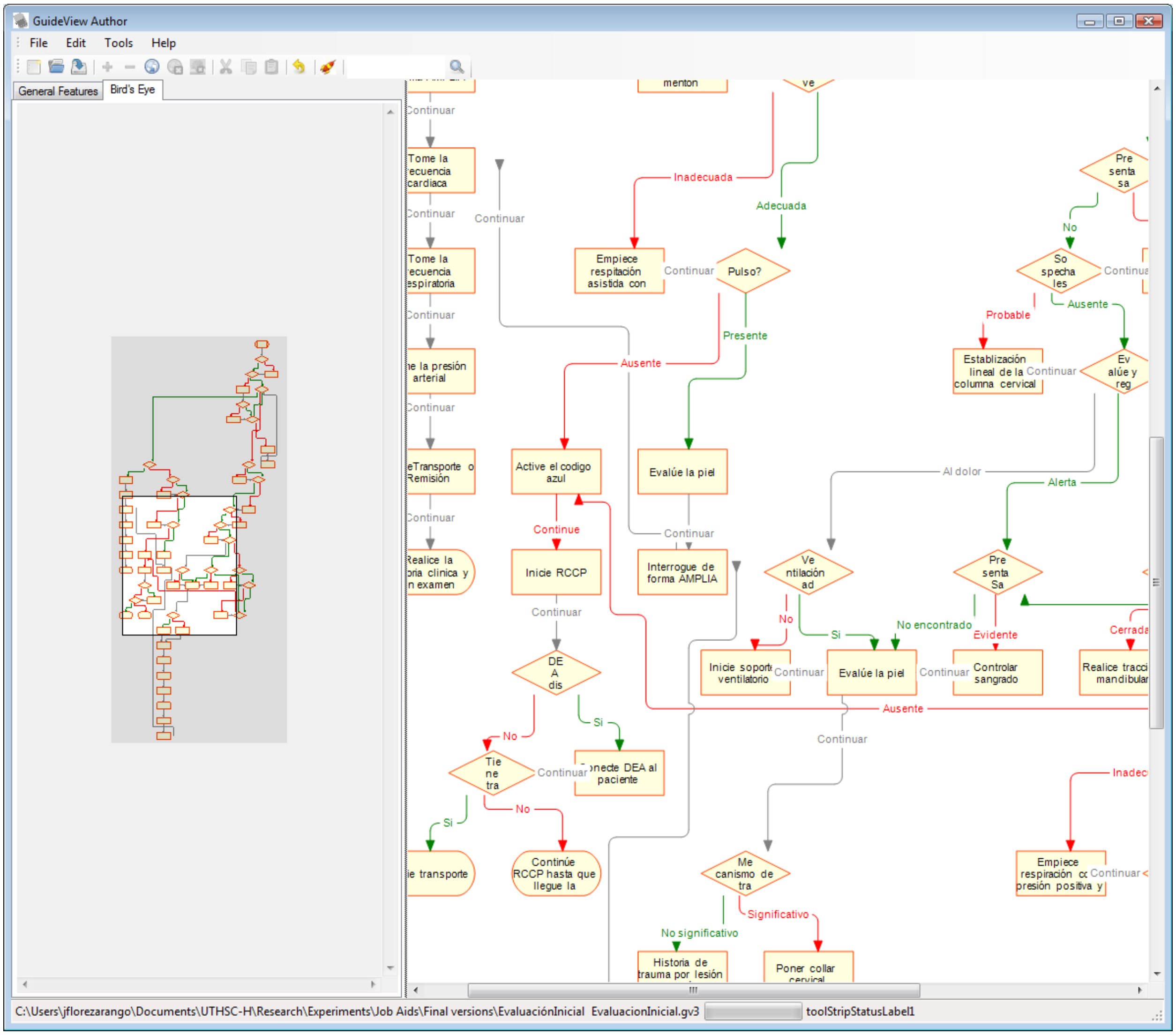


Figure 8. Detail guideview for Adult Initial Assesment in GuideView Author

It is important to note that there are granularity differences between pediatrics and adult guideviews. Pediatric guideviews contains more nodes, since tasks are more granular than in the adult guideview. Adult guideview contains more plans than actions, as consequence is less specific.

GuideViews were presented to participants using CGV running in a HTC TyTNII cellphone. This phone has a touch screen with a resolution of 320x240 pixels (See Figure 9). Audio was provided by a Jabra BT8040 Bluetooth earphone.

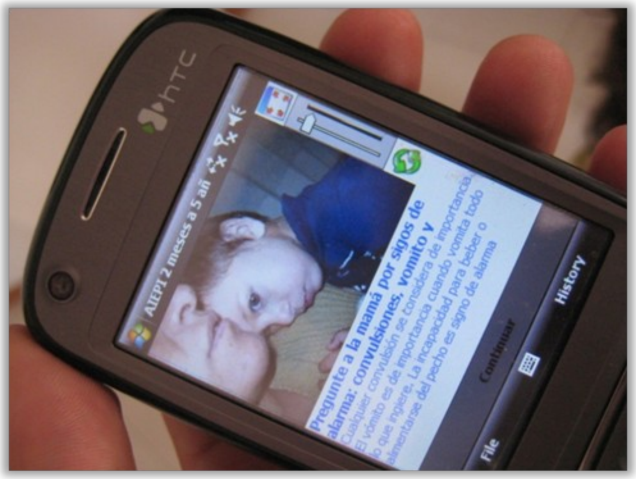


Figure 9. CGV running in a HTC cellphone

Experimental subjects training in using CGV was minimal. Training consisted in hands on explanation, with duration less than 5 minutes. During training subjects were instructed how to turn on the cellphone, use touch screen, recover cellphone from sleep mode, select and navigate a guideview, and undo errors.

8 Location

This Study was conducted at the Simulation Center and Clinical Skills lab (CS) at the Faculty of Medicina – Universidad de Antioquia (UdeA). Colombia is a tropical country located in north western South America, has a population close to 45 million, distributed in mestizo (white+amerindian) 58%, white 20%, mulatto (white+black) 14%, black 4%, zambo (black + ameridian) 3% and amerindian 1%. Most (74%) of the population is concentrated in urban areas. Colombia has an estimated GDP per capita of US$9000 for 2009. Literacy level is close to 93% in population of age 15 and over (Central Intelligence Agency, 2009; World Bank, 2009). Given the population income and distribution CHWs play an important role in health care in Colombia. UdeA was founded in 1803 it is a public University, that has research and teaching as the main foci, integrated closely with community outreach activities. This is the premier research university in Colombia, accounting for 60% of the biomedical research output of Colombia. There is a close relationship existing between the University of Texas and the CS at UdeA that facilitated to select the site as performance location.

Figure 10. Standarzied Examination Room with a Simman

The CS is equipped with 3 Simman® (adult size) and 1 SimBaby® (infant size) human patient simulators (Laerdal Medical, http://www.laerdal.com) that can simulate a wide range of medical disorders and physiological conditions. One (1) Simman and once (1) Simbaby were programmed with selected cases(see section 10.3). Cases were presented to subjects in an standardized exam room. The standardized room has a microphone and a videocamera (Figure 10) that allows recording all activity performed inside the room. Additionally a telephone, a thermometer, a stethoscope, an adult sphygmomanometer, cotton swabs, cotton balls, gauze, bandage, scissors, adhesive tape, immobilization devices, simulated medication among other elements were available to all subjects.

9 Human Subjects

Subjects were selected from a group of personnel attending SENA (Colombian National Service for Learing) for certification or re-certification of competences as public health assistannt, nursing assistant or technician in prehospital care. These subjects were considered CHWs given their competences.

Pre-hospital care technician is somebody with 500 up to 2500 hours of certified training. The maximum of 2500 hours is equivalent to a technician degree in Colombia. The emphasis of pre-hospital training is to develop competences in health promotion and disease prevention in case of emergency and disaster, competences to support and manage healthcare related activities derived from a catastrophic event, and competences to support the design of institutional emergency plans that articulates with the National System for Disaster Preparedness and Management in Colombia (Zuleta, 2009). The training as pre hospital care provider has emphasis in trauma life support. Pre hospital care technicians are located all over the country and include policemen that provide care in remote locations.

Nursing assistant is a subject with 2640 hours of certified training. The training as nursing assistant permits individuals to develop skills to provide health care in timely, personalized, humanized, integral, uninterrupted manner and based on standardized procedures within a health care institution (Zuleta, 2009). It is important to notice that they are trained to work unsupervised under experts instructions or protocols. Nursing assistants are used in Colombia as a primary contact with communities in health posts.

Public Health Assistant training is 2640 hours. The training develops skills to provide health care in timely, personalized, humanized, integral, uninterrupted manner and based on standardized procedures within a community (Zuleta, 2009). This is a relatively new occupation and certification (less than 5 years). Newly certified personnel in public health assistance is expected to provide some of the services traditionally offered by lay workers, nursing assistants or pre hospital personnel.

The study considered for inclusion any subject that can be classified as community health worker within any of previous categories. This means they can provide health care services in a given community. Additional inclusion criteria included not been enrolled in a post-secondary human health degree as dentistry, nursing or medicine at the moment of the evaluation. Individuals were required to have previous experience using to human patient simulators. In order to be enrolled all subjects needed to sign the informed consent (See Appendix 1. Informed Consent)

A reference population of more than 750 community health workers was identified. Subjects were distributed in the whole territory of Colombia, mainly in the Department (State) of Antioquia. A by-convenience systemic approach was used to enroll subjects given the dispersion of the population and the limitations on transportation.

Study took advantage of the mandatory certification and re-certification process for Health Care providers in Colombia. The institution in charge of such certification is SENA. All participants of the certification or re-certifiaction program at SENA –Medellin were invited to participate in the study. Once they were in town for certification/recertification process they received an introduction of the study and were invited to participate.

This project was reviewed by the Univesidad de Antioquia's IRB and is considered exempt. IRB at University of Texas Health Sciences reviewed the project and provided IRB number HSC-SHIS-08-0100 the IRB expires 02/28/2010. Both IRBs required to provide informed concent to subjects.

An electronic informed consent was presented to all subjects invited to participate, this document was written in Spanish and approved by UT and UdeA . UdeA required to keep a printed. The approved informed consent is available at the end of this document as Appendix 1. Informed Consent. Informed consent specifies how participation in the study will not affect subjects work positions or be reflected in the certification provided by SENA. There were no direct risks associated with the participation of the subjects in the study. Also the informed consent states that no information will be available to people external to the research group and never will be used with aims to evaluate their academic or professional performance. Benefits are related with the collateral learning that can generate the exposure to the standardized cases and feedback provided in their performance.

There were no direct risks from this research since the intervention has no potential harming components. There may be some perceived risks for the participants that information can be used for academic or professional evaluation. In order to mitigate this risk all registries were anonymized by a system of subject identification by numbers; codes are kept locked at The University of Texas Health Science Center at Houston as well as in the Centro de Investigaciones Médicas (Medical research center) at

Universidad de Antioquia. Video tapes of experiment are available only to individuals directly involved in the present study for research purpose only, and will not be disclosed otherwise.

All subjects, which show up to the experimental session, were compensated for dedication to the study with gifts cards. The face value of the gifts cards was proportional to the length of participation in the study up to the equivalent in Colombian pesos of US$25.

10 Variables

This section presents different variables measured in the study. They are grouped in 7 topics of interest: Demographics, tasks, cases, performance, workload, acceptability and learning. Each topic contains detailed explanation of the variables used.

10.1 Demographics

This group of variables allows classification of subjects and intends to see if there is homogeneity in the sample. In addition demographic variables will constitute factors in detailed analysis. Variables of interest in demographics topic includes: Gender, age in years, level of education, type of training, and years of experience. Following are the definitions.

10.1.1 **Gender**

Gender was used with two categories: Male and Female. This variable was used to validate the randomization of subject between groups. Literature reports that CHWs are usually female.

10.1.2 **Age**

This is a continuous variable measured in years, allows comparing the study sample with others.

10.1.3 **Level of education**

This is a categorical variable to establish the level of education of the population under evaluation. It is important to establish the education, since it is commonly at low levels of education among CHWs.

10.1.4 **Type of Training**

This variable permits to classify the background of the subjects. It is a categorical variable to evaluate were pre-hospital care, nursing and public health. It is important to notice that this training is not equivalent to a title. These domains refer to the type of health services provided as explained in human subjects section.

10.1.5 **Years of experience**

Since experience can be related with change and adoption resistance of new technologies. This variable measures the amount of years that the subject has been providing health care services.

10.2 Clinical tasks

A clinical task is the atomic independent evaluation unit in this study. In order to evaluate performance in a standardized way it is necessary to identify clearly the task that will be evaluated. In general, a '**task** is part of a set of actions which accomplish a job, problem or assignment" (Wikipedia, 2008b) or as defined in this study of a case. For the purpose of this study a clinical task is then an action accomplished by a health care provider with the aim of solving a clinical case. Clinical Tasks can be determined by protocols. *"Clinical protocols* are agreed statements about a specific issue, with explicit steps based on clinical guidelines and/or organizational consensus. A protocol is not specific to a named patient" (Fox, Alabassi, Patkar, Rose, & Black, 2006).

A way to study tasks is called task analysis. A way to look the **task analysis** is as a method to analyze the mental and physical steps required to perform a task (Zhang & Butler, 2007). Even though there are different approaches to classify tasks and to perform tasks analysis, none of the identified methods are suitable to evaluate clinical task in relationship with sources of workload and performance in terms of errors and protocol compliance as expected in this particular study. Therefore, it was necessity to develop a system for task analysis and classification. Given the fact that tasks have multiple properties that can be related with dependent variables in the study, an ontological approach was selected to build the classification system. The aim of the classification system is to classify clinical tasks by properties that are in relationship with the sources of workload and error, including mental and physical steps of the traditional task analysis.

10.2.1 A clinical tasks ontology

It was necessary to construct an ontology for clinical tasks classification. The ontology intends to model clinical tasks required to be performed according with clinical protocols. The ontology is called the Clinical Task Ontology (CTO). The CTO is encoded in OWL/RDF using Protégé 3.3.1 and available at http://www.informaticaysalud.com/Ontology/ClinicalTasks.owl.

To create CTO two clinicians with experience with Protégé extracted and classified 447 clinical tasks from 6 published clinical guidelines for the pre-hospital domain in Colombia. Validation of the clinical task ontology was conducted by Methods of Task Assessment and Agreement Assessment, additionally logical consistence and individuals' classification was conducted in protégé 3.3.1 using pellet 1.5.1 as reasoner.

CTO takes into account aspects of theory of clinical guidelines as plans (Kumar, Smith, Pisanelli, Gangemi, & Stefanelli, 2006), a clinical tasks ontology for event oriented representation for collaborative activities (EROCA) (Pellegrin et al., 2007) and PROforma task model (Fox et al., 2006). The ontology contains 6 main axes that are related with the NASA workload sources (See Table 5).

Task property	**Workload Source**
Physical activity or skills required	Physical demands
Cognitive processes involved	Cognitive demands
Criticality	Frustration
Complexity	Effort
Temporal constrains	Temporal constrains
Goal	Performance

Table 5. Relationship among task's properties and sources of workload

Figure 11 shows a general overview of the Clinical Task Ontology. In this representation is possible to identify the different elements that constitute the ontology. Properties are represented by blue arrows, and classes by black boxes.

From the theory of clinical guidelines as plans, the new ontology consider that tasks can be divided in subtasks (*hasSubTask* and *isSubtaskOf*), tasks are defined by plans (in this case the protocol), task can be classified by clinical actions (Kumar et al., 2006). From EROCA, the new ontology is using the definition of type of actions within a task (*ActionInTask*) (Pellegrin et al., 2007) from the PROForma task model the new ontology of clinical task is using the type of task (*TypeOfTask*) as actions, plans or decisions and the goal ontology (*Goal*) (Fox et al., 2006). A task could be represented in every branch simultaneously.

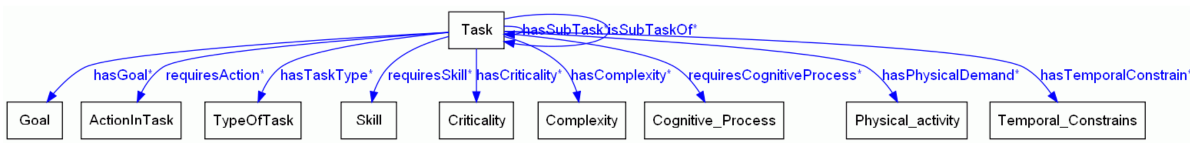


Figure 11. Main Axis for Clinical Tasks Ontology

Following each main axis of the ontology is explained in detail

10.2.1.1 ***Types of tasks***

Tasks can be classified in three major types: Plans, Decisions and Actions. Plans refer to a task that can contain a subset of task that is enunciated, but not granularly stated. An example of a plan can be "Evaluate Nutritional Status". This plan can be performed by measuring age, weight and height, then comparing data with standardized curves. Decision in the method to conduct previous plan is pretty much open to the evaluator. An action is basically a single task as "Measure rectal temperature" where method is specific. A decision is just a task where cognitive process needs to be done in order to choose from given options as "Is the patient responsive?".

10.2.1.2 ***Goals of tasks***

A goal talks about the ultimate intention of a task, it allows to evaluate if an error is present or not, because a task can be performed, but the goal not achieved, then there is an error in the task. These goals are derived from the goal ontology described in (Fox et al., 2006). In developing this ontology it was necessary to add the ActionGoals/Communicate/Referal class to be able to model the referral tasks. Details about classes of goals are provided in Figure 12.

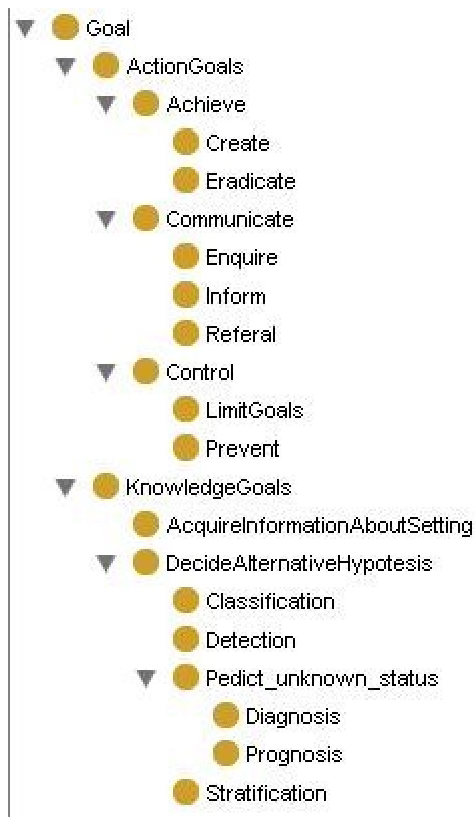


Figure 12. Fragment of Clinical Tasks Ontology showing the Goal Class

10.2.1.3 <u>Actions in Tasks</u>

ActionInTask provides information about the process that has been conducted in a given task and is different from the *Actions* as *TypeOfTask*. Every task requires at least one action. *Plans* and *Decisions* usually contain more than one action, while *Actions* requires only one. ActionsInTask were derived from the EORCA work (Pellegrin et al., 2007). Actions are close related with goals, and can be consistent in most of cases, e.g: most of the "Physical Examination" actions have as goal "Detection". A detailed of ActionsInTask is provided in Figure 13

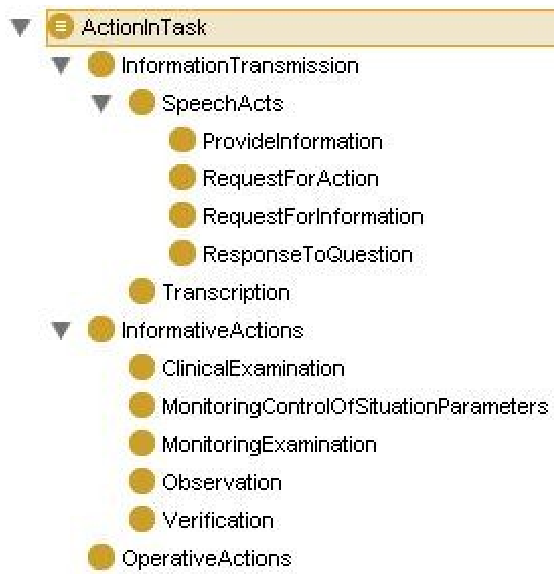


Figure 13. Fragment of Clinical Tasks Ontology showing Actions in Tasks

10.2.1.4 Skills

Skill is a new class developed during the task analysis. This class has two subclasses, but could be expanded once new knowledge is incorporated into the ontology. Subclasses are *KnowledgeOfDevice* and *ManualSkill*. The first class *KnowledgeOfDevice* refers to a task that requires any physical device to be accomplished, and as consequence the user need to be knowledgeable. For example in the task of "take quantitative temperature", a device is required to conduct such task, independent of the device in order to achieve the task user need to know how the device operates. The second subclass *ManualSkill* refers to those tasks where it is necessary to manipulate patient in any way, as cleaning a wound or conducting an immobilization.

10.2.1.5 Criticality of task

This class refers to the possible negative impact of the omission in performing the task, regardless the time constrain. Criticallities are Indispensable, Critical, No-Critical. A task with an*Indispensable* criticality is such task that if not done the patient will die or result in a complication, an example is "Evaluate airway permeability". A Critical Task is a task that if not performed properly has a consequence that a patient will be misdiagnosed or mistreated with a potential damage but not immediately life threatening, an example is "Look for rash" in a child with fever.

10.2.1.6 Complexity of Tasks

There are only two classes of complexities, those with multiple tasks (Complex) and those that are single tasks(Simple) by themselves.

10.2.1.7 Temporal Constrains of tasks

Temporal constrains permit to identify the level of urgency of a task. Three levels were established: Emergency, Urgency and Non-urgent. Main difference with criticality is the role of the time. Usually Emergency tasks are Indispensable or Critical.

Emergency: If task is not performed right now patient will die momentarily or will be permanently disabled. e.g: Extract a foreign body from the airway.

Urgent: If task is not performed within reasonable time patient will die or will be disabled. e.g: Start antibiotics in a child with pneumonia.

Non-urgent: Task can wait to be performed and condition will not be modified, unless other intervention changes the condition. e.g: Immobilization of spine under suspicious of spine injury.

10.2.2 **Summary of Clinical Tasks**

As shown clinical tasks analysis played an important role both in the design of the experiments, as in the data collection and evaluation of this dissertation. Task analysis derived in a newly created ontology of clinical tasks, which recycled some of the previous knowledge in this topic. Tasks are the atomic units in the evaluation of this work. Tasks were framed within clinical cases that are presented in the following section.

10.3 Cases

Case refers to a standardized scenario. In the study each case was materialized as a simulated patient. Each simulation has a predefined clinical condition and context, which is expected to be identified by subjects under evaluation. It is expected that subjects provide care according to predefined protocols or guidelines. Since the study required repeated measurements, a set of 30 cases (calculation of sample size is explained in section 11.2.) was derived and adapted by the investigator from an existing case bank at CS.

Special care was taken to have comparable cases of each type in each treatment and period. The balance was achieve in creating pairs of cases according with type of case (pediatrics, trauma, no trauma), priority of attention, number of tasks, time constrains, expected cognitive load and physical demand.

Cases were denoted from A to F. Cases in A and D series correspond to pediatric cases; cases in B and E series correspond to trauma cases; and cases in C and F series correspond to non-traumatic cases. Additionally each case has an index to identify it. Case A-1 is equivalent to Case D-1, while B3 is equivalent to E-3 and so on. Table 6. Summary of experimental cases presents a description of the cases and allows comparing them.

Code	Description	Priority	Task index	Time Pressure	Cognitive Load	Physical Demand	Case Index
A-1	Mastoiditis	3	1.92	2	2	1	0.661333
A-2	Malaria	2	1.5	1	2	1	0.5
A-3	Pneumonia	2	1.74	2	2	1	0.582667
A-4	Diarrhea no dehydration	1	1.56	1	2	1	0.437333
A-5	Common Cold	1	1.2	1	2	1	0.413333
B-1	Femur Fracture	3	1.86	3	1	3	0.790667
B-2	Burning	2	2.46	2	1	2	0.630667
B-3	Brain Injury + open fx + closed fx	2	3	2	1	3	0.733333
B-4	Dirty Wound	1	2.58	1	1	2	0.505333
B-5	Closed Fracture forearm	1	2.22	1	1	2	0.481333
C-1	Airway obstruction	3	1.8	3	2	3	0.853333
C-2	Stroke	2	1.86	1	3	1	0.590667

Code	Description	Priority	Task index	Time Pressure	Cognitive Load	Physical Demand	Case Index
C-3	Myocardial Infarction	2	2.04	2	3	1	0.669333
C-4	Hypertension	1	1.5	1	2	1	0.433333
C-5	Food poisoning	1	1.14	1	1	1	0.342667
D-1	Severe disease + malnutrition	3	1.74	2	2	1	0.649333
D-2	Supurated pharingoamigdalitis	2	1.44	1	2	1	0.496
D-3	Severe Pneumonia	3	1.92	3	2	1	0.728
D-4	Diarrhea with some dehydration	1	1.08	2	2	1	0.472
D-5	Healthy	1	0.78	1	2	1	0.385333
E-1	Poly-trauma	3	2.7	3	2	3	0.913333
E-2	Distal leg Amputation	2	2.34	2	2	2	0.689333
E-3	Hand amputation	2	2.64	2	1	2	0.642667
E-4	ligament elongation grade II	1	1.26	1	1	2	0.417333
E-5	Superficial wound	1	1.8	1	1	2	0.453333
F-1	Heart arrest	3	1.62	3	2	3	0.841333
F-2	Atrial fibrillation	2	2.28	2	3	1	0.685333
F-3	Brain transitory ischemia	2	1.68	1	3	1	0.578667
F-4	Cephalea	1	2.1	1	2	1	0.473333
F-5	Osteochondritis	1	1.56	1	2	1	0.437333

Table 6. Summary of experimental cases

Each of the variables priority, time pressure, cognitive load, physical demand and case index presented in Table 6 to compare cases are explained below. Measurements for these variables are derived from the task analysis of each case.

Priority: this reflect the need for an expert support scale reflects: 1= Low, 2 = Medium, 3 = High.

Task index: this is an index of number of tasks to be performed in each case. Range is 0 to 3. Task index was calculated as described in Equation 1.

$$Task\ Index_i = \frac{n_i * 3}{\max(n)}$$

Equation 1. Task index calculation for cases comparison

Time Pressure: this reflects the time factor that is pressuring the health care provider to solve patient condition before negative outcomes. 1 = Non urgent (Patient can wait), 2 = Urgent (Patient is at risk of death or disability), 3 = Emergent (Patient is at risk of death within minutes)

Cognitive Load: is used to represent the amount of cognitive processed are necessary to accomplish the task. Cases that require learning development and long term memory are classify as 3, Cases where calculation is important and working memory needed are classify as 2, others are classify as 1

Physical demand: If provider is required to use muscular strength, the whole body, or constantly switch positions is classified as 3. If provider requires using both hands to manipulate and solve patient problem this case is classified as 2. Any other condition is classified as 0.

Case Index: it is a summarization measurement of the previous observations. Case index ranges 0 to 1. It is calculated as explained in Equation 2

$$Case\ Index_i = \frac{Priority_i + Task\ Index_i + Time\ Pressure_i + Cognitive\ Load_i + Physical\ Demmand_i}{15}$$

Equation 2. Case index calculation

Previous values allow creating a radar graph for visual examination of the multidimensional relationship of the cases.

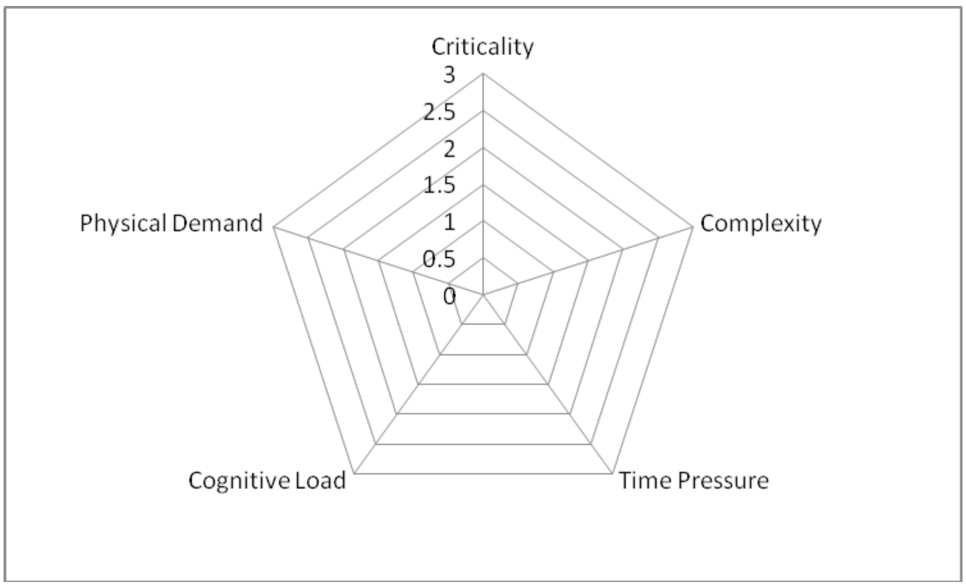
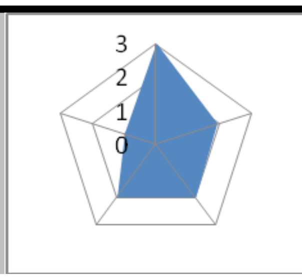
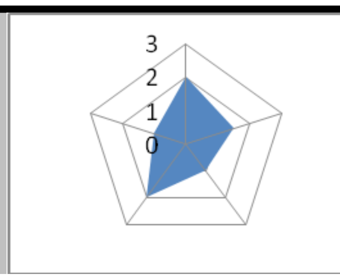
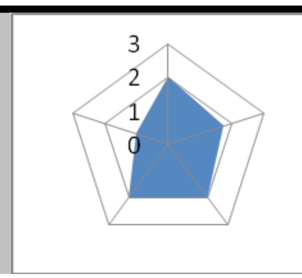
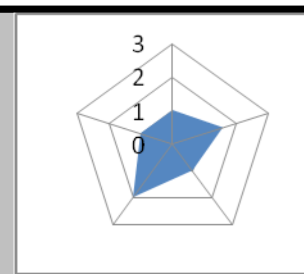


Figure 14. Radar visualization Axis for Cases

Block	Case	1	2	3	4
Pediatrics	A				

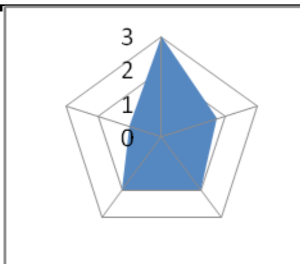
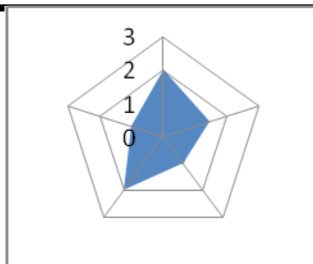
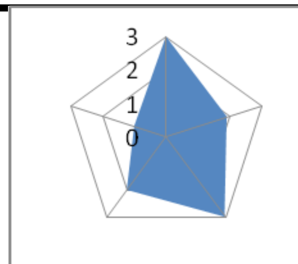
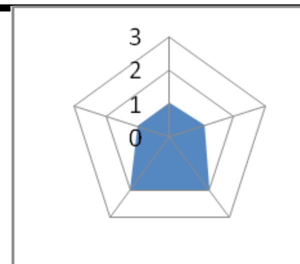
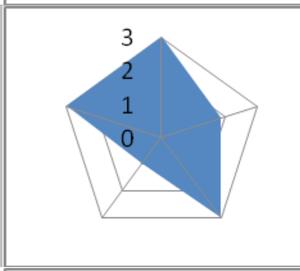
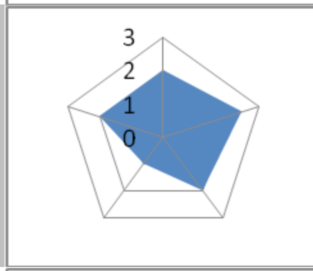
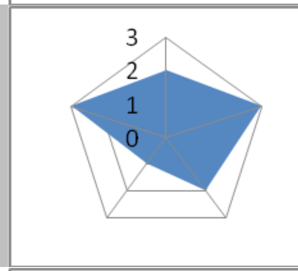
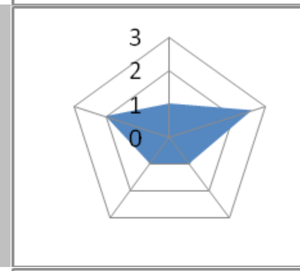
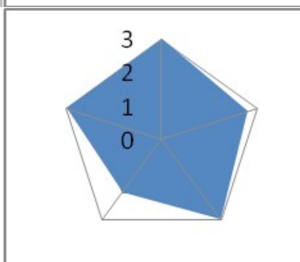
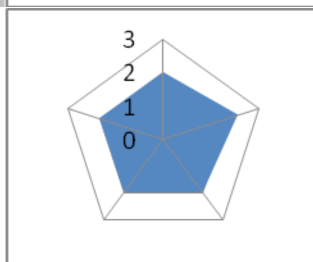
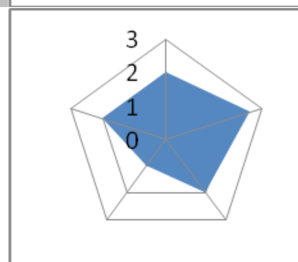
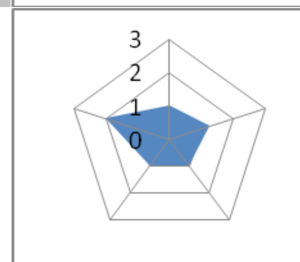
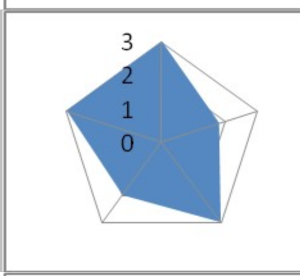
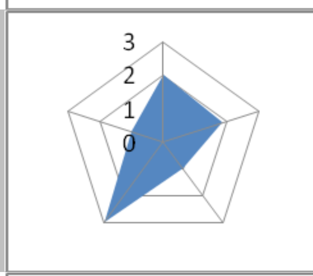
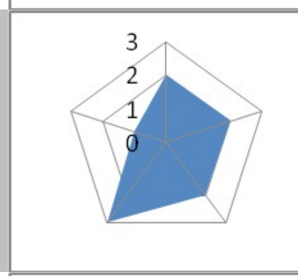
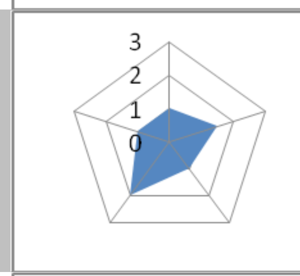
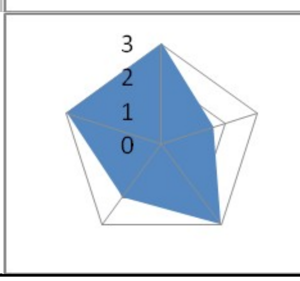
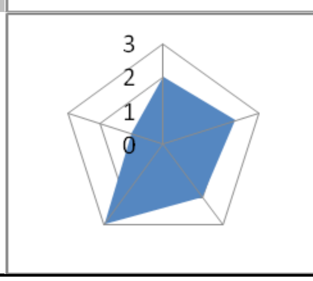
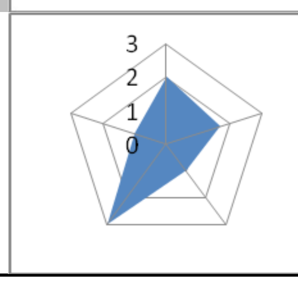
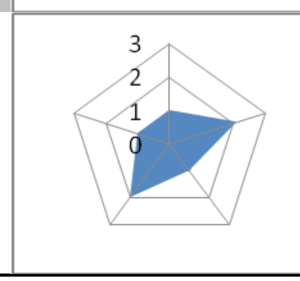
Block	Case	1	2	3	4
Trauma	D				
	B				
	E				
No Trauma	C				
	F				

Table 7. Visual comparison of cases using radar graphs

All cases were presented in a randomized order keeping the type of case grouping. A table with the randomization is available as Appendix 6. Cases Randomization Table.

10.4 Performance

Performance is defined as the **activity of an organizational unit** intended to **accomplish a desired result** (Wikipedia, 20(
Performance measurement is to **evaluate a subject** performance **given the goal** or desired result.
There are several ways to measure performance in health care, some of the commonly used performance measurements are outcomes, patient length of stay, mortality and morbidity.

In the context of this research the performance elements can be mapped as:

- Organizational unit is the healthcare provider.
- The organization is the healthcare process.
- And the goals, or desired results, are low time consumption, high protocol compliance and low error rate.

10.4.1 **Time**

Time of case resolution was evaluated In order to have an approach to the learning curve of the tool under evaluation. Time of case resolution was counted since the moment that evaluator told the subject that “this patient has arrived…” or “you arrived to the scene”. To finalize the case there were variety of outcomes: 1. in case that patient requires a referral the finishing time is when the phone call is performed. 2. If the patient does not require referral then the time is considered in the moment that subject presents the plan to the subject. And 3. If the subject is following a wrong path, then case is suspended when reaching a no return point (as in a failed resuscitation). Time is handled as a continuous variable measured in a resolution of seconds that allows computing minutes for data presentation.

10.4.2 **Errors**

Errors were defined as any intervention or omission which a subject performed on a simulated patient, which in the context of the protocol to be follow given the case could potentially harm a real patient,. In order to have a point of comparison the classification provided in Rowe's study will be used (S. Y. Rowe et al., 2007).

This classification includes:

- No error: A complete absence of error during the intervention.
- Minor error: An intervention or an omission that doesn't risk the life of patient, e.g., taking temperature with the hand when a thermometer is available or offering more liquids than necessary, by mouth, to a child with moderate dehydration.
- Major error: An intervention or omission that risks the life of the patient, e.g., misclassification of a severe patient that requires immediate attention as moderate.

Additionally to the Rowe's classification, a taxonomy for patient safety is available and will be used. This taxonomy has three levels. In the first level are two major error types related with the source of error: 1. Errors related with the process of health or 2. Errors related with skills and knowledge of the provider. The second level in composed by themes. The final level allows to have 35 descriptors (Makeham et al., 2008). This taxonomy was represented in OWL/RDF using protégé and integrated with the OCT to make possible automatic computation of type of error related with tasks.

If in a given task at least one observer reported an error it is counted as a task with errors. The direct count is corrected by the number of tasks performed in the case error/tasks. It is considered that a task was performed if at least one observer reported the task as done. In order to calculated error rate

$$error\ rate_i = \frac{count\ of\ errors\ in\ task_i}{times\ task_i\ was\ performed}$$

Equation 3. Error rate per task calculation

10.4.3 **Protocol compliance**

There are several approaches to measure protocol compliance. In order to avoid ambiguity, three approaches are defined and explained. Approaches used in this study include task compliance, protocol compliance and weighted protocol compliance.

10.4.3.1 Task compliance

This is a common way to look into compliance, often confound as protocol compliance. It accounts for compliance with X task across cases. This approach is useful to target evaluation of specific tasks independently. This unit does not summarize the global compliance. It is important to notice that when using more than one observer a task is considered performed if at least one observer reported the task as such.

$$Task\ Compliance_i = \frac{n_i}{m_i}\begin{cases} n_i = number\ of\ cases\ where\ task\ i\ was\ performed\ properly \\ m_i = number\ of\ cases\ with\ task\ i \end{cases}$$

Equation 4. Task Compliance

References use this approach, that is limited to task evaluation, but does not allow to evaluate overall compliance is a case(S. Y. Rowe et al., 2007).

10.4.3.2 General Protocol Compliance

The literature contains, implicitly, at least, a number of views about protocol compliance. In general protocol compliance is expressed as the level of agreement of one individual executing a protocol with the expected path determined by experts. In this study *Protocol compliance* is defined as the traditional approach (Equation 5) to protocol compliance (PC) is defined as the number of tasks performed in a path to achieve a goal that corresponds with the protocol path, over the total of number of tasks required to be performed in the path.

$$PC = \frac{tasks_{performed \in path}}{tasks_{path}}$$

Equation 5. Standard protocol compliance

Figure 15 will be used to illustrate an example. Let's say that the flowchart represents the protocol X. Each node in the flowchart represents a task. In the particular analyzed case the goal that the user need to achieve is the node w9; this means that from node w1 to node w9 it is expected that the user perform 6 tasks (nodes w1, w2, w3, w4, w8 and w9); suppose that in the particular case the user only visited 5 nodes (skipped w4). Applying the Equation 5 PC is 5/6 = 0.83.

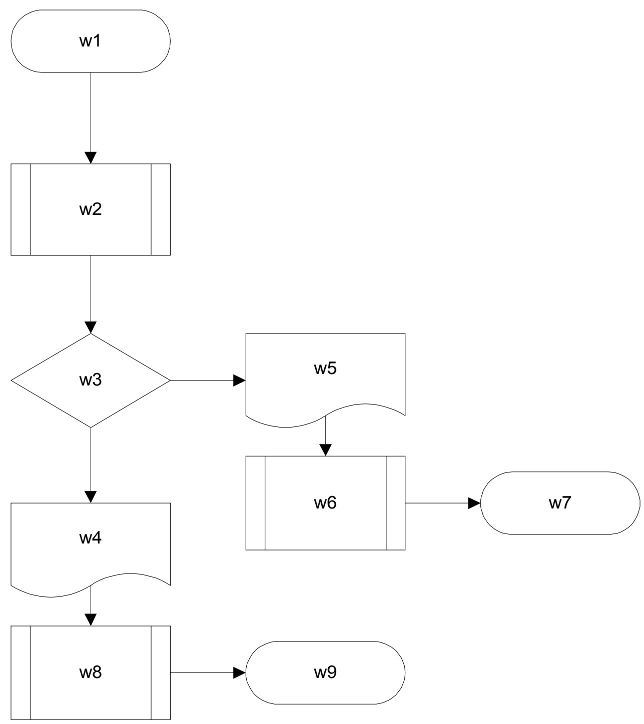


Figure 15. Representation of a protocol path

The general protocol compliance metric requires establishing meaning for the numerical results. Protocol compliance is considered good above 80% while fair is between 70% and 80% (S. Y. Rowe et al., 2007). In the example the protocol compliance is good.

10.5 Workload

Workload represents the amount of work that is necessary to accomplish a task. It is important to consider that changing the traditional workflow of a task modifies the associated workload. Changes in performance can appear as consequence of modifying the workload (D. A. Norman, 1991). Since this study is modifying the workflow with the inclusion of a new representation with the rich media job aids, it was anticipated that changes in performance could be related to changes in workload. Then it is an important measurement to be taken into account for the analysis.

Workload has been evaluated with different techniques; one well studied is the *NASA Task Load Index (NASA TLX)*. *NASA TLX* subjective evaluation instrument that measures workload in six dimensions: physical, mental, frustration, effort, temporal, and performance. NASATLX index allows for the quantification of the magnitude of the work load and establishes the sources of workload (Hart & Staveland, 1988).

Dimension	Question
Mental	How mentally demanding was the task?
Physical	How physically demanding was the task?
Effort	How hard did you have to work to accomplish your level of performance?
Time	How hurried or rushed was the pace of the task?
Performance	How successful were you in accomplishing what you were asked to do?
Frustration	How insecure, discouraged, irritated, stressed and annoyed were you?

Table 8. NASA TLX dimensions and questions

NASA TLX has two parts, the first part consist in an analog scale to grade the level of workload generated per each of the six sources. The second part has 15 pair-wise comparisons where individuals identify the source of workload most important for each pair. To calculate the NASA TLX it is necessary to count the number of times that a specific dimension was chosen in the pair-wise comparison. The counting for each dimension ranges from 0 to 5. There are 6 dimensions then there will be six counts, which are denoted by w_i in Equation 6. To obtain the value of workload for a single dimension the w_i value is multiply by an arbitrary measurement of the analog scales from the first part of the questionnaire (denoted as x_i). In this study all measurements from the scale where normalized to 1. Possible values for each dimension workload ranges then 0 to 5. To obtain the global NASA TLX all six values of workload are sum and divided by 15 (Equation 6), this value ranges from 0 to 1.

$$NASA\,TLX = \frac{\sum_{i=1}^{6} w_i * x_i}{15}$$

Equation 6. NASA TLX calculation for a single case

NASA TLX instrument has been translated into Spanish an validated with health workers and workers of the electronic industry (González Gutierrez, Moreno Jiménez, Garrosa Hernández, & López López, 2005; González Muñoz & Gutierrez Martínez, 2006). These two populations are comparable with CHWs in the domain of work (health workers) and the level of scholarship (workers of the electronic industry). An electronic version of the Spanish NASA TLX has created for this study, which was presented to each subject after each case that they faced, a paper version was available as well as a contingence(Appendix 2. Spanish NASA TLX form).

10.6 User acceptability and usability

A questionnaire of usability derived from the Post Study System Usability Questionnaire from IBM computer usability satisfaction questionnaires was used (Lewis, 1993). This questionnaire was used in previous studies with GuideView (M. S. Iyengar et al., 2005) and translated into Spanish. Questionnaire has a total of 17 questions, 12 related with usability, appeal, utility and intention of adoption. Remaining 5 questions are about media preferences. All questions used 5-options likert scales that allowed to group variables, a proportion analysis of the answers is used to produce results.

10.7 Summary of Variables

Name	Type	Data Type	Values
Treatment	Independent	Nominal	1= Use GV (Experimental) 2=No Use GV (Traditional)
Group	Independent	Nominal	A = Period 1 Treatment 1 B= Period 1 Treatment 2 GV
Case	Independent	Nominal	A-1…F-5
Case Type	Independent	Nominal	Pediatric Trauma No Trauma
Task	Independent	Nominal	By Ontology
Experience	Independent	Scale	Years of Experience in years
Type of Training	Independent	Nominal	EMT Nursing Public Health Other
Time	Dependent	Scale	Time in seconds
Error	Dependent	Scale	Counting/Proportion
Compliance	Dependent	Scale	Counting/Proportion
Mental Workload	Dependent	Scale	0-5
Physical Workload	Dependent	Scale	0-5
Effort	Dependent	Scale	0-5
Time constraint	Dependent	Scale	0-5
Perceived Performance/Failure	Dependent	Scale	0-5
Frustration	Dependent	Scale	0-5
General Workload	Dependent	Scale	0-1

Table 9. Table Summary of Variables

11 Sample Size

11.1 Number of Subjects

Difference of means between groups was used to calculate number of subjects required for the study. Assumptions included a confidence of 95% (alpha 2-tails 0.05), a power of 0.9 (beta = 0.1), and a difference of 1 or more standard deviations between groups. Calculations required 25 subjects per group, for a total of 50 subjects.

11.2 Number of Cases

To establish differences in performance comparing Use of GuideView vs No Use of GV, sample size was calculated using proportion of errors differences. Assuming a significance level of 95% (alpha = 0.05), a beta error of 0.8, with expected of a minimum of 5% of proportion differences, and equal sample sizes for both groups, the experiment required at least 725 observations per group; this means that it is necessary 15 comparisons of cases per subject between treatments. In total, a subject was expected to solve 30 cases in the length of the study. In total 1500 cases (750 per treatment) was the targeted sample size.

12 Data Collection

A typical experimental session was composed by four parts. The first part had the intention of obtain informed consent from the subject and basic demographic data, in this moment subjects were scheduled for the experimental session. The second part took place at the beginning of an experimental session consisted in a briefing of the activity that will take place, in this part subjects received a basic instruction about the sequence of events, the use of CGV and the use of the data collection tools. The third and fourth parts were the interventions with each treatment; all subjects had a break between both interventions.

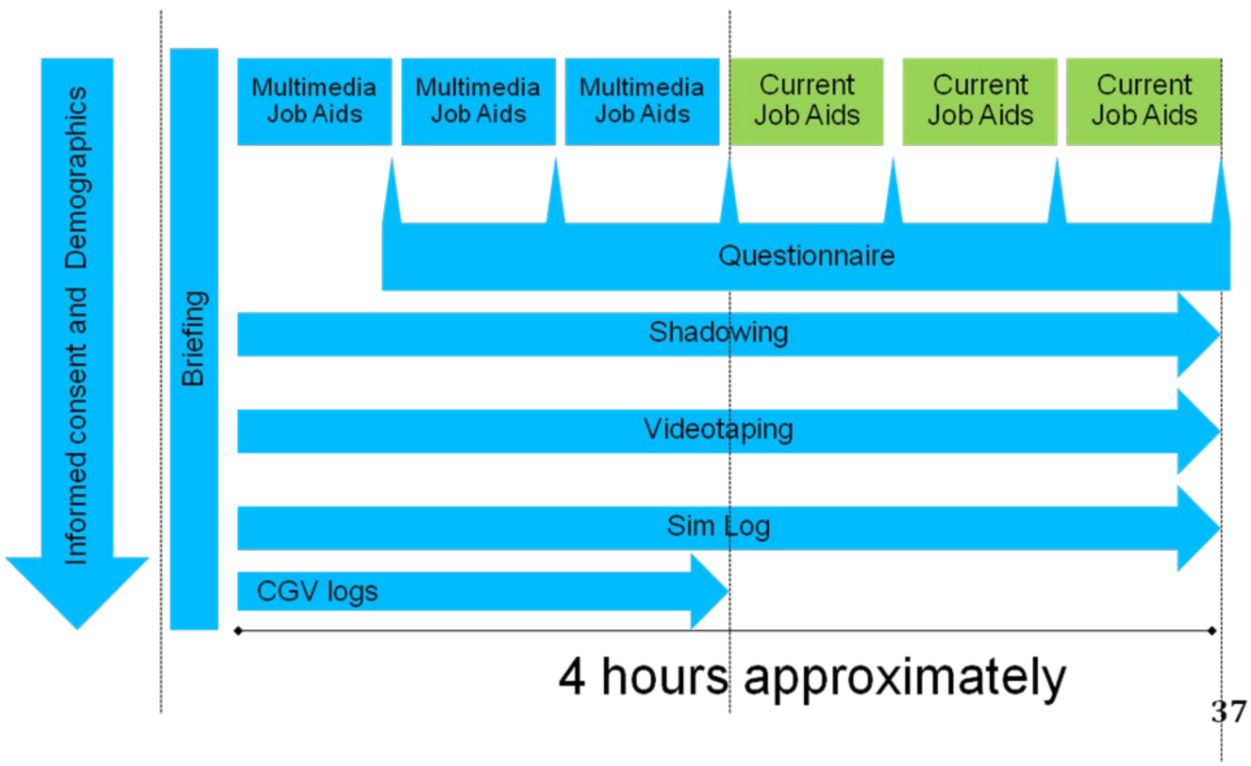


Figure 16. Typical experimental session

Figure 16 shows a typical session for an individual in Group "A" is used as example to summarize the data capture process. Total duration of an experimental session was between 4 and 10 hours. Variablity was provided by speed to answer cases, questionnaires and brakes. Data capture process implied observational and self reported data capture strategies. In the figure in noted the time when initial demographics collection, post treatment self-assessment questionnaires for NASATLX, the continuous videotaping for errors and protocol compliance, as well as the electronic registry of the CGV.

A web-based tool named GuideViewEval was written in PHP by the research team in order to handle subject's enrollment, scheduling, data collection and shadowing. Paper copies of all questionnaires were in place as backup in case that connectivity was lost. There was an open area where subjects could read informed consent, register as participants and schedule a session. GuideViewEval automatically randomized subjects once they were registered; an email confirmation was send with the experimental session scheduling and a system-generated login and password.

Login and password were used to identify subjects and evaluators. A proved identification allowed access to the data collection process, no data entry was allow if subject logins out of scheduling. Self reported data provided by subjects included a pre-test, the NASA TLX questionnaire and the usability questionnaire. Each form was presented by the system timely according with the state of progress in the experiment. Evaluators performed a data collection based on shadowing subject.

Shadowing was conducted by physicians trained in the use of data collection tool. Most of the cases at least two physicians where present during the evaluation with the aim of increasing sensitivity in error detection. Subjects were asked to think aloud when performing tasks to facilitated shadowing process.

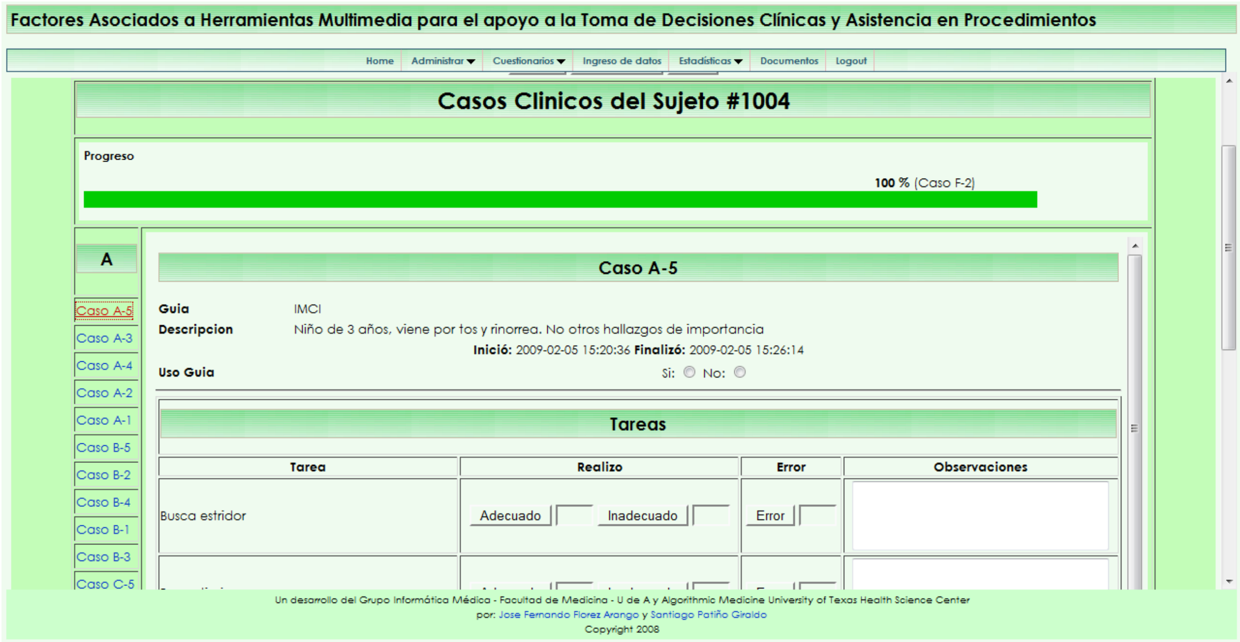


Figure 17. Typical Screen for Evaluator Data Entry

Shadowing allowed to actively record errors, performed tasks, and the quality of the action. If a task was not recorded by any observer it is considered not performed. When any protocol deviations were detected the tasks performed by CHWs in those cases could be mapped into the ontology of tasks, and sources of deviation could be identified. Figure 17 presents a screen shot for the shadowing tool used by the evaluators. Figure 18 shows the typical setup for the data collection, including control of simulation, direct visualization of simulation, data entry forms and video recording controls.

Figure 18. General data collection setup

Shadowing was complemented by videotaping and the use of the human patient simulator records. Video tape was obtained with the closed television circuit available at CS.

Results

Data was stored in a MySQL database. Proper queries were written to allow data extraction and analysis. SPSS 17.0 licensed to The University of Texas was used as the main tool for data analysis. An active search for data errors was performed based on queries to the study data base looking for missing, repeated and inconsistent data. Manual review was conducted for suspicious data and video recordings, simulation records, and additional information used to clarify situations. All cases where consolidated information was not possible where dropped from the analysis.

13 Subjects

A total of 67 subjects responded to the invitation and signed the informed consent. Seven of them were participated in a pilot study to review and adjust data collection process. Pilot study results are not reported in this document. Three (3) subjects did not provide enough contact information to schedule their session in the study. Seven (7) subjects did not show to the scheduled session. A total of 50 subjects showed to scheduled sessions. A total of 44 subjects completed 30 cases. Additional 6 completed partially the study. Figure 19 shows a detailed distribution of subject during the study.

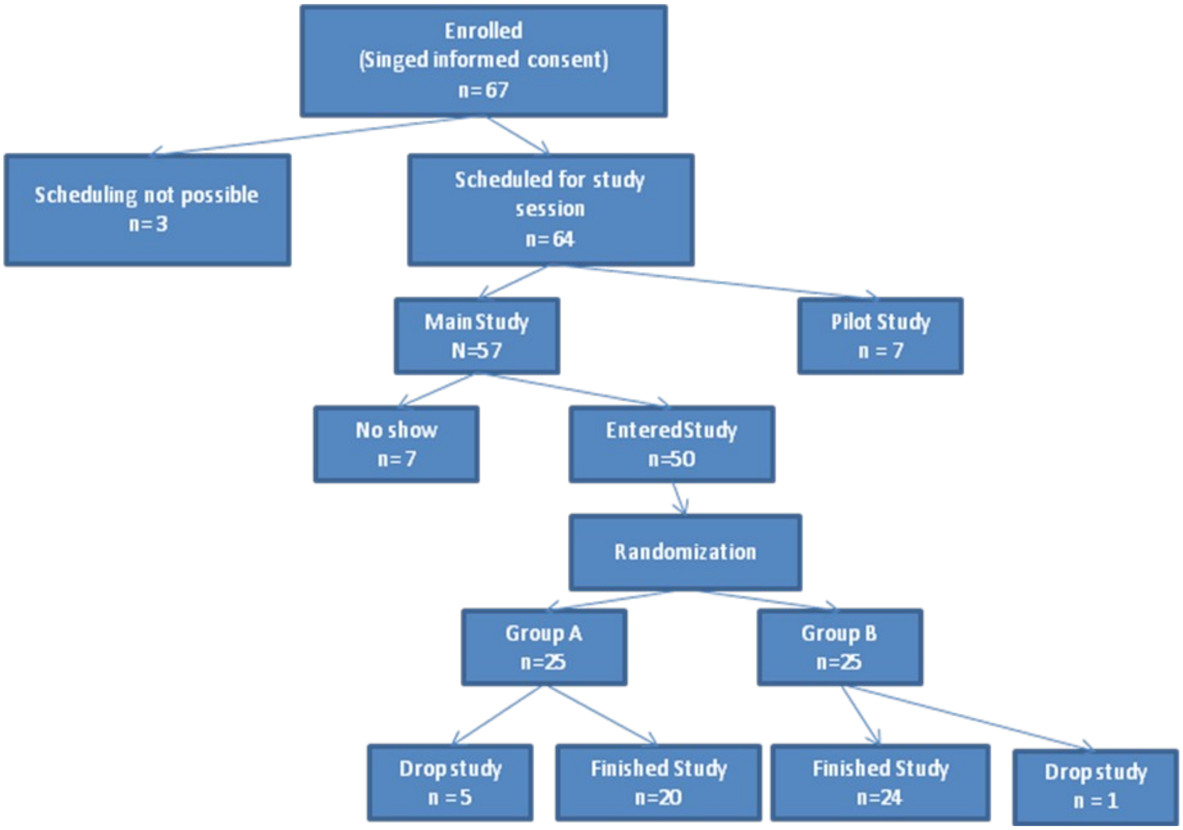


Figure 19. Enrollment and subject retention tree

Results regarding subjects include all (50) randomized subjects. The gender distribution in the study is predominately male (58%). This apparent contradictory observation can be explained by the convenience sampling. By change during the enrolling process a group of policemen was enroll in the process of certification, and sampling choose many of them. None of the participants had a level of education below high school. This is an important difference in Colombia given that it is mandatory to complete high school before applying to a training/certification program with SENA. Even though a large amount (44%) of the subject did not have a previous formal certification as CHWs they qualify as such by training or experience. It is important that almost half (48%) of the CHWs reported to be unemployed, regardless of this condition they serve their communities. Subjects were predominantly pre-hospital care technicians (38%). Subjects with training different to expected categories included an engineer, a physicist and a psychotherapist whom are volunteers as community health workers with more than 10 years of experience. In general the subjects group can be in the range of novice since mean is around 3.24 years of experience.

		Group A		Group B			Total	
Variable	Value	n	% within group	n	% within group	Sig.	n	%
Gender	Male	14	56	15	60	0.5	29	58
	Female	11	44	10	40		21	42
Level of Education	High School	10	40	12	48	0.256	22	44
	Certified Technician	6	24	7	28		13	26
	Undergrad Student	5	20	3	12		8	16
	Other		0		0			0
	NR	4	16	3	12		7	14
Training	Pre Hospital Tech	12	48	7	28	0.221	19	38
	Nursing Asst	6	24	6	24		12	24
	Public Health Asst	3	12	7	28		10	20
	Other	3	12	1	4		4	8
	NR	1	4	4	16		5	10
Employment	Yes	11	44	9	36	0.832	20	40
	No	11	44	13	52		24	48
	NR	3	12	3	12		6	12
Zone of Service	Urban	11	44	6	24	0.501	17	34
	Rural	5	20	7	28		12	24
	Both	1	4	2	8		3	6
	NR	8	32	10	40		18	36
		Value (DS)		Value(DS)				
Age in years	Mean	26 (6.40)		27.6 (10.56)		0.52		
	Range							
Years of experience	mean	2.92(5.330)		3.56(5.71)		0.684	3.24(5.4)	
	Range							

Table 10. Subjects Demographic Data

Table 10. presents a summary of the subjects distributed by groups. It is important to notice that there are no statistical differences between experimental groups in any of the demographic variables. This makes both groups comparable and statistically homogeneous.

14 Performance Measurements

Three performance measurements were used: time to complete a case, error rate and protocol adherence. All variables were analyzed at the case level; additional analysis in the task level was conducted for errors and protocol adherence.

A total of 17 cases out of 1401 complete evaluations needed at least one correction. Cases with data errors were dropped. All missing data were related with missing data or wrong timestamps (Start time and End time of cases) in the data base of evaluation windows. The error rate is equal to 0.012134 or 1.2%. When the error is evaluated at the variable level 22 data point corrections out of 2802 (2 variables affected in each evaluated case) were necessary, this rate is 0.007852 or 0.7%. This error was easily corrected in most cases using information provided by video recordings and simulation registries. Errors in time stamping are related with lack of attention by evaluator or lost of connection at the moment of data storage. Time stamping cases was a manual procedure. An additional factor that could be related is evaluator fatigue and overload.

Number of Evaluators per case	Number of cases
1	594
2	704
3	86
4	10

Table 11. Number of Evaluators per Case

After dropping cases with errors a total of 1394 distinct cases were included for analysis. Each case consists of a certain number of clinical tasks (Table 6). As seen in Table 11, the subjects' performance was evaluated task by task by one or more physicians. Thus 45,685 distinct task evaluations were performed by evaluators. Scheduling and time availability was a limitation for evaluators. In addition, shadowing subjects created an added burden on evaluators. In total eight hundred cases (57.4%) were evaluated by at least two physicians.

14.1 Case based Performance Model

A mixed linear model of the form on Equation 7 was used for all case-based analysis (SPSS Inc., 2005). This model is able to take into account repeated measurements and multiple factors at the same time. Methods in (West, Welch, & Galecki, 2007) were used for particular modeling decisions..

$$X_{i,j} = (1-\delta)\mu_0 + \delta\mu_1 + (1-\delta)\beta_{o,j} + \delta\beta_{1,j} + \varepsilon_{i,j} \qquad \delta = \begin{cases} 0 \; if \;\; GV \\ 1 \; if \; No \; GV \end{cases}$$

Equation 7. Model of data distribution

Additional models explored included General Linear Model with repeated measures but this resulted in too many observations being dropped due to missing data and other model assumptions. On the other hand (West et al., 2007), the Mixed Linear model approach that was used can handle these issues better and maximize data utilization.

14.2 Time to complete a case

Time is presented in minutes; a case time was established from the moment when a patient was presented to a subject until the time that an outcome decision was made. More details are available at section10.4.1. Figure 20 shows additionally a cyclic peak every 5 cases. These peaks are reflecting the change in the type of case that subject is exposed during the evaluation. It is easy to see how first case of each set requires slightly more time that remaining 4 to be completed this reflect the effect of learning to navigate in a new guideview.

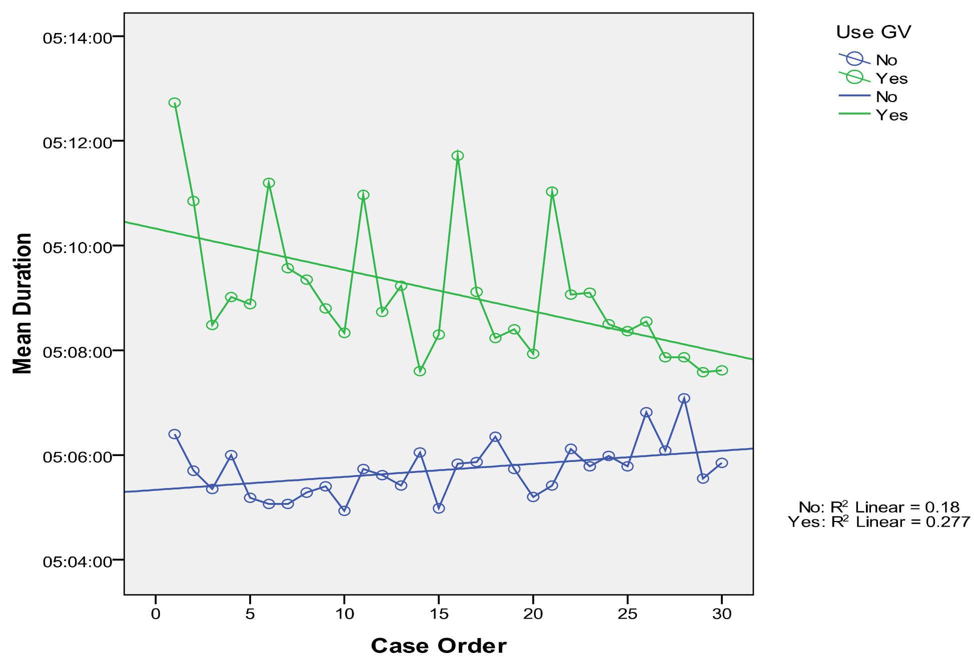


Figure 20. Time to complete a case by case order

The model used in SPSS is like:

```
MIXED duration BY treatment group WITH order   /CRITERIA=CIN(95) MXITER(100)
MXSTEP(5) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
LCONVERGE(0, ABSOLUTE) PCONVERG    E(0.000001, ABSOLUTE)   /FIXED=treatment
group order treatment*orden group*order | SSTYPE(3)   /METHOD=REML
/PRINT=SOLUTION   /RANDOM=order | SUBJECT(idsubject) COVTYPE(VC).
```

	Intercept	Group	Treatment	Case Order	Treatment*Order	Group*Order	Error
Time Mean (lower,upper)	**694,35 (625,59,763,11)**	-69.16 (-140.46,2.14)	**-377.24 (-448.55,-305.95)**	**-7.72 (-11.219,-4.22)**	**10.16 (6.20,14.13)**	1.72 (-2.99,6.43)	0.000264
P	**0**	0.57	**0**	**0**	**0**	0.474	

Table 12. Coefficients for Case Duration Model

Coefficients in Table 12 are expressed in seconds. Differences in mean time to complete (duration) a case between treatments were statistically different (p<0.001). Subjects expend in average 6.28 minutes (377.24 seconds) more in each case when treatment use of CGV is in place. Same treatment presented a learning effect that is absent in the control treatment represented as the reduction in the duration of cases when order is taken into account (p <0.001), there is a gain of 10 seconds with each subsequent case. There is no evidence that subjects group had an effect in the change in the learning effect.

14.3 Errors

Errors were evaluated at the case level and at the task level. In both situations cases with treatment use of CGV presented a positive impact in the performance (reduction of errors). The treatment with CGV reduces errors in 33.15% (p=0.001). There are differences in the type of training (p=0.004) and the order in which cases are presented (p= 0.020). Additionally there are interactions between treatment and order (p=0.037), type of training and treatment (p= 0.001) , and type of training and experience (p < 0.001). A more detailed explanation of factors interaction is presented in

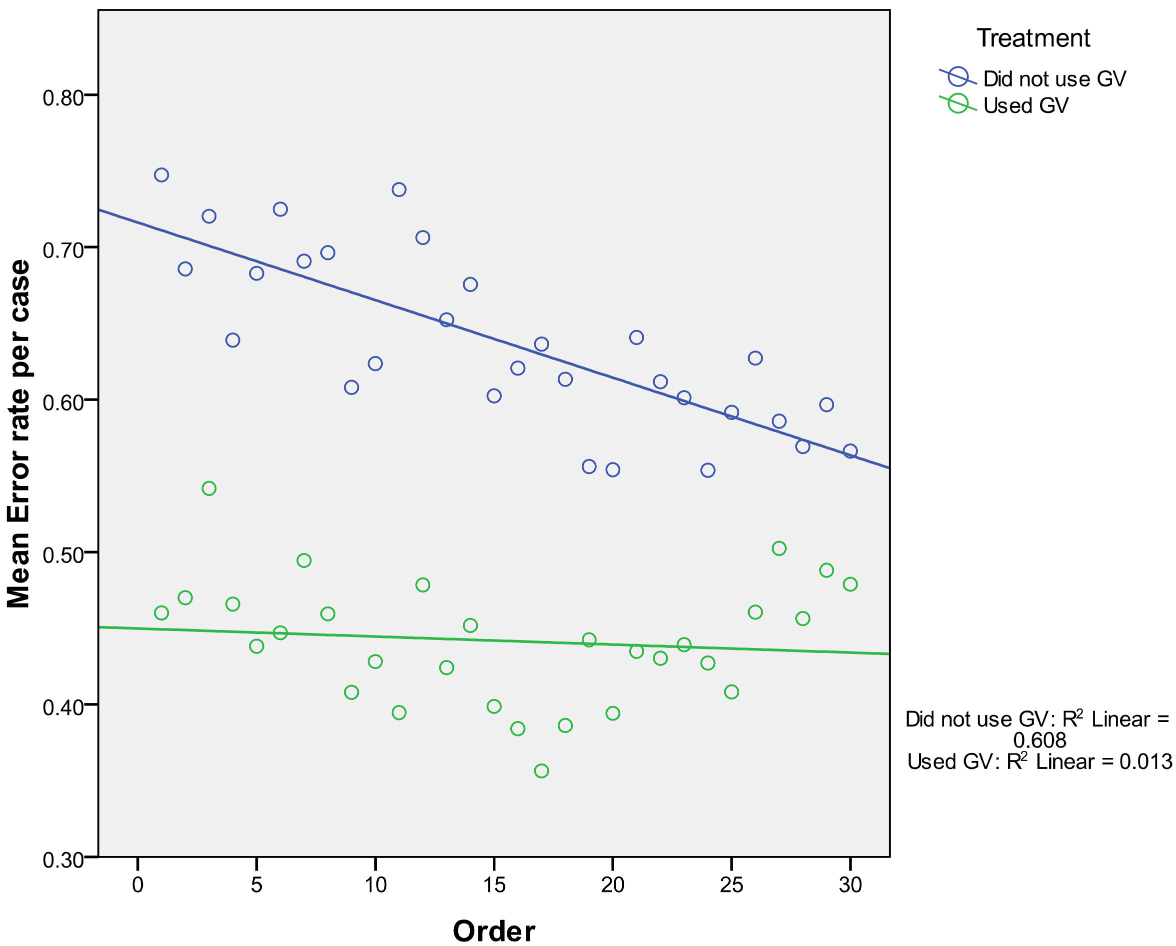


Figure 21. Error rate per case over time per treatment

Extrapolating the upper curve discloses that to achieve the misperfromance rate of 0.45 using GV, subjects would have to perform about 60 cases without GV.

	Intercept	Group	Treatment	Case Order	Type of training	Treatment*Order	Treatment*experience	Group* Order	Training *treatment	Training *experience	Error
P	**0**	0.122	**0.001**	**0.020**	**0.004**	**0.037**	0.623	0.98	0.001	0.000	

Table 13. Significance of factors in mixed linear model for error rate per case

For the case based analysis a mixed linear model was used. The SPSS syntax is:

```
MIXED errorrate BY group training treatment WITH yearsofexperience order
/CRITERIA=CIN(95) MXITER(100) MXSTEP(5) SCORING(1) SINGULAR(0.000000000001)
HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERG    E(0.000001,
ABSOLUTE)   /FIXED=group training treatment yearsofexperience order
group*order training*treatment training*yearsofexperience
treatment*yearsofexperience     treatment*order | SSTYPE(3)
/METHOD=REML   /PRINT=SOLUTION   /RANDOM=order | SUBJECT(subjectid)
COVTYPE(VC).
```

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	.339581	.066501	645.168	5.106	.000	.208995	.470166
[group=A]	.061189	.039516	1219.979	1.548	.122	-.016337	.138715
[group=B]	0[a]	0	.	.	.	.	.
Pre-hospital technician	.008133	.057832	395.806	.141	.888	-.105563	.121830
Nursing assitants	.101987	.058060	426.732	1.757	.080	-.012133	.216106
Public health assitants	.050828	.059717	377.174	.851	.395	-.066592	.168249
Other	0[a]	0	.	.	.	.	.
[Treatment=No CGV]	.351540	.108568	381.519	3.238	.001	.138073	.565006
[Treatment=CGV]	0[a]	0	.	.	.	.	.
yearsofexperience	-.000834	.003312	455.588	-.252	.801	-.007342	.005674
order	.002228	.001981	316.556	1.125	.262	-.001669	.006126
[group=A] * order	-.004348	.002622	272.356	-1.658	.098	-.009510	.000815
[group=B] * order	0[a]	0	.	.	.	.	.
Pre-hospital technician * [Treatment=No CGV]	-.066722	.095986	325.828	-.695	.487	-.255553	.122109
Pre-hospital technician * [Treatment=CGV]	0[a]	0	.	.	.	.	.
Nursing assitants * [Treatment=No CGV]	-.087543	.100524	338.956	-.871	.384	-.285272	.110186
Nursing assitants * [Treatment=CGV]	0[a]	0	.	.	.	.	.
Public health assitants * [Treatment=No CGV]	.026718	.099986	315.418	.267	.789	-.170006	.223443
Public health assitants * [Treatment=CGV]	0[a]	0	.	.	.	.	.
Other * [Treatment=No CGV]	0[a]	0	.	.	.	.	.
Other * [Treatment=CGV]	0[a]	0	.	.	.	.	.
Pre-hospital technician * yearsofexperience	.037419	.009700	450.498	3.858	.000	.018357	.056482
Nursing assitants * yearsofexperience	-.009539	.006481	311.797	-1.472	.142	-.022292	.003214
Public health assitants * yearsofexperience	-.023464	.010026	187.802	-2.340	.020	-.043242	-.003686
Other * yearsofexperience	0[a]	0	.	.	.	.	.
[Treatment=No CGV] * yearsofexperience	-.002523	.005123	352.517	-.492	.623	-.012600	.007553
[Treatment=CGV] * yearsofexperience	0[a]	0	.	.	.	.	.

Table 14. Estimates of fixed effects of mixed linear model for error rate by case

Parameter		Estimate	Std. Error
Residual		.027663	.001128
order [subject = idsujeto]	Variance	2.288384E-5	6.044687E-6

Table 15. Estimates of radom effects for mixed linear model for error rate by case

A total of 174 independent task where evaluated (see Appendix 5. Table of tasks errors). Not all tasks are present in all cases. Seventy five (43.1%) of tasks where considered critical tasks. Error rate was calculated per task as the times a tasks was performed with error divided by the number of times the task was evaluated (Equation 3). A paired t-test was used to calculate error rate differences between treatments. Mean of error rate for control treatment was 68.26%, while mean or error rate for CGV treatment was 49.98%, the mean difference is significantly 18.28% ($p<0.001$). An analysis of variance showed that treatment with CGV has significant ($p=0.011$) differences in reducing error rate in critical tasks (45.24%) compared with non critical tasks (53.56%). Distribution of error rate by criticality of task in the no CGV treatment did not show differences ($p=0.72$), this is an error rate of 70.73% for non critical tasks and 65.01% for critical tasks. A comparison of error rate per task between treatments shows that the use of CGV reduced errors significantly in 52.87% of tasks, while increased errors in 1.72%.

	N	Error rate with CGV	Error rate w/o CGV
Type of error		p < 0.001	p = 0.834
History Taking	41	42.01%	66.23%
Physical Examination	55	48.55%	67.53%
Classification	20	43.40%	72.73%
Treatment-intervention	38	60.79%	68.39%
Other	20	56.28%	69.71%

		p= 0.011	p=0.072
Critical	75	53.57%	65.01%
Non-critical	99	45.24%	70.73%
Total	174	49.98%	68.26%

Table 16. Summary of error rate by type of error and treatment

14.4 Protocol compliance

Protocol compliance was measured according with definition in section 5510.4.3.2. Figure 22 shows how protocol adherence with CGV was statistically different since beginning. Also learning effect is observed in the control treatment without reaching performance supported by CGV. Learning effect could be explained by the feedback provided to subjects after completion of each case.

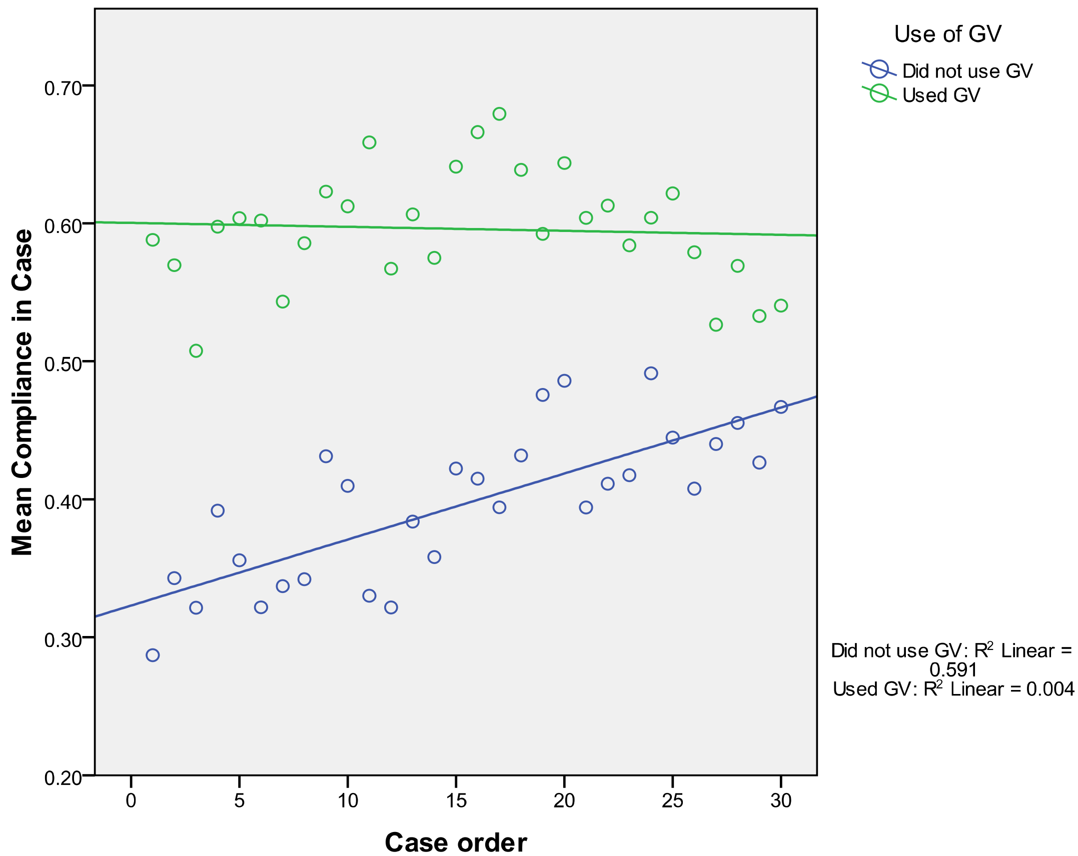


Figure 22. Protocol compliance by case order and treatment

Following the mixed linear model results are as in Table 17. Estimates of fixed effect of mixed linear model for protocol adherence by case

	Intercept	Group	Treatment	Case Order	Treatment*Order
Mean	**0.6448**	-0.036	**-0.3081**	**-0.0166**	**0.006198**
P	**0**	0.187	**0**	**0.228**	**0.002**

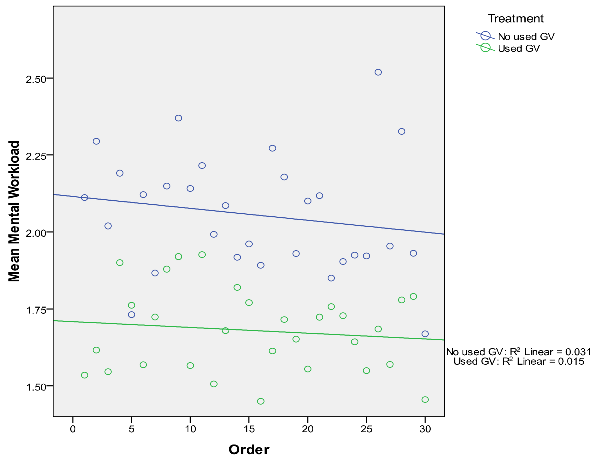
Table 17. Estimates of fixed effect of mixed linear model for protocol adherence by case ***15 Workload***

Out of a total of 1399 NASA TLX questionnaires 28 were incomplete and discarded from the analysis. A total of 1371 valid evaluations where used to the analysis following the mixed linear model presented in section 14.1. Errors were caused mainly by bugs in the identification of transactions by the data collection tool not detected in the pilot study or by loss of internet connectivity. These errors were detected in the interim data reviews.

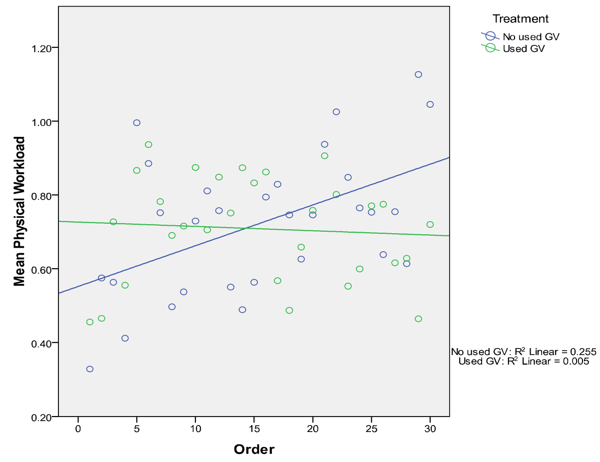
Results are summarized graphically in Figure 23. Each workload factor is presented separately. All green points and lines represent treatment that use CGV, while blue lines and points represents control treatment. Workload factor indexes were calculated as in Equation 6 (W_i*X_i). Every point in a graph represents the mean of the factor for a given case (by order of presentation) at a given treatment.

Workload factor with most noticeable differences between treatments is the mental workload (.a). Treatment with CGV reduces workload significantly (p = 0.011) in average 10% (0.52 in a range from 0 to 5). Additional factor that are modified by the treatment is frustration, which is significantly lower (p = 0.015) . Reduction in frustration is equivalent to 10% (0.5 in a range from 0 to 5).

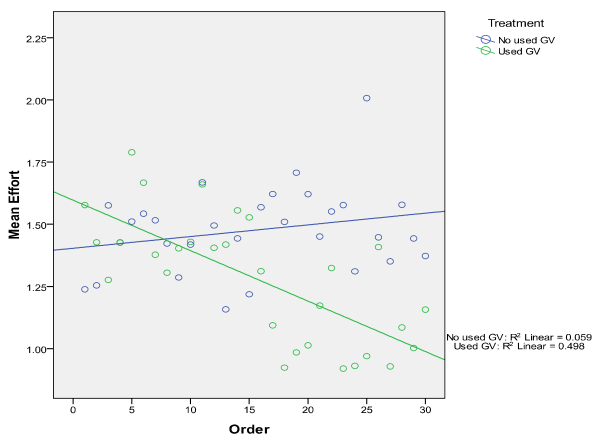
The order of cases as independent variable is playing an important role especially in the frustration factor. There is a progressive reduction in frustration regardless the treatment or other variable, which can be explained by the gain in confidence of the subject after repeated exposure to the simulated cases.



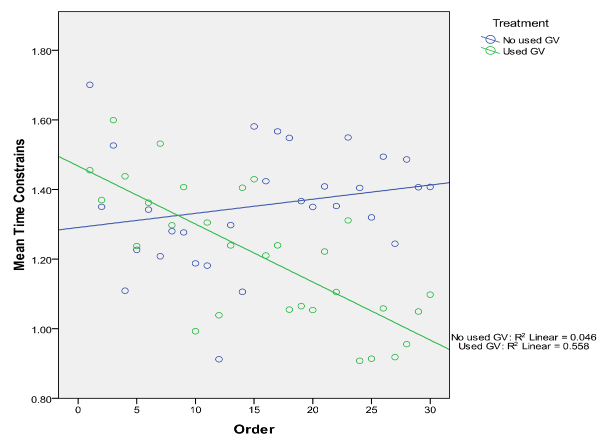
a- Mental



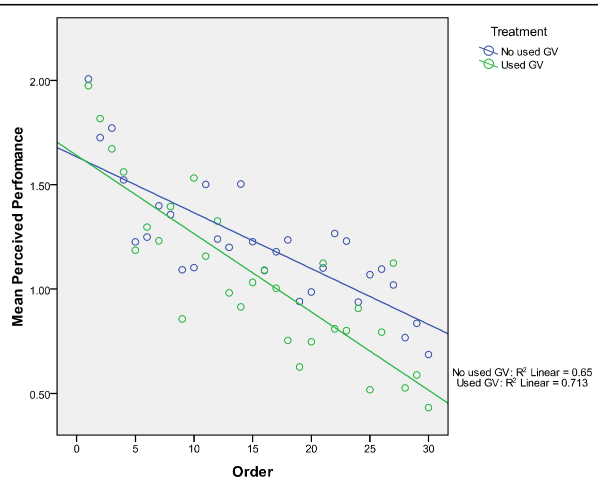
b- Physical



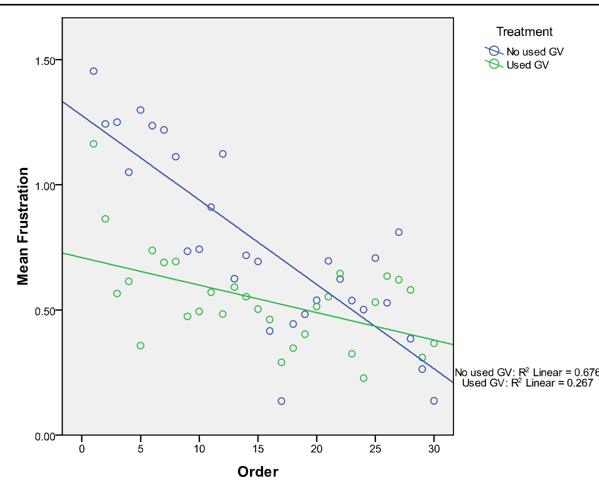
Effort



Constrain



Performance



Frustration

Figure 23. Workload sources by treatment and case order

Overall workload (NASA TLX) was calculated as Equation 6. In the graphical representation each dot represents the mean of the NASA TLX for each case and treatment. Same color convention as is individual factors is used. It is important to notice that both treatment and order of cases are significant variables for the result. Interaction between treatment and order does not present any difference. Previous observation indicates that both effects are independent. There is an effect that is given by the treatment (p=0.009). The effect of treatment is expressed as a 9% reduction in the overall workload (NASA TLX values range from 0 to 1) when using CGV. Additionally is an order effect that reduces the workload 1% per case. This can be explained as in other factors by learning and confidence in the tasks.

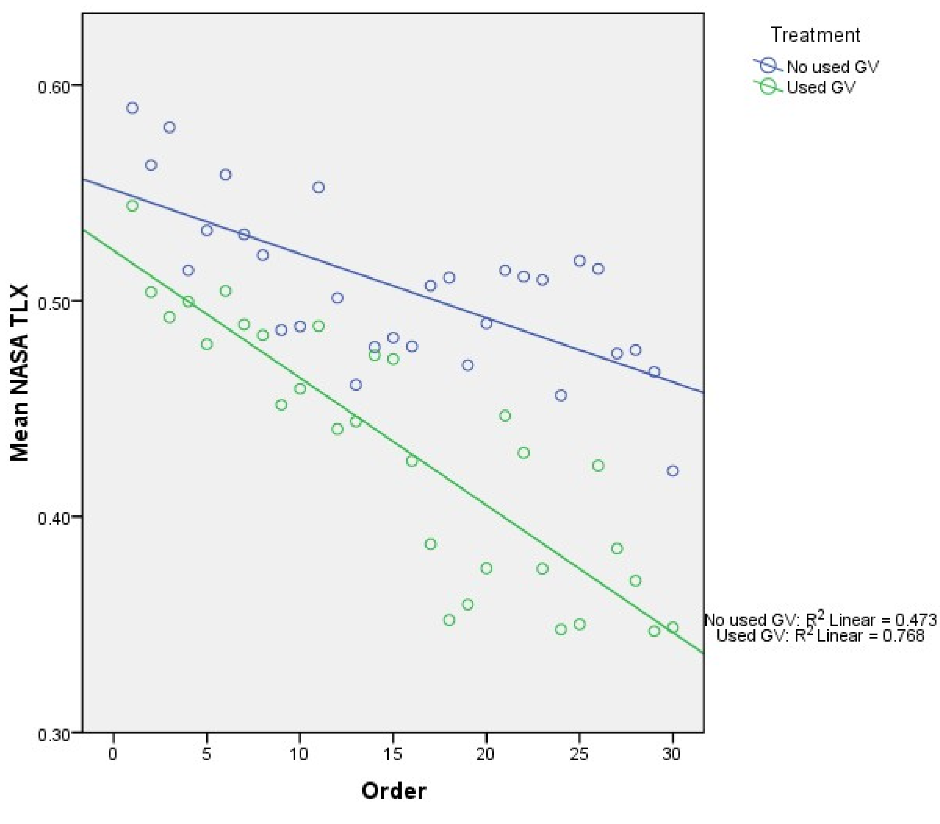


Figure 24. Differences in Overall Workload by Treatment and Case Order

The mixed linear model used to calculate significances follows a structure as:

```
MIXED X WITH group treatment order experience typeoftraining
  /CRITERIA=CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE)
PCONVERG
   E(0.000001, ABSOLUTE)
  /FIXED=group treatment order treatment*order group*orden experience
typeoftraining | SSTYPE(3)
  /METHOD=REML
  /PRINT=CPS  SOLUTION
  /RANDOM=order treatment*order | SUBJECT(subjectId) COVTYPE(VC).
```

“X” represents each workload factor (mental, physical, effort, time, performance, frustration, global) under evaluation. Results of the operation are presented in Table 18. Numbers in gray rows show the mean of the coefficient and between parenthesis upper and lower limits for a 95% confidence interval. White rows represent the significance (p value).

	Intercept	Group	Treatment	Case Order	Treatment*Order	Group*Order	Experience	Type of Training
Mental	2.6 (2.24,2.96)	0.01 (-0.39,0.42)	**-0.52 (-0.93,-0.12)**	-0.01 (-0.04,0.03)	0.01 (-0.02,0.04)	0 (-0.05,0.04)	0.14 (-0.04,0.31)	**-0.32 (-0.49,-0.16)**
	0	0.955	**0.011**	0.704	0.497	0.944	0.128	**0**
Physical	0.58 (0.31,0.85)	-0.14 (-0.48,0.19)	0.14 (-0.2,0.47)	0 (-0.02,0.02)	0 (-0.02,0.02)	0.02 (-0.01,0.04)	**0.15 (0.02,0.27)**	-0.05 (-0.16,0.07)
	0	0.398	0.423	0.979	0.729	0.198	**0.022**	0.432
Effort	1.4 (1.08,1.72)	0.31 (-0.04,0.66)	-0.26 (-0.61,0.09)	-0.01 (-0.03,0.02)	0 (-0.02,0.03)	0 (-0.04,0.03)	-0.05 (-0.21,0.1)	0.05 (-0.09,0.19)
	0	0.085	0.142	0.715	0.766	0.839	0.485	0.47
Time	1.71 (1.4,2.01)	0.04 (-0.32,0.39)	-0.02 (-0.37,0.33)	-0.01 (-0.04,0.01)	-0.01 (-0.03,0.02)	0.01 (-0.02,0.04)	-0.01 (-0.16,0.13)	**-0.16 (-0.3,-0.03)**
	0	0.835	0.897	0.274	0.661	0.611	0.878	**0.019**
Performance	1.43 (1.12,1.74)	0.23 (-0.13,0.59)	-0.13 (-0.49,0.23)	**-0.03 (-0.05,0)**	-0.01 (-0.03,0.02)	0 (-0.04,0.03)	-0.05 (-0.2,0.1)	0.14 (0,0.28)
	0	0.215	0.47	**0.025**	0.56	0.803	0.505	0.052
Frustration	0.84 (0.5,1.18)	-0.05 (-0.45,0.35)	**-0.49 (-0.89,-0.09)**	**-0.04 (-0.07,-0.01)**	0.02 (0,0.05)	0 (-0.03,0.04)	0.01 (-0.15,0.17)	**0.25 (0.1,0.4)**
	0	0.806	**0.015**	**0.005**	0.103	0.794	0.904	**0.001**
NASA TLX	0.56 (0.5,0.62)	0.03 (-0.04,0.09)	**-0.09 (-0.15,-0.02)**	**-0.01 (-0.01,0)**	0 (0,0.01)	0 (-0.01,0.01)	0.01 (-0.01,0.04)	0 (-0.03,0.02)
	0	0.416	**0.009**	**0.02**	0.601	0.698	0.33	0.805

Table 18.. Coeffiecients of Mixed Linear Model Results for Workload

Effort seems to be modified by none of the considered variables. Additional experience and type of training are playing a role in the physical demand, time constrain and frustration.

16 User acceptability

User acceptability was measured as explained in section 10.6. Figure 25 and Figure 26 shows a high acceptance of GuideView. All usability variables scored good or excellent more than 90% of the time. All media modes were found to be very useful or indispensable in more than 85% of the users.

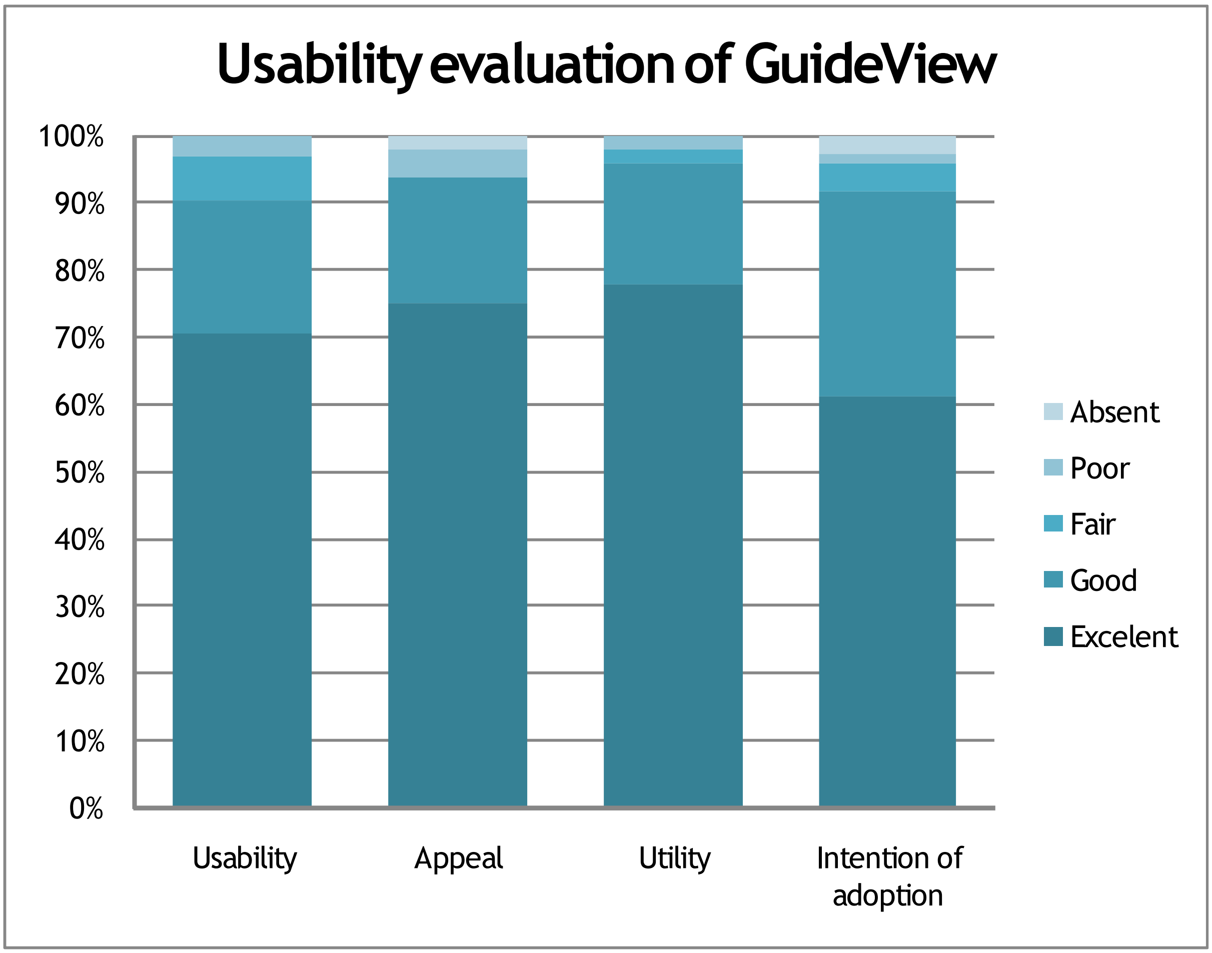


Figure 25. Usability evaluation of GuideView

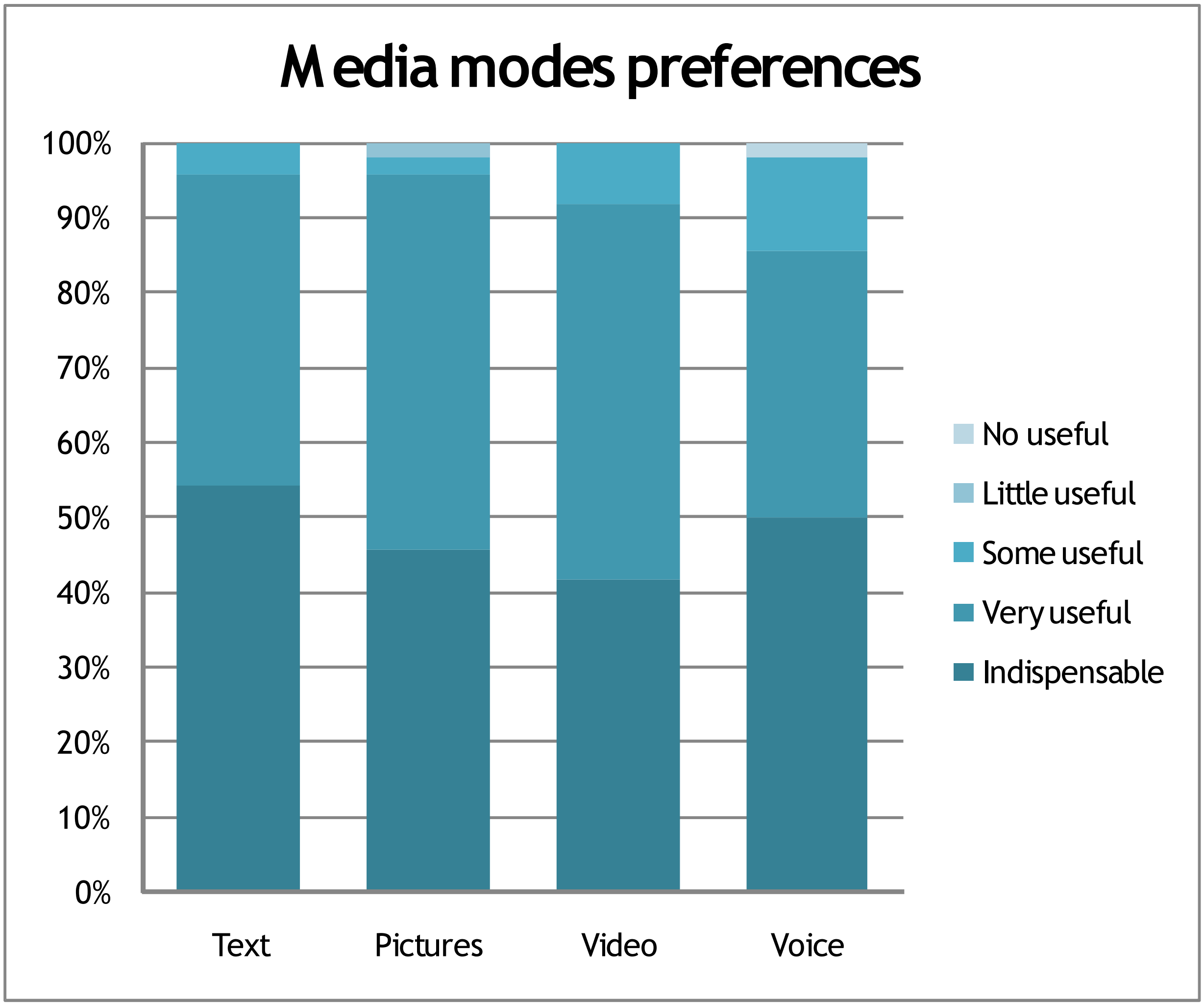


Figure 26. Media modes preferences of CGV users

Discussion

This discussion will be focus on differences found in the study regarding published literature. This section will expose new findings, possible explanation that will generate new hypothesis and opportunity of research. The discussion follows the overall structure presented for variables and results, starting with subjects, moving to different elements in the study.

Gender distribution of subjects is different to the one reported in literature. Usually CHWs are women but in this study men represent the majority. This can be explained by the process to enroll subjects by convenience to secure its participation, given that by change there was a group of police medics, all of them men, been certified as EMTs that were included in the study. Even though there is no evidence that women perform in a different way that men, it is important to take into account for populations homogeneity and do not generalize in the distribution of the community health workers population in Colombia.

With respect to the educational levels of CHWs in our study it should be noted that due to requirements in Colombia, all CHWs have completed high school. This differs from CHWs in certain other countries, where usually a mean of 7 grade completed is reported (S. Y. Rowe et al., 2007). This is important given that higher level of basic education is not making a change in the performance rate supported by traditional job aids. The use of multimedia job aids need to be evaluated in lower education level to be able to generalize results to other populations.

CGV had a positive impact in all performance and workload measurements. The increased time to complete cases using CGV is compensated by the increased protocol adherence. Since subjects are performing more tasks, then time to complete them is longer. A more detailed analysis of task duration could be performed, but this study did not collect such data. A possible approach to this will be a revision of video tapes and time stamping of tasks.

Error rate without CGV are comparable to those reported in the literature (S. Y. Rowe et al., 2007). Reduction in error rate of the CGV treatment is important, but not optimal. There is a need to improve even more the reduction of errors provided by CGV, a more detailed analysis of each node will help to illustrate the causes. It is important to notice that most of the errors remain in task that are*Plans*, while reduction of error in *Actions* is very important.

CGV is providing a protocol compliance rate significantly higher that the control treatment. A recent paper shows how comparable performance is reached after multiple interventions with CHWs. Interventions to enhance performance include training, supervision and motivational plans (A. K. Rowe et al., 2009). This finding will require an additional cost analysis in order to recommend CGV as a replacement for such interventions.

Use of CGV demonstrated that has a positive effect in problematic sources of workload as mental workload and frustration. This is in consistence with smaller studies (unpublished) conducted with CGV. The overall reduction of workload by CGV is important, especially when avoiding fatigue.

Statistics model showed that there in an interaction of control treatment and time in variables from workload (increasing in perceived performance) and performance (reduction of errors and enhancing performance). Previous finding suggest that there is a procedural learning in the control treatment. This can be explained by the fact that every subject received feedback after each case by physicians. A more detailed evaluation could be performed based on auxiliary data to confirm such finding. In addition it is necessary to evaluate if that knowledge persist over time.

Differences in preferences of media can de influences by the training of the subjects, and the type of task. Based on the heuristics is understandable that tasks with high physical demand and requirement of use of both hands are in favor of audible instructions. It is necessary to conduct a more detailed analysis of the relationship between content of each node and the task performance in order to understand the effect of the media in the task. e.g: When a subject is performing CPR and needs to stop after 1 minutes of maneuvers to check respiration and pulse in the patient.

Methods developed for this dissertation are easily extrapolated to different domains. Methods can be applied to any evaluation of interventions that try to demonstrate enhancing human performance, reducing workload. e.g: devices, job aids, workflows among others.

Conclusions

The use of multimedia job aids, as CGV, enhances performance and reduces workload factors for community health workers. Multimedia job aids reductes error rate and frequency of cases with wrongly performed tasks. An increment in protocol adherence was observed in the multimedia job aid intervention. Additionally the use of multimedia job-aids increases case resolution time. Increased time in cases is a permitted trade-off taking into account the better protocol adherence and the reduction of errors. Mental workload, frustration and overall workload are reduced by the use of multimedia job-aids. User acceptance, intention of adoption, perceived usability and appeal of the tool is highly scored. Multimedia job aids are potentially useful for CHWs.

Future Work

Since this study focused on simulated cases next natural step is to produce a Clinical Trial. A clinical trial will help to understand: acceptance by patients of the technology and some factors associated with the environment that would affect the utilization of multimedia job aids. Occasionally a clinical trial could allow to measure health outcomes after intervention. Other possible outcome of clinical trial is to establish how the usage of CGV is changed when CHWs gain expertise.

As stated in the discussion a study to compare Procedural vs Conceptual learning is necessary in order to explain the effect accumulated by the order of cases. This study will be designed to determine if that knowledge is sustainable over time.

A third direction of this research is to understand the cognitive mechanisms that can explain the results especially in the modification of work load, and the relationship of the enhanced representation.

Acknowledgements

This research was possible by value contributions of Microsoft Research, TATRC, U de A, Schull Institute and the participation of:

- Santiago Giraldo Patiño MD. Udea. Ontology of Clinical Tasks development. Data Collection. Programming of Data Collection Tool (GuideViewEval).

- Carlos A Garcia MD. SHIS. Delphi methodology for guideviews and cases creation. Pilot study and data collection.
- James P Turley, RN PhD. SHIS. Review and feedback in the proposal. Constant support and encouragement.
- Jorge I Lopez-Jaramillo MD. UdeA. Facilitating access to Simulation Lab facilities and providing all logistic and technical support.
- Sandra P Libreros MD.UdeA. Delphi analysis and voice recording for guideviews
- Juan C Caro EMT. UdeA. Logistics and photography model.
- Anderson David MD. UdeA. Data collection
- Diana Cárdenas MD. UdeA. Data collection.
- Cesar Pulgarín MD. UdeA. Data collection.
- Juan A Sucerquia MD. UdeaA. Data collection.
- Maria E Zuleta RN. SENA. Subjects Enrollment.
- Silvia M Echeverri MD. UdeA. Subjects Enrollment and administrative support
- Paola A Cortés. Cruz Roja Colombiana. Subjects Enrollment
- Alan Feiveson PhD, NASA JSC , Statistical Expert
- Craig Johnson PhD. SHIS. Statistical models consultant.

References

Ashwell, H. E. S., & Freeman, P. (1995). The clinical competency of community health workers in the eastern highlands province of papua new guinea. Papua New Guiena Medical Journal, 38, 198-207.

ASOCEL. (2006). *Distribución mercado tecnología móvil en colombia* No. 2007)ASOCEL.

Central Intelligence Agency. (2009). *The world factbook.* Retrieved 03/24, 2009, from https://www.cia.gov/library/publications/the-world-factbook/geos/co.html.

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load* Pfeiffer.

Clayton, M. J. (1997). Delphi: A technique to harness expert opinion for critical.. *Educational Psychology, 17*(4), 373. Retrieved from *http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=49870&site=ehost-live*

Dallal, G. E. (2007). ***The computer-aided analysis of crossover studies***. Retrieved 02/07, 2008, from *http://www.tufts.edu/~gdallal/crossovr.htm*

Dower, C., Knox, M., Linder, V., & O'Neil, E. (2006). *Advancing community health workers practice and utilization: The focus on financing*

Elsenheimer, J. (1998). Job aids in the technology age. *Performance Improvement, 37*(8), 32-35.

Farmer, A., Gibson, O., Hayton, P., Dudley, C., Neil, A., & Tarassenko, L. (2005). A real-time, mobile phone-based telemedicine system to support young adults with type 1 diabetes. *Informatics in Primary Care, 13*, 171-177.

Florez-Arango, J. F., & Iyengar, M. S. (2007). Delivering structured, multi-modal clinical guidelines via cell phones. *Medinfo,* Brisbane.

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. San Francisco: Morgan Kaufmann.

Folley, J. D. J., & Munger, S. J. (1961). A review of the literature on design of informational job performance aids [Abstract]. Retrieved from *http://stinet.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0270867*

Fox, J., Alabassi, A., Patkar, V., Rose, T., & Black, E. (2006). An ontological approach to modelling tasks and goals. *Computers in Biology and Medicine, 36*(7-8), 837-856.

González Gutierrez, J. L., Moreno Jiménez, B., Garrosa Hernández, E., & López López, A. (2005). Carga mental y fatiga en servicios especiales de enfermería *Revista Latinoamericana De Psicología, 37*(3), 477-492.

González Muñoz, E. L., & Gutierrez Martínez, R. E. (2006). La carga de trabajo mental como factor de riesgo de estrés en trabajadores de la industria electrónica *Revista Latinoamericana De Psicología, 38*(2), 259-270.

Gove, S. (1997). Integrated management of childhood illness by outpatient health workers: Technical basis and overview. the WHO working group on guidelines for integrated management of the sick child. *Bull World Health Organ, 75 Suppl* 1, 7-24.

Grol, R., Dalhuijsen, J., Thomas, S., Veld, C. i. '., Rutten, G., & Mokkink, H. (1998). Attributes of clinical guidelines that influence use of
guidelines in general practice: Observational study. *British Medical Journal, 317*(7162), 858-861.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task loas index) : Results of a theoretical and empirical research. In P. A. Hancock, & N. Meshkati (Eds.), *HUMAN MENTAL WORKLOA*D (pp. 139-83). North-Holland: Elsevier Science Publishers.

Iyengar, M. S., & Florez-Arango, J. F. (2007). Demonstration of GuideView, a multi-platform system for interactive, multimodal presentation of clinical advice. *Medinfo,* Brisbane.

Iyengar, M. S., Sarkar, S., Bacal, K., Defow, G., McCulley, P., & Hurst, V. (2005). GuideView: Structured multimodal delivery of clinical guidelines. *IMIA,* San Francisco.

Kelly, D. (2006). Sofisticated and savvy. *PROFIT,*

Kim, H., Kim, N., & Abn, S. (2006). Impact of a nurse short message service intervention for patients with diabetes. *Journal of Nurse Care Quality, 2*1(3), 266-271.

Kim, Y. M., Kols, A., Martin, A., Silva, D., Rinehart, W., Prammawat, S., et al. (2005). Promoting informed choice: Evaluating A decision-making tool for family planning clients and providers in mexico. *International Family Planning Perspectives, 3*1(4), 162-171.

Knebel, E. (2000). *The use of manual job aids by health care providers: What do we know*? No. 1)The Quality Assurance Project - USAID. Retrieved from *http://www.qaproject.org/pubs/PDFs/ISSUESJA.PDF*

Kumar, A., Smith, B., Pisanelli, D. M., Gangemi, A., & Stefanelli, M. (2006). Clinical guidelines as plans: An ontology theory. *Methos Inf Med, 4*5(2), 204-210.

Lewis, J. R. (1993). *IBM computer usability satisfaction questionnaires: Psychomentric evaluation and instructions for us*e (Technical Report No. 54.786). Boca Raton, FL: IBM Corporation.

Makeham, M. A., Stromer, S., Bridges-Webb, C., Mira, M., Saltman, D. C., Cooper, C., et al. (2008). Patient safety events reported in general practice:A taxonomy. *QUALITY & SAFETY IN HEALTH CARE, 1*7(1), 53-57.

Mayer, R. E. (2001). *Multi-media learnin*g. United States of America: Cambridge University Press.

Migration Dialogue. (2006). India: IT, health. *Migration News, 1*3(4)

Morrato, E.,H., Dillon, P., & Ziegler, M. (2008). Surgical outcomes research: A progression from performance audits, to assessment of administrative databases, to prospective risk-adjusted analysis - how far have we come?. *Current Opinion in Pediatrics, 2*0(3), 320-325.

Netsize Group. (2006). *The netsize guide*

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. New York: Addison-Wesley.

Norman, D. A. (1991). Cognitive artifacts., 17-38.

Patel, V. L., Arocha, J. F., Diermeier, M., Greenes, R. A., & Shorliffe, E. H. (2001). Methods of cognitive analysis to support the design and evaluation of biomedical systems: The case of clinical practice guidelines. *Journal of Biomedical Informatics, 34*(52-66)

Patel, V. L., Arocha, J. F., Diermeier, M., Greenes, R. A., & Shortliffe, E. H. (2001). Methods of cognitive analysis to support the design and evaluation of biomedical systems: The case of clinical practice guidelines. *Journal of Biomedical Informatics, 34*(1), 52-66.

Pellegrin, L., Bonnardel, N., Antonini, F., Albanese, J., Martin, C., & Chaudet, H. (2007). Event oriented representation for collaborative activities (EORCA). A method for describing medical activities in severely-injured patient management. *Methods of Information in Medicine, 46*(5), 506-515.

Quaglini, S., & Ciccarese, P. (2006). Models of guideline representation. *Neurol Sci, 27*, S240.

Rice, R., & Katz, J. (2003). Comparing internet and mobiel usage: Digital divides of usage, adoption and dropouts. *Telecommunications Policy, 27*, 597-623.

Rossett, A., & Gautier-Downes, J. (1991). *A handbook of job aids*. San Francisco: Pfeiffer.

Routtio, P. (2004). *Theory of design*. Retrieved 07/11, 2007, from http://www.uiah.fi/projects/metodi/123.htm

Rowe, S. Y., Kelly, J. M., Olewe, M. A., Kleinbaum, D. G., McGowan, J. E.,Jr, McFarland, D. A., et al. (2007). Effect of multiple interventions on community health workers' adherence to clinical guidelines in siaya distric,kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 101*, 188-202.

Rowe, A. K., Onikpo, F., Lama, M., Osterholt, D. M., Rowe, S. Y., & Deming, M. S. (2009). A multifaceted intervention to improve health worker adherence to integrated management of childhood illness guidelines in benin. *American Journal of Public Health, 99*(5), 837-846. doi:10.2105/AJPH.2008.134411

Schickendantz, S., Pillekamp, F., Emel, M., Sreeram, N., & Brockmeier, K. (2006). Wireless holter trasmission in suspected dysrhytmias. *Journal of Electrocardiology, 39*, S54-S56.

Simoes, E. A., Desta, T., Tessema, T., Gerbresellassie, T., Dagnew, M., & Gove, S. (1997). Performance of health workers after training in integrated management of childhood illness in gondar, ethiopia *Bull World Health Organ, 75 Suppl* 1, 43-53.

SPSS Inc. (2005). *Linear mixed-effects modeling in SPSS: An introduction to the MIXED procedure* (Technical Report No. LMEMWP-0305). Chicago: SPSS.

Victora, C. G., Huicho, L., Amaral, J. J., Armstrong-Schellenberg, J., Manzi, F., Mason, E., et al. (2006). Are health interventions implemented where they are most needed? district uptake of the integrated management of childhood illness strategy in brazil, peru and the united republic of tanzania *Bull World Health Organ, 84*(10), 792-801.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models*. Boca Raton, FL: Chapman & Hall/CRC.

Wikipedia. (2008a). Performance (disambiguation) --- wikipedia, the free encyclopedia. Retrieved from *http://en.wikipedia.org/w/index.php?title=Performance_%28disambiguation%29&oldid=206746507*

Wikipedia. (2008b). Task --- wikipedia, the free encyclopedia. Retrieved from *http://en.wikipedia.org/w/index.php?title=Task&oldid=192875463*

World Bank. (2009). *Low to middle income country*. Retrieved 03/26, 2009, from *http://siteresources.worldbank.org/DATASTATISTICS/Resources/CLASS.XLS*

World Health Organization. (2006). *The world health report 2006: Working together for health*. No. 978 92 4 156317 8). France: World Health Organization.

Zhang, J., & Butler, K. A. (2007). UFuRT: A work-centered framework and process for design and evaluation of information systems. *Proceedings of HCI International 2007,*

Zuleta, M. E. (2009). In Florez-Arango J. F. (Ed.), *Perfiles profesionales y competencias del personal que presta servicios de salud comunitario en Colombia: APH, Auxiliares de Enfermería y Auxiliares de Salud Públic*a [Competences and profiles of community health workers in Colombia: pre-hospital care, nursing assistants and public health auxiliaries] (J. F. Florez-Arango Trans.).

Appendices

Appendix 1. Informed Consent

Appendix 2. Spanish NASA TLX form

Appendix 3. Acceptability Questionnaire

Appendix 4. Paper-based job aids

Appendix 5. Table of tasks errors

Task name	Error proportion without GV	Error proportion with GV	Error differences	Error differences proportion	Chi	p
Look anaemia	0.88	0.16	0.71	0.81	233.49	0.00
Ask for blood in stools	0.81	0.13	0.68	0.84	40.93	0.00
Classify as mastoiditis	0.86	0.20	0.66	0.77	19.71	0.00
Classify anaemia severity	0.96	0.30	0.65	0.68	20.99	0.00
Classify as malaria	1.00	0.40	0.60	0.60	18.70	0.00
Classify dehydration level	0.79	0.22	0.57	0.73	15.50	0.00
Ask for sore throat	0.72	0.15	0.57	0.79	151.94	0.00
Ask for ear problems	0.71	0.15	0.56	0.79	150.03	0.00
Ask for thick blood smear	1.00	0.44	0.56	0.56	16.91	0.00
Review vaccination card	0.75	0.20	0.54	0.73	137.18	0.00
Idnetifys irritability and pain source	0.88	0.33	0.54	0.62	13.98	0.00
Look for alarm signs in fever	0.88	0.34	0.54	0.62	74.33	0.00
Evaluate nutritional status	0.74	0.21	0.53	0.71	128.58	0.00
Classify denutrition degree	0.87	0.35	0.52	0.60	13.14	0.00
Starts antibiotics for amigdalitis	0.92	0.41	0.51	0.55	13.46	0.00
Ask for seizures	0.73	0.23	0.50	0.69	105.84	0.00
Ask for clinical background	1.00	0.50	0.50	0.50	1.50	0.22
Search for rash	0.91	0.41	0.50	0.55	65.92	0.00
Classify as bacterial faringitis	0.71	0.23	0.48	0.68	10.65	0.00
Start malaria treatment	1.00	0.52	0.48	0.48	13.64	0.00
Classify as diarrhea without dehydration	0.81	0.33	0.48	0.59	10.29	0.00
Search for bleeding	0.95	0.48	0.47	0.50	11.11	0.00
Stablishes malaria risk	0.87	0.40	0.47	0.54	56.52	0.00
Ask for hemoglobin availability	0.99	0.52	0.47	0.47	66.74	0.00
Recomends vitamin A	1.00	0.54	0.46	0.46	15.90	0.00
Ask for Asthma symptoms	0.80	0.38	0.42	0.52	34.77	0.00
Evaluate awakeness	0.59	0.19	0.40	0.68	76.58	0.00
Ask for vomiting	0.50	0.11	0.39	0.77	83.06	0.00
Evaluation of skin	0.82	0.43	0.39	0.47	59.17	0.00
Ask for coughing	0.47	0.08	0.38	0.82	86.09	0.00
Listen for stridor	0.78	0.40	0.38	0.49	35.95	0.00
Wash hands	1.00	0.62	0.38	0.38	9.47	0.00

Task name	Error proportion without GV	Error proportion with GV	Error differences	Error differences proportion	Chi	p
Ask for alarm signs	0.99	0.63	0.36	0.37	75.58	0.00
Ask for chronic respiratory symptoms	0.80	0.44	0.36	0.45	20.41	0.00
Look for infection foci	0.91	0.56	0.35	0.39	21.13	0.00
Ask for allergies	0.70	0.36	0.34	0.48	83.30	0.00
Measure temperature	0.56	0.23	0.32	0.58	32.67	0.00
Evaluate abdomen	0.78	0.47	0.31	0.40	55.02	0.00
Evaluate head and neck	0.79	0.48	0.31	0.39	47.72	0.00
Classify as severe disease	0.78	0.48	0.30	0.39	4.57	0.03
Evaluate pelvis	0.89	0.59	0.30	0.34	61.42	0.00
Draw growth graphs	0.86	0.56	0.30	0.35	23.81	0.00
Ask for medications	0.61	0.32	0.29	0.47	61.59	0.00
Ask for history of seizures	0.89	0.61	0.29	0.32	14.59	0.00
Ask for diarrhea	0.34	0.07	0.27	0.79	52.80	0.00
Evaluate back	0.88	0.61	0.27	0.31	50.82	0.00
Ask for stools properties	0.81	0.54	0.27	0.33	7.40	0.01
Measures vital signs	0.65	0.39	0.27	0.41	55.32	0.00
Ask for last ingestion of food or liquids	0.72	0.46	0.27	0.37	49.90	0.00
Evaluate for cervical-section	0.64	0.38	0.26	0.41	16.24	0.00
Ask and document temporal evolution of symptoms	0.66	0.40	0.26	0.39	12.67	0.00
Starts rescue ventilations	0.51	0.26	0.26	0.50	6.38	0.01
Puts cervical immobilization	0.66	0.40	0.26	0.39	12.64	0.00
Evaluate Thorax	0.66	0.40	0.26	0.39	37.33	0.00
Call for help	0.65	0.40	0.25	0.39	11.39	0.00
Ask for capability of feeding	0.35	0.10	0.25	0.71	42.25	0.00
Classify as common cold	0.57	0.32	0.25	0.44	2.94	0.09
Educates mother	0.84	0.60	0.24	0.29	30.32	0.00
Verify respiration	0.49	0.25	0.24	0.49	30.88	0.00
Ask for history of diseases	0.53	0.30	0.24	0.45	37.37	0.00
Evaluate ventilation	0.76	0.53	0.24	0.31	6.34	0.01
Conduct SAMPLE history	0.94	0.70	0.23	0.25	53.87	0.00
Measures blood pressure	0.50	0.27	0.23	0.46	42.45	0.00
Verify thorax expansion	0.62	0.39	0.23	0.37	7.01	0.01
Evaluate pharynx	0.76	0.53	0.22	0.30	5.00	0.03

Task name	Error proportion without GV	Error proportion with GV	Error differences	Error differences proportion	Chi	p
Classify as severe trauma	0.65	0.42	0.22	0.35	11.78	0.00
Look for dehydration signs	0.65	0.43	0.22	0.34	6.76	0.01
Ask for palliation of pain	0.86	0.64	0.22	0.25	23.74	0.00
Ask for pain properties	0.70	0.48	0.21	0.31	2.94	0.09
Evaluate the scene	0.77	0.56	0.21	0.28	21.25	0.00
Evaluate superior and inferior limbs	0.66	0.45	0.21	0.32	25.58	0.00
Measure pulse	0.48	0.28	0.21	0.43	2.84	0.09
Evaluates airway	0.71	0.50	0.21	0.29	2.18	0.14
Classify as pneumonia	0.79	0.58	0.21	0.26	4.97	0.03
Look for blowing wound on thorax	0.61	0.42	0.20	0.32	4.11	0.04
Starts antibiotics	0.95	0.76	0.19	0.20	3.27	0.07
Align spine	0.90	0.72	0.18	0.20	10.19	0.00
Classify as hypertension emergency	0.76	0.58	0.18	0.23	1.61	0.20
Ask for vomit properties	0.86	0.68	0.18	0.21	1.97	0.16
Starts rehydration plan A	0.62	0.45	0.18	0.28	2.84	0.09
Conduct history for chest pain	0.80	0.63	0.18	0.22	3.45	0.06
Evaluate mental status	0.73	0.55	0.17	0.24	20.73	0.00
Evaluate cardiac frequency	0.90	0.73	0.17	0.19	2.03	0.15
Look for chest indrawing	0.64	0.47	0.17	0.27	5.85	0.02
Ask for personal background	0.70	0.53	0.17	0.24	4.14	0.04
Schedule control visit	0.96	0.79	0.16	0.17	5.73	0.02
Classify event as trauma	0.38	0.22	0.16	0.42	11.21	0.00
Secure the scene	0.64	0.48	0.16	0.25	21.27	0.00
Ask for symptoms onset	0.69	0.53	0.16	0.23	10.78	0.00
Evaluates airway	0.63	0.47	0.16	0.25	4.73	0.03
Ask for chief complaint	0.48	0.32	0.16	0.33	1.17	0.28
Provide ventilation support	0.77	0.62	0.15	0.19	5.22	0.02
Complete clinical history	0.98	0.84	0.15	0.15	9.11	0.00
Provide antibiotics	0.96	0.81	0.15	0.15	5.20	0.02
Ask for fever	0.21	0.07	0.15	0.68	20.75	0.00
Consider cervical spine stabilization	0.85	0.71	0.14	0.16	2.51	0.11
Look for hemorrhages	0.43	0.30	0.14	0.31	6.51	0.01
Ask for symptoms severity	0.86	0.72	0.13	0.16	10.80	0.00
Establishes response level	0.65	0.52	0.13	0.20	4.98	0.03

Task name	Error proportion without GV	Error proportion with GV	Error differences	Error differences proportion	Chi	p
Dress thorax with one way air valve	0.50	0.38	0.13	0.25	0.76	0.38
Evaluates airway	0.66	0.54	0.12	0.19	9.62	0.00
Evaluate stability of thorax	0.77	0.66	0.12	0.15	6.33	0.01
Complete physical examination	0.82	0.71	0.12	0.14	8.58	0.00
Connects Automatic external defibrillator	0.79	0.68	0.11	0.14	0.72	0.40
Starts transportation	0.84	0.74	0.11	0.13	6.27	0.01
Identify need to referral	0.51	0.40	0.11	0.21	2.58	0.11
Calls referral center	0.70	0.60	0.10	0.15	4.81	0.03
Prepares for referral	0.65	0.55	0.10	0.15	4.67	0.03
Refers properly	0.61	0.52	0.10	0.16	4.45	0.03
Identify lower limb swelling	0.29	0.19	0.10	0.33	0.53	0.47
Starts chest compressions	0.73	0.64	0.10	0.13	0.96	0.33
Establish trauma mechanics	0.28	0.19	0.09	0.32	5.34	0.02
Ask for symptom properties	0.70	0.61	0.09	0.13	1.64	0.20
Measures respiratory rate	0.63	0.54	0.09	0.14	6.67	0.01
Evaluate limbs perfusion	0.94	0.86	0.09	0.09	8.59	0.00
Report case	1.00	0.92	0.08	0.08	1.76	0.19
Explain alarm signs	0.98	0.91	0.07	0.07	6.20	0.01
Establish a secure scene	0.77	0.69	0.07	0.09	2.41	0.12
Jaw sub-luxation	0.99	0.92	0.07	0.07	3.75	0.05
Secure proper ventilation	0.73	0.66	0.07	0.09	2.71	0.10
Cover wound	0.44	0.38	0.07	0.15	0.83	0.36
Control hemorrhages	0.54	0.48	0.06	0.12	0.97	0.32
Urgent referral	0.49	0.43	0.06	0.13	1.29	0.26
Cincinnati test	0.86	0.80	0.06	0.06	0.74	0.39
Examine ear	0.33	0.28	0.05	0.16	0.15	0.70
Ask for radiation of pain or relation with other symptoms	0.69	0.63	0.05	0.08	1.40	0.24
Look for pulse	0.33	0.28	0.05	0.16	3.38	0.07
Evaluate bleeding	0.51	0.46	0.05	0.10	0.83	0.36
Start urgent referral	0.72	0.67	0.05	0.07	1.05	0.31
Evaluate general patient aspect	0.71	0.66	0.04	0.06	1.18	0.28
Fixed unstable thorax, removes restrictions	0.58	0.54	0.04	0.07	0.08	0.77
Immobilize wound	0.71	0.67	0.04	0.06	0.10	0.76

Task name	Error proportion without GV	Error proportion with GV	Error differences	Error differences proportion	Chi	p
Evaluate circulation	0.79	0.75	0.04	0.05	0.21	0.65
Classify event as no trauma	0.43	0.39	0.04	0.09	0.62	0.43
Establish contact with patient	0.06	0.02	0.04	0.62	7.66	0.01
Evaluate ventilation	0.55	0.51	0.04	0.07	0.62	0.43
Ask for temporal evolution of symptom	0.50	0.47	0.03	0.06	0.45	0.50
Continues CPR	0.67	0.64	0.03	0.05	0.05	0.83
Start fast trauma assessment	0.54	0.51	0.03	0.05	0.18	0.67
Handles pain	0.87	0.84	0.03	0.03	0.49	0.48
Elevation of affected limb	0.71	0.69	0.02	0.03	0.05	0.82
Open Airway	0.55	0.53	0.02	0.03	0.07	0.79
Selects high complexity hospital	0.69	0.68	0.02	0.02	0.06	0.80
Ask for associated events	0.74	0.73	0.01	0.02	0.11	0.74
Control bleeding	0.43	0.42	0.01	0.03	0.03	0.87
Evaluate duration of current disease	0.68	0.68	0.00	0.00	0.00	0.99
Identify proper ventilation and breathing	0.46	0.47	-0.01	-0.02	0.01	0.92
Establishes life threatening conditions	0.67	0.68	-0.01	-0.02	0.02	0.88
Treats fever	0.68	0.70	-0.02	-0.03	0.06	0.81
Evaluate shock	0.92	0.94	-0.02	-0.02	0.44	0.51
Verifies tetanus vaccination	0.87	0.89	-0.03	-0.03	0.16	0.69
Physical examination from head to toe	0.85	0.88	-0.03	-0.04	0.27	0.61
Initial wound management	0.43	0.46	-0.03	-0.08	0.05	0.82
Procedure to stop bleeding	0.53	0.56	-0.04	-0.07	0.33	0.57
Listen thorax and document findings	0.85	0.90	-0.05	-0.05	1.79	0.18
Ask chief complaint	0.11	0.16	-0.05	-0.43	4.77	0.03
Pressure on the wound	0.45	0.51	-0.06	-0.12	0.58	0.45
Identify type of wound	0.14	0.21	-0.07	-0.46	0.33	0.57
Identify proximal and distal articulation	0.56	0.63	-0.07	-0.13	0.46	0.50
Immobilize fracture	0.73	0.80	-0.07	-0.10	1.93	0.16
Identify wounded limb	0.06	0.15	-0.08	-1.33	1.79	0.18
Compressive bandage	0.40	0.48	-0.08	-0.21	0.68	0.41
Cleans wound properly	0.48	0.57	-0.09	-0.19	1.97	0.16

Task name	Error proportion without GV	Error proportion with GV	Error differences	Error differences proportion	Chi	p
Uses biosafety devices	0.43	0.52	-0.09	-0.22	3.26	0.07
Identify a contaminated wound	0.65	0.75	-0.10	-0.16	1.24	0.27
Suspects dentition as cause of pain	0.65	0.77	-0.12	-0.18	0.80	0.37
Selects a reference hospital	0.53	0.66	-0.13	-0.25	2.67	0.10
Extracts Foreign Body from airway	0.52	0.68	-0.16	-0.30	1.17	0.28
Detects Arrhythmia	0.78	0.96	-0.17	-0.22	3.07	0.08
Classify wound as clean/contaminated	0.52	0.71	-0.19	-0.37	3.62	0.06
Start IV fluids	0.67	0.96	-0.29	-0.44	6.52	0.01
Put patient in Lateral Secure Position	0.57	0.96	-0.39	-0.68	10.13	0.00

Appendix 6. Cases Randomization Table

Subject	Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1001	A	F-1	F-4	F-2	F-3	F-5	D-1	D-5	D-4	D-3	D-2	E-1	E-4	E-5	E-2	E-3	A-2	A-5	A-4	A-3	A-1	B-3	B-4	B-1	B-5	B-2	C-3	C-4	C-5	C-1	C-2
1002	B	C-3	C-4	C-5	C-1	C-2	A-3	A-5	A-1	A-2	A-4	B-4	B-2	B-5	B-1	B-3	D-4	D-1	D-2	D-3	D-5	E-1	E-4	E-5	E-2	E-3	F-1	F-5	F-2	F-3	F-4
1003	A	E-1	E-4	E-5	E-3	E-2	F-1	F-3	F-5	F-4	F-2	D-4	D-2	D-3	D-5	D-1	C-1	C-3	C-4	C-5	C-2	A-5	A-3	A-2	A-1	A-4	B-1	B-4	B-5	B-2	B-3
1004	A	A-5	A-3	A-4	A-2	A-1	B-5	B-2	B-4	B-1	B-3	C-5	C-1	C-4	C-3	C-2	D-2	D-3	D-4	D-1	D-5	E-2	E-5	E-3	E-4	E-1	F-3	F-5	F-4	F-1	F-2
1005	A	B-2	B-3	B-4	B-1	B-5	A-2	A-5	A-1	A-4	A-3	C-5	C-3	C-4	C-2	C-1	F-4	F-5	F-3	F-2	F-1	D-2	D-5	D-1	D-4	D-3	E-2	E-3	E-4	E-5	E-1
1006	A	E-1	E-2	E-3	E-5	E-4	D-2	D-5	D-3	D-4	D-1	F-5	F-4	F-3	F-1	F-2	C-3	C-2	C-4	C-5	C-1	A-4	A-2	A-1	A-5	A-3	B-3	B-1	B-5	B-4	B-2
1007	A	C-1	C-2	C-3	C-4	C-5	A-2	A-5	A-3	A-4	A-1	B-2	B-4	B-3	B-1	B-5	E-1	E-2	E-4	E-5	E-3	D-3	D-5	D-1	D-4	D-2	F-1	F-4	F-5	F-2	F-3
1008	A	C-2	C-3	C-4	C-5	C-1	A-3	A-1	A-2	A-4	A-5	B-1	B-3	B-2	B-4	B-5	D-4	D-1	D-5	D-3	D-2	E-3	E-5	E-4	E-2	E-1	F-1	F-5	F-2	F-3	F-4
1009	A	E-4	E-5	E-2	E-1	E-3	D-3	D-1	D-2	D-5	D-4	F-5	F-3	F-4	F-2	F-1	A-2	A-1	A-5	A-3	A-4	C-4	C-2	C-1	C-5	C-3	B-1	B-2	B-3	B-5	B-4
1010	A	E-3	E-5	E-2	E-4	E-1	F-3	F-2	F-1	F-4	F-5	D-2	D-4	D-3	D-1	D-5	B-5	B-3	B-1	B-2	B-4	A-2	A-4	A-1	A-5	A-3	C-5	C-1	C-4	C-2	C-3
1011	A	A-4	A-3	A-5	A-1	A-2	B-1	B-3	B-2	B-4	B-5	C-2	C-3	C-5	C-4	C-1	D-4	D-1	D-3	D-5	D-2	E-5	E-2	E-4	E-3	E-1	F-5	F-3	F-2	F-4	F-1
1012	A	B-1	B-3	B-4	B-5	B-2	C-3	C-4	C-1	C-2	C-5	A-3	A-4	A-1	A-2	A-5	E-4	E-3	E-2	E-5	E-1	F-1	F-2	F-5	F-4	F-3	D-4	D-1	D-3	D-5	D-2
1013	B	A-4	A-1	A-5	A-2	A-3	C-5	C-2	C-1	C-4	C-3	B-5	B-2	B-3	B-1	B-4	D-2	D-5	D-4	D-3	D-1	E-5	E-4	E-3	E-1	E-2	F-5	F-1	F-3	F-4	F-2
1014	B	C-5	C-2	C-4	C-3	C-1	A-5	A-3	A-1	A-2	A-4	B-3	B-4	B-2	B-1	B-5	E-2	E-5	E-1	E-3	E-4	F-1	F-4	F-2	F-3	F-5	D-4	D-2	D-3	D-5	D-1
1015	A	F-1	F-2	F-3	F-4	F-5	E-3	E-5	E-2	E-1	E-4	D-2	D-3	D-1	D-5	D-4	A-1	A-2	A-3	A-4	A-5	B-4	B-3	B-5	B-2	B-1	C-3	C-5	C-2	C-4	C-1
1016	B	E-2	E-5	E-4	E-3	E-1	D-5	D-2	D-3	D-4	D-1	F-2	F-1	F-3	F-4	F-5	C-3	C-2	C-1	C-4	C-5	B-4	B-5	B-3	B-1	B-2	A-5	A-4	A-2	A-3	A-1
1017	B	D-1	D-2	D-4	D-3	D-5	F-4	F-2	F-1	F-5	F-3	E-4	E-3	E-1	E-5	E-2	A-5	A-1	A-4	A-2	A-3	B-5	B-3	B-1	B-2	B-4	C-1	C-5	C-2	C-3	C-4
1018	B	E-2	E-1	E-5	E-3	E-4	F-2	F-1	F-4	F-3	F-5	D-4	D-5	D-1	D-3	D-2	B-1	B-3	B-4	B-5	B-2	A-3	A-4	A-1	A-2	A-5	C-3	C-5	C-4	C-1	C-2
1019	B	B-4	B-2	B-3	B-1	B-5	C-4	C-1	C-3	C-2	C-5	A-3	A-1	A-2	A-4	A-5	D-5	D-4	D-2	D-1	D-3	E-2	E-1	E-5	E-3	E-4	F-1	F-4	F-2	F-5	F-3
1020	A	E-4	E-1	E-3	E-5	E-2	F-5	F-2	F-3	F-1	F-4	D-4	D-5	D-1	D-3	D-2	B-4	B-1	B-5	B-3	B-2	C-3	C-4	C-2	C-5	C-1	A-2	A-1	A-5	A-4	A-3
1021	A	F-5	F-4	F-3	F-1	F-2	D-5	D-1	D-2	D-3	D-4	E-5	E-3	E-2	E-1	E-4	B-1	B-4	B-3	B-5	B-2	C-2	C-5	C-4	C-1	C-3	A-2	A-3	A-4	A-1	A-5

Subject	Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1022	B	C-3	C-2	C-1	C-4	C-5	B-1	B-2	B-3	B-4	B-5	A-3	A-1	A-5	A-2	A-4	E-3	E-2	E-4	E-1	E-5	F-2	F-4	F-3	F-5	F-1	D-4	D-5	D-3	D-1	D-2
1023	B	F-5	F-4	F-3	F-2	F-1	D-4	D-1	D-3	D-5	D-2	E-3	E-1	E-4	E-5	E-2	C-5	C-2	C-1	C-4	C-3	A-3	A-2	A-1	A-5	A-4	B-1	B-3	B-4	B-2	B-5
1024	A	B-5	B-3	B-4	B-1	B-2	A-3	A-1	A-2	A-4	A-5	C-1	C-2	C-4	C-3	C-5	E-1	E-5	E-3	E-4	E-2	F-3	F-2	F-1	F-5	F-4	D-5	D-2	D-3	D-1	D-4
1025	A	F-4	F-2	F-5	F-3	F-1	D-4	D-3	D-1	D-5	D-2	E-2	E-1	E-3	E-4	E-5	B-4	B-5	B-2	B-3	B-1	A-4	A-5	A-1	A-2	A-3	C-1	C-2	C-5	C-4	C-3
1026	B	B-4	B-1	B-3	B-5	B-2	C-1	C-2	C-4	C-3	C-5	A-2	A-4	A-3	A-5	A-1	D-2	D-5	D-1	D-3	D-4	E-2	E-4	E-3	E-5	E-1	F-5	F-4	F-1	F-2	F-3
1027	A	D-5	D-4	D-3	D-1	D-2	E-5	E-3	E-2	E-1	E-4	F-3	F-4	F-2	F-5	F-1	C-2	C-3	C-4	C-1	C-5	B-2	B-4	B-3	B-1	B-5	A-5	A-3	A-2	A-1	A-4
1028	B	E-5	E-4	E-1	E-2	E-3	F-2	F-1	F-3	F-5	F-4	D-3	D-4	D-1	D-2	D-5	B-2	B-3	B-4	B-5	B-1	C-1	C-5	C-3	C-4	C-2	A-4	A-3	A-2	A-1	A-5
1029	A	B-2	B-1	B-4	B-3	B-5	C-5	C-2	C-1	C-3	C-4	A-1	A-5	A-4	A-2	A-3	F-2	F-1	F-4	F-3	F-5	D-5	D-2	D-3	D-4	D-1	E-3	E-5	E-2	E-4	E-1
1030	B	F-5	F-2	F-4	F-3	F-1	D-5	D-2	D-4	D-1	D-3	E-3	E-1	E-4	E-2	E-5	B-5	B-4	B-1	B-3	B-2	C-3	C-1	C-5	C-4	C-2	A-3	A-1	A-4	A-2	A-5
1031	B	C-3	C-4	C-5	C-2	C-1	B-1	B-4	B-2	B-5	B-3	A-2	A-5	A-4	A-1	A-3	D-4	D-3	D-1	D-5	D-2	E-3	E-2	E-1	E-4	E-5	F-1	F-2	F-4	F-5	F-3
1032	B	C-4	C-5	C-2	C-3	C-1	A-5	A-2	A-1	A-3	A-4	B-3	B-2	B-5	B-4	B-1	D-3	D-5	D-4	D-2	D-1	F-3	F-2	F-4	F-5	F-1	E-5	E-3	E-1	E-4	E-2
1033	A	A-5	A-4	A-3	A-2	A-1	C-4	C-2	C-3	C-1	C-5	B-3	B-4	B-2	B-5	B-1	E-3	E-2	E-1	E-4	E-5	D-1	D-5	D-3	D-4	D-2	F-4	F-5	F-3	F-1	F-2
1034	B	C-5	C-1	C-2	C-4	C-3	A-2	A-1	A-4	A-3	A-5	B-1	B-4	B-2	B-3	B-5	D-2	D-4	D-1	D-5	D-3	F-5	F-4	F-1	F-3	F-2	E-2	E-5	E-4	E-3	E-1
1035	A	D-1	D-5	D-3	D-4	D-2	F-3	F-4	F-5	F-2	F-1	E-3	E-2	E-5	E-1	E-4	B-3	B-4	B-2	B-1	B-5	C-4	C-5	C-3	C-1	C-2	A-2	A-3	A-1	A-5	A-4
1036	B	E-3	E-1	E-4	E-5	E-2	D-3	D-1	D-2	D-5	D-4	F-1	F-4	F-5	F-2	F-3	B-3	B-4	B-5	B-2	B-1	A-2	A-3	A-1	A-4	A-5	C-1	C-3	C-5	C-2	C-4
1037	B	F-2	F-5	F-1	F-4	F-3	E-3	E-5	E-1	E-2	E-4	D-1	D-2	D-5	D-3	D-4	C-5	C-2	C-4	C-1	C-3	B-1	B-5	B-2	B-3	B-4	A-2	A-1	A-5	A-3	A-4
1038	A	B-2	B-3	B-4	B-5	B-1	C-1	C-2	C-5	C-3	C-4	A-4	A-3	A-2	A-5	A-1	E-1	E-4	E-2	E-5	E-3	F-1	F-3	F-5	F-2	F-4	D-1	D-2	D-3	D-5	D-4
1039	B	A-3	A-5	A-1	A-2	A-4	B-5	B-4	B-3	B-1	B-2	C-1	C-4	C-5	C-3	C-2	F-5	F-1	F-4	F-2	F-3	D-3	D-4	D-2	D-5	D-1	E-3	E-5	E-4	E-2	E-1
1040	B	E-5	E-1	E-2	E-4	E-3	F-3	F-4	F-2	F-5	F-1	D-2	D-1	D-4	D-3	D-5	C-3	C-4	C-1	C-2	C-5	A-5	A-3	A-2	A-4	A-1	B-3	B-5	B-4	B-1	B-2
1041	B	D-2	D-1	D-5	D-4	D-3	E-2	E-4	E-5	E-3	E-1	F-1	F-2	F-5	F-4	F-3	C-4	C-1	C-2	C-3	C-5	A-4	A-2	A-5	A-1	A-3	B-1	B-5	B-4	B-3	B-2
1042	B	D-1	D-5	D-4	D-2	D-3	F-1	F-4	F-5	F-2	F-3	E-1	E-5	E-3	E-2	E-4	B-2	B-3	B-4	B-1	B-5	A-3	A-2	A-4	A-5	A-1	C-3	C-2	C-4	C-5	C-1
1043	A	C-4	C-5	C-2	C-3	C-1	A-4	A-3	A-2	A-5	A-1	B-3	B-2	B-1	B-4	B-5	E-1	E-2	E-3	E-4	E-5	D-3	D-2	D-1	D-4	D-5	F-4	F-1	F-5	F-3	F-2

Subject	Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1044	B	A-2	A-4	A-1	A-3	A-5	C-4	C-3	C-1	C-2	C-5	B-5	B-2	B-3	B-1	B-4	E-5	E-4	E-3	E-2	E-1	F-1	F-5	F-4	F-3	F-2	D-3	D-1	D-5	D-4	D-2
1045	A	C-1	C-2	C-3	C-4	C-5	B-1	B-2	B-4	B-3	B-5	A-2	A-4	A-1	A-3	A-5	D-4	D-2	D-1	D-5	D-3	F-1	F-4	F-2	F-3	F-5	E-1	E-2	E-3	E-5	E-4
1046	A	F-5	F-2	F-4	F-3	F-1	E-4	E-1	E-3	E-5	E-2	D-1	D-5	D-3	D-4	D-2	B-4	B-1	B-2	B-3	B-5	A-1	A-4	A-3	A-5	A-2	C-5	C-1	C-2	C-3	C-4
1047	B	E-3	E-2	E-5	E-1	E-4	F-5	F-2	F-4	F-3	F-1	D-4	D-2	D-5	D-1	D-3	A-1	A-5	A-3	A-4	A-2	C-4	C-1	C-5	C-3	C-2	B-5	B-3	B-4	B-2	B-1
1048	B	A-3	A-4	A-1	A-5	A-2	C-5	C-1	C-4	C-2	C-3	B-5	B-1	B-4	B-3	B-2	F-1	F-2	F-4	F-3	F-5	D-4	D-1	D-2	D-3	D-5	E-4	E-5	E-2	E-3	E-1
1049	A	B-5	B-4	B-3	B-2	B-1	A-2	A-4	A-3	A-5	A-1	C-2	C-3	C-4	C-1	C-5	E-5	E-3	E-2	E-1	E-4	D-4	D-3	D-2	D-1	D-5	F-4	F-3	F-5	F-2	F-1
1050	B	D-2	D-5	D-3	D-4	D-1	E-4	E-2	E-3	E-1	E-5	F-2	F-5	F-1	F-3	F-4	C-5	C-3	C-2	C-1	C-4	A-5	A-1	A-3	A-2	A-4	B-1	B-4	B-2	B-3	B-5