Preparation, Exchange, and Utilization:
A Three-Phase Approach Toward Improving Resident Physician Handoffs

---

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Systems Engineering

by

Justin M. DeVoge

May 2014

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Systems Engineering

Justin M. DeVoge

The dissertation has been read and approved by the examining committee:

Robert R. Bailey (Advisor)

Stephanie A. Guerlain (Chair)
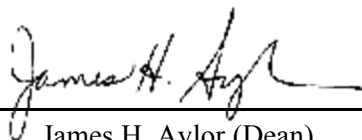

Donald E. Brown

Stephen M. Borowitz

Wendy F. Cohn

Accepted for the School of Engineering and Applied Science:

James H. Aylor (Dean)
School of Engineering and Applied Science

May 2014

Abstract

Physician handoffs are a mechanism for transferring patient information and care responsibility from one set of caregivers to another at shift change. Handoff, by nature, is a complex process characterized by factors such as the structure and volume of information, individual and group cognition, environmental characteristics, and organizational policies. These factors make handoffs a critical point of vulnerability in health care organizations. There is a growing body of evidence suggesting links between errors originating at handoffs and adverse outcomes in patient care.

Teaching residents what information to exchange with other physicians, both verbally and in clinical documentation, has been suggested as a means for reducing errors related to the information discussed during the handoff process. While most medical residency programs generally provide some form of medical communication training, most are not focused on teaching specific information content to exchange at handoff. The critical information to exchange is essentially "learned along the way". Furthermore, few studies have employed performance evaluation in assessing transfer of handoff training to the work environment.

Some view electronic medical record (EMR) systems as solutions to directly support patient care tasks through storing, aggregating, and making accessible a vast amount of patient data. However, most are not designed to support communication and information exchange between people or between people and other information systems. Because of this, it is critical to understand the underlying roles and embedded responsibilities of health care providers and how they serve as human information systems addressing the information needs of others.

This research studied the process of resident physician handoff at shift change in the pediatric acute care setting, and in particular, focused on: 1) further characterizing the handoff process in terms of the information residents' exchange, 2) evaluating the effect of training on the information residents exchange during handoff, and 3) characterizing how information residents exchange during handoff supports their role as information systems for other providers engaged in patient care activities. In addition, this work provides a perspective on the increasingly significant role of information systems in health care, with a particular focus on the outcomes of employing user-centered design principles in the development of an information and decision support system for resident handoff.

Acknowledgements

Table of Contents

List of figures

List of tables

Chapter 1

Introduction

Overview

This work involves the study of the process of patient handoff between resident physicians at shift change in the pediatric acute care setting. In particular, this work focuses on: 1) evaluating the effect of training on the information residents exchange during handoff, and 2) describing the relationship between information residents exchange during handoff and information they are asked to provide in response to questions they receive from other patient care providers overnight.

Background

In hospitals, failures in communication have been linked to both medical errors and adverse events [1-3], and account for 60% of the root (primary) causes of sentinel events reported to The Joint Commission [4]. Miscommunication during patient handoffs, in particular, has been linked to a variety adverse clinical outcomes as well, including longer patient hospital stays, increases in the number of laboratory tests ordered and performed, and increases in the number of self-reported preventable adverse events [5-11]. Beginning in 2003, the Accreditation Council for Graduate Medical Education (ACGME) released the first in a series of successive duty hour restrictions for resident physicians [5]. As duty hour restrictions have necessitated more frequent shift changes for residents, so have they necessitated an increase in the frequency of patient care handoffs and thus, increased opportunities for miscommunication. The primary purpose of handoff is the transfer of patient information and care responsibility from one set of providers to another in order to prevent discontinuity in patient care across shifts [12-14]. In general, handoffs are opportunities for synthesizing, organizing, socializing, maintaining an error-free system; information which may not be available once the outgoing shift has gone.

While residency programs generally provide some form of medical communication skills training, many do not officially focus on instructing residents what information to discuss for each patient [15, 16]. Lack of formal handoff training for residents and the accuracy and availability of relevant patient information are examples of factors which contribute to the complexity of the handoff

process and increase the chance of process related errors resulting in negative patient care outcomes [10, 17].

The handoff process

Handoff is a multi-stage process through which information about patients and their treatment is prepared by a healthcare provider (going off shift), shared with another healthcare provider (coming on shift), and utilized by the new provider in patient care [18]. Building on this emerging perspective in the literature, this work seeks to understand the characteristics and role of information as it exists and flows between three sequential phases of the overall process of handoff (Figure 1).

Ideally, preparation for handoff for leaving residents involves gathering, organizing, summarizing, and updating patient information, checking on the status of tasks in progress, and planning for the immediate future. For incoming residents, preparation involves reviewing and assessing the current state of available patient information, and the preparation of questions to ask during handoff to ensure that the discussion addresses the information they feel they need. The outcome of the preparation phase is exhibited by the information discussed by residents during the exchange phase. In this phase, leaving residents exchange information and transfer patient care responsibility and authority to incoming residents through a communicative event. From the end of the handoff exchange, through the subsequent preparation phase at the end of their shift, residents are in the utilization phase. In this phase, residents utilize information available to them (which they may or may not have received during the exchange) to make medical decisions, care for patients, and to address the requests and concerns of other care providers, as well as patients, their family members, and other individuals they may encounter over the course of their shift. Information residents require in this phase may be internally known, externally available, pending, not pending, missing, unknown, or impossible to determine.

Figure 1. The three phases of the handoff process: preparation, exchange, and utilization. Information is represented by the encircled letter **i's**, arrows depict information flow

Primary research aims

The handoff problem

There remains a great deal of debate over what patient information should be discussed during handoffs, how to implement training interventions which address these characteristics, the extent to which training interventions result in measurable changes in the information discussed during handoffs, or whether such information, if discussed, helps maintain continuity in patient care. As yet, prior work has not developed an empirical basis for what information should be exchanged during handoff, the impact of information exchanged on information-related patient care activities, or how to train residents to discuss critical information during handoff in order to support subsequent patient care activities.

Aim 1
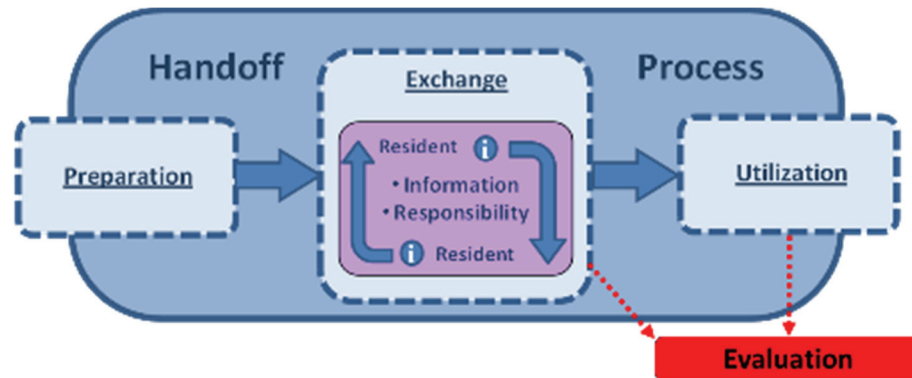
Figure 2. Evaluating the use of information after handoff

As previously discussed, there is a lack of evidence to support exactly what information is the right information to exchange during handoff, therefore, the first aim of this work seeks to address this gap. The purpose is to determine whether information discussed at handoff has an impact on patient care. More precisely, this aim focuses on the information which best prepares residents to provide responses to the information requests of other care providers, and the information residents seek in external sources available to them when they cannot initially provide a response. The specific research questions addressed in relation to this objective are:

- Research Question 1.1: What are the characteristics of the information requests residents receive from other care providers (including patients' family members) following handoff?
- Research Question 1.2 What factors of the information discussed at handoff correlate with residents' ability to respond to the information requests of other care providers?

Aim 2
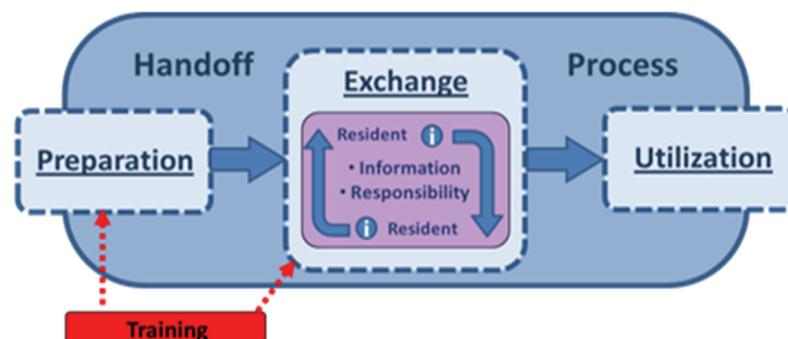
Figure 3. Assessing the impact of training on the information discussed during handoff

The second aim of this work focuses on evaluating the efficacy of training aimed at teaching what information to discuss during the exchange phase of handoff employing a behavioral evaluation strategy to assess the transfer of training to job performance. The specific research questions addressed in relation to this objective are:

- ▪ Research Question 2.1: Does training residents to discuss a specific set of information at handoff correlate to changes in the appearance, frequency, or amount of time they spend discussing that information during handoff?

Ancillary Objective 1

Considerable attention has been given to improving information technology which supports the handoff process. For the most part, these solutions have been developed without engaging physicians in the design process. As information technology has vital role in the health care industry, an ancillary objective of this work is to evaluate the impact of a user-centered design process for building an information system to support handoff and to provide guidelines for the development of future information support systems for handoff.
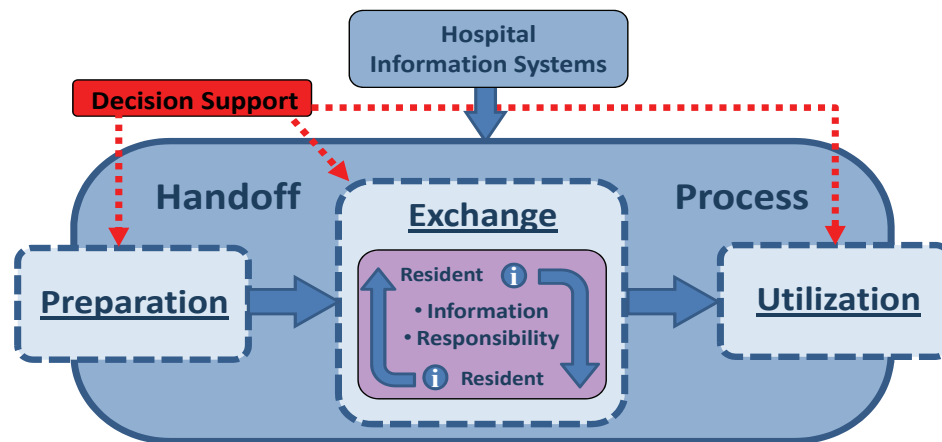
Figure 4. Leveraging decision support to provide additional structure to the information content of handoff in all three phases

Significance

Gaps in the literature
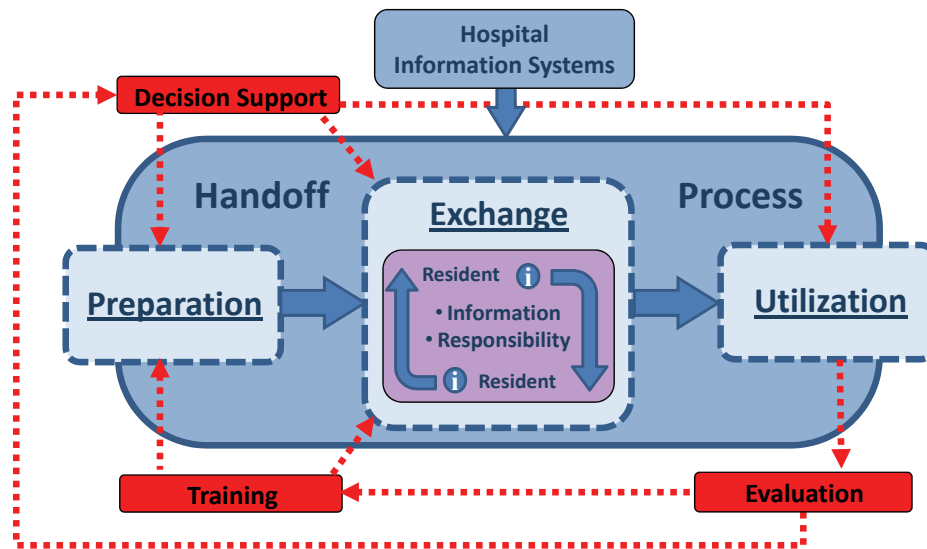
Figure 5. Gaps in the literature on handoff exist in all three aspects relevant to this work: evaluation, training, and decision support

On the utilization of information after handoff

In health care, problems with information management are often associated with problems with technology, and proposed solutions frequently involve the move from one technologically-based system to another [19, 20] as illustrated by the emerging transition from paper to electronic medical records (EMR) [21]. While technology provides a means for augmenting human cognitive capacity in domains faced with a burgeoning set of pertinent information [22], as a whole, managing information is not simply a matter of storage and retrieval of data. Though the storage and retrieval of information in health care are problematic issues in their own right, complex processes such as communication, aggregation, summarization, and conversion are necessary in order to turn information into knowledge that is useful for health care providers [23].

While technology to support processes such as communication is improving, its capability to replace many of the intricacies that exist within the complexity of such processes remains limited [21, 23, 24]. Furthermore, there is a growing body of evidence in health care of the inefficiencies and hazards resulting from the implementation of inappropriate, "technology-heavy" solutions which exacerbate human information management problems [25-27]. Understanding cognitive and

behavioral as well as technological aspects of problems related to data storage, search, retrieval, and dissemination, is critical to designing solutions, given that both humans and technology frequently perform such function [19, 23]. For example, humans interact with other information systems, both human and non-human, to find, share, and exchange information. Despite being critical aspects of patient care, such responsibilities are generally embedded within the health care system, are not clearly defined or outlined by organizational policies, which tend to prioritize the more obvious, explicable responsibilities associated with specific occupational roles (e.g. physicians diagnose, pharmacists fill orders, etc.), and have remained relatively unsupported by information technology [23, 24]. Many electronic medical record (EMR) systems are designed to directly support patient care tasks by storing, aggregating, and making accessible a vast amount of patient data but are not designed to support communication and information exchange between people or between people and other information systems.

Resident physicians are an example of a role where further examination of the embedded responsibilities which exist in health care is needed, particularly within their role as a human information system serving the information needs of others. This seeks to understand two specific aspects related to this role: first, to understand who asks residents for information, and second, to understand the methods residents employ in response to questions they are asked by others. A greater understanding of these issues can help inform the design of the EMR and other interventions that support the information needs of health care providers.

Training residents what to discuss at handoff

There have been relatively few in-situ examinations of the impact of handoff training on real-world handoff behaviors of physicians, particularly in acute care settings. The complexity of the acute care environment and the consequences of interfering with work activities make controlled experimentation (with the expectation of collecting valid data) difficult not to mention a detriment to patient safety. This is particularly relevant in research where study populations, such as resident physicians, are often physically or emotionally depleted, under rigorous time constraints to make high-consequence decisions, complete required work-tasks, address unexpected and frequent interruptions and alarms, and remain in compliance with duty hour restrictions. Time, consequently, is a valuable commodity to residents. Unfortunately, professional training can take a great deal of time and behavioral performance assessment can be intrusive. Two reasons, historically speaking, why many medical skills, are generally learned through practice, on-the-job, over time. Furthermore, to this point, the most common methods for evaluating training outcomes with respect

to handoffs have shown little more than resident satisfaction or a positive attitude toward the value of training and rarely discuss the transfer of training to job performance. Of the few studies of handoff training interventions, only [28] has employed objective measures of performance with respect to resident handoff skills, albeit through simulated tasks, and only [29] has evaluated the impact of training structured information content for handoff.

This work addresses time constraints involved in training handoff to residents by developing a web-based training module which could be completed in approximately one hour from any location with an internet connection and on any device running an Adobe Flash®-compatible web-browser. This work also brings observational assessment to the forefront of handoff education research in an effort to understand how handoff training influences real handoff behaviors. Using observation of real handoffs as an assessment strategy it addresses the problems of subjectivity in self-reports, and the challenges associated with creating realistic representations of the work environment in simulated or job-like tasks. Evaluating residents on the job through observation provides the added benefits of learners who are engaged and invested in their work (critical to job performance), requires no additional time commitments, and is minimally invasive. This work addresses a frequent question in the literature with respect to handoff training and performance, that is: can training residents to focus on specific types of information influence what they discuss during handoff?

Designing systems to support the handoff process

Supporting handoff with technology can be critical to the efficiency of handoff though not necessarily the most important way to make it a more effective process [30]. However, to ensure that such systems are effective, designers must collaborate with end users. Collaboration can be difficult when working with end users who are busy and have irregular schedules and engaging them in the design process is both art and science [103]. End-users may not be able to articulate their problems or point to their root causes. They may jump to potential solutions before truly understanding underlying issues. They may not be available at typical work hours to discuss design concerns. This work discusses a multifaceted approach combining focus groups, interviews, iterative prototyping, testing, and feedback for engaging resident physician end-users in the design process of an electronic handoff support too, potential improvements from an existing tool to a prototype based on task analytic methods, and offers practical strategies for increasing resident physician engagement in the design process for handoff decision support systems.

A normative view of the handoff process

Overall, this work seeks to provide insight in three critical areas that influence the handoff process: conducting observation in order to evaluate what information is critical to residents in the utilization phase of handoff, assessing the effect of training on the information that is discussed during the handoff exchange (and ultimately integrating knowledge of what information is critical to discuss into the content of training), and analyzing how decision support can be leveraged to support the information critical to the handoff process across all three phases (Figure .
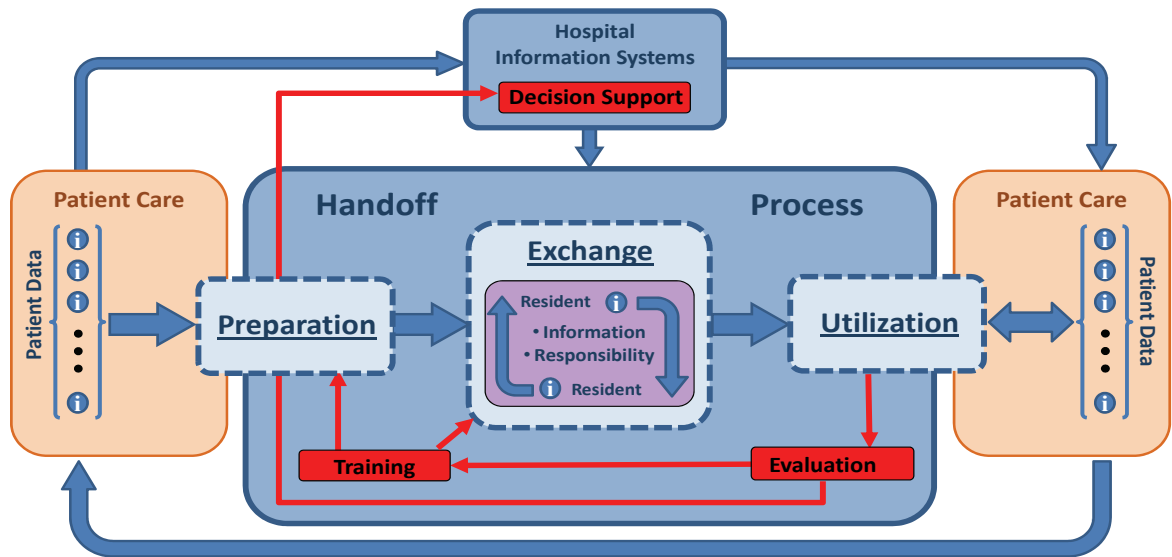
Figure 6. Training, decision support, and evaluation as embedded elements in a normative view of the handoff

Organization of remaining chapters

Chapter two reviews the literature relevant to the handoff topics of this dissertation. Chapter three provides an overview of the research methods. Chapter four describes a study aimed at linking information exchanged at handoff to the behavior of residents in response to information requests they receive following handoff (Research Questions 1.1 and 1.2). Part one of chapter four describes the content of information requests residents receive during this phase and characterizes residents' information seeking behaviors in response to such requests (Research Question 1.1). Part two of chapter four uses observational data from the handoffs preceding these events to assess whether residents responses were due to the information which was exchanged during handoff (Research Question 1.2). Chapter five describes a study conducted on the efficacy of training information

content to residents (Research Question 2.1). Chapter six discusses an information support system for resident handoff developed in collaboration with pediatric residents and faculty iteratively over the course of this work (Ancillary Objective 1), and based on the perspectives gained from engaging end-user resident physicians throughout the design process, also presents a set of basic guidelines for designing handoff decision support tools. Chapter seven summarizes major findings and presents general conclusions drawn from this work.

Chapter 2

Literature Review

Problems in the handoff process

The primary objective of any resident physician handoff is the accurate transfer of information about a patient's state and plan of care [31]. During handoffs, the resident physician who is assuming patient care responsibilities should receive the essential information needed to care for patients during his/her shift [32]. Experienced physicians, however, understand that handoff is not merely a sharing of vital signs or a transfer of quantitative data; it is a discussion of each patient as well as a transfer of responsibility [33-36]. Exclusions and misunderstandings of critical patient information during handoff discussions may lead to deteriorating care of patients [1-3, 31].

Deficiencies in handoffs have been shown to result in longer patient hospital stays, an increase in the number of laboratory tests ordered and performed, and an increase in the number of self-reported preventable adverse events [11]. Although handoffs are an essential process serving a variety of functions [12-14], it is generally informal, and is traditionally not an explicit component of medical training [8, 37], one of many skills which physicians are expected to learn along the way [12, 13, 38-40]. Additionally, there has been little research focused on objectively defining requirements for the procedures and content of handoff [35]. Numerous methods for structuring the content of handoffs have been suggested [41], yet few have approached defining handoff requirements based on the actual needs of residents. For example, assessing work events to characterize the actual needs of physicians and, in turn, using the results as a basis for infusing structure into handoffs. The post-handoff work shift, in particular, can be studied as one area for gaining insight into the information that is most necessary to physicians following handoffs [42].

Handoffs occur in a variety of settings [31] but vary significantly given the information needs of the situation and the individuals involved [43]. Handoff as a method for maintaining continuity of operations at shift change has been studied in domains outside of health care [44, 45]. In [45] the authors describe two implications of maintaining continuity in space shuttle mission control: 1) the seamless transfer of physical presence/activity (as if the same person had retained responsibility across shifts), and 2) shared understanding, or "shared situation-awareness", of the current situation

(between those handing over responsibility and those accepting it). Reaching a shared situation awareness, or "co-orientation" [46] includes communicating the perception and assessment of a system (i.e. patient) state and plan (i.e. plan of care) as well as "just in case" (or contingency) information [47, 48] to support events that may occur during the upcoming shift. Collaboration has also been described as a key factor in mitigating errors that occur during team related communications. In particular, the ability of individuals to evaluate the accuracy of each other's beliefs or behaviors, or collaborative cross-checking [49], has been shown to be a significant factor in error detection during team-based communications [50].

[51] qualitatively examined the handoff process on a general medicine service at the University of Chicago and found that communication failures during the handoff process frequently led to uncertainty during patient care decisions [51]. They identified two major types of communication failure during the handoff process. The most common communication failures they identified were content omissions in which critical information needed to care for a patient was not communicated during the handoff. Common examples included failure to report an active medical problem, failure to report a medication or other treatment, or failure to report pending or ordered diagnostic tests or consults. The other major type of communication failure they identified was a failure-prone communication process such as unclear or illegible written communication or lack of face-to-face communication. For the most part, the information that is exchanged and strategies employed during handoff are based on consensus, tradition, and personal or subjective opinion [6, 52].

Examinations of the content of resident handoffs

The "relevant information" to discuss at handoff may in fact vary greatly depending on characteristics of the individuals involved in the handoff (such as medical experience and knowledge of the patients), the condition of patients at the time of handoff (acuity, recency of admission, existence of new problems), and the amount, availability, and accessibility of patient data. Furthermore, one particular gap in the literature discussing relevant information are studies which look at the information needs of resident physicians during the course of patient care activities during their shifts following handoff.

Some research has shown that using simple qualitative rules is one of the best means to mitigate complex tasks and knowledge. Such simple rules may also be helpful in the presentation of information with respect to patient handoff. For example, researchers have made suggestions such as using mnemonics for the sharing of handoff information. [53] conducted a systematic review of

studies aimed at improving physician handoffs through mnemonics intended to introduce structure to the information presented during handoffs. The review uncovered 46 unique handoff mnemonics being used as a general outline for each patient [53]. However, only 4 out of 46 of the handoff mnemonics were accompanied by data to show their effect on the handoff process. The most frequently reported mnemonic for handoff information in the literature is SBAR, standing for Situation, Background, Assessment, and Recommendation [54].

SBAR is one particular model that is derived from theories of team cognition, and provides less structured approach to organizing information for each patient discussed [54]. According to [55], one of the most useful characteristics of the SBAR is that it can be applied in any clinical domain, making it valuable as means of information sharing among teams, particularly disconnected teams engaging in patient handoffs such as those that occur between departments (such as the ED to the inpatient setting) or between individuals with varying roles such as handoffs between nurses and physicians or handoffs that occur at discharge.

Beyond mnemonics: prescriptive handoff content models

[56] as well as [47], outline a minimum core set of information that should be handed off for each patient. These include identifying the patient, a short summary of the diagnosis (the "big sentence") and the patient's current condition, with emphasis placed on the plan of care, immediate tasks to complete, and contingencies for anticipated events. The last three elements (plan of care, tasks, and contingencies) are given emphasis as they cannot easily be determined from common data sources alone such as the patient chart. They are abstract representations of patient data, are based on medical experience, and are developed over time through discussions and increased knowledge of each patient. Not including such information in the handoff consequently may prevent work from being accomplished by the incoming team during their shift (assuming they were not present to help define these information elements).

Beyond mnemonics: information needs based handoff content models

Assessing information needs of residents to gain insight into what information should be discussed at handoff has also been attempted in an effort to guide the content of handoff discussions. Two studies in particular, have utilized observation of the questions residents are asked in an effort to characterize the information needs of residents [57, 58]. Of these two studies, however, neither linked data from actual handoff discussions to the information which later came up in questions residents received.

Training the content of handoff

Teaching residents what information to exchange with other physicians, both verbally and in clinical documentation, has been suggested as a means for reducing errors related to the information discussed during the handoff process [6, 31, 33, 59]. While most medical residency programs generally provide some form of medical communication training, most are not focused on teaching specific information content to exchange at handoff [8, 15, 37, 60]. The critical information to exchange is essentially "learned along the way" [12, 38-40]. Of the studies which have investigated the efficacy of handoff training specifically, few have employed on the job evaluation to assess transfer of training to residents' behavior during handoff. In particular, their ability to aggregate, summarize, or verbally communicate critical patient care information.

Transfer of training

The transfer of knowledge from training or instruction to on the job performance is the transfer of training. According to [61] the transfer of training is determined by three components: training inputs, training outputs, and conditions of transfer. The training inputs include training content, trainee characteristics, and work environment. Training outputs are elements that are measured by knowledge gained through training (learning) and the ability to recall the training material at a later time (retention). Conditions of transfer refers to the individual's ability to use material learned in training to improve performance on the job immediately (generalization), and over time (maintenance). In [61]'s model, the training inputs each have an effect on the training outputs, and the training outputs directly influence the conditions of transfer. They also suggest that, aside from the learning that occurs, trainee characteristics such as motivation, and work environment factors such as the opportunity to present learned skills, directly affect the conditions of transfer.

Chapter 3

Methodological Overview

This chapter focuses on the overall methods used to in this research and how the methods are linked to the research questions. Greater attention is given to the methods for each aim within their respective chapters.

Research setting and study population

This work focuses on the general pediatric wards at the University of Virginia Children's Hospital. The University of Virginia Children's Hospital is a university-based tertiary care hospital with a medium sized pediatric residency training program of 36 residents (12 per residency year). The Children's Hospital is comprised of approximately 100 pediatric faculty members, 70 of whom have their primary appointment in the Department of Pediatrics. It consists of 37 beds on two Acute Care Wards, one for infants and another for older children and adolescents. There is a Newborn Intensive Care Unit with 31 beds served by an air and ground neonatal transport system. There is also a 20 bed Newborn Nursery as well as a Transitional Nursery with eight beds. The hospital has a Pediatric Intensive Care Unit with 14 beds and a high acuity unit with 700 to 800 annual admissions. There are approximately 6,000 admissions to the inpatient units of the Children's Hospital each year.

The institutional review board of the University of Virginia has approved all methods in this study and all participants provided informed consent.

Coding audio recordings of resident handoff

Coding and evaluation of handoff audio recordings is a method common to the work supporting each aim: studying if/how information exchanged during handoff impacts residents' responses to questions received on the ensuing shift (Research Questions 1.1 and 1.2) and evaluating whether handoff behaviors of interns change after receiving handoff training (Research Question 2.1). The general method of collecting and coding these recordings is presented in this chapter. To address the specific aims of this research, four independent sets of audio recordings of verbal handoff

discussions were collected, along with observer handwritten notes and printed, de-identified copies of handoff sheets used by pediatric residents to organize patient data for the verbal handoff discussion and as an artifact for handwritten notes. Table 1 provides a summary of the overall characteristics each of the handoff audio recording data sets organized by the title of the study in which the data set was analyzed.

Table 1. Summary and description of handoff session audio recording data sets collected.

Handoff sessions observed and audio recording data		Aim 1	Aim 2		
		Pre QA[1] handoffs	Baseline	Pre-training	Post-training
Dates of collection	Start	6/21/2007	7/25/2005	1/27/2009	1/27/2009
	End	7/31/2007	2/16/2006	6/22/2009	6/22/2009
Setting[3]	Time	4pm handoff	4pm handoff	10am handoff	10am handoff
	Location	peds library	peds library	peds library	peds library
Care providers observed	Leaving[4]	PL-1 PL-2, PL-3	PL-1 PL-2, PL-3	PL-1	PL-1
	Incoming[4]	PL-1 PL-2, PL-3	PL-1 PL-2, PL-3	PL-1 PL-2, PL-3	PL-1 PL-2, PL-3
	Others	med. students chief residents attendings	med. students chief residents attendings	chief residents attendings	chief residents attendings
Audio recordings	Handoffs recorded	7	15	16	15
	Total patients	138	202	88	75
	M	19.7	14.2	5.5	5.0
	Sd	10.3	5.8	3.27	2.2
	Min	7	3	2	2
	Max	37	21	15	10
	Total duration[2]	4:10:11	9:08:58	5:57:29	3:28:27
	M	0:35:44	0:36:36	0:22:21	0:21:54
	Sd	0:24:22	0:12:03	0:18:11	0:07:15
	Min	0:16:24	0:09:09	0:04:25	0:03:13
	Max	1:28:07	0:53:55	1:22:44	0:30:39

[1] QA, *questions asked*, refers to questions on-call residents were asked overnight following handoff

[2] all measures of durations shown in Table 1 are in the format *h:mm:ss* (read *hours:minutes:seconds*)

[3] the *peds library* was primarily a room with a conference table, lockers, the chief resident offices, and a number of computer terminals, phones, and served as a common area, lounge, meeting place and the preferred location for almost every handoff observed during the course of this work

[4] PL-1, *first year residents,* PL-2, *second year residents*, PL-3, *third year residents*

Content analysis[62]

Content analysis is a structured approach to assessing and quantifying verbal interactions which has been applied in a variety of research settings including research involving the observation and characterization of various process elements of resident physician handoffs [62, 63]. In general, content analysis is a blanket term referring to any technique used for making inferences by identifying specific characteristics of messages contained within a communication process [64] . This broad definition covers textual analysis techniques such as Verbal Protocol Analysis (VPA), but also generalizes to the coding of actions observed, as well as non-verbal communication such as gesturing or drawing. The core purpose, however, remains constant regardless of the application. That is, content analysis seeks to answer the basic questions of who says what, to whom, why, to what extent, and with what effect [63]. The primary applications of content analysis are threefold: 1) to make inferences about the antecedents of communication, 2) to describe and make inferences about characteristics of communication, 3) to make inferences about the effects or consequences of communication [62, 63] .

As content analysis studies are often at risk to observer or experimenter bias due to their involvement in the development of a coding scheme or in the coding process of observational data, multiple raters are generally employed (subject matter experts whenever possible) in order to assess inter-rater reliability [65]. A common mistake in the calculation of inter-rater reliability is the use of percent agreement between raters which fails to account for inter-rater agreement which may have occurred by chance alone. Therefore, inter-rater reliability is best measured using a Kappa statistic, generally either Cohen's or Fleiss' [65]. Cohen's kappa measures the agreement between two raters who each classify N items into C mutually exclusive categories [65]. Fleiss' kappa works for any number of raters greater than two who give categorical ratings to a fixed number of items [66]. Table 2 summarizes the generally accepted value system for determining level of agreement between raters based on a calculated value for kappa.

Table 2. Standardized Kappa Scores [65]

Kappa	Relative Agreement
0.01 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

The analyses of coded audio data in this work include summary and descriptive statistics to compare handoff discussions based on the proposed aims of this work and are based on a software tool described in [67]. Specifically, frequency data to show how often each particular coding category was observed, as well as the percent of time each of the coding categories contributes to overall handoff session times. The use of duration as a measure has been used in empirical studies of handoff: 1) to describe the negative implications of the effects of increased handoff duration [45], and 2) as a measure of the difference between a pre-intervention handoff and a post-intervention handoff [68].

Questions residents receive on-call

The work conducted to address Aim 1 consists of an additional set of observational data which did not involve audio recordings. These data were collected during the time period beginning with the 4:00 p.m. week-day handoff sessions until approximately midnight of the ensuing "on-call" shift. In this study "observational data" refers to all data collected between a single 4:00 p.m. handoff and midnight of the ensuing shift. For each observation session, a graduate research assistant silently observed and audio recorded the 4:00 p.m. handoff and then shadowed the on-call senior resident until approximately midnight of their on-call shift. During the on-call shift, all verbal questions asked to the senior resident, the role of the question asker, how the resident responds to the question, and whether the resident provides an actionable answer to the question are documented. Any external sources the resident solicits for information are also noted as well as the sequence in which the resident solicits each source if the resident solicits more than one. To structure the data collection process, all observations were recorded using a paper template designed specifically for this type of data collection (Appendix F). The data were then reviewed and categorized using a content categorization scheme from prior work (Appendix H).

Frequency data are summarized for question categories, question askers, and external sources solicited. In addition, pair-wise frequency data are reported for question categories by question askers and question categories by information sources solicited. Sequence data for questions where multiple information sources were solicited are also reported.

Keystroke-level modeling in the assessment of handoff decisions support

The primary method for comparing decision support systems for handoff utilizes Keystroke-Level Modeling (KLM) [69], a task-analytic modeling technique for predicting the total time required to complete a common task across varying system designs. Based on task analyses, KLM is a method

for quantifying the time required to complete a task, or set of tasks, on the given system assessed in the task analysis, by dividing each main task into the most appropriate set of lowest-level sub-tasks which, when completed in a given sequence, result in completion of the main task. Sub-tasks may include such items as pointing with the mouse, pressing or releasing the mouse button, moving hands between keyboard and mouse, or pressing a key on the keyboard. The most common and frequently used sub-tasks (e.g. moving hand from keyboard to mouse) have been extensively studied in the literature and thus the amount of time to complete them is a standard value, most of which can be researched in text books or peer reviewed publications with a task-analytic theme.

Summary of Analyses Conducted within each Research Aim

Table 3 offers a graphical summary of the analyses conducted to address the research questions of each research Aim. In addition, it lists the manuscripts related to the work discussed in this dissertation. In total, four archival journal articles have either been published, are in preparation, or are in revision based on the studies described in the remaining chapters.

Chapter Summary

This chapter has described the general methods employed and the data collected in order to address the aims of this research. Methodological aspects such as study setting, participants, procedures and experimental apparatus, study hypotheses and statistical analysis methods for Aims 1 and 2 will be addressed more specifically in Chapters Four and Five, respectively. Chapter Six provides further detail regarding the specific research methods employed in addressing the Ancillary Research Objective.

Table 3. Analyses within and among the major data sets collected (and status of related publications)

Aim	Research Question	Critical information for handoff (Chapter 4)		Resident Handoff Training Study (Chapter 5)						
				Baseline		Pre-Training		Post-Training		
		coded Q's asked	coded audio recordings	coded audio recordings	knowledge assessment	coded audio recording set 1	coded audio recording set 2	knowledge assessment	coded audio recording set 3	coded audio recording set 4
Aim 1	RQ 1.1: Coding and descriptive analysis of post-handoff questions asked (1)	●								
Aim 1	RQ 1.2: Linking questions asked to patients in prior handoffs and modeling of combined handoff discussion/questions data (2)		●							
Aim 2	RQ 2.1: comparison of coded baseline audio recordings to coded pre-training recordings (to mitigate threat of maturation effect given no true control group) (3)			●———————————●		●				
Aim 2	RQ 2.1: Comparison of coded pre-training recording set 1 to coded pre-training recording set 2 (to mitigate threat of maturation effect given no true control group) (3)					●———●	●			
Aim 2	RQ 2.1: Comparison of coded pre-training recordings to coded post-training recordings (to test hypotheses regarding impact of training module on handoff behaviors) (3)					●	●—————————————————●		●	

(1) **Bass, E.J., DeVoge, J.M., Waggoner-Fountain, L.A., & Borowitz, S.M. (2012). Resident physicians as human information systems: sources yet seekers. *Journal of the American Medical Informatics Association.***

(2) **DeVoge, J.M, Bass, E.J., Borowitz, S.M., & Waggoner-Fountain, L.A. Continuity of care through continuity of content: how the information residents exchange during handoff is utilized on call. *Unpublished manuscript.***

(3) **DeVoge, J.M, Bass, E.J., Waggoner-Fountain, L.A., & Borowitz, S. Objective evaluation of a handoff training curriculum. *Unpublished manuscript***

Chapter 4

Evaluating critical content to discuss at handoff

Introduction

In health care, problems with information management are often associated with problems with technology, and proposed solutions frequently involve the move from one technologically-based system to another [19, 20] as illustrated by the emerging transition from paper to electronic medical records (EMR) [21]. While technology provides a means for augmenting human cognitive capacity in domains faced with a burgeoning set of pertinent information [22], as a whole, managing information is not simply a matter of storage and retrieval of data. Though the storage and retrieval of information in health care are problematic issues in their own right, complex processes such as communication, aggregation, summarization, and conversion are necessary in order to turn information into knowledge that is useful for health care providers [23].

While technology to support processes such as communication is improving, its capability to replace many of the intricacies that exist within the complexity of such processes remains limited [21, 23, 24]. Furthermore, there is a growing body of evidence in health care of the inefficiencies and hazards resulting from the implementation of inappropriate, "technology-heavy" solutions which exacerbate human information management problems [25-27]. Understanding cognitive and behavioral as well as technological aspects of problems related to data storage, search, retrieval, and dissemination, is critical to designing solutions, given that both humans and technology frequently perform such function [19, 23]. For example, humans interact with other information systems, both human and non-human, to find, share, and exchange information. Despite being critical aspects of patient care, such responsibilities are generally embedded within the health care system, are not clearly defined or outlined by organizational policies, which tend to prioritize the more obvious, explicable responsibilities associated with specific occupational roles (e.g. physicians diagnose, pharmacists fill orders, etc.), and have remained relatively unsupported by information technology [23, 24].

Resident physicians are an example of a role where further examination of the embedded responsibilities which exist in health care is needed, particularly within their role as a human information system serving the information needs of others.

The goal of this study is to understand two specific aspects related to this role: first, to understand who asks residents for information, and second, to understand the methods residents employ in response to questions they are asked by others. A greater understanding of these issues can help inform the design of the EMR and other interventions that support the information needs of health care providers. The specific research question addressed in relation to this objective are:

- <u>Research Question 1.1</u>: What are the characteristics of the information requests residents receive from other care providers (including patients' family members) following handoff?

There is a lack of empirical evidence to support exactly what information *should be discussed* during handoff, this study addresses this gap. More precisely, it seeks to address the question of how the information discussed at handoff relates to the information requests of other care providers, the sources residents used to seek information they needed to provide a response, and whether a response was given.

- <u>Research Question 1.2:</u> What factors of the information discussed at handoff correlate with residents' ability to respond to the information requests of other care providers?

Research Question 1.1

Method

Study Setting

The study was conducted on two adjacent general pediatric acute care wards at the University of Virginia Children's Hospital, a university-based tertiary care hospital with a medium sized pediatric residency training program.

At the time of the study, the general pediatrics service consisted of three first year pediatric and/or family medicine residents (PGY-1), also called interns, two third year pediatric residents (PGY-3), also called senior residents, and a pediatric attending physician. At the time of the study, each patient admitted to the ward was assigned to both an intern and third year senior resident. The intern served as the patient's primary care provider while the senior resident supervised. The daytime care team consisted of two second or third year pediatric residents (PGY-2 or PGY-3), three interns (PGY-1), and the attending physician. The day team rounded together every morning. Night coverage was shared by eight residents: five on the primary care team (three interns and two PGY-3 senior residents on general ward rotations), and the cross-cover team (two PGY-2 senior residents and one intern on other pediatric rotations).

Data Collection

This study focused on the time period beginning with the 4:00 p.m. week-day handoff sessions until approximately midnight of the ensuing "on-call" shift. Henceforth, an "observation session" means all data collected between a single 4:00 p.m. handoff and midnight of the ensuing shift.

For each observation session, a graduate research assistant silently observed the 4:00 p.m. handoff and then shadowed the on-call senior resident until approximately midnight of their on-call shift. Observation sessions also partially overlapped with the hospital's visiting hours for patients (9am to 9pm daily).

During this time family members also have the opportunity to interact with residents and ask questions. For each observation session, the graduate research assistant documented all verbal questions the senior resident was asked, the role of the question asker, how the resident answered the question, and whether the resident provided an actionable answer or did not provide an answer.

Any external sources residents solicited for information were also noted. If the resident solicited more than one source, the sequence in which the resident solicited each source was also noted.

Descriptions and abbreviations for a subset of the external sources available to residents for soliciting information are provided in Appendix E. To structure the data collection process, the graduate research assistant recorded all observations using a data collection worksheet designed specifically for this study (Appendix F). The worksheet data were subsequently entered into a custom Microsoft Access database.

Question Categorization

Two medical faculty members and one systems engineer reviewed and categorized each question independently using a content categorization scheme (Appendix H) from prior work [70]. After independently categorizing the set of questions, the group convened to determine final "consensus categories" for each question (i.e. agreement between at least two of the three reviewers). Group discussion was encouraged to help reach consensus categories for questions on which reviewers initially disagreed.

Data Analysis

Data in the following section are presented based on the method in which residents responded to questions, specifically: no response, responded by doing/showing, responded with clinical judgment only, and responded by soliciting information from external sources. Summary and count data are reported for question categories, question askers, and external sources solicited. In addition, pair-wise summary and count data are presented for question category by question asker and question category by information source solicited. Order data for questions in which multiple information sources were solicited are also presented.

Results

Observations

30 observations involving ten individual senior residents (five PGY-2 and five PGY-3) were conducted during the study. Some residents were observed on multiple occasions due to varying schedules over the six-week observation period.

Total questions

Over the course of the thirty observations, senior residents were asked a total of 399 questions; total questions asked per observation ranged from 4 to 30 (M = 13.3, Mdn = 13.5). The analysts determined a consensus category for 397 of the 399 questions. The two remaining questions were categorized as "unknown" and are not included in subsequent analyses.

Question content

Figure 7 shows a summary of the categories of questions which residents were asked across all 30 observations.

The greatest proportion of questions residents were asked concerned the plan of care for patients (158, 39.8%). More plan of care questions were asked than the next three most frequently occurring question categories combined: medical knowledge (53, 13.4%), systems-based practices (48, 12.1%), and the current condition of patients (44, 11.1%).

Total Questions Asked by Question Category (N = 397)

	plan of care	medical knowledge	system-based practice	current condition	past condition	task in-progress	rationale for plan of care	medical practice	technical knowledge	request
Total	158	53	48	44	21	18	16	14	13	12
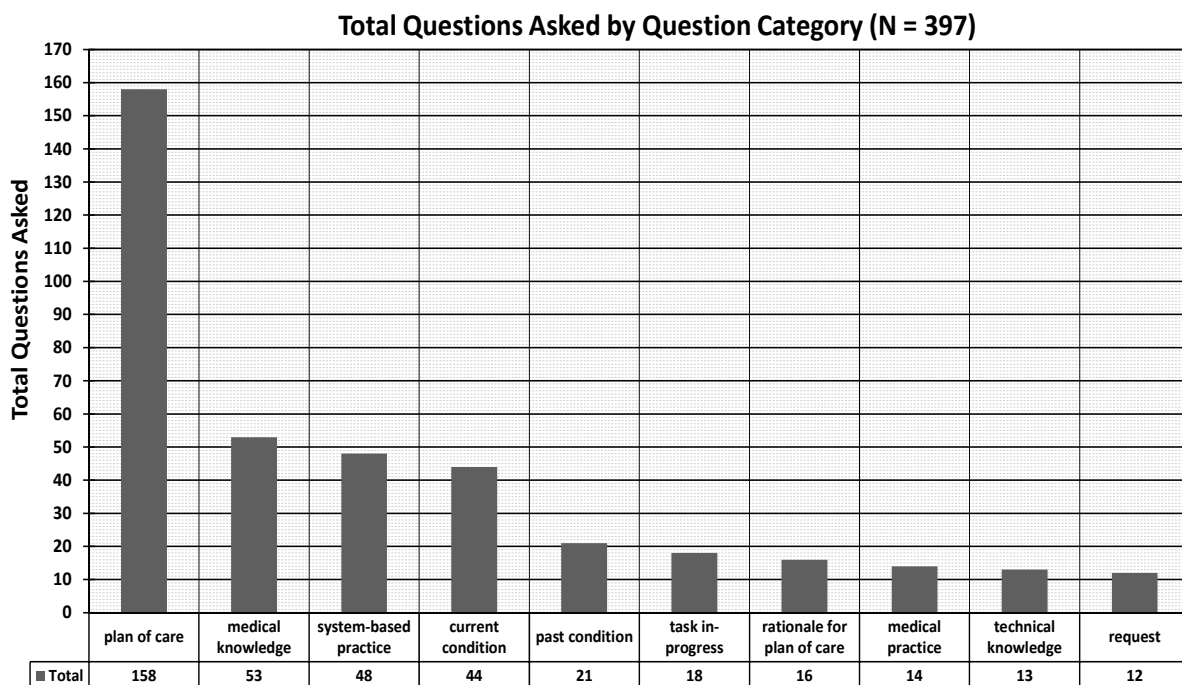
Figure 7. Total questions asked by question category

Question askers

Figure 8 provides a summary of the roles of those who asked residents questions across all 30 observations. The majority of questions were asked by individuals with prior medical training and patient care responsibilities, namely interns (115, 29.0%) and nurses (103, 25.9%). The patient's family accounted for 82 (20.7%) of the remaining questions asked followed by medical students (45, 11.3%), members of the primary care team (30, 7.6%), cross-covering care providers (18, 4.5%), and other hospital roles (4, 1.0%).

Total Questions Asked by Question Asker (N = 397)

	intern	nurse	family	medical student	primary care team	cross-cover team	other
Total	115	103	82	45	30	18	4
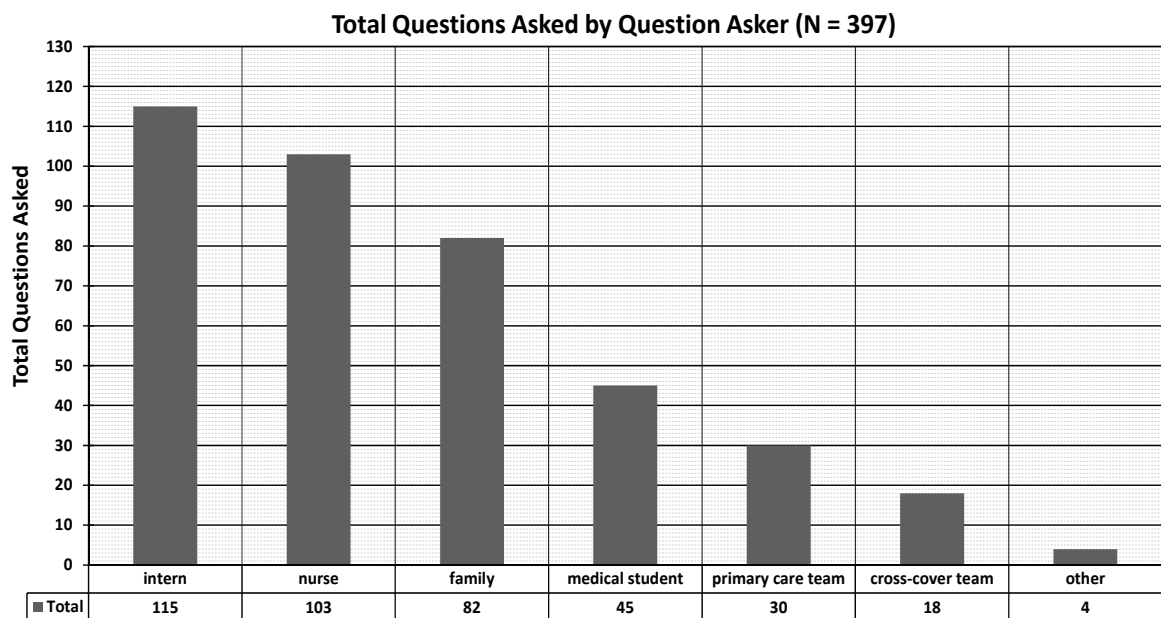
Figure 8. Count of questions asked by question asker

Question content by question asker

Figure 9 summarizes question categories across types of question askers across the 397 questions residents were asked. Nurses and interns asked the most plan of care questions, 63 (39.9%) and 44 (27.8%) of the 158 plan of care questions, respectively. A Z-test comparing the proportion of plan of care questions asked by nurses to that of interns showed that the 63 asked by nurses was a significantly larger proportion of plan of care questions than those asked by interns ($Z = 2.28, p < 0.05$). Overall, nurses, interns and the patient's family accounted for 137 of the 158 (86.7%) plan of care questions asked.

Patients' family members and medical students asked the greatest number of medical knowledge questions as well as questions concerning the rationale for patients' plan of care. Interns asked the majority of questions pertaining to SBP as well as the most technical questions while nurses made five times more requests of residents than all of the other question askers combined (Figure 9, following page).

Resident responses to questions

Residents responded in some way to all of the questions they received. Overall, residents provided an actionable response (labeled as *answered*) for 343 of the 397 questions they were asked (86.4%), and were unable to provide an actionable response, answer, or indicated they did not know an answer (categorized as *unanswered*) for the remaining 54 (13.6%) (Table 4).

Residents provided actionable responses for a majority of the questions in each question category, however, the proportion of questions for which residents did not provide actionable responses (i.e. questions labeled as *unanswered*) was comparatively high for questions concerning the current condition of a patient (11 of 44 *unanswered,* 25.0%) and SBP questions (8 of 48 *unanswered*, 16.7%). In each of these two cases, the questions were primarily evenly distributed with respect to question askers (Table 4).

Table 4. Summary counts of all categorized questions by category, question asker, and question status (answered or unanswered)

question category	all question askers			intern			nurse			family			student			primary care team			cross-cover team			other		
	Σ	a^1	u^2	Σ	a^1	u^2	Σ	a^1	u^2	Σ	a^1	u^2	Σ	a^1	u^2	Σ	a^1	u^2	Σ	a^1	u^2	Σ	a^1	u^2
plan of care	158	141	17	44	38	6	63	59	4	30	24	6	7	6	1	9	9	--	5	5	--	--	--	--
current condition	44	33	11	10	8	2	7	5	2	6	5	1	8	5	3	7	7	--	5	2	3	1	1	--
past condition	21	19	2	8	8	--	3	2	1	3	3	--	5	5	--	2	1	1	--	--	--	--	--	--
rationale	16	13	3	2	2	--	3	2	1	4	4	--	4	3	1	1	--	1	2	2	--	--	--	--
Σ	239	206	33	64	56	8	76	68	8	43	36	7	24	19	5	19	17	2	12	9	3	1	1	--
medical knowledge	53	46	7	8	6	2	5	5	--	23	21	2	13	12	1	1	1	--	2	--	2	1	1	--
SBP	48	40	8	20	17	3	6	4	2	8	7	1	5	4	1	5	4	1	2	2	--	2	2	--
medical practice	14	13	1	4	4	--	2	2	--	6	5	1	2	2	--	--	--	--	--	--	--	--	--	--
technical knowledge	13	11	2	10	8	2	--	--	--	--	--	--	1	1	--	1	1	--	1	1	--	--	--	--
Σ	128	110	18	42	35	7	13	11	2	37	33	4	21	19	2	7	6	1	5	3	2	3	3	--
task in-progress	18	16	2	8	8	--	4	4	--	1	--	1	--	--	--	4	3	1	1	1	--	--	--	--
request	12	11	1	1	1	--	10	9	1	1	1	--	--	--	--	--	--	--	--	--	--	--	--	--
Σ	30	27	3	9	9	--	14	13	1	2	1	1	--	--	--	4	3	1	1	1	--	--	--	--
Σ	397	343	54	115	100	15	103	92	11	82	70	12	45	38	7	30	26	4	18	13	5	4	4	0

[1] *answered*, resident provided an answer to the question
[2] *unanswered*, resident did not provide an answer to the question

Questions Asked by Category and Question Asker (N = 397)

	plan of care	medical knowledge	system-based practice	current condition	past condition	task in-progress	rationale for plan of care	medical practice	technical knowledge	request
intern	44	8	20	10	8	8	2	4	10	1
nurse	63	5	6	7	3	4	3	2	0	10
family	30	23	8	6	3	1	4	6	0	1
medical student	7	13	5	8	5	0	4	2	1	0
primary care team	9	1	5	7	2	4	1	0	1	0
cross-cover team	5	2	2	5	0	1	2	0	1	0
other	0	1	2	1	0	0	0	0	0	0

Total Questions Asked

Figure 9. Summary of question categories by question asker across all 397 questions asked

Responding with clinical judgments

For 242 of the 397 questions asked (61.0%), residents' responded to the question immediately using their clinical judgment without soliciting information from additional, external information sources (these 242 questions are labeled "clinical judgment only" below for simplicity). 99 of these 242 questions (41.0%) concerned the plan of care, 41 (16.9%) concerned medical knowledge, and 28 (11.6%) concerned local system-based practices (SBP). Overall, using only their clinical judgment, residents provided actionable responses (i.e. *answered* the question) for 227 of the 397 questions they were asked (57.2%).

For 16 of the 397 questions (4.3%), residents responded by modeling or completing a task for the asker of the question. These questions were primarily technical questions (7 of 16, 43.8%), and requests (3 of 16, 18.8%). Six of the seven (85.7%) technical questions were asked by an intern and all three requests arose from a nurse.

For 21 of the 397 questions (5.3%), residents immediately responded by saying they did not know an answer or would not be able to answer the question and made no further attempts to address the question. The majority of these questions concerned the plan of care for a patient (14 of 21, 66.7%) with the remainder related to patients' current conditions (7 of 21, 33.3%).

Consultation of external information sources

118 of the 397 questions (29.7%) led residents to solicit information from external sources in the course of responding. 50 of these 118 questions concerned the plan of care (42.4%), 19 concerned the current condition of patients (16.1%), and 16 concerned SBP (13.6%). Of the remaining seven question categories, none accounted for more than 10% of the 118 questions where residents solicited external information sources (Table 5).

Residents solicited one external information source in response to 77 of 118 questions (65.3%). 54 of these 77 questions (70.1%) related to either the plan of care, the current condition of a patient, or SBP. For these 54 questions, the most frequently solicited external sources were phone calls to other physicians (22 of 54 questions, 40.7%), and handoff sheets (10 of 54 questions, 18.5%), followed by progress notes, nurses, and calls to other non-physicians. Residents provided an actionable answer for 67 of the 77 questions (87.0%) when only one external source was solicited for information, most often after calling other physicians (24 of 77 questions, 31.2%), checking handoff sheets (11 of 77 questions, 14.3%), checking progress notes (7 of 77 questions, 9.1%), or

asking a nurse (6 of 77 questions, 7.8%). Residents did not provide an answer for three of seven (42.9%) questions after soliciting information by calling other non-physicians, respectively.

Table 5. Categorical totals for questions where residents consulted external sources of information in total and by consultation of a single external information source and more than one external information source

question category	external sources > 0*			external sources = 1*			external sources > 1*		
	Total (%)	An¹ (%)	Un² (%)	Total (%)	An¹ (%)	Un² (%)	Total (%)	An¹ (%)	Un² (%)
plan of care	50 (42.4)	47 (39.8)	3 (2.5)	26 (22.0)	25 (21.2)	1 (0.8)	24 (20.3)	22 (18.6)	2 (1.7)
current condition	19 (16.1)	16 (13.6)	3 (2.5)	14 (11.9)	12 (10.2)	2 (1.7)	5 (4.2)	4 (3.4)	1 (0.8)
systems-based practice (SBP)	16 (13.6)	11 (9.3)	5 (4.2)	14 (11.9)	10 (8.5)	4 (3.4)	2 (1.7)	1 (0.8)	1 (0.8)
medical knowledge	10 (8.5)	8 (6.8)	2 (1.7)	5 (4.2)	4 (3.4)	1 (0.8)	5 (4.2)	4 (3.4)	1 (0.8)
past condition	9 (7.6)	8 (6.8)	1 (0.8)	7 (5.9)	6 (5.1)	1 (0.8)	2 (1.7)	2 (1.7)	--
rationale	6 (5.1)	4 (3.4)	2 (1.7)	4 (3.4)	3 (2.5)	1 (0.8)	2 (1.7)	1 (0.8)	1 (0.8)
task in-progress	5 (4.2)	5 (4.2)	--	5 (4.2)	5 (4.2)	--	--	--	--
request	1 (0.8)	1 (0.8)	--	--	--	--	1 (0.8)	1 (0.8)	--
technical knowledge	1 (0.8)	1 (0.8)	--	1 (0.8)	1 (0.8)	--	--	--	--
medical practice	1 (0.8)	1 (0.8)	--	1 (0.8)	1 (0.8)	--	--	--	--
Total	**118 (100)**	**102 (86.4)**	**16 (13.6)**	**77 (65.3)**	**67 (56.8)**	**10 (8.5)**	**41 (34.8)**	**35 (29.7)**	**6 (5.1)**

* Note, data in the first column, *external sources > 0*, represent combined data from the two rightmost columns (*external sources = 1*, and *external sources > 1*), alternatively, view the two rightmost columns as mutually-exclusive sub-groups of the left.

[1] An, *answered*, resident provided an answer to the question or an actionable response to the asker, or the resident performed or modeled how to perform a task or completed the requested action

[2] Un, *unanswered*, resident did not provide an answer to the question or no actionable response given or overtly indicated they did not know the answer or could not find an answer at that time

Multiple external information sources

In response to 41 of the 118 questions (34.7%), residents solicited information from more than one external source (Table 5), and provided an actionable response for 35 of the 41 (85.4%).

With respect to content, 24 of the 41 (58.5%) questions related to the plan of care, and residents provided answers for 22 of these (91.7%). Residents most often called other physicians, or checked the EMR (see Appendix E) or handoff sheet as the first source when soliciting more than one source of information. Nurses were never the first of multiple external sources. Residents frequently called other physicians as a first preliminary source, as well as second or third final sources.

With respect to the order in which residents consulted external sources for information, they most often called other physicians or checked the computerized provider order entry system (CPOE; see Appendix E) second, and most often called other physicians third.

Soliciting more than three external sources occurred only once (Table 6). Phone calls to other physicians and the EMR were most often the final source when residents solicited more than one external source for information. They rarely accessed handoff sheets, progress notes, the EMR or attending physician notes as the final source for information. Nursing notes and phone calls to non-physicians were never the final source of information they solicited. In general, residents rarely solicited information from attending physician notes (Table 6).

Table 6. External sources by order consulted for more than one external source (N = 41 questions)

	preliminary source						final source					
	1st n = 41 questions			2nd n = 15 questions			3rd n = 1 question			2nd n = 26 questions			3rd n = 14 questions			4th n = 1 question		
external source of information	Σ	%n , %N		Σ	%n	%N	Σ	%n	%N	Σ	%n	%N	Σ	%n	%N	Σ	%n	%N
phone call - other physician	7	17.1%		1	6.7%	2.4%	--	--	--	5	19.2%	12.2%	6	42.9%	14.6%	--	--	--
mis	1	2.4%		7	46.7%	17.1%	--	--	--	5	19.2%	12.2%	2	14.3%	4.9%	1	100%	2.4%
carecast	7	17.1%		1	6.7%	2.4%	1	100%	2.4%	3	11.5%	7.3%	--	--	--	--	--	--
chart - progress note	5	12.2%		2	13.3%	4.9%	--	--	--	2	7.7%	4.9%	2	14.3%	4.9%	--	--	--
sign-out sheet	7	17.1%		--	--	--	--	--	--	2	7.7%	4.9%	--	--	--	--	--	--
other non-chart document	1	2.4%		1	6.7%	2.4%	--	--	--	4	15.4%	9.8%	1	7.1%	2.4%	--	--	--
nursing chart	4	9.8%		--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
chart - attending note	1	2.4%		2	13.3%	4.9%	--	--	--	1	3.8%	2.4%	--	--	--	--	--	--
phone call - other non-physician	2	4.9%		--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
nurse	--	--		1	6.7%	2.4%	--	--	--	--	--	--	1	7.1%	2.4%	--	--	--
clinical judgement	6	14.6%		--	--	--	--	--	--	4	15.4%	9.8%	2	14.3%	4.9%	--	--	--
Σ	41	100%		15	100%	36.6%	1	100%	2.4%	26	100%	63.4%	14	100%	34.1%	1	100%	2.4%

For six of the 41 questions (14.6%), residents first provided information using their clinical judgment but followed this by soliciting information from an external information source. In all six of these instances, residents provided an answer immediately after consulting the first external information source. For a separate six questions (14.6%), residents first solicited information from either one or two external sources prior to using their clinical judgment to formulate a final response (Table 6). Actionable responses were provided for five of these six questions (83.3%).

Research Question 1.2

Review

There are many reasons why handoff content is critical to patient care, though very few studies have been successful in objectively determining "essential content elements" for handoff based on, for example, information residents most need or use while caring for patients. There are cautions in the literature against "over-standardization" of handoff [71], although this is an area still largely under debate [32]. Given that the primary objective of handoff is the accurate transfer of information relevant to patient care [31], this section looks at how information discussed or not discussed at handoff related to how resident physicians addressed the information requests of other providers.

- ▪ Research Question 1.2: What factors of the information discussed at handoff correlate with residents' ability to respond to the information requests of other care providers?

Method

Audio-recordings and written transcriptions for 7 of the 30 handoffs which immediately preceded the on-call observation sessions in Research Question 1.1 provide the foundation for the analyses of Research Question 1.2. Written transcriptions of handoff audio recordings as opposed to directly coded audio recordings were used as typed transcriptions afforded the ability to quickly and easily search through a large amount of verbal data for keywords and phrases which could be utilized to match each question asked to a patient discussed during handoff as true patient identifiers for the questions asked were not available.

Links to a patient discussed at handoff could not be established for every question. There are many reasons for this including questions being too general, questions in the form of requests unrelated to the care of a patient, the majority of the knowledge questions such as medical, technical, and SBP (systems-based practice) questions, which frequently resulted in residents responding with a teaching event. In other cases, knowledge questions were clearly in response to care activities and if enough information was available to determine the associated patient, the link was established. Printed handoff sheets from the electronic handoff information system used by residents were also collected at the time of audio recording. The handoff sheets offered additional information used to guide the process of linking patients to questions. The handoff sheets were also the only available data source listing the acuity of each patient. In general the linking process involved successive iterations of determining keywords in questions, scanning and using the search functions of text editors on the audio recording transcripts and handoff sheets, and eliminating ties based on

identifiable exclusion criteria such as the appearance of the words "he" or "she", patient age, names of rare illnesses, specific medication doses, and other unique data, words, or phrases.

Determining if the content of the question was discussed at handoff

After a link had been established between a patient discussed at handoff and a question received by the resident overnight, it was necessary to determine whether the content of the question had been discussed for the patient during handoff.

A three-level categorization was used to denote whether the information content of the question was discussed at handoff. If the specific content of the question was also specifically discussed during handoff, it was given the value *discussed specifically*. If the general content category of the question (e.g. *plan of care*) was an element of the discussion for the patient during handoff but the specific content of the question was not a part of that discussion, it was given the value *discussed generally*. If the content of the question was not discussed either generally or specifically, it was given the value of *not discussed* at handoff.

Presentation of results

Results summarizing the handoff discussions are presented first with a summary of the total questions asked overnight following the discussion, the number of questions which were linked to a patient, a summary of the outcome of the linked questions (response – yes or no), and the total questions which were responded to by a resident using their clinical judgment.

Next, summary data of the question outcome, information source, and whether the information content of the question was discussed during handoff are presented for each information category. This is followed by a presentation of key results related to the linked data. Where appropriate, Chi-square tests of proportions were conducted in addition to descriptive data summaries.

Results

Summary of linked questions

Table 7 summarizes data about these seven handoffs, presents results from the question to patient linking process, and gives a breakdown of the binary outcome for the questions which were linked to a patient (the resident provided an *actionable response* or *no actionable response*).

In total, 100 questions were asked during the shifts following the seven handoffs which were observed, recorded, and transcribed. A link to a specific patient was established for 82 of the 100 questions. Of those 82 questions, residents provided actionable responses to 65 and no actionable response to 17. Residents responded to 47 linked questions using their clinical judgment (Table 7).

Table 7. Summary of handoff recording times and corresponding questions asked for RQ 2.2

Handoff date	Total patients	Handoff duration[1]			Questions asked overnight				
		Total	*M*	*sd*	*Total*	*link[2]*	*resp[3]*	*no resp[4]*	*respCJ[5]*
6-28-2007	16	30.63	1.91	1.41	14	14	8	6	6
7-02-2007	24	34.30	1.43	0.72	15	11	9	2	8
7-03-2007	37	88.12	2.38	1.86	18	16	16	0	5
7-05-2007	20	23.55	1.18	1.06	13	8	7	1	6
7-06-2007	25	37.67	1.51	1.39	14	12	10	2	8
7-09-2007	9	16.40	1.82	1.21	20	16	11	5	10
7-20-2007	7	19.52	2.79	1.52	6	5	4	1	4
Total	**138**	**250.18**	**1.81**	**1.46**	**100**	**82**	**65**	**17**	**47**

[1] discussion times shown in *minutes*
[2] *link,* questions for which a link could be established to a patient discussed during the preceding handoff
[3] *resp,* questions which the on-call resident responded to
[4] *no resp,* questions which the on-call resident did not respond to
[5] respCJ, the resident responded to the question using their clinical judgment

Analysis of linked handoff discussion/questions asked data

Table 8 summarizes the number of questions which were asked by category, response type, and level of discussion of the question content at handoff.

Close to half of the linked questions asked (40 of 82) were discussed specifically at handoff, a significantly greater proportion than those discussed generally (22 of 82; $\chi^2 = 7.495, p < 0.01$) or not discussed at handoff (20 of 82; $\chi^2 = 9.488, p < 0.005$).

Overall, residents responded to 47 linked questions using their clinical judgment, 24 of which (29.3% of the linked questions) were discussed specifically at handoff, a significantly greater proportion than those which were discussed generally at handoff (11 of 82; $\chi^2 = 5.231, p < 0.05$) as well as those which were not discussed at handoff (12 of 82; $\chi^2 = 4.306, p < 0.05$).

Plan of care/rationale for plan of care questions

39 linked questions concerned the plan of care or rationale for the plan of care. Residents responded with their clinical judgment for 23 of the 39 plan of care/rationale for the plan of care questions, 16 of which were discussed specifically at handoff, a significantly greater proportion than the 7 which were either not discussed or discussed only generally at handoff ($\chi^2 = 4.096, p < 0.05$).

Table 8. Summary of the number of questions which were asked and whether the information from the question was discussed at handoff grouped by response or the information source consulted by the resident to provide a response

Responded / info source	Information category	Level of discussion at handoff			
		generally	specifically	not discussed	Total
Clinical judgment	Plan of care	2	13	5	20
	Local SBP	3	3	1	7
	Medical knowledge	5	--	2	7
	Rationale for plan of care	--	3	--	3
	In process	--	2	1	3
	Medical practice	1	1	--	2
	Past condition	--	1	1	2
	Current condition	--	--	2	2
	Request	--	1	--	1
	Subtotal	11	24	12	47
No response	Plan of care	5	2	1	8
	Local SBP	1	1	1	3
	Current condition	--	--	2	2
	Past condition	--	1	--	1
	Request	--	1	--	1
	Medical knowledge	--	--	1	1
	Rationale for plan of care	--	--	1	1
	Subtotal	6	5	6	17
Phone another physician	Plan of care	1	3	1	5
	Rationale for plan of care	--	2	--	2
	Past condition	1	--	--	1
	Medical knowledge	--	1	--	1
	Local SBP	--	1	--	1
	Subtotal	2	7	1	10
Phone non-physician	Technical knowledge	--	--	1	1
	Local SBP	--	1	--	1
	Subtotal	--	1	1	2
Nurse	In process	2	--	--	2
Performed task	Technical knowledge	1	--	--	1
MIS	In process	--	1	--	1
Handoff sheet	Local SBP	--	1	--	1
Other non-chart document	Request	--	1	--	1
Grand Total		**22**	**40**	**20**	**82**

For 9 of the 39 plan of care/rationale for plan of care questions the resident did not provide an actionable response, seven of which were not discussed specifically at handoff (five discussed generally, two not discussed). In contrast, the opposite trend was true for medical knowledge questions. For example, of the seven questions residents responded to using clinical judgments

which concerned medical knowledge, none were specifically discussed at handoff (5 generally, 2 not discussed).

Phone calls to other physicians

Residents made a phone call to another physician in response to 10 of the linked questions, seven of which were discussed specifically at handoff, five of which concerned the plan of care or rationale for the plan of care.

Discussion

Residents as Information Systems

In addition to providing patient care, resident physicians serve as an information system that receives and responds to questions from other health care providers as well as from patients' families. Residents do not always respond to clinical questions, however. Similar studies have shown that residents do not always attempt to answer clinical questions, though most have focused on residents' response to their own clinical questions as opposed to the questions they are asked by others [19, 72-74]. In this study, more often than not, residents did provide actionable responses to questions. The majority of the time residents were able to respond using their own clinical judgment without soliciting external sources for information. When residents did not immediately respond using a clinical judgment, they sometimes completed a task implied by the question or showed the asker how to do it on their own, indicated that they did not know the answer to the question, or attempted to solicit information from one or more external sources. Completion or modeling of tasks or indicating they did not know were rarely observed as responses from the residents, however. Residents most often chose to solicit information from external sources if they could not immediately respond using their clinical judgment.

With respect to handoff, there appeared to be a clear effect of the level at which questions were discussed and the responses residents provided to specific types of questions later in the night. That is, responses to certain question categories appear to be influenced heavily by the level at which they are discussed at handoff. For example, residents responses to plan of care questions appeared to be heavily influenced by whether the content was discussed specifically at handoff or not. When the specific content for a plan of care question was discussed specifically at handoff residents frequently provided actionable responses using their clinical judgments. In contrast, when not discussed specifically (i.e. either generally or not at all) residents most frequently did not provide actionable responses to plan of care questions. The data make a strong case for inclusion of the plan

of care as an essential element to be communicated during handoffs. First, the majority of the questions posed to residents related to a patient's plan of care and were most often asked by someone whose responsibilities would include direct patient care (e.g. interns, nurses, medical students, and family members). The majority of these individuals, in addition, could have been responsible for patient care activities at the time the question was asked. Furthermore, the results of this study suggest that the plan of care is a critical item which should be discussed *at a detailed level* at handoff. In particular, as general discussion of the plan of care had the same result as not discussing the plan of care at all with respect to residents' ability to address the plan of care questions of others.

On the other hand, even specific discussions of critical elements such as the plan of care do not always yield successful outcomes. There were a number of questions which residents did not respond to even though the content was discussed specifically at handoff. Unfortunately, there are infinite reasons why this may have occurred which were not accounted for in the study such as time constraints, task prioritization by the resident, the information may have changed since it was handed handoff, or possibly the resident did not remember that it had been discussed at handoff by the time the question was asked. It is beyond the scope of this study to say whether issues related to memory or attention affected residents' retention of information exchanged during handoffs, though this is a direction for future work. Future work could utilize an on the spot two minute drill to evaluate the incoming resident (receiver of handoff) immediately after handoff in order to gauge their understanding of details such as the plan of care. Such a quiz or quick drill could be repeated multiple times over the course of a shift to develop a time-line or even a metric representing the rate of information decay following handoff or simply to highlight possible differences between the information verbally communicated at handoff and residents' internal representations of that information.

Verbal communication between providers

The plan of care is information which can be abstract, role-specific, dynamic, and while overtly shared between providers internalized very differently, not to mention, difficult to document in paper or electronic medical records [75]. This is a possible explanation why residents, when necessary, consulted other human sources to find information about the plan of care, and perhaps why such questions were sometimes left unanswered. Even when topics such as the plan of care were discussed specifically at handoff, residents still phoned other physicians in search of the information. Though the telephone conversations were not recorded in this study, the data suggest

that these calls frequently led to the information residents were seeking. While the data do not say for certain what was discussed during these phone calls, it is clear that residents frequently seek information related to the plan of care through verbal interactions with other physicians, such as those which occur during handoff. In general, verbal communication appears to be a preferred method for residents when exchanging information relevant to patient care.

Predicting questions asked as a preparatory activity for handoff

Family members and medical students frequently asked questions concerning medical knowledge. This is not surprising as medical students are still learning medicine and the patient's family may not have formal medical training. In a similar fashion, interns asked the most SBP and technical questions. Though they are expected to have a greater knowledge of medicine than medical students, for example, interns, particularly those from outside the organization, must become familiar with the way work is conducted in a specific organization. With respect to medical students' and interns' questions, these are also examples of the informal nature in which teaching and learning occur on the job. Some authors have suggested that teaching is an integral function of handoffs and have advocated for its inclusion in them [86]. The data do not support or negate this claim, but provide evidence that teaching opportunities seem to occur as necessitated by the characteristics of the situation, that is, the people involved and the subject matter in question.

Nurses and interns asked the most plan of care questions presumably because they are performing care duties but may not have been involved in the formulation of the patient's plan of care. This is evidence which suggests making the plan of care accessible across hospital roles in addition to information which is discussed at handoffs. Finally, nurses made the majority of requests, perhaps because of the hierarchical nature of responsibility/authority in the hospital (e.g. requiring a physician signature or an order to be written). Further analysis of the nature of these requests may provide evidence for organizations to increase the responsibility given to certain hospital roles such as nurses. In general, it appears there is an advantage with respect to preparation if residents can anticipate the roles of those they will encounter during their shift.

The current condition of patients and infrequently utilized information sources

Residents frequently consulted external sources of information when they were asked questions about the current condition of patients. For 25% (11 of 44) of questions that related to the current condition of a patient, residents did not use their own clinical judgment to answer the question nor did they check on the patient as a source for that information. Excluding clinical judgments,

residents consulted nine other answer sources at least once when answering questions about the current condition of a patient (ignoring only attending notes and the patient's electronic chart, i.e. CareCast, see Appendix E). Checking on the patient, however, was never observed in this study as a source of information utilized by residents. This has been previously described in academic medical settings where members of the care team frequently exchange patient information between each other in order to avoid unnecessary disturbance to their patients [19].

Discussing the current condition of patients is generally encouraged during handoffs but by the time the resident must make a decision regarding the patient, the data may no longer be current. Interestingly, none of the current condition questions linked to handoff were actually discussed (either generally or specifically). If handoffs were conducted at patients' bedsides (or close by) the incoming resident would presumably begin the shift with an up to date picture of the patients' current condition. As checking on the patient frequently may not be efficient, perhaps handoffs could occur at or near the patient so that their current condition can be verified and updated if necessary.

For questions they did not answer, residents rarely used their handoff sheet as an external source for information. Further investigation of the utility of printed or written handoff sheets is necessary in order to ascertain why such items were so infrequently utilized. It is possible such artifacts are less useful in responding to other providers' questions as opposed their utility as a reference for residents in completing their own work-related tasks. On the other hand, residents may have simply committed the data on the handoff sheet to memory and therefore would have been assimilated into their clinical judgments.

Study limitations

Unidentified External Information Sources

It is possible that there are many other sources from which residents solicit information that were not observed during this study. Thus, as some authors have suggested [20], future studies in this area should account for a broad spectrum of potential information sources available to residents.

Clinical Judgments

The greatest limitation of this study was the inability to clarify details related to clinical judgments and lack of a more complete set of data linking questions to patients discussed at handoff. There are many ways to speculate as to the source of the information supporting residents' clinical

judgments. For example, it is important to consider whether information was discussed during handoff and ultimately forgotten, not discussed at handoff but acquired in other ways not observed in this study, or if other, less conspicuous sources of information provide residents' with critical information. In the latter case, it may even be practical to include these additional sources as elements of the handoff process.

Incomplete set of linked handoff discussion/questions asked data

Future work, in order to avoid problems in the linking process, will find it useful to collect identifiable information about the patient for every question asked observed in order to more easily establish the link between questions and the patients discussed during handoff. Larger, more complete linked data sets would provide a more representative sample and ultimately allow for more robust set of statistical analyses.

Chapter 5

Training the content to discuss at handoff

Introduction

Prior research suggests that educational interventions should be developed to train residents how to communicate effectively during handoffs [6, 31, 33, 59]. A number of training programs have been designed for general medical team training which address communication issues between members of medical teams using simulations as well as classroom training [15, 16]. Of these few training programs, however, only four specifically teach physicians how concepts such as verbal communication strategies or models for structuring content are applicable specifically to patient handoff discussions [28, 37, 76, 77]. While there have been many studies which have evaluated the impact of using structured checklists to guide the handoff discussion, there have been none which have evaluated training as an alternate method to the checklist for structuring the discussion of information during handoff.

The transfer of knowledge from training or instruction to on the job performance is the transfer of training. According to [61], transfer of training is determined by three components: training inputs, training outputs, and conditions of transfer. Training inputs include training content, trainee characteristics, and work environment. Training outputs are elements that are measured by knowledge gained through training (learning) and the ability to recall the training material at a later time (retention). Conditions of transfer refers to the individual's ability to use material from training to improve performance on the job immediately (generalization), and over time (maintenance). In [61]'s model, the training inputs each have an effect on the training outputs, and the conditions of transfer. This suggests that handoff behaviors can be influenced by training. It also suggests that the work environment provides learners a chance to present learned skills and an opportunity for measuring the impact of training on the conditions of transfer. This implies that evaluations of the impact of handoff training must provide an opportunity for residents to carry out training concepts in the work environment.

It has been suggested in the literature that handoff training interventions must be accompanied by thorough behavioral evaluation methods [78]. However, no studies have evaluated the effect of

training a specific set of information to discuss during handoff on the information residents actually discuss during handoffs. The effects of training handoff-specific concepts to residents have typically been measured using only, satisfaction-based questionnaires, knowledge assessments to gauge conceptual learning or simulated handoffs [28], none of which are true measures of the conditions of transfer.

This work investigated the characteristics of the information residents discuss during handoff before and after a training intervention designed to teach residents what information to discuss at handoff. If training residents to discuss a specific set of information influences the information residents actually discuss during handoff, it would be best measured by observing the characteristics of the information they discuss during actual handoffs both before and after training.

- ▪ Research Question 2.1: Does training residents to discuss a specific set of information at handoff correlate to changes in the appearance, frequency, or amount of time they spend discussing that information during handoff?

Method

Setting and participants

This study was conducted on two contiguous general pediatric acute care wards (7-West and 7-Central) at the University of Virginia Children's Hospital. At the onset of data collection, the general pediatrics service consisted of three first-year pediatric residents (interns), two pediatric residents in their second or third year (senior residents), two fourth-year "chief" residents, and an attending physician. Year of residency was commonly labeled as PL for "pediatric level" or PGY for "post-graduate year" followed by a dash and a number indicating year of residency (e.g. PL-1 or PGY-1).

At the time of the study, each patient admitted to the ward was assigned to both an intern and third year senior resident. The intern served as the patient's primary care provider while the senior resident supervised. The daytime care team consisted of two second or third year pediatric residents (PGY-2 or PGY-3), three interns (PGY-1), and the attending physician. The day team rounded together every morning. Night coverage was shared by eight residents: five on the primary care team (three interns and two PGY-3 senior residents on general ward rotations), and the cross-cover team (two PGY-2 senior residents and one intern on other pediatric rotations). Night coverage was shared by eight residents during a rotation – two of which (one intern and one senior resident) would be on-call overnight, responsible for caring for all of the patients on both wards from 4pm until 7am

the next morning. This schedule repeated every fourth night, to ensure each pair of on-call residents did not exceed the ACGME 80 duty hours per week regulation.

Handoffs took place throughout the day most notably at 7am and 4pm when day and night coverage shifts began, respectively. Handoffs most frequently took place in the Pediatric Library located on one of the acute care units. The library was a central, highly trafficked room serving many functions, including housing the chief resident offices, lockers for all of the residents, multiple computers, printers, phones, a small reference library, a small kitchenette, and a conference table for meetings such as handoff and rounds. During week-days, there were three separate resident handoff sessions.

The first session occurred at 7:00 a.m. Cross-covering night shift residents met with members of the primary care team to review overnight events and transfer patient care responsibilities back to the primary care team. The second weekday handoff sessions occurred invariably throughout the morning from approximately 10:00am to 11:45a.m. During these handoffs, the intern and/or senior who had been on call overnight met with members of the primary care team to transfer care responsibilities for any patients whom they had continued to care for until the remaining members of the primary care team arrived for their shift. The third weekday handoff sessions began at approximately 4:00 p.m. in the afternoon. During these handoffs, primary care team members met with the two residents who were scheduled to be on-call that night. This was the longest and most comprehensive of the three weekday handoff sessions as members of general ward team transferred care of all patients on the two wards to the two overnight residents (who may or may not have been members of the primary team).

On weekends, there was a single handoff at noon each day. At this time, the pair of residents who had been caring for all of the patients during the previous 24 hour shift transferred patient care responsibilities to a new pair of residents who would be assuming patient care responsibilities for the subsequent 24 hour shift.

Procedure

Training

A one-hour training module was developed to teach pediatric interns a core set of information to discuss for each patient handed-off. The development of the training is discussed in greater detail in [79]. The initial version of the training system was developed with Adobe Captivate 4.0 and ColdFusion. The training was designed as a slide by slide presentation designed to allow for self-

paced learning. The content slides were populated with text, images, video, audio clips, and provided learners with optional interactive learning opportunities. With respect to multimedia, the training system included five instructional videos and ten audio examples of handoff to supplement the text-based training content. A consistent design template was used to give the system interface an overall professional and consistent look and feel. In summary, the training was divided into five major sections: a pre-training knowledge assessment, three content presentation sections, and a post-training knowledge assessment. The three content sections were: an overview of handoff, the five critical information components to discuss during handoff, and basics of the handoff process. The overview of handoff was designed to motivate the learner and included a history of medical handoffs, current definitions of handoff in the literature, types of handoff errors, and an overview of the training objectives. The content for the five information components of handoff was based on a literature review and prior work [32-34] (Appendix A). The section was divided into seven subsections: an overview, sub-sections for each of the five components (identifying the patient, the "big sentence", the current condition of patients, the plan of care, and contingency plans), and a review section. The handoff process section had five subsections: an overview, the importance of following a process, logistical guidelines, methodological guidelines, and a review section.

The training module also included introductory and concluding interactive case studies as a way of modeling situations where poor handoff led to critical errors in patient care and to promote learner engagement. The training also included optional interactive learning items throughout the training sections such as series of facts and data from the literature on handoffs. These items could be accessed via a large colorful button labeled "Did You Know?" which would open a pop-up window containing the additional information. While the learner was not required to view the "Did You Know?" facts, the purpose was to make them available for the interested learner and to support multiple learning styles.

Training Module Heuristic Evaluation

An initial heuristic evaluation of the training system was conducted in order to ensure that the design adhered to the requirements and did not exhibit any unforeseen content, design, or technical issues. Overall, results from the heuristic evaluation of the prototype suggested that the initial design adequately implemented the requirements for the system. As a result of the evaluation some of the content was modified or shortened from its original version prior to the onset of the study. A summary of the results of the heuristic evaluation is provided in Appendix C. Detailed results of this heuristic evaluation are described in [79].

Data collection

Baseline data

First, second, and third year residents were observed and audio-recorded while conducting 4pm handoffs over a seven month period from July 2005 to February 2006. In total, seven different rotations of all three levels of residents were observed during the baseline data collection, all of whom provided consent for their involvement. In total, fifteen handoff sessions comprised of a total of 202 patient discussions were observed and recorded.

Training data

Data collection for the training study focused on the 10am handoffs given by pediatric interns. Data were collected during a six month period from January to June, 2009 for a group of 12 interns each of whom had entered the pediatric residency program in the summer of 2008. Observations were conducted during each interns' second general ward rotation to avoid overloading them with supplemental learning material at the onset of residency (i.e. during their first ward rotation). One intern was recorded giving handoff per observation session (the overnight/on-call intern whose shift was ending). The data collection proceeded as follows: twice during the first two weeks of each intern's second general ward rotation, each intern was observed and audio recorded giving handoff at the end of their shift (observations 1 and 2). The training module was given to each intern following completion of the first two observations.

Table 9. Schedule for collection of pre-training recordings, date of the training, and collection of post-training recordings for participants in the handoff training study.

Participant	Pre-Training		Training Date	Post-Training	
	Observation 1	Observation 2		Observation 3	Observation 4
Participant 1	02/05/2009	02/06/2009	02/21/2009	02/23/2009	02/24/2009
Participant 2[*]	None		02/20/2009	02/22/2009	None
Participant 3	02/05/2009	02/06/2009	02/18/2009	02/18/2009	02/20/2009
Participant 4	03/02/2009	03/04/2009	03/14/2009	03/18/2009	03/20/2009
Participant 5	02/26/2009	03/02/2009	03/17/2009	03/18/2009	03/26/2009
Participant 6	02/25/2009	02/26/2009	03/24/2009	03/25/2009	03/29/2009
Participant 7	02/27/2009	03/07/2009	03/12/2009	03/23/2009	03/27/2009
Participant 8	04/03/2009	04/07/2009	04/15/2009	04/23/2009	04/24/2009
Participant 9	04/04/2009	04/08/2009	04/18/2009	04/23/2009	04/24/2009
Participant 10	04/04/2009	04/05/2009	04/17/2009	04/18/2009	04/19/2009
Participant 11[*]	None		None	None	05/21/2009
Participant 12	05/27/2009	06/11/2009	06/14/2009	06/15/2009	06/19/2009

[*]Only partial data were collected for participants 2 and 11 due to scheduling errors and therefore were not included in subsequent analyses. All analyses, therefore, are based on data from only ten of the twelve interns observed (participants 1, 3 – 10 and 12)

Once the training module had been completed, during the remaining weeks of each intern's second general ward rotation, two additional observations and audio recordings of each intern giving handoff were collected (observations 3 and 4). **Error! Reference source not found.** summarizes the dates of each observation before and after training (pre-training observations 1 and 2; post-training observations 3 and 4) and the date each resident completed the training module.

Audio recording of handoffs

Each handoff session observed was audio recorded using a digital recorder with a tie-clip microphone placed facing the middle of the handoff conference table. The research assistant collected audio recordings while silently observing the handoff meeting, taking notes as necessary to assist cataloguing of the audio recordings.

Coding handoff audio recordings

After all observations and recordings were collected, the research assistant coded each verbal utterance based on its content using the coding hierarchy in Table 10.

Table 10. Information elements used in coding the utterances of each handoff audio recording

Information Code	Description
Patient Identification	▪ Patient name ▪ Patient medical record number ▪ Patient description (e.g. age, gender etc.)
Big Sentence	▪ History of present illness ▪ Past medical history ▪ Reason(s) for hospitalization ▪ Current diagnosis(es) ▪ Evaluation and treatment performed prior to current hospitalization
Current Condition	▪ Discussion of patient's current physical and mental state relevant to diagnostic and/or treatment plan
Plan of Care	▪ Diagnostic tests performed during this hospitalization (laboratory tests, radiology procedures etc.) ▪ Results and pending results ▪ Consultations ▪ Any prescribed medications ▪ Recent changes in medicines ▪ Non-medicine orders that have been entered into the computerized order entry system (e.g. diet, activity, nursing orders) ▪ A summation of the overall treatment plan ▪ Explicitly stated tasks that need to be accomplished overnight (e.g. review the chest radiograph, obtain a complete blood count at midnight, review intake and output and adjust fluids)
Contingency Plans	▪ Any contingency instructions for overnight care (e.g. if child develops fever, obtain blood culture)
Off-Task	▪ Any part of the conversation which could not be reasonably associated with any of the above categories**

** Discussion segments which were coded as "Off-Task" may or may not contain "relevant" information related to patient care under this particular categorization scheme.

The content codes in Table 10 were determined prior to the study during the development of the training and are based on the training content and learning objectives. Specifically, codes were based on the five information categories as defined in the objectives of the handoff training module as well as off-task discussion.

To facilitate the coding process, the graduate research assistant employed a modified version of the software tool described in detail in [67]. Separately, two medical experts were provided with a randomly selected subset of the entire catalog of audio recordings (representing approximately 10% of the entire audio recording time) in order to assess inter-rater reliability.

Rater reliability

Fleiss' Kappa was used to determine reliability between the codes of the graduate research assistant and the two medical experts (inter-rater reliability). Separate kappa scores were calculated for intra-rater reliability within the baseline and training data sets for the research assistant who coded all of the audio recordings. The purpose was to assess the internal consistency of the research assistant's codes within each of the three data sets. To calculate these scores the research assistant randomly selected a recording as the first recording to code, then coded the recording at two additional points in time over the course of coding the entire set of data, once after half of the recordings in each data set had been coded, and once after all of the recordings in each data set had been coded. This method provided three independent sets of codes for the three randomly selected baseline, pre-training, and post-training recordings. Three independent sets of codes allowed for the calculation of a unique value for Fleiss' Kappa for the baseline, pre-training, and post-training data sets.

For inter-rater reliability across the three raters (the graduate research assistant and the two medical experts) the calculated fixed-marginal value for Fleiss' kappa (κ) was 0.57. The calculated fixed-marginal values for Fleiss' kappa (κ) for the research assistant's codes were 0.59 for the coded baseline data, 0.55 for the coded pre-training data, and 0.56 for the coded post-training data.

Independent variables

There were two primary independent factors in this study. The first factor was based on the time of each observation and had two levels: *pre-training* and *post-training*. The second factor was a blocking factor on resident and had ten levels, one for each subject in the study. This was a within subjects repeated measures design in which each of the ten subjects were observed at both the *pre-training* and *post-training* observation times.

Dependent variables

The following 15 dependent measures are based on the available output data from coding audio recordings using the software tool described in [67]. There were three main categories of dependent measures in the study: *mean discussion times*, *proportion of patients in which content categories were discussed,* and *mean counts of utterances by content category*.

1. Mean discussion time (mm:ss)
    a. ***on-task*** per patient (discussion related to the five information categories)
    b. ***off-task*** per patient
2. Proportion of patients (for which):
    a. ***patient id*** was discussed
    b. the ***big sentence*** was discussed
    c. ***current condition*** was discussed
    d. ***plan of care*** was discussed
    e. ***contingency plans*** were discussed
    f. ***all five components*** were discussed
        i. "all five components" means each of the five information components was discussed at least once for the patient
    g. the discussion went ***off-task***
3. Mean utterances (of):
    a. ***patient id*** per patient
    b. ***the big sentence*** per patient
    c. ***current condition*** per patient
    d. ***plan of care*** per patient
    e. ***contingency plans*** per patient
    f. ***off-task*** per patient

Data analysis

Comparison of mean discussion times

Mean discussion times on and off task per patient are presented first. These include aggregated mean discussion times for each of the three groups (baseline, pre-training, and post-training), as well as for individual sets of observations within the pre and post-training groups (observations 1 and 2 for the pre-training, observations 3 and 4 for the post-training). Kruskal-Wallis One-Way ANOVAs on Ranks were conducted to test differences in on-task and off-task discussion times per

patient across the baseline, pre-training and post-training groups. Bonferroni post-hoc pairwise multiple comparisons were conducted in the event a significant difference between two or more groups. The hypotheses were that mean on-task discussion time per patient and mean off-task discussion time per patient (behavioral dependent measures 1 and 2) would be significantly higher in the pre-training data as compared to both baseline and post-training.

Comparison of pre-training and post-training observational data

The primary set of analyses which address Research Question 2.1 compared residents across the two levels of the observation time factor: pre-training and post-training.

First, multinomial Chi-Square tests were conducted to assess differences in paired proportions across the two levels of the observation time factor. For each dependent measure, a separate multinomial test was conducted. As such, seven multinomial Chi-Square tests were conducted to compare the dependent measures of patient proportions across the pre-training and post-training levels of the observation time factor. Second, two-factor repeated measures ANOVAs were used to assess main effects of observation time (pre or post training) and training subject (the ten subjects in the study) as well as interaction effects between the two factors for dependent measures related to mean utterance counts. Six separate repeated measures ANOVAs were conducted, one for each of the six dependent measures related to mean utterance counts. Tukey-Kramer multiple comparison tests were used to assess all pairwise differences between means when significant effects were observed.

Bonferroni Correction

Due to the large number of individual comparisons which were conducted, the significance level employed for every statistical comparison was adjusted using the Bonferroni Correction [80]. Specifically, seven multinomial Chi-Square tests and six repeated measures ANOVAs were conducted, resulting in 13 separate comparisons. As such, the standard level of $\alpha = 0.05$ was adjusted to $\alpha = \frac{0.05}{13} = 0.004$.

Addressing internal validity

Two sets of group pairwise comparisons across the 13 dependent measures were conducted in order to address issues concerning internal validity of the study. In particular, the baseline group was compared to the pre-training group in the first set of comparisons. In the second set, comparisons were conducted between the first and second sets of pre-training data. The particular analyses for

each set of comparisons are described below. For the sake of brevity and clarity, the results of these analyses are presented in Appendix J.

Results

Characteristics of information discussed at handoff

Baseline

Fifteen handoff sessions were observed and audio recorded. Across the 15 sessions, residents discussed a total of 202 patients. On average, 14 patients were discussed per session with a minimum of three and a maximum of 21. In sum, 9h 8m 58s of handoff time was observed and audio recorded during the baseline collection. Mean on-task and off-task discussion time per handoff and per patient for the baseline data are summarized in Table 11.

Pre-training

Of the 20 pre-training handoff sessions observed, four resulted in audio recordings in which the sound quality was too poor to distinguish the details of the conversation. Of the remaining 16 handoff sessions nine were from observation 1, and seven from observation 2. Across the sixteen sessions, interns discussed a total of 88 patients, 56 in observation 1 and 32 in observation 2.

On average, six patients were discussed per handoff session with a minimum of two and a maximum of 15. In sum, 5h 57m 29s of handoff time was observed and audio recorded during the pre-training data collection, 4h 10m 47s in observation 1 and 1h 46m 42s in observation 2. Again, mean on-task and off-task discussion time per handoff and per patient for the pre-training data are summarized in Table 11.

Post-training

Of the 20 post-training handoff sessions observed, five resulted in audio recordings in which the sound quality was too poor to distinguish the details of the conversation. Of the remaining 15 handoff sessions nine were from recording set three, and six from recording set four. Across the fifteen sessions, interns discussed a total of 75 patients, 48 in recording set three and 27 in recording set four.

On average, five patients were discussed per handoff session with a minimum of two and a maximum of 10. In sum, 3h 28m 27s of handoff time was observed and audio recorded during the

post-training data collection, 2h 12m 23s in observation 3, and 1h 16m 4s in observation 4. Lastly, mean on-task and off-task discussion time per handoff and per patient for the post-training data are summarized in Table 11.

Table 11. Mean duration of on-task and off-task discussions by handoff session and by patient across baseline, pre-training, and post-training observations

Time	$\mu_{\text{on-task/session}}$	$\mu_{\text{on-task/patient}}$	$\mu_{\text{off-task/session}}$	$\mu_{\text{off-task/patient}}$
Baseline	0:36:36 (0:12:03)	0:02:35 (0:01:59)	0:01:11 (0:01:01)	0:00:41 (0:01:01)
Pre-training	0:22:21 (0:17:36)	0:04:04 (0:03:27)	0:01:04 (0:01:58)	0:00:36 (0:02:00)
Observation 1	0:27:52 (0:21:24)	0:04:29 (0:03:19)	0:00:38 (0:00:34)	0:00:29 (0:01:13)
Observation 2	0:15:15 (0:05:28)	0:03:20 (0:03:37)	0:01:36 (0:02:55)	0:00:49 (0:02:53)
Post-training	0:13:54 (0:07:21)	0:02:47 (0:02:01)	0:00:26 (0:00:27)	0:00:19 (0:00:38)
Observation 3	0:14:43 (0:06:50)	0:02:45 (0:01:55)	0:00:26 (0:00:22)	0:00:18 (0:00:31)
Observation 4	0:12:41 (0:08:35)	0:02:49 (0:02:13)	0:00:26 (0:00:33)	0:00:20 (0:00:47)

Time spent on-task and off-task per patient

Unless otherwise noted, all statistical tests in this section were conducted at the $\alpha = 0.05$ level for significance.

On-task discussion time per patient

Figure 10 shows the mean times (in total seconds) spent on-task per patient across the three groups in the study (baseline, pre-training, and post-training). Visual inspection indicated residents spent a greater amount of time on-task per patient pre-training than post-training and the baseline group less time on-task per patient than interns both before and after training.

A Kruskal-Wallis test on the median times spent-on task per patient revealed that at least one group median was significantly different from one of the other two groups ($H = 21.548$, $df = 2$, $p < 0.001$). In particular, at $\alpha = 0.05$, post-hoc Bonferroni multiple comparison tests revealed that pre-training interns spent significantly more time on-task per patient ($M = 207.38$ s) than the baseline group ($M = 121.98$ s) and significantly more time on-task per-patient pre-training vs post-training ($M = 148.04$ s), $F(2,362) = 2.405$, $p < 0.05$.
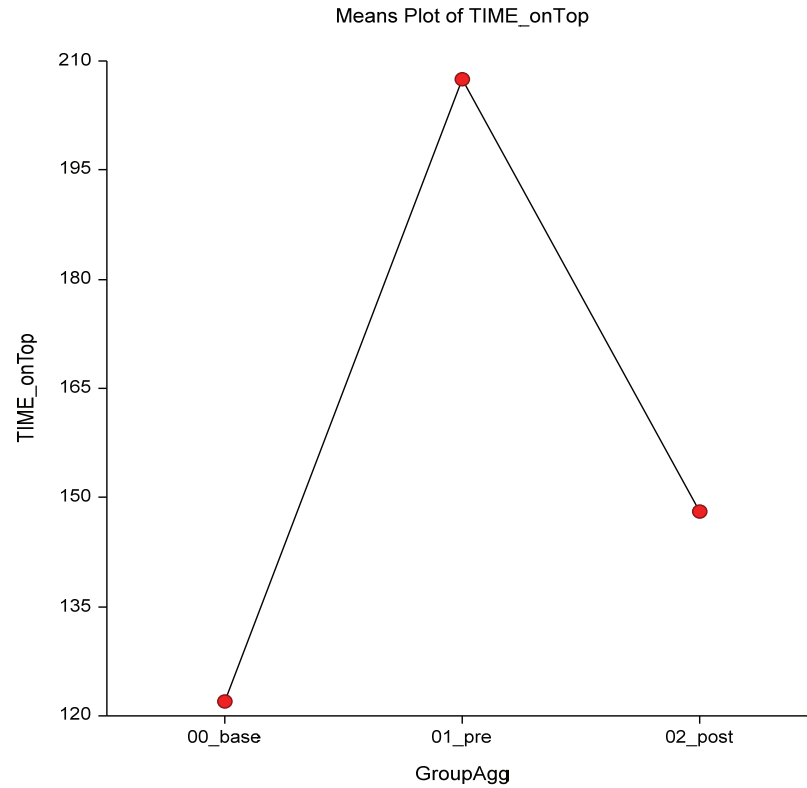
Figure 10. Mean time spent on task per patient by group

Time spent off-task per patient

Figure 11 shows the mean times (in total seconds) spent off-task per patient across the three groups in the study (baseline, pre-training, and post-training). An initial visual inspection indicated that pre-training interns spent slightly less time off-task per patient than the baseline group and a smaller amount of time off-task per-patient post-training.

A Kruskal-Wallis test on the median times spent off-task per patient, however, no significant differences between the groups with respect to discussion time spent off-task per-patient ($H = 3.328$, $df = 2$, p = 0.189).

Means Plot of TIME_offTop
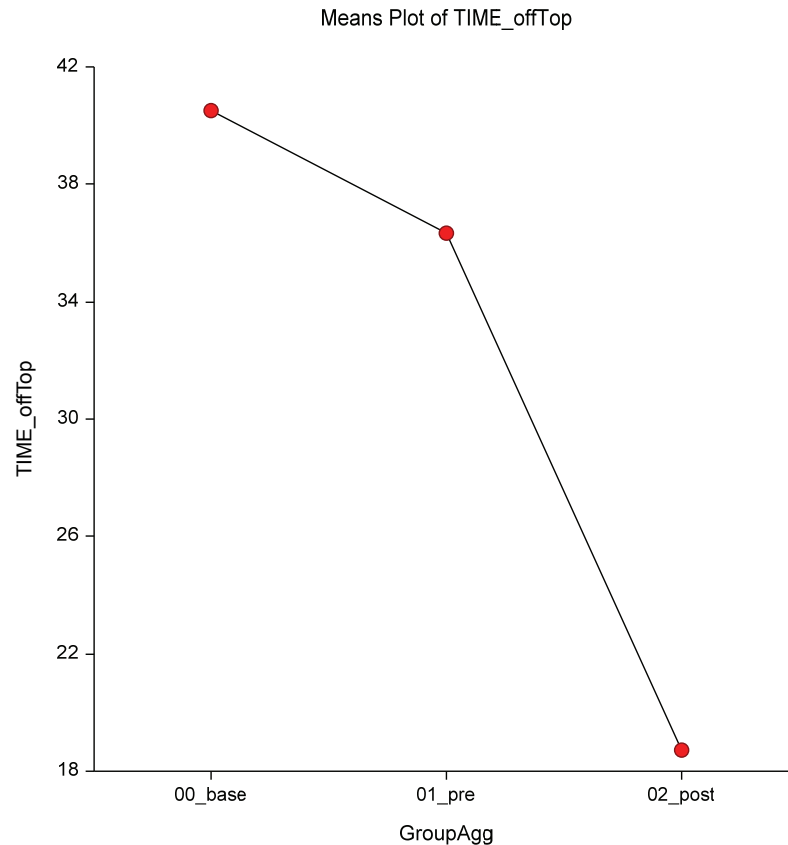
Figure 11. Mean time spent off-task per patient by group

Pre-training vs Post-training comparisons

Unless otherwise noted, all statistical tests in the following section were conducted at the $\alpha = 0.004$ level for significance as a result of the Bonferroni Correction with an initial value of $\alpha = 0.05$ and for 13 separate statistical comparisons. All statistical comparisons were conducted using the NCSS statistical software package. Complete statistical printouts for each of the analyses conducted are provided in Appendix K.

Proportions of patients

One significant difference was observed from pre-training to post-training with respect to the dependent measures of proportions: a significantly higher proportion of patients for which all five information categories were discussed post-training. Two separate Chi-square tests of proportions were conducted to check the accuracy of this result. First, a multinomial test of proportions was

conducted which revealed a that the proportion of patients for which residents discussed all five components was significantly higher post-training (67.1%) than pre-training (32.9%) ($\chi^2 = 8.562, p < 0.004$). A McNemar test of paired proportions was conducted to confirm this result under the condition of testing the difference in paired proportions. The McNemar test confirmed that the difference in proportions was significant from pre to post-training, that is, was significantly higher post-training ($\chi^2 = 21.929, p < 0.001$).

No significant differences were found for the resident factor across all dependent measures related to the proportion of patients.

Mean utterances per-patient for each information category and *off-task*

Patient ID

There was a significant effect of observation time (i.e. pre training vs post training) on the mean utterances of Patient ID per patient ($F=34.82, p<0.001$). In particular, a post-hoc Tukey-Kramer comparison showed that prior to training mean utterances of the Patient ID were significantly higher per patient than after training ($p < 0.004$).

The Big Sentence

There were significant main effects of both the observation time factor ($F=9.42, p<0.004$) and resident factor . Specifically, Tukey-Kramer tests showed there were significantly fewer mean utterances of *the big sentence* per-patient by interns post-training as opposed to pre-training.

There was also a significant interaction effect between the observation time and resident factors . Specifically, residents 3 and 12 differed significantly in their mean utterances of the big sentence per patient as well as the direction of change for their means from pre to post training. That is, there was a significant increase in the mean utterances of the big sentence per patient for resident 3 from pre to post training while also a significant decrease in the mean utterances of the big sentence per patient for resident 12 from pre to post-training (Figure 12).

Current Condition, Plan of Care, Contingency Plans, and mean Off-Task utterances per patient

Overall, there were no observed significant main effects of observation time or resident or significant interactions across mean utterances per-patient of the current condition, plan of care, contingency plans, or mean utterances off-task per-patient.
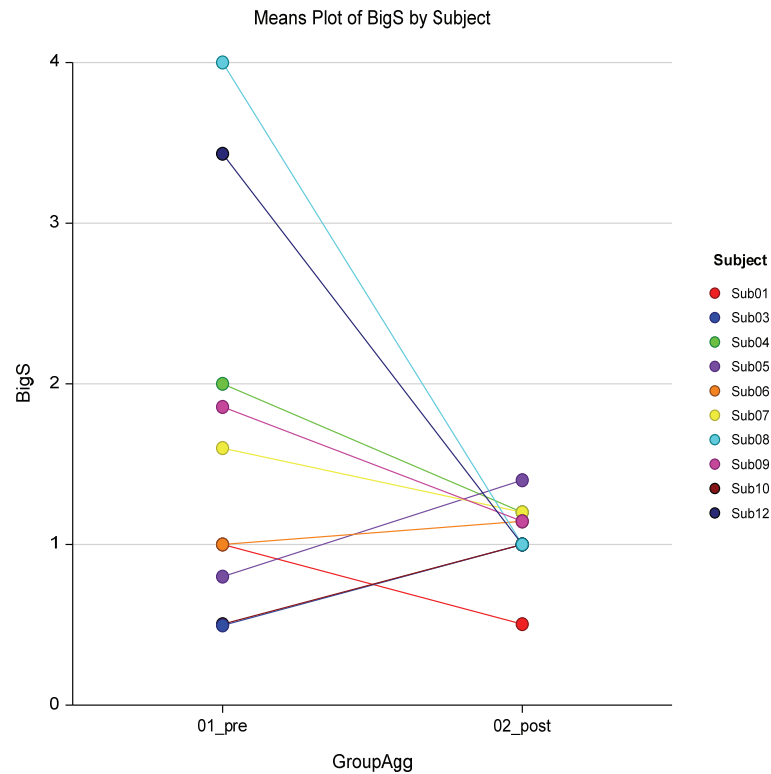
Figure 12. Mean utterances of the big sentence per patient by observation time and resident. A significant interaction effect exists for residents 3 and 12 (Sub03 and Sub12, respectively)

Discussion

Given the lack of outcomes that support the translation of knowledge, skills and attitudes into behavior change and reductions in adverse events [81], this study sought to answer the question "does training residents to discuss a specific set of information at handoff correlate to specific changes in the characteristics of the information they discuss during handoff?"

With respect to content, the most important finding in this study was that training residents to discuss a critical or core set of information components for each patient resulted in a significant increase from pre to post-training in the proportion of patients for which *all* of those critical information components were discussed. This suggests that with proper training, it may be possible to mitigate problems related to the transfer of incomplete patient information at handoff. The idea of completeness with respect to the information discussed for each patient is important as incomplete information has been shown to be a critical factor in residents' negative attitudes with respect to their preparation for patient care [47]. Furthermore, missing information or incomplete

transfer of information at handoff can have a negative impact on the success of patient care activities after handoff, as was shown here in Chapter Four.

Two interesting results of the study were the significant decreases in the mean utterances of Patient ID and the Big Sentence from pre to post-training data. Identifying patients more than once is a possible reason why residents spent more time on-task per-patient prior to training. The big sentence (meant to set the stage for the discussion by painting an overall picture of the patient) was discussed significantly less often per patient (closer to once per patient) after training, and is another possible factor that explains why residents spent more time on-task per-patient prior to training.

The reduction of mean utterances of each of these two information components post-training is one explanation for the significantly shorter post-training on-task discussion times per patient. The repetition of these two components specifically may also be redundant. As such, reducing unnecessary redundancy in the information discussed for each patient at handoff is a plausible way of reducing the overall time required to conduct handoff. Perhaps teaching residents to limit the information they discuss for each patient to a manageable set of critical components is ultimately an effective way to improve the efficiency of the process. Alternatively, there are many reasons why redundancy could be beneficial to the process such as in repeating the patient ID for a handoff participant who arrived late or could not hear what was said due to background noise or some other type of interruption. Teaching residents to minimize utterances related to items such as the big sentence may also result in more time for discussion related to items such as the plan of care.

Studies have shown that longer handoffs are subject to more frequent interruptions and increases in off-task discussion [82]. It is possible that the reduction of redundant information in post-training handoffs observed in this study (e.g. patient ID and the big sentence) also had an effect on the amount of time residents spent off-task. However, there were no significant changes from pre to post-training with respect to any of the off-task dependent measures and therefore this conclusion cannot be made. It is possible there may have been factors not accounted for in the data such as outside interruptions which prevented a significant change in the amount of time or frequency of off-task discussions during handoffs. Future work may want to consider factors such as outside interruptions and their impact on the handoff discussion. Off-task discussion was given a very broad definition in this study as "any part of the conversation which could not be reasonably associated with any of the [five information] categories". Under this definition, discussion which was off-task could very easily have contained information relevant to patient care, or interruptions or other types of communication which served to benefit the discussion in general. As [83] suggests, some types

of interruptions may actually be beneficial to the handoff process. It is clear that the definition of what constitutes "off-task" discussion should be reformulated with more attention to factors such as logistics, teaching, and socializing. In fact, it may even be appropriate to define a separate coding hierarchy devoted entirely to the structure of off-task discussions. An initial hierarchy for re-coding off-task discussions and a preliminary analysis conducted on the baseline data from this study are presented in Appendix D.

Limitations and Future Work

Lack of a true control group hindered the power of the results for this study and shed doubt on its validity. While significant changes occurred in some cases, there are still factors unaccounted for which could have unintentionally influenced the results. In addition, the comparisons were based on loosely similar groups of participants (in the case of the baseline versus the pre-training comparisons) or on a small number of repeated measures (in the case of comparing pre-training observation 1 to pre-training observation 2). Future work on assessing the behavioral impact of handoff training interventions will benefit from a larger population sample than was available in this study, as well as an identical or characteristically very similar control group and random assignment of participants to groups. Conducting a study on a larger residency service or across multiple institutions would be a way to secure a larger sample size and might also provide greater insight into the generalizability of both the training itself as well as the results of this study.

Residents scored high on the knowledge assessment both before and after training. This result was not surprising as the knowledge assessment was short, misaligned from pre to post-training, and not pre-validated as an assessment tool. However, the results of the observational assessment do suggest that residents' gained a working knowledge of the concepts from the training which they were able to transfer to their handoff behaviors in some cases. An improved knowledge assessment will provide future work with the ability to test conceptual learning in addition to behavioral changes. Use of both behavioral and knowledge assessments is a more robust method for evaluating the impact of a training intervention.

The relevancy of the training content was also a concern in this study. While the choice to train five critical information components to discuss at handoff was not entirely arbitrary and is, for the most part, supported by prior research and medical expert opinion, these particular five components may not represent *the* most important information to discuss for each patient at handoff. It may also be the case that the particular categorization of the five components was all-inclusive but ill-defined.

Even at an expert level, care providers disagree on how to categorize the same part of a discussion, and it is likely that many different categorization schemes could be just as inclusive but provide vastly different results when used in coding handoff discussions. It is not to say that one is more correct than another, but future work should extensively research and validate the particular categorizations given to the information discussed during handoffs.

Finally, the choice to focus training only on the content of handoff discussions, although acceptable as a constraint in this study, is a very limited view of the process as a whole. Handoff is a more complex process than simply that of information exchange. It involves factors of interpersonal communication, situational awareness, sense-making, co-orientation, and many others. The initial goal of the particular training system developed for this study was to start with simple concepts for newer residents followed by increasingly abstract concepts as residents gained experience over the course of their second and third years. Future work will continue on this course having gained valuable knowledge from the methods in this study. Ultimately, while training the content of handoff is critical, future training efforts should also focus on more abstract concepts and skills beneficial to handoff. .

Chapter Summary

This study developed and evaluated a handoff training system designed to teach first-year pediatric residents critical information components to discuss at handoff. It focused on using observational methods to conduct a behavioral evaluation of the impact of the training intervention in a real work setting as opposed to the use of surveys, knowledge assessments or simulated handoffs. The training was shown to be influential primarily in reducing redundancy and ensuring completeness of the information discussed for each patient during handoff. While some of the results are quite promising, it is clear that future efforts should address the methodological limitations of this study and should consider a broader perspective of factors beyond simply the content of handoff discussions.

Chapter 6

Development of a Handoff Information Support System

Introduction

Support tools are critical to the success of handoff. To ensure that a tool is effective, designers must collaborate with end users. Collaboration can be difficult when working with end users who are busy and have irregular schedules. This chapter reports on a collaborative effort between physicians and engineers to redesign a handoff support tool. Strategies included focus groups, interviews, "on the fly" feedback, and an iterative design process. Task analytic methods were used to compare the original tool with the prototype in order to quantify any differences in functionality. Last, I discuss general conclusions and offer practical techniques for engaging end-users in the design process.
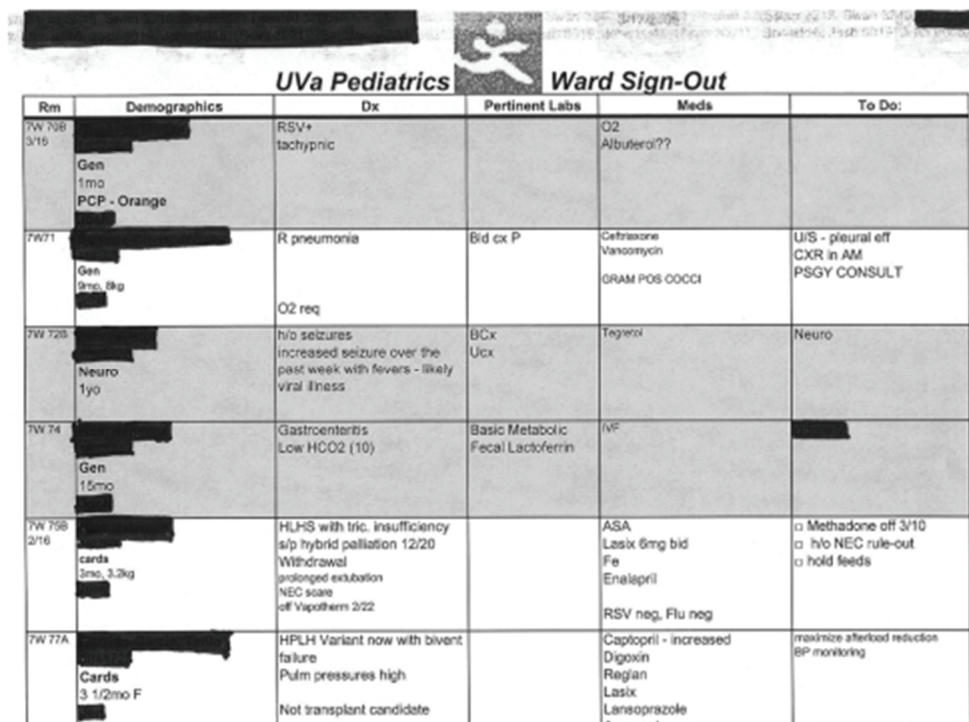
User-centered design: an example from healthcare

While a user-centered approach is critical to the development of new process support tools, engaging end users can be both art and science [84]. Users may not always be able to articulate their problems or point to their root causes. They may jump to potential solutions before truly understanding the underlying issues. They may not be available at typical work hours to discuss issues with their tools. This chapter discusses a multifaceted approach combining focus groups, interviews, iterative prototyping, testing, and feedback for engaging one such group of end users, resident physicians, in the design process of a handoff support tool. It discusses potential improvements based on task analytic methods comparing workflow with the existing tool with the re-designed tool, and offers practical strategies for increasing end-user involvement in the design process.

Background

Handoff (i.e. when outgoing physicians handoff patient care responsibilities to incoming physicians) is a critical process in hospital settings [7]. Resident duty hour restrictions are 30 continuous hours and 80 total hours per week, accompanied by 10 hour rest periods away from patient care [5]. As a result of duty hour restrictions, the number of handoffs between resident physicians has increased along with concern for its effect on continuity of care [85-87].

This chapter is based on the study of handoff in the pediatric service of a university-based tertiary care hospital with a medium sized residency training program. The pediatric service consisted of the general wards, and two intensive care units. Specifically, this chapter focuses on handoff between residents on the general pediatric wards. On the general pediatric wards, handoff occurred in the morning and in the afternoon. The brief morning handoff allowed the post-call residents to update the incoming residents on the overnight events and status of each patient. The more comprehensive afternoon handoff allowed the outgoing residents to review with the night staff what care is required until the following morning. Prior to this work, residents had developed an Excel-based spreadsheet to support handoff (Figure 13).

Figure 13. Original Handoff Spreadsheet

The residents considered this system a "reasonable" means for storing and summarizing patient data for handoff. However, there were issues associated with it. The handoff tool was manually updated by residents as they had time, and was difficult to access as it was installed on a single, inconveniently located computer. As a result, data in the handoff tool were not always up to date and information for recently admitted patients was often not included at all.

There were usability issues as well. Some residents were not familiar with Excel or had never received formal training on how to use the handoff tool. Residents who were familiar with the Excel tool's functionality complained that adding, deleting, or modifying records could be "tedious, error-prone, and time consuming". Part of the problem involved the process of adding a patient to the spreadsheet while maintaining proper ordering by room number. This required finding the correct location for the record to be added, inserting a new row, and then ensuring that the room number was typed in the correct format. As Figure 10 shows, data entry over time has yielded inconsistent room numbering formats. Simple errors such as these, however, can have serious consequences such as patients being overlooked. In addition, the columns' fields and sizes prohibited residents from providing information according to their needs, personal style, the experience of the receiving physician, or the receiving physician's familiarity with the patients on the ward (e.g. some residents prefer to include more detailed information than others).

Method

A Collaborative Design Strategy

To begin the re-design of the handoff tool, effort initially focused on defining requirements for information content and functionality of the tool. The approach included focus groups, interviews, task analyses, iterative prototyping, testing, and feedback. Focus groups and interviews were conducted to help residents define the information and functional requirements for the prototype handoff tool [88]. As residents defined additional requirements, they were incrementally implemented until a working prototype was put into use on the general wards. During this time, residents interacted with the prototype and provided feedback and additional requirements that were iteratively incorporated into the prototype's design.

Defining Requirements and Design of a Prototype

Information Requirements

In the first round of focus groups, residents identified six general and 29 specific information requirements for each patient discussed during handoff. The information requirements were refined via interviews with two pediatric chief residents. The final list of information categories consisted of seven general information categories and twenty-one sub-categories (Table 12).

Table 12. Final Information Requirements

Information Category	Information Sub-Category	
Patient Demographics	Name	Date of admission, current hospital
	Age, weight	day
	Medical record number	Service
	Unit/room number	Intern
	Acuity	Student
		Primary care provider (PCP)
Problem List	Current	
	Past	
Medications & Treatments	General	
	Fluids/Electrolytes/Nutrition (F/E/N)	
Test Results	A.M. Only	
General Plans	To do list	
	Discharge	
Contingency Plans	Fluids & Medications (IV's)/Fever/Pain/Parameters	
	Allergies/Other	
Immediate Plans	To do list	
	Current status/Notes	

In terms of the tool's functionality, the chief residents outlined two general areas for the tool's functional requirements: data entry/modification requirements and data reporting/printing requirements. Detailed requirements were then identified in focus groups with all of the pediatric residents and additional meetings with only the chief residents. Based on the functional requirements, a "working" prototype could not be developed using a spreadsheet format. Thus, Microsoft Access® was chosen as the platform for the prototype. First, Microsoft Access offers separation of the data entry, storage, and output processes making prototype design a "divide and conquer" activity. Second, Microsoft Access includes design tools that help simplify functionality for end-users.

Data Entry Requirements

The main focus for data entry was to reduce complexity of the process and provide flexibility not found in the existing spreadsheet tool. The following list includes requirements followed by design elements in parentheses:

1. Allow patients to be added to the system (using Microsoft Access® a blank patient record is always available at the bottom of the page; Figure 14).

2. Allow patients to be deleted from the system (by placing the Active Record Identifier on a record and pressing the delete key; Figure 14).

3. Provide a separate entry space for each of the twenty-one information sub-categories (see multiple data input areas in Figure 14).

4. Employ features such as drop-down menus for information sub-categories that have a limited set of choices such as Demographics - Room Number, Demographics - Acuity, and General Plans – Discharge (Figure 14).

5. Allow free text input when drop-down menus are not applicable (Figure 14).

6. Limit the number of characters available for free text input to maintain simplicity (Microsoft Access® provides this ability when specifying data types in the database).

7. Initially sort records by one of the common ways residents discuss patients: acuity or room (note that the data entry page initially sorts patients by room number; Figure 14).

8. Allow records to be sorted by any of the information categories for flexibility (Figure 14).

Figure 14. Prototype Data Entry Page

"On the Fly" Feedback

Part of the approach was to get feedback from as many end users as possible. To help achieve this goal, an interactive user feedback form was built into the data entry page (Figure 14). The goal of the feedback form was to provide residents an opportunity to give feedback "on-the-fly" while using the tool, as well as to provide an outlet for comments when interviews could not be scheduled. The comments were used throughout the process to define additional requirements and modify the prototype.

Data Output/Printing Requirements

With respect to data output, the tool needed to be able to generate and print customized reports based on resident preferences and predetermined values (e.g. the time of day - morning/AM or afternoon/PM; and the sorting method for patients - by acuity or by room number). While Microsoft Excel® spreadsheets can only be printed in the format of the data entry interface, Access allows the creation of multiple reports linked to the same data table. Therefore four reports were designed corresponding to each of the four possible combinations of the time of day and the desired sorting method. Each report could be viewed or printed via the print menu on the data entry form (Figure 15).
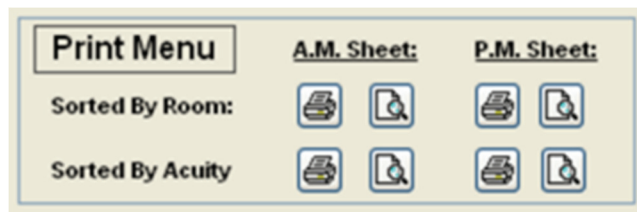
Figure 15. Menu for printing reports

The design of each report (Figure 16) was based on formatting requirements residents defined below:

1. Reports must be able to display at least five patients per page

2. The tool must be able to shrink or expand data columns in a report based on the amount of data entered

3. The text for a patient's name, room, and acuity must be salient (e.g. appear in bold)

4. The report must highlight (using text shading) the acuity of any patient listed as "high acuity" as well as the discharge plan for any patient listed for 9:00 a.m.

5. The tool must provide pager numbers of the current rotating physicians in the footer of all handoff reports

Figure 16. Prototype Handoff Report (AM by acuity)

While footer information was initially included at the bottom of each printed report, the chief residents pointed out that names and numbers needed to be updated when a new rotation was underway. Therefore, functionality to modify the report footer was added (Figure 17).

Figure 17. Modify Report Footer Page

Additional Requirements

Patients on the pediatric wards were comprised of those on the general wards service, but also those on the hematology/oncology (Heme/Onc) service as well as the pediatric intensive care unit (PICU). While the handoff process for all three services was similar, the location, time, and residents involved varied. Comments from the feedback forms and interviews suggested that the prototype

support handoff for Heme/Onc and PICU patients. Therefore separate data entry forms, database tables, and reports, all based on the original designs, were added to the prototype for the two additional services. Residents also requested a patient archive and a patient dictation list. The archive and dictation list were added to the tool as separate data entry forms, and database tables. With new data entry forms, residents also specified the need to transfer patients between the service lists, the archive, and the dictation list. Transfers were implemented in the prototype so residents could transfer one patient to one location or multiple patients (from one list) to multiple locations at one time. In terms of reports, multiple reports were designed for the Heme/Onc service identical to those for the general wards. However, residents required only one report for the PICU service, as few patients are typically on the service at a given time. Residents did not request a report for the patient archive and requested only one report for the dictation list (similar to the PICU report).

During interviews, resident supervisors and attending physicians requested support for procedures other than handoff. In particular, they described how the data stored in the prototype could be used to generate morning reports as well as a portion of each patient's daily progress note. Both of these functions were added to the prototype in subsequent revisions. The morning report function displays a single report of the current and past problems for all patients admitted between two specified dates, sorted by date and service, and formatted with large text for output to a projector. The daily progress note required the implementation of tabs (one for handoff, one for the daily progress note) for each patient on the data entry page of the general wards, Heme/Onc, and PICU services.

 With the addition of six new reports (excluding morning report as residents asked it to be a separate function), it was no longer appropriate to associate viewing or printing reports with separate buttons. Therefore the "print menu" was removed in favor of a "report menu" that utilized drop-downs instead of buttons to view and print reports. To organize all of the functions implemented in the prototype, I developed a main menu that opened with the database (Figure 18).

Figure 18. Prototype Main Menu

Results – Keystroke-Level Models for the Handoff Tools

Based on the task analyses, the Keystroke-Level Model (KLM) [69] was used to quantify the time required to complete the tasks on each of the two systems. Table 13 - Table 16 provide an example comparison of two tasks common to both tools: adding a patient and modifying the report footer. For the KLM, "P" refers to pointing with the mouse, "B" to pressing or releasing the mouse button, "H" to moving hands between keyboard and mouse, and "K" to pressing a key on the keyboard. For the KLM, timing data were collected and averaged for one participant.

Table 13. Task analysis for adding a patient to the spreadsheet

KLM Model	Task 1.1: Add a patient to the spreadsheet
	⇨ PLAN: Do 1.1.1 then 1.1.2 then 1.1.3 (for name column) and 1.1.4 (for all other columns), repeat 1.1.1 – 1.1.4 (until all patients added)
	1.1.1: Place the patient in the correct order by room
	⇨ PLAN: Do 1.1.1.1 then 1.1.1.2
P – B – P – B	1.1.1.1: Scroll to the existing patient row whose room number is one higher than the patient to be added
P – B – P - B	1.1.1.2: In the far left column showing the row number of the existing patient, right click and select "insert"
P – BBBB	1.1.2: Select and double click the desired box to edit text column of the new row to edit
	1.1.3: Add patient identification information
	⇨ PLAN: Do 1.1.3.1 then 1.1.3.2 then 1.1.3.3 then (if needed) do 1.1.3.4 and/or 1.1.3.5, repeat 1.1.3.2 – 1.1.3.5 until all lines entered
H – K	1.1.3.1: Type the name of the patient
K	1.1.3.2: Press Alt + Enter to move to the next line
K	1.1.3.3: Enter additional text
H - P – BB	1.1.3.4: Click the "increase font size" button
H - P – BB	1.1.3.5: Click the "B" button to make the text bold
	1.1.4: Add data to the text field
	⇨ PLAN: Do 1.1.4.1, if desired do 1.1.4.2 and repeat until all data entered
H – P – BBBB – K	1.1.4.1: Type text into box
K	1.1.4.2: To make a new line, press Alt + Enter

Table 14. Task analysis for adding a patient to the prototype

KLM Model	Task 1.2: Add a patient to the prototype
	⇨ PLAN: Do 1.2.2, then 1.2.3, then 1.2.4 or 1.2.5 (repeat 1.2.3 – 1.2.5 until all desired data has been added)
P – BB	1.2.2: At the bottom of the page, click the →* icon
P – BB	1.2.3: Select a data field to edit
H - K	1.2.4: (If text box) Modify text of the data field
P	1.2.5: (If drop-down) Change the selection of the drop down menu

Table 15. Task analysis for modifying the footer of the report in the spreadsheet

KLM Model	Task 2.1: Modify the footer of a handoff report (Spreadsheet)
	⇨ PLAN: Do 2.1.1 then 2.1.2 then 2.1.3 then 2.1.4 then 2.1.5 then 2.1.6 then 2.1.7
P – B – P – B	2.1.1: Click View menu and select "Header and Footer"
P – BB	2.1.2: Click the "custom header" or "custom footer" button
P – BB	2.1.3: Click the "left section" text box
H - P – BB – K	2.1.4: Edit the text in the "left section" text box
H - P – BB	2.1.5: Click the "ok" button to close the editing window
P – BB	2.1.6: Click the "ok" button to close the header and footer window
P – B – P – B	2.1.7: Save changes: Select File -> Save from the top menu

Table 16. Task analysis for modifying the footer of the reports in the prototype

KLM Model	Task 2.2: Modify the footer of a handoff report
	⇨ PLAN: Do 2.2.1 then 2.2.2 then 2.2.3
P – BB	2.2.1: Click "modify footer" button on main menu
H - P – BB – K	2.2.2: Edit text in the text box of the "edit footer" window
H - P – BB	2.2.3: Click "return to main menu" button

When typing patient identifying information, the number of keys pressed is identical. Therefore it was modeled as a single keystroke given the difference in time would be the same no matter how many keys were pressed. For adding a patient, the spreadsheet required completion of four plans involving 33 separate Keystroke-Level actions, while the prototype required only one plan and nine actions. To modify the footer of the printed handoff reports, the spreadsheet required one plan and 26 Keystroke-Level actions while the prototype required one plan and only 12 actions.

Table 17. Predictive KLM-based models for two handoff tasks

Task	Tool	Predictive Equation	Total Time
Add Patient	Spreadsheet	$T_{execute} = 8T_P + 16T_B + 4T_H + 5T_K$	13.4 s.
	Prototype	$T_{execute} = 3T_P + 4T_B + T_H + T_K$	4.4 s.
Modify Footer	Spreadsheet	$T_{execute} = 9T_P + 14T_B + 2T_H + T_K$	12.4 s.
	Prototype	$T_{execute} = 3T_P + 6T_B + 2T_H + T_K$	5.0 s.

TP (move mouse to a point) = 1.1 s.
TB (time to press or release mouse button) = 0.1 s.
TH (time to home hands to keyboard or mouse) = 0.4 s.
TK (time to press a key on the keyboard) = 0.28 s.

Both tasks could be completed in a much shorter time using the prototype (Table 17). There was a larger discrepancy between the two tools for adding a patient but, overall, modifying the report footer was the most complex of the two tasks. Results such as these may explain why some residents found working with the spreadsheet to be "tedious, error-prone, and time consuming".

Discussion

Access to the Tools

One of the problems expressed with the original handoff tool was that it was only stored on a single computer away from the location of most patients. Therefore updates did not occur as often as they should have and patient information was often missing or completely left off the handoff sheet. The prototype was made available on a secure hospital server that could be accessed from any computer on the pediatric wards. This change allowed residents to update the database without leaving their current activities. Combining improved accessibility with the diminished complexity of common tasks improved efficiency.

Prototype Yoga – Flexibility through Increased Functionality

A number of tasks implemented in the database were not available with the spreadsheet. These unique aspects of the prototype involved two areas: support for handoff on additional pediatric services, and support for pediatric processes beyond handoff. The prototype provided five unique data entry/modification forms for patient data (handoff and daily progress notes for three pediatric services, a patient archive, and a patient dictation list, and the ability to transfer patients) compared to one in the spreadsheet with no specific way for transferring patients. The prototype provided eleven unique pre-sorted reports selected via drop-down menus for viewing and printing. Conversely, the spreadsheet could only print the current view of the data and required manual sorting to obtain different views. In addition, reports in the prototype automatically resize based on the amount of information entered. In the spreadsheet, a user would have to manually format the text of any overloaded cell in order to adjust the size of the printed sheet.

Creating Successful Collaboration with End-Users

Collaboration can be difficult when working with "hard to reach" end users, let alone gaining their interest in the process. However, a few practical strategies that are universally applicable in situations where collaboration is necessary but difficult to achieve were useful to this work.

A multifaceted strategy is important. Limiting collaboration to a single strategy may isolate end-users, gain the interest of only those who thrive on the method chosen, and skew feedback. The strategy included focus groups, interviews, feedback "on the fly", and an iterative design process. The focus groups were a central aspect for creating buy-in. In addition, the focus groups helped create a sense of teamwork not only between the end-users and the systems engineers, but also among the end-users themselves. Interviews were an extension of the focus groups that allowed the engineers to gain a more detailed and expert perspective of the process the tool was meant to support, a critical step toward its overall improvement.

Providing users with the opportunity to give feedback and comments in the context of using the tool was beneficial in a number of ways. First, it was a way to gather feedback "on-the-fly" and minimize the time in which a user had to commit his or her comments to memory before reporting them. Second, it helped users gain a sense of personal involvement and interaction in the design process and the feeling that every comment was principally important to the final product. Finally, allowing users to provide feedback during the design of the tool was an easy way to note necessary improvements for each version that was developed, and ultimately allowed the systems engineers to iteratively incorporate user suggestions into the prototype's design.

An iterative design process was essential in creating an atmosphere that did not pressure the system designers to develop a "perfect" tool on the first try, or force the end users to "get it right" the first time either. Over time, it allowed modifications to be incrementally adapted in the prototype in an effort to shape the tool to the end-users' needs. In addition, it allowed the end-users to feel as if they could provide negative (as well as positive) feedback without the concern that it would be detrimental to the implementation of the tool or to their collaborations with the systems engineers. An iterative process also helped reinforce and maintain end-user interest by periodically providing new and interesting designs to review instead of making them wait for the final product.

When designing a computerized process support tool, it is critical to clearly understand the process and develop requirements that support that process. End-users must be involved early to peak interest, and continuously, because additional requirements become apparent after the tool or prototype is put into operation. Collaboration with end-users throughout the design activities enhances "buy-in" to the system and encourages the end-users to modify and improve the process on their own.

Conclusion

The potential of computerized hand-off systems appears limited only by technology and the imaginations of the developers of these systems. A number of novel improvements are already under consideration for inclusion in existing systems. For example, a very important concept for these systems is the ability to dynamically present information based on the characteristics of the user. This type of capability adds a level of "intelligence" to an information system that can help reduce inefficiency. A topic in this area that has been widely discussed in the literature is automated decision support. While hand-off reports are not necessarily meant to completely scaffold decision making, there is certainly an opportunity for reducing and filtering information in order to direct physicians toward a particular set of unusual or hard to find data, or convert raw data into more useful graphical displays. Intelligent displays of information can also derive suggestions, send prompts, or highlight cues for providers which may help redirect care plans or behaviors, for example, in the event a patient's condition or trajectory has changed based on newly available or updated test results.

System integration issues such as the integration of hand-off systems with other clinical information systems is critical for the success of intelligent data displays in medicine. Some systems already have these capabilities and it is important for emergent systems to learn from them, both in how information is exchanged between systems and the impact of integrated systems on workflow and patient care. Intelligent, information systems that share and integrate data between medical services, within an entire hospital system, and across the boundaries of healthcare organizations, have become a foundation for ongoing efforts to improve patient safety and the overall quality of patient care.

Chapter 7

Conclusion

A context for organizing this work: the phases of handoff

This work involved the study of the process of resident physician handoff at shift change in the pediatric acute care setting. It focused on three major aspects of the handoff process: the preparation phase, the exchange phase and the utilization phase. The process was first discussed in Chapter 1 and is again presented here for reference (Figure 19).

Ideally, preparation for handoff for leaving residents involves gathering, organizing, summarizing, and updating patient information, checking on the status of tasks in progress, and planning for the immediate future. For incoming residents, preparation involves reviewing and assessing the current state of available patient information, and the preparation of questions to ask during handoff to ensure that the discussion addresses the information they feel they need. The outcome of the preparation phase is exhibited by the information discussed by residents during the exchange phase. In this phase, leaving residents exchange information and transfer patient care responsibility and authority to incoming residents through a communicative event. From the end of the handoff exchange, through the subsequent preparation phase at the end of their shift, residents are in the utilization phase.

In this phase, residents utilize information available to them (which they may or may not have received during the exchange) to make medical decisions, care for patients, and to address the requests and concerns of other care providers, as well as patients, their family members, and other individuals they may encounter over the course of their shift. Information residents require in this phase may be internally known, externally available, pending, not pending, missing, unknown, or impossible to determine.

Figure 19. The three phases of the handoff process: preparation, exchange, and utilization. Information is represented by the encircled letter **i's**, arrows depict information flow

Review of major findings

Chapter 4 focused on the utilization phase of handoff. Specifically, characterizing how information residents exchange during handoff supports their role as information sources for other providers engaged in patient care activities, the results are published in [89]. Verbal communications continue to be an important way that information is transferred between healthcare providers in the inpatient setting. Residents are asked many questions during their overnight shifts and are able to answer many of them without consulting external sources. When external sources are consulted, synchronous communications such as phone calls to other healthcare providers are common in addition to checking written and electronic sources. As synchronous communications such as phone calls continue to be important pathways for information flow, informatics researchers and designers need to consider the relationship between such communications and workflow in the development of healthcare decision support tools.

As previously discussed, there is a lack of evidence to support exactly what information *should be discussed* during handoff, therefore, though there has been data presented thus far in this body of work which suggest at least the need for further research. Chapter 4 looked at how information discussed at handoff played a role in the outcome of questions residents received on the ensuing shift. In particular, whether there was a relationship between discussing information at handoff and

residents' responses to the questions they received. The results of Research Question 1.2 are planned for publication.

Chapter 5 focused on the exchange phase of handoff. Specifically, on further characterizing the handoff process in terms of the information residents' exchange and on evaluating the efficacy of training aimed at teaching what information to discuss during the exchange phase of handoff employing a behavioral evaluation strategy to assess the transfer of training to job performance.

With respect to the characteristics of the information residents discuss during handoff, the most important finding in Chapter 5 was that training a set of information to discuss for each patient resulted in an increase in the proportion of patients for which *the entire set* of that information gets discussed, suggesting that training is an effective way to teach residents to discuss the complete picture of information for each patient. The results of the study also suggest that training can be an effective way to teach residents how to limit the volume of information they discuss for each patient, and therefore reduce the amount of time it takes to conduct handoffs. While this study did not assume that the content of the training was the most optimal for improving patient care, the results suggest that if a core set of information were to be defined that are correlated with improved patient care, then a training program such as that in this study would be an effective way to change handoff from a process that puts patient safety at risk to a process that improves it. Ultimately it may be possible to reduce the frequency of adverse events related to incomplete information resulting from poorly conducted handoffs, simply by training a minimum core set of the information to discuss.

Considerable attention has been given to improving information technology which supports the preparation phase of handoff, this was discussed in detail in Chapter 6. For the most part, these solutions have been developed without engaging physicians in the design process. As information technology has vital role in the health care industry, an ancillary objective of this work is to evaluate the impact of a user-centered design process for building an information system to support handoff and to provide guidelines for the development of future information support systems for handoff. While hand-off reports are not necessarily meant to completely scaffold decision making, there is certainly an opportunity for reducing and filtering information in order to direct physicians toward a particular set of unusual or hard to find data, or convert raw data into more useful graphical displays. Intelligent displays of information can also derive suggestions, send prompts, or highlight cues for providers which may help redirect care plans or behaviors, for example, in the event a patient's condition or trajectory has changed based on newly available or updated test results.

System integration issues such as the integration of hand-off systems with other clinical information systems is critical for the success of intelligent data displays in medicine. Some systems already have these capabilities and it is important for emergent systems to learn from them, both in how information is exchanged between systems and the impact of integrated systems on workflow and patient care. Intelligent, information systems that share and integrate data between medical services, within an entire hospital system, and across the boundaries of healthcare organizations, have become a foundation for ongoing efforts to improve patient safety and the overall quality of patient care.

Limitations and future work

Scope

One goal of this work was to remain as objective as possible, and thus maintained a narrow scope with respect to the handoff process. That is, viewing handoff only from the perspective of information flow and utilization as opposed to a more comprehensive view involving other relevant aspects of the highly complex handoff process (e.g. team interaction, coordination, and communication strategies for co-orientation). While this limited scope may have been beneficial to the methodological aspects of this work, it also limits the generalizability of the results and the conclusions which can be drawn to other organizations or hospital services, particularly the procedural specialties (e.g. surgery).

Furthermore, this work does not attempt to measure what would typically be considered in the health care domain as "traditional" patient care outcomes. However, alternative outcome measures are utilized in this work to in order to achieve similar goals: a better understanding of the impact of handoffs on patient care.

Finally, this data gathered during the course of this research are from a one very specific setting (one hospital, one service, one type of care provider), thus the generalizability of the outcomes of this work will be limited primarily to the methods employed, such as the procedures utilized to gather and analyze data. On the other hand, the results should at least be generalizable to resident handoffs in pediatric acute care settings with a similarly sized residency program with the ability to care for a similar volume of patients.

Handoff training

Lack of a true control group hindered the power of the results discussed in Chapter 5. While significant results were shown, there are still factors not accounted for which could have had unintended influences. In addition, the comparisons were based on loosely similar groups of participants (in the case of the baseline versus the pre-training comparisons) or on a small number of repeated measures (in the case of comparing pre-training observation 1 to pre-training observation 2). Future work on assessing the behavioral impact of handoff training interventions will benefit from a larger population sample than was available in this study, as well as an identical or characteristically very similar control group (in addition to an experimental group) and a randomization method which determines which group participants are assigned to. Conducting a study on a larger residency service or across multiple institutions would be a way to secure the large sample sizes which would be necessary. The latter, in particular, would also provide valuable insight into how handoffs are conducted across organizations and would likely result in the development of a more robust, generalizable method for collecting observational data on handoffs specifically.

While a behavioral change was measured on most accounts following training, this study provides little evidence to support conceptual learning. Results of the learning assessments indicated residents generally scored high both pre-and post-training. This result was not surprising as the knowledge assessment was short, poorly aligned pre to post-training, and poorly validated as a knowledge elicitation tool. However, the results of the observational assessment do suggest that residents' have a working knowledge of the concepts from the training (perhaps reflected by their generally high scores on the knowledge assessments) which they were able to transfer to their handoff behaviors. An improved knowledge assessment will provide future work with the ability to test whether conceptual learning from training material occurs in addition to behavioral change. Use of both behavioral and knowledge assessment methods would be a very robust method for evaluating the impact of a training intervention particularly in the area of graduate medical education.

The relevancy of the training content was also a concern in this study. While the choice to train five critical information components to discuss at handoff was not entirely arbitrary and is, for the most part, supported by prior research and medical expert opinion, these five components may not represent *the* most important information to discuss at handoff, in general. It may also be the case that the particular categorization of the five components was all-inclusive but ill-defined. Even at an expert level, care providers disagree on how to categorize the same part of a discussion, and it is

likely that many different categorization schemes could be just as inclusive but provide vastly different results when used in characterizing handoff discussions.

Finally, the choice to focus training only on the content of handoff discussions, although acceptable as a constraint in this study, is a limited view of how training could address the process as a whole. Handoff is a vastly more complex process than simply that of information exchange. It involves factors of interpersonal communication, situational awareness, sense-making, co-orientating, and many more factors discussed frequently in the literature on handoffs. The initial goal of the particular training system developed for this study was to start with simple concepts for newer residents followed by increasingly abstract concepts as residents gained experience over the course of their second and third years. Future work will continue on this course having gained knowledge from the methods in this study. Ultimately, while training the content of handoff is critical in itself, future training efforts will likely need to focus on training handoff concepts and skills throughout the spectrum of medical experience. Specifically, handoff training and evaluations need to be developed for medical students, physicians in their later years of residency, as well as continuing education for post-residency care providers so they maintain proficiency throughout their careers.

Residents as information systems

It is possible that there are many other sources from which residents solicit information that were not discussed in Chapter 4. Thus, as some authors have suggested [20], future studies in this area should account for a broad spectrum of potential information sources available to residents. The greatest limitation of the study discussed in Research Question 1.1 was an inability to clarify details related to clinical judgments. There are many ways to speculate as to the source of the information supporting residents' clinical judgments. In addition, it will be important to consider whether information is being discussed during handoffs and ultimately forgotten, not being discussed at handoffs, or if other, less conspicuous sources of information are supporting residents' clinical judgments. In the latter case, it may even be practical to include these additional sources as elements of handoff.

References

[1] Brown PJ, Borowitz SM, Novicoff W. Information exchange in the NICU: what sources of patient data do physicians prefer to use? International journal of medical informatics. 2004;73:349-55.

[2] Coiera E. When conversation is better than computation. Journal of the American Medical Informatics Association. 2000;7:277-86.

[3] Parker J, Coiera E. Improving clinical communication. Journal of the American Medical Informatics Association. 2000;7:453-61.

[4] Commission J. Standards Project Focuses Attention on High-Value Standards, Cuts 16 EPs. The Joint Commission Perspectives. 2010;30:1-6.

[5] ACGME | Accreditation Council for Graduate Medical Education.

[6] Arora V, Johnson J. A model for building a standardized hand-off protocol. Joint Commission journal on quality and patient safety / Joint Commission Resources. 2006;32:646-55.

[7] Borowitz SM, Waggoner-Fountain LA, Bass EJ, Sledd RM. Adequacy of information transferred at resident sign-out (in-hospital handover of care): a prospective survey. Qual Saf Health Care. 2008;17:6-10.

[8] Lee LH, Levine JA, Schultz HJ. Utility of a standardized sign-out card for new medical interns. Journal of general internal medicine. 1996;11:753-5.

[9] Nemeth CP, Kowalsky J, Brandwijk M, Kahana M, Klock PA, Cook RI. Before I Forget: How Clinicians Cope with Uncertainty Through ICU Sign-Outs. p. 939-43.

[10]   Van Eaton EG, Horvath KD, Lober WB, Pellegrini CA. Organizing the transfer of patient care information: the development of a computerized resident sign-out system. Surgery. 2004;136:5-13.

[11]   Lofgren RP, Gottlieb D, Williams RA, Rich EC. Post-call transfer of resident responsibility: its effect on patient care. J Gen Intern Med. 1990;5:501-5.

[12]   Patterson ES. Structuring flexibility: the potential good, bad and ugly in standardisation of handovers. Qual Saf Health Care. 2008;17:4-5.

[13]   Patterson ES, Wears RL. Patient handoffs: standardized and reliable measurement tools remain elusive. Joint Commission journal on quality and patient safety / Joint Commission Resources. 2010;36:52-61.

[14]   Behara R, Wears R, Perry S, Eisenberg E, Murphy L, Vanderhoef M, et al. A Conceptual Framework for Studying the Safety of Transitions in Emergency Care. Advances in Patient Safety: From Research to Implementation. 2005;2.

[15]   Baker DP, Gustafson S, Beaubien JM, Salas E, Barach P. Medical team training programs in health care: Citeseer; 2005.

[16]   Chakraborti C, Boonyasai R, Wright S, Kern D. A systematic review of teamwork training interventions in medical student and resident education. J Gen Intern Med. 2008;23:846 - 53.

[17]   Petersen L, Brennan T, O'Neil A, Cook E, Lee T. Does Housestaff Discontinuity of Care Increase the Risk for Preventable Adverse Events? Annals of Internal Medicine. 1994;121:866-72.

[18]   Yang X, Koh RYI, Tay BTC, Siah KTH, Donchin Y, Park T. Clinical shift handoffs in Singapore A three-phase prospective. Proceedings of the Human Factors and Ergonomics Society Annual Meeting: SAGE Publications; 2011. p. 778-82.

[19]   Gorman PN. Information needs of physicians. Journal of the American Society for Information Science. 1995;46:729-36.

[20]   Forsythe DE, Buchanan BG, Osheroff JA, Miller RA. Expanding the concept of medical information: an observational study of physicians' information needs. Computers and biomedical research, an international journal. 1992;25:181-200.

[21]   Stead W. Computational technology for effective health care : immediate steps and strategic directions. Washington D.C.: National Academies Press; 2009.

[22]   Anand SX, Kim MC, Kamran M, Sharma SK, Kini AS, Fareed J, et al. Comparison of platelet function and morphology in patients undergoing percutaneous coronary intervention receiving bivalirudin versus unfractionated heparin versus clopidogrel pretreatment and bivalirudin. The American journal of cardiology. 2007;100:417-24.

[23]   Detmer DE. Engineering information technology for actionable information and better health. Information, Knowledge, Systems Management. 2009;8:107-18.

[24]   Davenport TH. Saving IT's Soul: Human-Centered Information Management. Harvard business review. 1994;72:119-31.

[25]   Sittig DF, Ash JS. Clinical information systems: overcoming adverse consequences: Jones & Bartlett Learning; 2009.

[26]   Sittig DF, Classen DC. Safe Electronic Health Record Use Requires a Comprehensive Monitoring and Evaluation Framework. JAMA: The Journal of the American Medical Association. 2010;303:450-1.

[27]   Karsh B-T, Weinger MB, Abbott PA, Wears RL. Health information technology: fallacies and sober realities. Journal of the American Medical Informatics Association. 2010;17:617-23.

[28]   Mayer D, Klamen D, Gunderson A, Barach P, Telluride IR. Designing a patient safety undergraduate medical curriculum: the Telluride Interdisciplinary Roundtable experience. Teaching and learning in medicine. 2009;21:52.

[29]   Ahmed J, Mehmood S, Rehman S, Ilyas C, Khan LUR. Impact of a structured template and staff training on compliance and quality of clinical handover. International Journal of Surgery. 2012;10:571-4.

[30]   Borowitz SM, Waggoner-Fountain LA, Bass EJ. Impact of a Computerized System on Resident Sign-Out. Biomedical Instrumentation & Technology. 2013;47:68-72.

[31]   Patterson ES, Roth EM, Woods DD, Chow R, Gomes JO. Handoff strategies in settings with high consequences for failure: lessons for health care operations. International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua. 2004;16:125-32.

[32]   Cohen MD, Hilligoss PB. Handoffs in Hospitals: A review of the literature on information exchange while transferring patient responsibility or control. 2009.

[33]   Solet DJ, Norvell JM, Rutan GH, Frankel RM. Lost in translation: challenges and opportunities in physician-to-physician communication during patient handoffs. Acad Med. 2005;80:1094-9.

[34]   Charap M. Reducing resident work hours: Unproven assumptions and unforeseen outcomes. Annals of Internal Medicine. 2004;140:814-5.

[35]   Perry S. Transitions in Care: Studying Safety in Emergency Department Signovers. Focus on Patient Safety. 2004;7:1-3.

[36]   Horwitz LI, Krumholz HM, Green ML, Huot SJ. Transfers of patient care between house staff on internal medicine wards: a national survey. Arch Intern Med. 2006;166:1173-7.

[37]   Horwitz LI, Moin T, Green ML. Development and implementation of an oral sign-out skills curriculum. J Gen Intern Med. 2007;22:1470-4.

[38]   Horwitz LI, Moin T, Krumholz HM, Wang L, Bradley EH. What are covering doctors told about their patients? Analysis of sign-out among internal medicine house staff. Quality and Safety in Health Care. 2009;18:248.

[39]   Bernau S, Aldington S, Robinson G, Beasley R. From medical doctor to junior doctor: The medical handover -- a good habit to cultivate. studentBMJ. 2006;14:188-9.

[40]   Roughton VJ, Severs MP. The junior doctor handover: current practices and future expectations. J R Coll Physicians Lond. 1996;30:213-4.

[41]   Riesenberg LA, Leitzsch J, Massucci JL, Jaeger J, Rosenfeld JC, Patow C, et al. Residents' and attending physicians' handoffs: a systematic review of the literature. Academic Medicine. 2009;84:1775.

[42]   Foster S, Manser T. The effects of patient handoff characteristics on subsequent care: a systematic review and areas for future research. Acad Med. 2012;87:1105-24.

[43]   Berner ES, Moss J. Informatics challenges for the impending patient information explosion. Journal of the American Medical Informatics Association. 2005;12:614-7.

[44]   Lardner R. Effective Shift Handover1996.

[45]   Patterson ES, Woods DD. Shift changes, updates, and the on-call architecture in space shuttle mission control. Computer supported cooperative work : CSCW : an international journal. 2001;10:317-46.

[46]   McLeod JM, Chaffee SH. Interpersonal approaches to communication research. American Behavioral Scientist. 1973;16:469.

[47]   Borowitz SM, Waggoner-Fountain LA, Bass EJ, Sledd RM. Adequacy of information transferred at resident sign-out (in-hospital handover of care): a prospective survey. Quality and Safety in Healthcare. 2008;17:6-10.

[48]   Patterson ES. Communication strategies from high-reliability organizations: translation is hard work. Ann Surg. 2007;245:170-2.

[49]   Patterson E, Woods D, Cook R, Render M. Collaborative cross-checking to enhance resilience. Cognition, Technology & Work. 2007;9:155-62.

[50]   Brown JP. Collaborative Cross-Checking. Improving Health Care Team Communication: Building on Lessons from Aviation and Aerospace. 2008:155-78.

[51]   Arora V, Johnson J, Lovinger D, Humphrey HJ, Meltzer DO. Communication failures in patient sign-out and suggestions for improvement: a critical incident analysis. Qual Saf Health Care. 2005;14:401-7.

[52]   Enfield KB, Hoke G. Discontinuity of care: further thoughts on standardized processes. Journal of Hospital Medicine. 2007;2:115-6.

[53]   Riesenberg LA, Leitzsch J, Little B. Systematic Review of Handoff Mnemonics Literature. American Journal of Medical Quality. 2009;24:196-204.

[54]   Haig KM, Sutton S, Whittington J. SBAR: A shared mental model for improving communication between clinicians. Joint Commission Journal on Quality and Patient Safety. 2006;32:167-75.

[55]   Leonard M, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. Quality and Safety in Health Care. 2004;13:i85-i90.

[56]   Carbo A, Li J. The hospitalist to primary care handoff. In: LaValley D, editor. Reducing risks during handoff. Cambridge, MA: CRICO/RMF; 2007.

[57]   Fogerty RL, Schoenfeld A, Salim Al-Damluji M, Horwitz LI. Effectiveness of written hospitalist sign-outs in answering overnight inquiries. Journal of Hospital Medicine. 2013;8:609-14.

[58]   McSweeney ME, Landrigan CP, Jiang H, Starmer A, Lightdale JR. Answering questions on call: Pediatric resident physicians' use of handoffs and other resources. Journal of Hospital Medicine. 2013;8:328-33.

[59]   Sharit J, McCane L, Thevenin DM, Barach P. Examining issues in communicating patient care information across shifts in a critical care setting. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 11 ed: SAGE Publications. p. 1062-6.

[60]   Chakraborti C, Boonyasai RT, Wright SM, Kern DE. A systematic review of teamwork training interventions in medical student and resident education. J Gen Intern Med. 2008;23:846-53.

[61]   Baldwin TT, Ford JK. Transfer of training: A review and directions for future research. Personnel Psychology. 1988;41:63-105.

[62]   Neuendorf KA. The content analysis guidebook: Sage Publications; 2002.

[63]   Krippendorff K. Content analysis an introduction to its methodology. Thousand Oaks, Calif.: Sage; 2004.

[64]   Weber RP. Basic content analysis. Newbury Park, Calif.: Sage Publications; 1990.

[65]   Cohen J. A coefficient of agreement for nominal scales. Educational and psychological measurement. 1960;20:37-46.

[66]   Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33:159-74.

[67]   Sledd RM, Bass EJ, Borowitz SM, Waggoner-Fountain L. Supporting the characterization of sign-out in acute care wards. Systems, Man and Cybernetics, 2006 SMC'06 IEEE International Conference on: IEEE; 2006. p. 5215-20.

[68]   Catchpole KR, de Leval MR, McEwan A, Pigott N, Elliott MJ, McQuillan A, et al. Patient handover from surgery to intensive care: using Formula 1 pit-stop and aviation models to improve safety and quality. Paediatric anaesthesia. 2007;17:470-8.

[69]   Card SK, Moran TP, Newell A. The psychology of human computer interaction: Routledge; 1983.

[70]   Waggoner-Fountain LA, Tomney BA, DeVoge JM, Bass EJ, Borowitz SM. Questions asked and answers located: What really happens on-call after sign-out. Pediatric Academic Societies Annual Meeting.

[71]   Perry S, Wears R, Patterson E. High-Hanging Fruit: Improving Transitions in Health Care2008.

[72]   Chase HS, Kaufman DR, Johnson SB, Mendonca EA. Voice capture of medical residents' clinical information needs during an inpatient rotation. Journal of the American Medical Informatics Association. 2009;16:387-94.

[73]   Coumou HCH, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. Journal of the Medical Library Association. 2006;94:55-60.

[74]   Covell DG. Information needs in office practice: are they being met? Annals of Internal Medicine. 1985;103:596.

[75]   Mueller FF, Kethers S, Alem L, Wilkinson R. From the certainty of information transfer to the ambiguity of intuition. Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments: ACM. p. 63-70.

[76]   Chu ES, Reid M, Schulz T, Burden M, Mancini D, Ambardekar AV, et al. A structured handoff program for interns. Acad Med. 2009;84:347-52.

[77]   Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM Computing Surveys (CSUR). 2009;41:1-58.

[78]   Arora VM, Johnson JK, Meltzer DO, Humphrey HJ. A theoretical framework and competency-based approach to improving handoffs. Qual Saf Health Care. 2008;17:11-4.

[79]   De Voge JM, Bass EJ, Atia M, Bond M, Waggoner-Fountain LA, Borowitz SM. The development of a Web-based resident sign-out training program. Systems, Man and Cybernetics, 2009 SMC 2009 IEEE International Conference on: IEEE; 2009. p. 2509-14.

[80]   Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. Bmj. 1995;310:170.

[81]   Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance patient safety: a systematic review. Med Educ. 2012;46:1042-54.

[82]   Bass EJ. Characterization of pediatric resident sign-out to inform process improvement. Applied Ergonomics.

[83]   Keyes C. Coordination of care provision: The role of the 'handoff'. International Journal for Quality in Health Care. 2000;12:519-.

[84]   Hooper K. Architectural design: an analogy. In: Norman DA, Draper SW, editors. User centered system design. Hillsdale, NJ: Erlbaum; 1986.

[85]   Horwitz LI, Krumholz HM, Huot SJ, Green ML. Internal medicine residents' clinical and didactic experiences after work hour regulation: a survey of chief residents. J Gen Intern Med. 2006;21:961-5.

[86]   Okie S. An Elusive Balanceâ€"Residents' Work Hours and the Continuity of Care. New England Journal of Medicine. 2007;356:2665-.

[87]   Vidyarthi AR, Arora V, Schnipper JL, Wall SD, Wachter RM. Managing discontinuity in academic medical centers: strategies for a safe and effective resident sign-out. Journal of hospital medicine : an official publication of the Society of Hospital Medicine. 2006;1:257-66.

[88]   Sledd R, Bass E, Borowitz S, Waggoner-Fountain L. Supporting the characterization of sign-out in acute care wards.

[89]   Bass EJ, Devoge JM, Waggoner-Fountain LA, Borowitz SM. Resident physicians as human information systems: sources yet seekers. Journal of the American Medical Informatics Association : JAMIA. 2012.

Appendices

Appendix A. The five information components of sign-out

PL-1 Module – Five Information Components of Sign-Out – **Identifying the Patient**

Identifying the patient is the most accurate way to differentiate the patients you discuss during sign-out. While the name and/or MRN are the typical ways of identifying the patient, other ways such as nicknames may also be useful. The most important thing is to ensure everyone in the conversation knows exactly what patient is currently being discussed.

Forgetting to or incorrectly identifying a patient can leave the incoming resident with the wrong information about a patient, or worse, result in a patient being overlooked entirely during the night.

Identifying each patient during sign-out:

- Helps differentiate between patients
- Ensures the right data is discussed for each patient ("everyone on the same page")
- Reduces the chance that a patient is overlooked or forgotten overnight

---

PL-1 Module – Five Information Components of Sign-Out – **The Big Sentence**

The purpose of the big sentence is to "paint a picture" of the patient. It is a visual aid that allows anyone to identify the patient based on a description. The big sentence includes the patient's age, primary diagnosis and chronic problems. The mention of meds is not required in the big sentence.

The important elements of The Big Sentence are:

- Age
- Primary Diagnosis
- Chronic Problems (pertinent to this admission/shift)

PL-1 Module – Five Information Components of Sign-Out – **The Big Sentence**

The big sentence includes the patient's age, primary diagnosis and chronic problems. These data are often scattered throughout the sign-out sheet so it is important to learn to recognize and assemble them for patient discussions.

Listen as a resident gives the big sentence for a patient and follow along as each of the elements are highlighted on the sign-out sheet to the left.

PL-1 Module – Five Information Components of Sign-Out **– The Current Condition**

The current condition identifies the patient's current physical and mental state relevant to the current diagnostic and/or treatment plan. It identifies a patient's baseline norms, VS, O2 sat, and rhythm. Knowing the current condition for a patient involves answering the questions…

- How does the patient look in comparison to the **big sentence**?
- What does the patient look like at the moment?
- What did I visually see about the patient *most recently*?
- Does the patient look fine now after having the seizure one hour ago?

In general, the **current condition** should tell you:

- How the patient looks *now*
- Baseline norms, VS, O2 sat, rhythm

PL-1 Module – Five Information Components of Sign-Out – **The Current Condition**

Like the big sentence, information about a patient's current condition may be scattered throughout the sign-out sheet so it is important to learn to recognize the current condition during sign-out discussions.

In some cases, it may be necessary to elicit the current condition from the clinician giving sign-out to you. Listen to the following example and note how the resident receiving sign-out must ask for the patient's current condition.

PL-1 Module – Five Information Components of Sign-Out – **The Plan of Care**

The plan of care gives a caregiver a set of instructions to perform. The caregiver recommends what to do about lab and test results, and I/O parameters; and plans a timely discharge that continues the continuity of healthcare and prevents readmissions. The caregiver also makes referrals for matters out of his/her scope.

Plan of Care Summary:

- Lab and test results pending/to check and what to do about results
- I and O parameters and what to do about them
- Knowing when and who to contact
- Discharge planning

PL-1 Module – Five Information Components of Sign-Out – **The Plan of Care**

The plan of care can be very detailed, or briefly stated. The important thing is recognizing it when it is discussed.

Listen to the following examples of residents giving the plan of care for their patients and note the differences between each example.

PL-1 Module – Five Information Components of Sign-Out – **Contingency Plans**

A contingency plan helps healthcare providers make better decisions. If the rationale for the plan of care is known, then more informed clinical decisions can be made during unexpected situations. The contingency plan should be in the form of an if/then statement. For example, if lab results show_____, then you may want to call _____.

Contingency Plans Summary

- Symptom management/action triggers
    - What to do, who to call and how to reach them
- Access (where, when/if to replace)
- Monitoring guidelines
- If-Then statements

In 49/158 (31%) surveys, resident physicians indicated something happened while they were on call for which they were not adequately prepared. In 40 of these 49 (82%) instances, they indicated there was information they did not receive during sign-out that would have been helpful to them in caring for a patient overnight, and in 33 of these 40 (82.5%) instances, they indicated the situation should have been anticipated and discussed during sign-out. (Borowitz et al, 2008)

''I would like to know who might bring me trouble or who is going to get sick. I think that it is helpful to prioritize sickness and what you are supposed to do.'' (Aurora, et al., 2005)

PL-1 Module – Five Information Components of Sign-Out – **Contingency Plans**

In some cases, it may be appropriate to acknowledge that there are no contingency plans associated with a particular patient. Listen to the following examples and note when the residents acknowledge that there are no important contingency plans.

Appendix B. Pre and Post-training Handoff Knowledge Assessment Questions

Scene Markers For PL-1 Assessment Questions

Scene Marker	Role In Strategy	Question Text	Solution (Feedback)	Performance Objective(s)
1-0-2	Pretest 1	Which statement is true? a) Healthcare professionals, who self-report and make erroneous assumptions, do not compromise the quality of care for patients. b) One-way communication in patient sign-out can assist a resident in conveying vital patient information c) Breakdown in communication is not a major problem in handover of care. d) Conducting patient sign-out over the phone can substitute for face-to-face communication.	B	Overall Training
1-0-3	Pretest 2	There are five information components that should be discussed during sign-out, choose the correct five from the following list: a) Detailed history b) Parent/guardian contact info c) Name and/or MRN d) Big sentence e) Covering intern f) Admit date g) Contingency plans h) Plan of care i) Current Condition j) Teaching points	C, D, G, H, I	Objective 1
1-0-4	Pretest 3	You are giving sign-out and need to discuss twins. What element would best identify each for the receiver? e) Full name f) Time of birth g) Sex h) Birth weight i) MRN	E	Objective 1
1-0-5	Pretest 4	The purpose of a big sentence is to_____ a) Outline what to do for the patient overnight b) Paint a picture of the patient c) Describe what the patient looked like when you last saw them d) Identify the patient e) List the patient's medications	B	Objective 1b
1-0-6	Pretest 5	Which three components constitute the big sentence? a) Age, room number, service b) Chronic problems, age, and diagnosis c) Admit, age, chronic problems d) Medication, diagnosis, chronic problems	B	Objective 1c

1-0-7	Pretest 6	In general, the purpose of discussing the current condition of the patient is to _____ a) relay the patient's lab results over the last 24 hours b) list the patients known allergies to medications c) describe the most recently known physical and mental state of the patient d) describe the patient's history and reason for the current admission e) list what to do if something happens with the patient overnight	C	Objective 1d
1-0-8	Pretest 7	What two elements should the current condition of a patient tell you? a) How the patient looks now; baseline norms b) Current medications; known allergies c) Diagnosis; treatment plan d) Lab results pending; lab results known	A	Objective 1e
1-0-9	Pretest 8	The plan of care describes procedures, tasks, and overall goals for caring for the patient. a) True b) False	A	Objective 1g
1-0-10	Pretest 9	Which statement (s) is (are) true about the plan of care when signing out patients? a) The plan of care may discuss what test results are pending and what to do about them b) The plan of care may discuss I/O parameters and what to do about them c) The plan of care may discuss when and who to contact d) The plan of care may discuss when the patient should be discharged e) All of the above. f) None of the above.	E	Objective 1f
1-0-11	Pretest 10	Contingency plans never prepare a resident for unexpected events a) True b) False	B	Objective 1h
1-0-12	Pretest 11	A complete sign-out for a patient means that you have discussed all of the information for the patient listed on the sign-out sheet a) True b) False	B	
1-5-1	Posttest 1	Which statement is true? a) Healthcare professionals, who self-report and make erroneous assumptions, do not compromise the quality of care for patients. b) One-way communication in patient sign-out can assist a resident in conveying vital patient information c) Breakdown in communication is not a major problem in handover of care. d) Conducting patient sign-out over the phone can substitute for face-to-face communication.	B	Overall Training

1-5-2	Posttest 2	There are five information components that should be discussed during sign-out, choose the correct five from the following list: a) Detailed history b) Parent/guardian contact info c) Name and/or MRN d) Big sentence e) Covering intern f) Admit date g) Contingency plans h) Plan of care i) Current Condition j) Teaching points	C, D, G, H, I	Objective 1
1-5-3	Posttest 3	Twins are born in the NICU with MRNs that are one digit apart. In giving sign-out, what element would best identify each to the receiver? j) Full name k) Time of birth l) Sex m) Birth weight n) MRN	E	Objective 1
1-5-4	Posttest 4	Which three components constitute the big sentence? a) Age, room number, service b) Chronic problems, age, and diagnosis c) Admit, age, chronic problems d) Medication, diagnosis, chronic problems	B	Objective 1c
1-5-5	Posttest 5	After listening to the patient sign-out, did the caregiver state the big sentence? a) Yes b) No	A	Objective 2a, b
1-5-6	Posttest 6	What two elements should the current condition of a patient tell you? a) How the patient looks now; baseline norms b) Current medications; known allergies c) Diagnosis; treatment plan d) Lab results pending; lab results known	A	Objective 1e
1-5-7	Posttest 7	After listening to the sign-out example, was the patient's current condition discussed? o) Yes p) No	B	Objective 2a, b
1-5-8	Posttest 8	After listening to the patient sign-out, did the clinician communicate a plan of care for the patient? q) Yes r) No	A	Objective 2a,b
1-5-9	Posttest 9	After listening to the sign-out example, did the clinician communicate a contingency plan(s)? s) Yes t) No	A	Objective 2a, b

1-5-10	Posttest 10	After listening to the sign-out example, which component(s) did the speaker discuss? a) Big sentence b) Name and/or MRN c) Contingency plans d) Plan of care e) Current Condition	A, B, C, D	Objective 2b
1-5-11	Posttest 11	After listening to the sign-out example, was the discussion for the patient complete (all five components given by the speaker)? a) Yes b) No, did not identify the patient c) No, did not say the big sentence d) No, did not describe the patient's current condition e) No, did not outline the plan of care f) No, did not give contingency plans g) No, did not discuss multiple elements	A	Objective 2c

The eleven test questions on the pre-test were not mapped identically in all cases to the eleven questions on the post-test but were designed to address identical content. The pre and post-test questions were reviewed by medical experts to ensure questions with similar content were of equal difficulty between the pre and post-test. After review, it was determined that only five of the eleven questions were of sufficient similarity to be included in the analysis.

Table 18. Revised handoff training knowledge assessment questions and grouped pre/post-test scores

Question	Ques. #		Total Responses	N correct (SD)		Pre/Post Differential
	Pre	Post		Pre	Post	
Which statement is true? a) Healthcare professionals, who self-report and make erroneous assumptions, do not compromise the quality of ☑ care for patients. b) One-way communication in patient handoff can assist a resident in conveying vital patient information c) Breakdown in communication is not a major problem in handover of care. d) Conducting patient handoff over the phone can substitute for face-to-face communication.	01	01	10	9 (0.3)	10 (0.0)	+ 1
There are five information components that should be discussed during handoff, choose the correct five from the following list: e) Detailed history f) Parent/guardian contact info ☑ g) Name and/or MRN ☑ h) Big sentence i) Covering intern j) Admit date ☑ k) Contingency plans ☑ l) Plan of care ☑ m)Current Condition n) Teaching points	02	02	10	9 (0.3)	10 (0.0)	+ 1
You are giving handoff and need to discuss twins. What element would best identify each for the receiver? a) Full name b) Time of birth c) Sex d) Birth weight ☑ e) MRN	03	03	10	5 (0.5)	8 (0.40)	+ 3
Which three components constitute the big sentence? a) Age, room number, service ☑ b) Chronic problems, age, and diagnosis c) Admit, age, chronic problems d) Medication, diagnosis, chronic problems	05	04	10	10 (0.0)	10 (0.0)	--
What two elements should the current condition of a patient tell ☑ ***you?*** a) How the patient looks now; baseline norms b) Current medications; known allergies c) Diagnosis; treatment plan d) Lab results pending; lab results known	07	06	10	10 (0.0)	10 (0.0)	--
Totals			**50**	**43 (0.35)**	**48 (0.20)**	**+ 5**

At the $\alpha = 0.05$ level, a two-tailed test of proportions revealed that the difference in the knowledge assessment scores between pre-training (86.0%) and post-training (96.0%) was not significant ($Z = -1.75$, $p = 0.081$).

Appendix C. Heuristic evaluation of the handoff training module [79]

Method

A heuristic evaluation questionnaire was designed to focus on content and human factors issues as well as general comments. The content section consisted of six questions, each covering a separate aspect of the training: audio examples, history and physical example, "did you know" facts, case studies, conceptual organization, and training length. Three human factors questions addressed navigation, the use of pop-up windows, font size and colors, and the overall "look and feel" of the training. The general comments section requested the evaluators to provide the three most useful and least useful aspects of the training, one aspect they would change, and one aspect they would emphasize in future designs.

The strategy was to employ both medical and human factors evaluators. The three evaluators included two third year pediatric residents and one undergraduate engineering student. The two pediatric residents had previously been identified by their peers as "good sign-out givers", and thus could provide the most relevant feedback with respect to the training content. The undergraduate had successfully completed a human factors course that focuses on user-centered design. The evaluators were contacted by email for their participation in the study.

At the start of the evaluation session, a graduate research assistant described the purpose of the training system and the purpose of the evaluation. Each evaluator was instructed to begin the training using the web-based prototype and was encouraged to stop at any point to discuss important issues or to ask questions. The graduate research assistant recorded the comments separately from the questionnaire. After finishing the training, the graduate research assistant stepped though questionnaire answers with the evaluator.

For each evaluator, the graduate research assistant recorded the amount of time required to take the training. If an evaluator stopped to ask a question or make a comment during the training, the time was not included in the total training completion time.

Heuristic evaluation results

All three heuristic evaluators were able to complete the training within the required time frame of 45-60 minutes (minimum = 45 min., maximum = 60 min.).

With respect to the training content, both pediatric resident heuristic evaluators commended the use of multimedia, particularly the audio examples of sign-out. Both recommended the use of additional audio examples in future versions of the training system. However, one suggested that the use of the History and Physical was "not necessarily helpful". That is, the History and Physical document did not help her grasp the concept of a "big sentence" as it relates to sign-out. In terms of the "Did You Know Facts", one of the residents commented that they "were an aspect of the system that provided entertaining, yet informative supplementary information." In discussing the case studies, one resident found them "helpful", while the other thought they could be improved by placing the learner in one of the roles from the case study, the goal being to make the case study "more personal and engaging". In terms of the conceptual organization and length of the training, both residents suggested that the content was "well organized" and could be completed within a "reasonable amount of time".

With respect to human factors concerns, one resident commented that there was a lag time when opening pop-up windows. None of the evaluators found the navigation of the system to be a problem and all three were satisfied with the general look and feel of the interface.

All of the evaluators provided suggestions in the general comments. Both pediatric resident evaluators listed the audio examples as one of the most useful aspects of the training. Other useful aspects given by the resident evaluators were the organization of the training concepts, and "Did You Know Facts" illustrating what can go wrong in certain situations.

Only the two resident evaluators provided feedback on the least useful aspects of the training. The first suggested that use of generic sign-out statistics in the "Did You Know Facts" was "boring" while the second suggested that there were too many audio examples in the "big sentence" section.

With regard to changing one aspect of the training, the first medical evaluator suggested making the training more "personal" by including additional "pediatric related examples" or by relating the case studies to more "personal experiences" of the residents. The engineering student evaluator suggested minimizing the amount of text on certain screens by including audio-based narration. In terms of emphasizing an existing aspect of the training, both resident evaluators said they would emphasize the training content while the engineering student said he would emphasize the overall user-friendliness of the system.

Discussion of results

The heuristic evaluation identified that in some cases, realistic examples were not necessary and will likely be removed to save time in future versions. For example, some audio files in the "big sentence" section as well as the History and Physical example can be trimmed or made optional. With this content removed, it is likely that the requirement for training length could be shortened-a good idea when the target population is busy residents.

With respect to the implementation, Adobe Captivate provided an acceptable platform for developing the prototype sign-out training system. However, there were limitations. The current version of Adobe Captivate does not provide an internal method for capturing and reporting assessment results. The workaround solution included sending email to an analyst account but in the long-term, this method is not acceptable. With regard to the assessments, another limitation is the lack of an individualized feedback mechanism. In the current prototype, learners receive feedback only for incorrect answer choices and are not provided with a detailed explanation for a specific response. Adobe Captivate is not capable of creating this critical type of feedback. Captivate also does support user login and tracking. As residents are busy, a better system would allow them to start training, logout and return later to complete the training. The system should allow completion of the training in multiple sittings as necessary. To enable the reporting of results through Adobe Captivate and the log-in and tracking aspects, Captivate training files must interface with a separate learning management system (LMS) which can be a costly and time consuming solution for training designers. Without a LMS, it is likely that a more robust development environment will be necessary to achieve these advanced functions.

Appendix D. Preliminary categorization and initial analysis of off-task discussion for the baseline training data

Table 15 lists newly defined sub-categories for off-task discussions and summarizes the number of occurrences of each type for the 87 patients in which off-task discussions occurred in the baseline data collected in the training study described in Chapter 4.

Table 19. Categorization of off-task discussion with patient counts and percentages for each category

Off-Task Category	Total	Percent	Description
Logistical	44	31%	Discussion about responsibility, documentation, or general hospital logistical issues (scheduling, beds, etc.)
Socializing	30	21%	Any conversation about matters unrelated to sign-out or general physician work, joking, etc.
Teaching	29	20%	Expounding on a current medical topic such as diagnoses, meds, procedures, etc. for the purpose of teaching
Interruption	18	13%	Any conversation spurred by the need to make a phone call, a page, an incoming phone call, or in-person interruption by a person not involved in sign-out
Extraneous Patient Data	15	11%	Patient background (9): • Acknowledging that too much background information was discussed for a patient (3) • Details about a prior hospitalization (2) • Clarifying details already discussed about background information about a patient (2) • Discussing chain of custody laws for a different state in terms of the patient's background • Discussing a decision made many days prior about a patient by an attending unrelated to the current state of the patient Patient family issues (5) Details about dosing for a particular medication (1)
Previous Patient	5	4%	Any discussion that reverted to a patient previously discussed during the sign-out
N Off-Task Conversations (For N = 87 patients)	**141**		

Appendix E. External Sources Used by Residents to Solicit Information

Table 20. External Sources Consulted By Residents

information source	description of content
other human sources (physicians, nurses, and others)	phone calls face to face communications
medical chart	progress notes nursing and therapy interactions with the patient attending notes admission notes
computerized provider order entry system (CPOE)	the electronic system used to place orders and view laboratory results
electronic medical record (EMR) system	the electronic system that contains historical data such as past laboratory results, previous care, and admissions. in this case CareCast.
sign-out sheet (handoff/handover sheet)	printed list of patients and sign-out specific information elements representative of data stored in an electronic sign-out database system

Appendix F. Standardized data collection worksheet

Date:			
Time question was asked:			
Question:			
Who asked the question:			
Where did resident look for information?	Order	Begin Time	End Time/comments
Chart - Progress Note			
Chart - Consult Note			
Chart - Attending Note			
Chart - Other			
Nursing Note			
CPOE (MIS)			
EMR			
Sign-out Database			
Sign-out Sheet			
Phone - Attending			
Phone - Other Resident			
Phone - Fellow			
Made it up/Clinical Judgment			
Was an answer provided?	Yes / No / Unclear		
Time question was answered:			
COMMENTS:			

Appendix G. Variable descriptions and labels for RQ 2.2

Name	Variable type	Variable Description
q_ID	Identifier	Unique numeric identifier for each item (row) in the data
q_Date	Identifier	Date of the overnight shift on which the question asked was observed (will match recordingDate exactly)
recordingDate	Identifier	Date the corresponding handoff was audio recorded
patientCODE	Identifier	A code corresponding to the patient in a de-identified copy of the handoff transcription (based on the *recordingDate*)
respYN	Dependent/ Response variable	**Binary – [1 = yes, 0 = no]** whether or not the resident gave an actionable response to the question asked (**NOT** whether the response was "correct" or if it "answered" the question
qAsker	Independent	**Categorical text-based variable described above:** **ptFam** The patient and/or their family **intern** Resident at the intern level **medStdnt** Medical student **nonPCTdoc** Any physician not part of the primary care team for the patient about whom the question was asked **nurse** Any type of nurse **othrHspRole** Any other person in a hospital employed role **primCareTeam** Any physician (other than an intern) serving on the primary care team for a patient about whom the question was asked
qTotMinBin	Independent	**Categorical - bins for the total time to complete a question** • < 1 LessThanOneMinute • = 1 OneMinute • 2-5 TwoToFiveMin (inclusive) • 6-10 SixToTenMin (inclusive) • > 10 MoreThanTenMin
qType	Independent	**Categorical - type of question asked, based on content:** **pOfCare** immediate plan for the next 12-24 hrs for a patient, as well as long term plan beyond 24 hours through the patient's entire hospital stay (e.g. "are we giving Ativan to the patient?") **pOfCareRtn** rationale for the plan of care, specific elements of it, or reason(s) they are under consideration (e.g. "why is the patient on telemetry?") **ptCndtnCurr** recent information concerning a patient's condition - within the last 4 hours (e.g. "does the patient still look pale?") **ptCndtnPast** older information concerning a patient's condition - prior to the last 4 hours (e.g. "how did the patient look when he came in yesterday?" **medKnwlg** general medical knowledge which could be found in a medical handbook, textbook, online (e.g. "what are the most common dosages for morphine?") **technical** related to the use of specific hospital computer systems such as order entry systems or EMR (e.g. "how do you put in an order for an MRI?") **medPractice** implementation of medical knowledge, such as how to carry out medical procedures (e.g. "how do you palpate a newborn's liver?")

		localSBP SBP (system-based practice) how to complete tasks relevant to the work at hand other than technical systems (e.g. "when does the next shift arrive?")
		inProcess activities or tasks which are currently in-progress for a patient & process-checks (e.g. "have we ordered blood cultures for the patient yet?"
		request requests to complete a task or related actions (e.g. "can you put in the order for his Ativan?")
totSrc	Independent	**Categorical – [1, 2, or 3],** the total number of information sources consulted by a resident in response to a question (including their own clinical judgment)
acuity	Independent	**Categorical [subjective] – The severity of the patient as given by the value on the residents' handoff sheet.** **Three levels:** 1 – low acuity (low severity of illness) 2 – medium acuity (average severity of illness) 3 – high acuity (high severity of illness)
discussNo	Independent	**Categorical, binary** **0 – False,** the content of the question WAS discussed in some way for the patient during handoff **1 – True,** the content or general category of the question was NOT discussed at handoff for the patient (or the patient was not discussed at all)
discussOne	Independent	**Categorical, binary –** **0 – False,** the general content category of the question was not discussed at handoff for that patient **1 – True,** the general content category of the question was discussed at handoff Example: Question Type was "Plan of Care" and it was discussed for the patient at handoff, but not related to the specific content of the question (receives a 1)
discussTwo	Independent	**Categorical, binary –** **0 – False,** the specific content of the question was NOT discussed at handoff for that patient **1 – True,** the content category of the question was discussed at handoff Example: Question Type was "Plan of Care", content was about the expected discharge time for the patient. At handoff, the transcript revealed that the residents discussed what time they thought the patient would be discharged

Appendix H. Question categorization hierarchy

Category label	Description
Plan of care	Both the immediate plan of care for the next 12-24 hrs for a patient, as well as the long term plan ("big picture") beyond 24 hours and for a patient's entire hospital course.
Rationale	The rationale for the overall plan of care or specific elements of the plan of care, or the reason(s) they are under consideration.
Current condition	The most recent information on the patient's current condition (within 4 hours)
Past condition	Concerns the patient's prior condition (prior to the last 4 hours)
Medical knowledge	General medical knowledge that could be found in a medical handbook, textbook, or online source, such as the dosages for morphine.
Technical knowledge	Related to the use of specific hospital computer systems such as computerized order entry systems or the EMR
Medical practice	Medical practice is the implementation of medical knowledge, such as how to carry out medical procedures. For example, "how do you palpate a newborn's liver?"
System-based practices (SBP)	SBP questions concern how to complete tasks relevant to the work at hand other than those concerning computer or technical systems. For example, "what time does the pediatric intern begin their call shift?"
Task in-progress	Activities or tasks which are currently in-progress for a patient. Process-check questions, e.g. "have we done _____ for the patient yet?"
Request	Requests to complete a task or some type of activity or action

Appendix I. GLM NCSS training data reports

Analysis of Variance Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Response   ptID

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.004)
A: GroupAgg	1	7.44629	7.44629	34.82	0.000000*	0.998081
B: Subject	9	3.250752	0.3611946	1.69	0.102383	0.376931
AB	9	3.561529	0.3957254	1.85	0.069169	0.435324
S	94	20.1	0.2138298			
Total (Adjusted)	113	37.65789				
Total	114					

* Term significant at alpha = 0.004

Means and Standard Error Section

Term	Count	Mean	Standard Error
All	114	1.321429	
A: GroupAgg			
01_pre	57	1.614286	0.06124867
02_post	57	1.028571	0.06124867
B: Subject			
Sub01	4	1.25	0.2312087
Sub03	16	1.25	0.1156043
Sub04	10	1.5	0.1462292
Sub05	10	1.2	0.1462292
Sub06	14	1.428571	0.1235862
Sub07	20	1.3	0.1033997
Sub08	4	1	0.2312087
Sub09	14	1.142857	0.1235862
Sub10	8	1.5	0.1634892
Sub12	14	1.642857	0.1235862
AB: GroupAgg,Subject			
01_pre,Sub01	2	1.5	0.3269784
01_pre,Sub03	8	1.5	0.1634892
01_pre,Sub04	5	2	0.2067993
01_pre,Sub05	5	1.4	0.2067993
01_pre,Sub06	7	1.571429	0.1747773
01_pre,Sub07	10	1.6	0.1462292
01_pre,Sub08	2	1	0.3269784
01_pre,Sub09	7	1.285714	0.1747773
01_pre,Sub10	4	2	0.2312087
01_pre,Sub12	7	2.285714	0.1747773
02_post,Sub01	2	1	0.3269784
02_post,Sub03	8	1	0.1634892
02_post,Sub04	5	1	0.2067993
02_post,Sub05	5	1	0.2067993
02_post,Sub06	7	1.285714	0.1747773
02_post,Sub07	10	1	0.1462292
02_post,Sub08	2	1	0.3269784
02_post,Sub09	7	1	0.1747773
02_post,Sub10	4	1	0.2312087
02_post,Sub12	7	1	0.1747773

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ptID

Plots Section

Means Plot of ptID

By GroupAgg

Means Plot of ptID

By Subject

Means Plot of ptID by Subject

Means Plot of ptID by GroupAgg

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ptID

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: ptID
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.2138298 Critical Value=2.9509

Group	Count	Mean	Different From Groups
01_pre	57	1.614286	02_post
02_post	57	1.028571	01_pre

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Bonferroni (With Control) Multiple-Comparison Test

Response: ptID
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.2138298 Critical Value=2.9509

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre	57	1.614286	02_post
02_post	57	1.028571	01_pre

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: ptID
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.2138298 Critical Value=4.1585

Group	Count	Mean	Different From Groups
01_pre	57	1.614286	02_post
02_post	57	1.028571	01_pre

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ptID

Planned Comparison: A1

Response: ptID
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.2138298

Comparison Value=-0.5857143   T-Value=6.7620   Prob>|T|=0.000000   Decision(0.004)=Reject
Comparison Std Error = 0.0866187   Comparison Confidence Interval = -0.8413156 to -0.330113

Group	Comparison Coefficient	Count	Mean
01_pre	-1	57	1.614286
02_post	1	57	1.028571

Notes:
This section presents the results of a standard set of planned comparisons. The first comparison
compares the first group with those below it (alphabetically). The second comparison compares the
second group with those below it (alphabetically) ignoring the first. This continues to the next
to the last group.

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: ptID
Term B: Subject

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.2138298 Critical Value=4.0964

Group	Count	Mean	Different From Groups
Sub01	4	1.25	
Sub03	16	1.25	
Sub04	10	1.5	
Sub05	10	1.2	
Sub06	14	1.428571	
Sub07	20	1.3	
Sub08	4	1	
Sub09	14	1.142857	
Sub10	8	1.5	
Sub12	14	1.642857	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ptID

Bonferroni (With Control) Multiple-Comparison Test

Response: ptID
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.2138298 Critical Value=3.6408

If Control Group Is	Count	Mean	Different From Treatment Groups
Sub01	4	1.25	
Sub03	16	1.25	
Sub04	10	1.5	
Sub05	10	1.2	
Sub06	14	1.428571	
Sub07	20	1.3	
Sub08	4	1	
Sub09	14	1.142857	
Sub10	8	1.5	
Sub12	14	1.642857	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: ptID
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.2138298 Critical Value=5.6741

Group	Count	Mean	Different From Groups
Sub01	4	1.25	
Sub03	16	1.25	
Sub04	10	1.5	
Sub05	10	1.2	
Sub06	14	1.428571	
Sub07	20	1.3	
Sub08	4	1	
Sub09	14	1.142857	
Sub10	8	1.5	
Sub12	14	1.642857	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ptID

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: ptID
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.2138298 Critical Value=4.4794

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre,Sub01	2	1.5	
01_pre,Sub03	8	1.5	
01_pre,Sub04	5	2	
01_pre,Sub05	5	1.4	
01_pre,Sub06	7	1.571429	
01_pre,Sub07	10	1.6	
01_pre,Sub08	2	1	
01_pre,Sub09	7	1.285714	
01_pre,Sub10	4	2	
01_pre,Sub12	7	2.285714	(02_post,Sub03), (02_post,Sub04) (02_post,Sub05), (02_post,Sub07) (02_post,Sub09), (02_post,Sub12)
02_post,Sub01	2	1	
02_post,Sub03	8	1	(01_pre,Sub12)
02_post,Sub04	5	1	(01_pre,Sub12)
02_post,Sub05	5	1	(01_pre,Sub12)
02_post,Sub06	7	1.285714	
02_post,Sub07	10	1	(01_pre,Sub12)
02_post,Sub08	2	1	
02_post,Sub09	7	1	(01_pre,Sub12)
02_post,Sub10	4	1	
02_post,Sub12	7	1	(01_pre,Sub12)

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ptID

Bonferroni (With Control) Multiple-Comparison Test

Response: ptID
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.2138298 Critical Value=3.8565

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre,Sub01	2	1.5	
01_pre,Sub03	8	1.5	
01_pre,Sub04	5	2	(02_post,Sub07)
01_pre,Sub05	5	1.4	
01_pre,Sub06	7	1.571429	
01_pre,Sub07	10	1.6	
01_pre,Sub08	2	1	
01_pre,Sub09	7	1.285714	(01_pre,Sub12)
01_pre,Sub10	4	2	
01_pre,Sub12	7	2.285714	(01_pre,Sub09), (02_post,Sub03) (02_post,Sub04), (02_post,Sub05) (02_post,Sub06), (02_post,Sub07) (02_post,Sub09), (02_post,Sub10) (02_post,Sub12)
02_post,Sub01	2	1	
02_post,Sub03	8	1	(01_pre,Sub12)
02_post,Sub04	5	1	(01_pre,Sub12)
02_post,Sub05	5	1	(01_pre,Sub12)
02_post,Sub06	7	1.285714	(01_pre,Sub12)
02_post,Sub07	10	1	(01_pre,Sub04), (01_pre,Sub12)
02_post,Sub08	2	1	
02_post,Sub09	7	1	(01_pre,Sub12)
02_post,Sub10	4	1	(01_pre,Sub12)
02_post,Sub12	7	1	(01_pre,Sub12)

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Analysis of Variance Report
```
Dataset            C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response           ptID
```

Tukey-Kramer Multiple-Comparison Test

```
Response: ptID
Term AB: GroupAgg,Subject

Alpha=0.004  Error Term=S(AB)  DF=94  MSE=0.2138298 Critical Value=6.1848
```

			Different From
Group	Count	Mean	Groups
01_pre,Sub01	2	1.5	
01_pre,Sub03	8	1.5	
01_pre,Sub04	5	2	
01_pre,Sub05	5	1.4	
01_pre,Sub06	7	1.571429	
01_pre,Sub07	10	1.6	
01_pre,Sub08	2	1	
01_pre,Sub09	7	1.285714	
01_pre,Sub10	4	2	
01_pre,Sub12	7	2.285714	(02_post,Sub03), (02_post,Sub04) (02_post,Sub05), (02_post,Sub07) (02_post,Sub09), (02_post,Sub10) (02_post,Sub12)
02_post,Sub01	2	1	
02_post,Sub03	8	1	(01_pre,Sub12)
02_post,Sub04	5	1	(01_pre,Sub12)
02_post,Sub05	5	1	(01_pre,Sub12)
02_post,Sub06	7	1.285714	
02_post,Sub07	10	1	(01_pre,Sub12)
02_post,Sub08	2	1	
02_post,Sub09	7	1	(01_pre,Sub12)
02_post,Sub10	4	1	(01_pre,Sub12)
02_post,Sub12	7	1	(01_pre,Sub12)

```
Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.
```

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Response BigS

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.004)
A: GroupAgg	1	8.07659	8.07659	9.42	0.002804*	0.549191
B: Subject	9	28.52218	3.169131	3.70	0.000532*	0.898561
AB	9	30.15526	3.350585	3.91	0.000302*	0.921444
S	94	80.58572	0.8572949			
Total (Adjusted)	113	145.6579				
Total	114					

* Term significant at alpha = 0.004

Means and Standard Error Section

Term	Count	Mean	Standard Error
All	114	1.363571	
A: GroupAgg			
01_pre	57	1.668571	0.1226387
02_post	57	1.058571	0.1226387
B: Subject			
Sub01	4	0.75	0.4629511
Sub03	16	0.75	0.2314755
Sub04	10	1.6	0.292796
Sub05	10	1.1	0.292796
Sub06	14	1.071429	0.2474578
Sub07	20	1.4	0.207038
Sub08	4	2.5	0.4629511
Sub09	14	1.5	0.2474578
Sub10	8	0.75	0.3273559
Sub12	14	2.214286	0.2474578
AB: GroupAgg,Subject			
01_pre,Sub01	2	1	0.6547117
01_pre,Sub03	8	0.5	0.3273559
01_pre,Sub04	5	2	0.414076
01_pre,Sub05	5	0.8	0.414076
01_pre,Sub06	7	1	0.3499581
01_pre,Sub07	10	1.6	0.292796
01_pre,Sub08	2	4	0.6547117
01_pre,Sub09	7	1.857143	0.3499581
01_pre,Sub10	4	0.5	0.4629511
01_pre,Sub12	7	3.428571	0.3499581
02_post,Sub01	2	0.5	0.6547117
02_post,Sub03	8	1	0.3273559
02_post,Sub04	5	1.2	0.414076
02_post,Sub05	5	1.4	0.414076
02_post,Sub06	7	1.142857	0.3499581
02_post,Sub07	10	1.2	0.292796
02_post,Sub08	2	1	0.6547117
02_post,Sub09	7	1.142857	0.3499581
02_post,Sub10	4	1	0.4629511
02_post,Sub12	7	1	0.3499581

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   BigS

Plots Section

Means Plot of BigS

Means Plot of BigS

Means Plot of BigS by Subject

Means Plot of BigS by GroupAgg

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   BigS

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: BigS
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.8572949 Critical Value=2.9509

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre	57	1.668571	02_post
02_post	57	1.058571	01_pre

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Bonferroni (With Control) Multiple-Comparison Test

Response: BigS
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.8572949 Critical Value=2.9509

If Control			**Different From**
Group Is	**Count**	**Mean**	**Treatment Groups**
01_pre	57	1.668571	02_post
02_post	57	1.058571	01_pre

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: BigS
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.8572949 Critical Value=4.1585

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre	57	1.668571	02_post
02_post	57	1.058571	01_pre

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset            C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response           BigS
```

Planned Comparison: A1

```
Response: BigS
Term A: GroupAgg
```

```
Alpha=0.004  Error Term=S(AB)  DF=94  MSE=0.8572949
```

```
Comparison Value=-0.61   T-Value=3.5171   Prob>|T|=0.000674   Decision(0.004)=Reject
Comparison Std Error = 0.1734374   Comparison Confidence Interval = -1.121793 to -0.09820729
```

Group	Comparison Coefficient	Count	Mean
01_pre	-1	57	1.668571
02_post	1	57	1.058571

Notes:
This section presents the results of a standard set of planned comparisons. The first comparison compares the first group with those below it (alphabetically). The second comparison compares the second group with those below it (alphabetically) ignoring the first. This continues to the next to the last group.

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: BigS
Term B: Subject
```

```
Alpha=0.004  Error Term=S(AB)  DF=94  MSE=0.8572949 Critical Value=4.0964
```

Group	Count	Mean	Different From Groups
Sub01	4	0.75	
Sub03	16	0.75	Sub12
Sub04	10	1.6	
Sub05	10	1.1	
Sub06	14	1.071429	
Sub07	20	1.4	
Sub08	4	2.5	
Sub09	14	1.5	
Sub10	8	0.75	
Sub12	14	2.214286	Sub03

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   BigS

Bonferroni (With Control) Multiple-Comparison Test

Response: BigS
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.8572949 Critical Value=3.6408

If Control Group Is	Count	Mean	Different From Treatment Groups
Sub01	4	0.75	
Sub03	16	0.75	Sub12
Sub04	10	1.6	
Sub05	10	1.1	
Sub06	14	1.071429	
Sub07	20	1.4	
Sub08	4	2.5	
Sub09	14	1.5	
Sub10	8	0.75	
Sub12	14	2.214286	Sub03

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: BigS
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.8572949 Critical Value=5.6741

Group	Count	Mean	Different From Groups
Sub01	4	0.75	
Sub03	16	0.75	Sub12
Sub04	10	1.6	
Sub05	10	1.1	
Sub06	14	1.071429	
Sub07	20	1.4	
Sub08	4	2.5	
Sub09	14	1.5	
Sub10	8	0.75	
Sub12	14	2.214286	Sub03

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset            C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response           BigS
```

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: BigS
Term AB: GroupAgg,Subject

Alpha=0.004  Error Term=S(AB)  DF=94  MSE=0.8572949 Critical Value=4.4794
```

Group	Count	Mean	Different From Groups
01_pre,Sub01	2	1	
01_pre,Sub03	8	0.5	(01_pre,Sub08), (01_pre,Sub12)
01_pre,Sub04	5	2	
01_pre,Sub05	5	0.8	(01_pre,Sub12)
01_pre,Sub06	7	1	(01_pre,Sub12)
01_pre,Sub07	10	1.6	
01_pre,Sub08	2	4	(01_pre,Sub03)
01_pre,Sub09	7	1.857143	
01_pre,Sub10	4	0.5	(01_pre,Sub12)
01_pre,Sub12	7	3.428571	(01_pre,Sub03), (01_pre,Sub05) (01_pre,Sub06), (01_pre,Sub10) (02_post,Sub03), (02_post,Sub06) (02_post,Sub07), (02_post,Sub09) (02_post,Sub12)
02_post,Sub01	2	0.5	
02_post,Sub03	8	1	(01_pre,Sub12)
02_post,Sub04	5	1.2	
02_post,Sub05	5	1.4	
02_post,Sub06	7	1.142857	(01_pre,Sub12)
02_post,Sub07	10	1.2	(01_pre,Sub12)
02_post,Sub08	2	1	
02_post,Sub09	7	1.142857	(01_pre,Sub12)
02_post,Sub10	4	1	
02_post,Sub12	7	1	(01_pre,Sub12)

```
Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.
```

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   BigS

Bonferroni (With Control) Multiple-Comparison Test

Response: BigS
Term AB: GroupAgg,Subject

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.8572949 Critical Value=3.8565

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre,Sub01	2	1	
01_pre,Sub03	8	0.5	(01_pre,Sub08), (01_pre,Sub12)
01_pre,Sub04	5	2	
01_pre,Sub05	5	0.8	(01_pre,Sub08), (01_pre,Sub12)
01_pre,Sub06	7	1	(01_pre,Sub08), (01_pre,Sub12)
01_pre,Sub07	10	1.6	(01_pre,Sub12)
01_pre,Sub08	2	4	(01_pre,Sub03), (01_pre,Sub05)
			(01_pre,Sub06), (01_pre,Sub10)
			(02_post,Sub03), (02_post,Sub07)
			(02_post,Sub12)
01_pre,Sub09	7	1.857143	
01_pre,Sub10	4	0.5	(01_pre,Sub08), (01_pre,Sub12)
01_pre,Sub12	7	3.428571	(01_pre,Sub03), (01_pre,Sub05)
			(01_pre,Sub06), (01_pre,Sub07)
			(01_pre,Sub10), (02_post,Sub01)
			(02_post,Sub03), (02_post,Sub04)
			(02_post,Sub06), (02_post,Sub07)
			(02_post,Sub09), (02_post,Sub10)
			(02_post,Sub12)
02_post,Sub01	2	0.5	(01_pre,Sub12)
02_post,Sub03	8	1	(01_pre,Sub08), (01_pre,Sub12)
02_post,Sub04	5	1.2	(01_pre,Sub12)
02_post,Sub05	5	1.4	
02_post,Sub06	7	1.142857	(01_pre,Sub12)
02_post,Sub07	10	1.2	(01_pre,Sub08), (01_pre,Sub12)
02_post,Sub08	2	1	
02_post,Sub09	7	1.142857	(01_pre,Sub12)
02_post,Sub10	4	1	(01_pre,Sub12)
02_post,Sub12	7	1	(01_pre,Sub08), (01_pre,Sub12)

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   BigS

Tukey-Kramer Multiple-Comparison Test

Response: BigS
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.8572949 Critical Value=6.1848

Group	Count	Mean	Different From Groups
01_pre,Sub01	2	1	
01_pre,Sub03	8	0.5	(01_pre,Sub08), (01_pre,Sub12)
01_pre,Sub04	5	2	
01_pre,Sub05	5	0.8	(01_pre,Sub12)
01_pre,Sub06	7	1	(01_pre,Sub12)
01_pre,Sub07	10	1.6	
01_pre,Sub08	2	4	(01_pre,Sub03)
01_pre,Sub09	7	1.857143	
01_pre,Sub10	4	0.5	(01_pre,Sub12)
01_pre,Sub12	7	3.428571	(01_pre,Sub03), (01_pre,Sub05)
			(01_pre,Sub06), (01_pre,Sub10)
			(02_post,Sub03), (02_post,Sub06)
			(02_post,Sub07), (02_post,Sub09)
			(02_post,Sub12)
02_post,Sub01	2	0.5	
02_post,Sub03	8	1	(01_pre,Sub12)
02_post,Sub04	5	1.2	
02_post,Sub05	5	1.4	
02_post,Sub06	7	1.142857	(01_pre,Sub12)
02_post,Sub07	10	1.2	(01_pre,Sub12)
02_post,Sub08	2	1	
02_post,Sub09	7	1.142857	(01_pre,Sub12)
02_post,Sub10	4	1	
02_post,Sub12	7	1	(01_pre,Sub12)

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Response   CurrCond

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.004)
A: GroupAgg	1	12.31418	12.31418	1.91	0.170361	0.063429
B: Subject	9	107.0674	11.89638	1.84	0.070290	0.432969
AB	9	118.1925	13.13249	2.04	0.043552	0.501295
S	94	606.3893	6.45095			
Total (Adjusted)	113	838.5263				
Total	114					

* Term significant at alpha = 0.004

Means and Standard Error Section

Term	Count	Mean	Standard Error
All	114	3.293036	
A: GroupAgg			
01_pre	57	3.669643	0.3364143
02_post	57	2.916429	0.3364143
B: Subject			
Sub01	4	3	1.269936
Sub03	16	2.4375	0.634968
Sub04	10	2.6	0.8031781
Sub05	10	2.2	0.8031781
Sub06	14	2.785714	0.6788093
Sub07	20	5.05	0.5679327
Sub08	4	4.5	1.269936
Sub09	14	4	0.6788093
Sub10	8	3	0.8979804
Sub12	14	3.357143	0.6788093
AB: GroupAgg,Subject			
01_pre,Sub01	2	4.5	1.795961
01_pre,Sub03	8	2.125	0.8979804
01_pre,Sub04	5	2	1.135865
01_pre,Sub05	5	1.6	1.135865
01_pre,Sub06	7	3.142857	0.9599814
01_pre,Sub07	10	5.4	0.8031781
01_pre,Sub08	2	7	1.795961
01_pre,Sub09	7	4	0.9599814
01_pre,Sub10	4	1.5	1.269936
01_pre,Sub12	7	5.428571	0.9599814
02_post,Sub01	2	1.5	1.795961
02_post,Sub03	8	2.75	0.8979804
02_post,Sub04	5	3.2	1.135865
02_post,Sub05	5	2.8	1.135865
02_post,Sub06	7	2.428571	0.9599814
02_post,Sub07	10	4.7	0.8031781
02_post,Sub08	2	2	1.795961
02_post,Sub09	7	4	0.9599814
02_post,Sub10	4	4.5	1.269936
02_post,Sub12	7	1.285714	0.9599814

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   CurrCond

Plots Section

Means Plot of CurrCond

By GroupAgg

Means Plot of CurrCond

By Subject

Means Plot of CurrCond by Subject

Means Plot of CurrCond by GroupAgg

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   CurrCond

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: CurrCond
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=2.9509

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre	57	3.669643	
02_post	57	2.916429	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Bonferroni (With Control) Multiple-Comparison Test

Response: CurrCond
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=2.9509

If Control			**Different From**
Group Is	**Count**	**Mean**	**Treatment Groups**
01_pre	57	3.669643	
02_post	57	2.916429	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: CurrCond
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=4.1585

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre	57	3.669643	
02_post	57	2.916429	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset            C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response           CurrCond
```

Planned Comparison: A1

```
Response: CurrCond
Term A: GroupAgg
```

`Alpha=0.004  Error Term=S(AB)  DF=94  MSE=6.45095`

```
Comparison Value=-0.7532143   T-Value=1.5832   Prob>|T|=0.116740   Decision(0.004)=Do Not Reject
Comparison Std Error = 0.4757616   Comparison Confidence Interval = -2.157129 to 0.6507007
```

	Comparison Coefficient	Count	Mean
Group			
01_pre	-1	57	3.669643
02_post	1	57	2.916429

Notes:
This section presents the results of a standard set of planned comparisons. The first comparison compares the first group with those below it (alphabetically). The second comparison compares the second group with those below it (alphabetically) ignoring the first. This continues to the next to the last group.

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: CurrCond
Term B: Subject
```

`Alpha=0.004  Error Term=S(AB)  DF=94  MSE=6.45095 Critical Value=4.0964`

Group	Count	Mean	Different From Groups
Sub01	4	3	
Sub03	16	2.4375	
Sub04	10	2.6	
Sub05	10	2.2	
Sub06	14	2.785714	
Sub07	20	5.05	
Sub08	4	4.5	
Sub09	14	4	
Sub10	8	3	
Sub12	14	3.357143	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   CurrCond

Bonferroni (With Control) Multiple-Comparison Test

Response: CurrCond
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=3.6408

If Control Group Is	Count	Mean	Different From Treatment Groups
Sub01	4	3	
Sub03	16	2.4375	
Sub04	10	2.6	
Sub05	10	2.2	
Sub06	14	2.785714	
Sub07	20	5.05	
Sub08	4	4.5	
Sub09	14	4	
Sub10	8	3	
Sub12	14	3.357143	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: CurrCond
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=5.6741

Group	Count	Mean	Different From Groups
Sub01	4	3	
Sub03	16	2.4375	
Sub04	10	2.6	
Sub05	10	2.2	
Sub06	14	2.785714	
Sub07	20	5.05	
Sub08	4	4.5	
Sub09	14	4	
Sub10	8	3	
Sub12	14	3.357143	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset            C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response           CurrCond
```

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: CurrCond
Term AB: GroupAgg,Subject

Alpha=0.004  Error Term=S(AB)  DF=94  MSE=6.45095 Critical Value=4.4794
```

Group	Count	Mean	Different From Groups
01_pre,Sub01	2	4.5	
01_pre,Sub03	8	2.125	
01_pre,Sub04	5	2	
01_pre,Sub05	5	1.6	
01_pre,Sub06	7	3.142857	
01_pre,Sub07	10	5.4	
01_pre,Sub08	2	7	
01_pre,Sub09	7	4	
01_pre,Sub10	4	1.5	
01_pre,Sub12	7	5.428571	
02_post,Sub01	2	1.5	
02_post,Sub03	8	2.75	
02_post,Sub04	5	3.2	
02_post,Sub05	5	2.8	
02_post,Sub06	7	2.428571	
02_post,Sub07	10	4.7	
02_post,Sub08	2	2	
02_post,Sub09	7	4	
02_post,Sub10	4	4.5	
02_post,Sub12	7	1.285714	

```
Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.
```

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   CurrCond

Bonferroni (With Control) Multiple-Comparison Test

Response: CurrCond
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=3.8565

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre,Sub01	2	4.5	
01_pre,Sub03	8	2.125	
01_pre,Sub04	5	2	
01_pre,Sub05	5	1.6	
01_pre,Sub06	7	3.142857	
01_pre,Sub07	10	5.4	
01_pre,Sub08	2	7	
01_pre,Sub09	7	4	
01_pre,Sub10	4	1.5	
01_pre,Sub12	7	5.428571	
02_post,Sub01	2	1.5	
02_post,Sub03	8	2.75	
02_post,Sub04	5	3.2	
02_post,Sub05	5	2.8	
02_post,Sub06	7	2.428571	
02_post,Sub07	10	4.7	
02_post,Sub08	2	2	
02_post,Sub09	7	4	
02_post,Sub10	4	4.5	
02_post,Sub12	7	1.285714	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   CurrCond

Tukey-Kramer Multiple-Comparison Test

Response: CurrCond
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=6.45095 Critical Value=6.1848

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre,Sub01	2	4.5	
01_pre,Sub03	8	2.125	
01_pre,Sub04	5	2	
01_pre,Sub05	5	1.6	
01_pre,Sub06	7	3.142857	
01_pre,Sub07	10	5.4	
01_pre,Sub08	2	7	
01_pre,Sub09	7	4	
01_pre,Sub10	4	1.5	
01_pre,Sub12	7	5.428571	
02_post,Sub01	2	1.5	
02_post,Sub03	8	2.75	
02_post,Sub04	5	3.2	
02_post,Sub05	5	2.8	
02_post,Sub06	7	2.428571	
02_post,Sub07	10	4.7	
02_post,Sub08	2	2	
02_post,Sub09	7	4	
02_post,Sub10	4	4.5	
02_post,Sub12	7	1.285714	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Response   PofC

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.004)
A: GroupAgg	1	15.43278	15.43278	1.03	0.312900	0.029724
B: Subject	9	242.7341	26.97046	1.80	0.078512	0.416675
AB	9	224.9532	24.9948	1.67	0.107879	0.369009
S	94	1409.225	14.99176			
Total (Adjusted)	113	1916.289				
Total	114					

* Term significant at alpha = 0.004

Means and Standard Error Section

Term	Count	Mean	Standard Error
All	114	4.819464	
A: GroupAgg			
01_pre	57	5.241071	0.5128482
02_post	57	4.397857	0.5128482
B: Subject			
Sub01	4	3	1.935959
Sub03	16	3.5625	0.9679797
Sub04	10	3.6	1.224408
Sub05	10	3.5	1.224408
Sub06	14	4.428571	1.034814
Sub07	20	7.55	0.8657874
Sub08	4	5.75	1.935959
Sub09	14	5.714286	1.034814
Sub10	8	5.375	1.36893
Sub12	14	5.714286	1.034814
AB: GroupAgg,Subject			
01_pre,Sub01	2	3.5	2.73786
01_pre,Sub03	8	2.625	1.36893
01_pre,Sub04	5	2.8	1.731575
01_pre,Sub05	5	1.8	1.731575
01_pre,Sub06	7	5.857143	1.463448
01_pre,Sub07	10	8.9	1.224408
01_pre,Sub08	2	6	2.73786
01_pre,Sub09	7	6.571429	1.463448
01_pre,Sub10	4	5.5	1.935959
01_pre,Sub12	7	8.857142	1.463448
02_post,Sub01	2	2.5	2.73786
02_post,Sub03	8	4.5	1.36893
02_post,Sub04	5	4.4	1.731575
02_post,Sub05	5	5.2	1.731575
02_post,Sub06	7	3	1.463448
02_post,Sub07	10	6.2	1.224408
02_post,Sub08	2	5.5	2.73786
02_post,Sub09	7	4.857143	1.463448
02_post,Sub10	4	5.25	1.935959
02_post,Sub12	7	2.571429	1.463448

Analysis of Variance Report

Dataset	C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response	PofC

Plots Section

Means Plot of PofC

By GroupAgg

Means Plot of PofC

By Subject

Means Plot of PofC by Subject

Means Plot of PofC by GroupAgg

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   PofC

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: PofC
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=14.99176 Critical Value=2.9509

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre	57	5.241071	
02_post	57	4.397857	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Bonferroni (With Control) Multiple-Comparison Test

Response: PofC
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=14.99176 Critical Value=2.9509

If Control			**Different From**
Group Is	**Count**	**Mean**	**Treatment Groups**
01_pre	57	5.241071	
02_post	57	4.397857	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: PofC
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=14.99176 Critical Value=4.1585

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre	57	5.241071	
02_post	57	4.397857	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   PofC

Planned Comparison: A1

Response: PofC
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176

Comparison Value=-0.8432143   T-Value=1.1626   Prob>|T|=0.247931   Decision(0.004)=Do Not Reject
Comparison Std Error = 0.7252768   Comparison Confidence Interval = -2.983419 to 1.29699

Group	Comparison Coefficient	Count	Mean
01_pre	-1	57	5.241071
02_post	1	57	4.397857

Notes:
This section presents the results of a standard set of planned comparisons. The first comparison
compares the first group with those below it (alphabetically). The second comparison compares the
second group with those below it (alphabetically) ignoring the first. This continues to the next
to the last group.

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: PofC
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176 Critical Value=4.0964

Group	Count	Mean	Different From Groups
Sub01	4	3	
Sub03	16	3.5625	
Sub04	10	3.6	
Sub05	10	3.5	
Sub06	14	4.428571	
Sub07	20	7.55	
Sub08	4	5.75	
Sub09	14	5.714286	
Sub10	8	5.375	
Sub12	14	5.714286	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   PofC

Bonferroni (With Control) Multiple-Comparison Test

Response: PofC
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176 Critical Value=3.6408

If Control Group Is	Count	Mean	Different From Treatment Groups
Sub01	4	3	
Sub03	16	3.5625	
Sub04	10	3.6	
Sub05	10	3.5	
Sub06	14	4.428571	
Sub07	20	7.55	
Sub08	4	5.75	
Sub09	14	5.714286	
Sub10	8	5.375	
Sub12	14	5.714286	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: PofC
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176 Critical Value=5.6741

Group	Count	Mean	Different From Groups
Sub01	4	3	
Sub03	16	3.5625	
Sub04	10	3.6	
Sub05	10	3.5	
Sub06	14	4.428571	
Sub07	20	7.55	
Sub08	4	5.75	
Sub09	14	5.714286	
Sub10	8	5.375	
Sub12	14	5.714286	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   PofC

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: PofC
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176 Critical Value=4.4794

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre,Sub01	2	3.5	
01_pre,Sub03	8	2.625	
01_pre,Sub04	5	2.8	
01_pre,Sub05	5	1.8	
01_pre,Sub06	7	5.857143	
01_pre,Sub07	10	8.9	
01_pre,Sub08	2	6	
01_pre,Sub09	7	6.571429	
01_pre,Sub10	4	5.5	
01_pre,Sub12	7	8.857142	
02_post,Sub01	2	2.5	
02_post,Sub03	8	4.5	
02_post,Sub04	5	4.4	
02_post,Sub05	5	5.2	
02_post,Sub06	7	3	
02_post,Sub07	10	6.2	
02_post,Sub08	2	5.5	
02_post,Sub09	7	4.857143	
02_post,Sub10	4	5.25	
02_post,Sub12	7	2.571429	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   PofC

Bonferroni (With Control) Multiple-Comparison Test

Response: PofC
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176 Critical Value=3.8565

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre,Sub01	2	3.5	
01_pre,Sub03	8	2.625	
01_pre,Sub04	5	2.8	
01_pre,Sub05	5	1.8	
01_pre,Sub06	7	5.857143	
01_pre,Sub07	10	8.9	
01_pre,Sub08	2	6	
01_pre,Sub09	7	6.571429	
01_pre,Sub10	4	5.5	
01_pre,Sub12	7	8.857142	
02_post,Sub01	2	2.5	
02_post,Sub03	8	4.5	
02_post,Sub04	5	4.4	
02_post,Sub05	5	5.2	
02_post,Sub06	7	3	
02_post,Sub07	10	6.2	
02_post,Sub08	2	5.5	
02_post,Sub09	7	4.857143	
02_post,Sub10	4	5.25	
02_post,Sub12	7	2.571429	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   PofC

Tukey-Kramer Multiple-Comparison Test

Response: PofC
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=14.99176 Critical Value=6.1848

Group	Count	Mean	Different From Groups
01_pre,Sub01	2	3.5	
01_pre,Sub03	8	2.625	
01_pre,Sub04	5	2.8	
01_pre,Sub05	5	1.8	
01_pre,Sub06	7	5.857143	
01_pre,Sub07	10	8.9	
01_pre,Sub08	2	6	
01_pre,Sub09	7	6.571429	
01_pre,Sub10	4	5.5	
01_pre,Sub12	7	8.857142	
02_post,Sub01	2	2.5	
02_post,Sub03	8	4.5	
02_post,Sub04	5	4.4	
02_post,Sub05	5	5.2	
02_post,Sub06	7	3	
02_post,Sub07	10	6.2	
02_post,Sub08	2	5.5	
02_post,Sub09	7	4.857143	
02_post,Sub10	4	5.25	
02_post,Sub12	7	2.571429	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Response   ContPlan

Analysis of Variance Table

Source		Sum of	Mean		Prob	Power
Term	DF	Squares	Square	F-Ratio	Level	(Alpha=0.004)
A: GroupAgg	1	3.235219	3.235219	0.79	0.376922	0.022320
B: Subject	9	27.66776	3.074196	0.75	0.663293	0.088789
AB	9	28.45072	3.161191	0.77	0.644124	0.093305
S	94	385.8464	4.104749			
Total (Adjusted)	113	446.6053				
Total	114					

* Term significant at alpha = 0.004

Means and Standard Error Section

			Standard
Term	Count	Mean	Error
All	114	1.650179	
A: GroupAgg			
01_pre	57	1.457143	0.2683527
02_post	57	1.843214	0.2683527
B: Subject			
Sub01	4	0.25	1.013009
Sub03	16	1.6875	0.5065045
Sub04	10	1.7	0.6406832
Sub05	10	1.8	0.6406832
Sub06	14	1.571429	0.5414761
Sub07	20	1.6	0.4530314
Sub08	4	1.75	1.013009
Sub09	14	2.857143	0.5414761
Sub10	8	1.5	0.7163056
Sub12	14	1.785714	0.5414761
AB: GroupAgg,Subject			
01_pre,Sub01	2	0.5	1.432611
01_pre,Sub03	8	1	0.7163056
01_pre,Sub04	5	1.2	0.9060628
01_pre,Sub05	5	1.6	0.9060628
01_pre,Sub06	7	1.857143	0.7657629
01_pre,Sub07	10	1.2	0.6406832
01_pre,Sub08	2	1	1.432611
01_pre,Sub09	7	2.142857	0.7657629
01_pre,Sub10	4	1.5	1.013009
01_pre,Sub12	7	2.571429	0.7657629
02_post,Sub01	2	1.076916E-14	1.432611
02_post,Sub03	8	2.375	0.7163056
02_post,Sub04	5	2.2	0.9060628
02_post,Sub05	5	2	0.9060628
02_post,Sub06	7	1.285714	0.7657629
02_post,Sub07	10	2	0.6406832
02_post,Sub08	2	2.5	1.432611
02_post,Sub09	7	3.571429	0.7657629
02_post,Sub10	4	1.5	1.013009
02_post,Sub12	7	1	0.7657629

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ContPlan

Plots Section

Means Plot of ContPlan

By GroupAgg

Means Plot of ContPlan

By Subject

Means Plot of ContPlan by Subject

Means Plot of ContPlan by GroupAgg

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ContPlan

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: ContPlan
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=2.9509

Group	Count	Mean	Different From Groups
01_pre	57	1.457143	
02_post	57	1.843214	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Bonferroni (With Control) Multiple-Comparison Test

Response: ContPlan
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=2.9509

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre	57	1.457143	
02_post	57	1.843214	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: ContPlan
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=4.1585

Group	Count	Mean	Different From Groups
01_pre	57	1.457143	
02_post	57	1.843214	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset          C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response         ContPlan
```

Planned Comparison: A1

```
Response: ContPlan
Term A: GroupAgg

Alpha=0.004  Error Term=S(AB)  DF=94  MSE=4.104749

Comparison Value=0.3860714   T-Value=1.0173   Prob>|T|=0.311624   Decision(0.004)=Do Not Reject
Comparison Std Error = 0.379508   Comparison Confidence Interval = -0.7338107 to 1.505954
```

	Comparison		
Group	Coefficient	Count	Mean
01_pre	-1	57	1.457143
02_post	1	57	1.843214

Notes:
This section presents the results of a standard set of planned comparisons. The first comparison compares the first group with those below it (alphabetically). The second comparison compares the second group with those below it (alphabetically) ignoring the first. This continues to the next to the last group.

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: ContPlan
Term B: Subject

Alpha=0.004  Error Term=S(AB)  DF=94  MSE=4.104749 Critical Value=4.0964
```

			Different From
Group	Count	Mean	Groups
Sub01	4	0.25	
Sub03	16	1.6875	
Sub04	10	1.7	
Sub05	10	1.8	
Sub06	14	1.571429	
Sub07	20	1.6	
Sub08	4	1.75	
Sub09	14	2.857143	
Sub10	8	1.5	
Sub12	14	1.785714	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ContPlan

Bonferroni (With Control) Multiple-Comparison Test

Response: ContPlan
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=3.6408

If Control Group Is	Count	Mean	Different From Treatment Groups
Sub01	4	0.25	
Sub03	16	1.6875	
Sub04	10	1.7	
Sub05	10	1.8	
Sub06	14	1.571429	
Sub07	20	1.6	
Sub08	4	1.75	
Sub09	14	2.857143	
Sub10	8	1.5	
Sub12	14	1.785714	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: ContPlan
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=5.6741

Group	Count	Mean	Different From Groups
Sub01	4	0.25	
Sub03	16	1.6875	
Sub04	10	1.7	
Sub05	10	1.8	
Sub06	14	1.571429	
Sub07	20	1.6	
Sub08	4	1.75	
Sub09	14	2.857143	
Sub10	8	1.5	
Sub12	14	1.785714	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset          C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response         ContPlan
```

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: ContPlan
Term AB: GroupAgg,Subject
```

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=4.4794

Group	Count	Mean	Different From Groups
01_pre,Sub01	2	0.5	
01_pre,Sub03	8	1	
01_pre,Sub04	5	1.2	
01_pre,Sub05	5	1.6	
01_pre,Sub06	7	1.857143	
01_pre,Sub07	10	1.2	
01_pre,Sub08	2	1	
01_pre,Sub09	7	2.142857	
01_pre,Sub10	4	1.5	
01_pre,Sub12	7	2.571429	
02_post,Sub01	2	1.076916E-14	
02_post,Sub03	8	2.375	
02_post,Sub04	5	2.2	
02_post,Sub05	5	2	
02_post,Sub06	7	1.285714	
02_post,Sub07	10	2	
02_post,Sub08	2	2.5	
02_post,Sub09	7	3.571429	
02_post,Sub10	4	1.5	
02_post,Sub12	7	1	

```
Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.
```

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ContPlan

Bonferroni (With Control) Multiple-Comparison Test

Response: ContPlan
Term AB: GroupAgg,Subject

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=4.104749 Critical Value=3.8565

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre,Sub01	2	0.5	
01_pre,Sub03	8	1	
01_pre,Sub04	5	1.2	
01_pre,Sub05	5	1.6	
01_pre,Sub06	7	1.857143	
01_pre,Sub07	10	1.2	
01_pre,Sub08	2	1	
01_pre,Sub09	7	2.142857	
01_pre,Sub10	4	1.5	
01_pre,Sub12	7	2.571429	
02_post,Sub01	2	1.076916E-14	
02_post,Sub03	8	2.375	
02_post,Sub04	5	2.2	
02_post,Sub05	5	2	
02_post,Sub06	7	1.285714	
02_post,Sub07	10	2	
02_post,Sub08	2	2.5	
02_post,Sub09	7	3.571429	
02_post,Sub10	4	1.5	
02_post,Sub12	7	1	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   ContPlan

Tukey-Kramer Multiple-Comparison Test

Response: ContPlan
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=4.104749 Critical Value=6.1848

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre,Sub01	2	0.5	
01_pre,Sub03	8	1	
01_pre,Sub04	5	1.2	
01_pre,Sub05	5	1.6	
01_pre,Sub06	7	1.857143	
01_pre,Sub07	10	1.2	
01_pre,Sub08	2	1	
01_pre,Sub09	7	2.142857	
01_pre,Sub10	4	1.5	
01_pre,Sub12	7	2.571429	
02_post,Sub01	2	1.076916E-14	
02_post,Sub03	8	2.375	
02_post,Sub04	5	2.2	
02_post,Sub05	5	2	
02_post,Sub06	7	1.285714	
02_post,Sub07	10	2	
02_post,Sub08	2	2.5	
02_post,Sub09	7	3.571429	
02_post,Sub10	4	1.5	
02_post,Sub12	7	1	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Response   OT

Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.004)
A: GroupAgg	1	4.174928	4.174928	1.38	0.242362	0.042083
B: Subject	9	14.1109	1.567878	0.52	0.856901	0.047787
AB	9	23.17469	2.574965	0.85	0.569133	0.112273
S	94	283.5214	3.016186			
Total (Adjusted)	113	326.2895				
Total	114					

* Term significant at alpha = 0.004

Means and Standard Error Section

Term	Count	Mean	Standard Error
All	114	0.9910714	
A: GroupAgg			
01_pre	57	1.210357	0.2300338
02_post	57	0.7717857	0.2300338
B: Subject			
Sub01	4	0.75	0.8683584
Sub03	16	1.125	0.4341792
Sub04	10	0.9	0.5491981
Sub05	10	0.4	0.5491981
Sub06	14	0.7142857	0.4641571
Sub07	20	1.45	0.3883417
Sub08	4	0.5	0.8683584
Sub09	14	1.357143	0.4641571
Sub10	8	1.5	0.6140221
Sub12	14	1.214286	0.4641571
AB: GroupAgg,Subject			
01_pre,Sub01	2	1	1.228044
01_pre,Sub03	8	0.875	0.6140221
01_pre,Sub04	5	0.6	0.7766834
01_pre,Sub05	5	0.8	0.7766834
01_pre,Sub06	7	1.285714	0.6564173
01_pre,Sub07	10	1.4	0.5491981
01_pre,Sub08	2	1	1.228044
01_pre,Sub09	7	2.142857	0.6564173
01_pre,Sub10	4	1	0.8683584
01_pre,Sub12	7	2	0.6564173
02_post,Sub01	2	0.5	1.228044
02_post,Sub03	8	1.375	0.6140221
02_post,Sub04	5	1.2	0.7766834
02_post,Sub05	5	$5.551115E{-17}$	0.7766834
02_post,Sub06	7	0.1428571	0.6564173
02_post,Sub07	10	1.5	0.5491981
02_post,Sub08	2	$4.551914E{-15}$	1.228044
02_post,Sub09	7	0.5714286	0.6564173
02_post,Sub10	4	2	0.8683584
02_post,Sub12	7	0.4285714	0.6564173

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   OT

Plots Section

Means Plot of OT

By GroupAgg

Means Plot of OT

By Subject

Means Plot of OT by Subject

Means Plot of OT by GroupAgg

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   OT

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: OT
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=2.9509

			Different From
Group	Count	Mean	Groups
01_pre	57	1.210357	
02_post	57	0.7717857	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Bonferroni (With Control) Multiple-Comparison Test

Response: OT
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=2.9509

If Control			Different From
Group Is	Count	Mean	Treatment Groups
01_pre	57	1.210357	
02_post	57	0.7717857	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: OT
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=4.1585

			Different From
Group	Count	Mean	Groups
01_pre	57	1.210357	
02_post	57	0.7717857	

Notes:
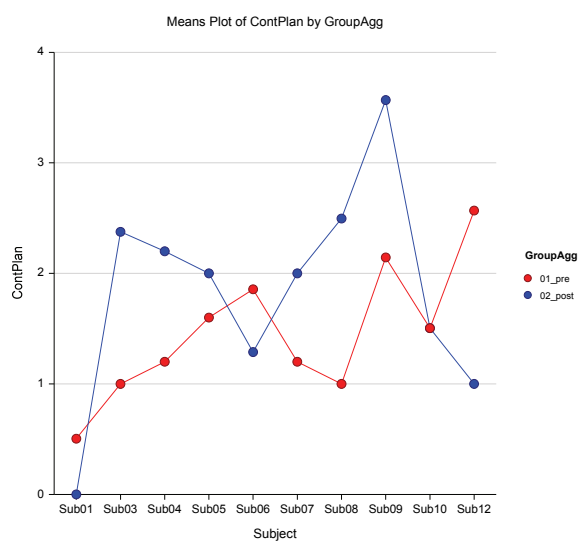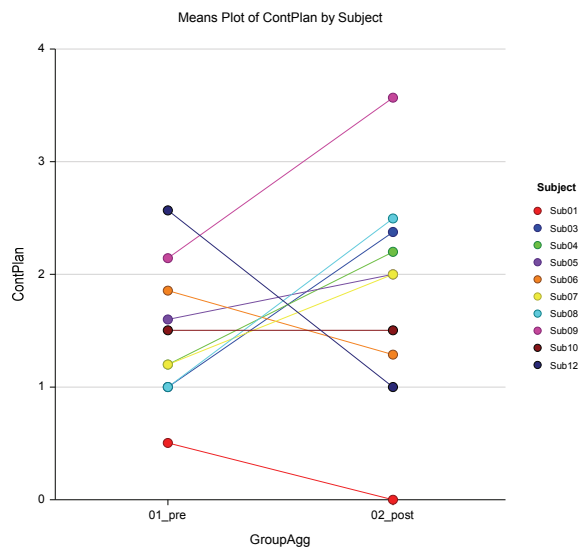This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
```
Dataset           C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response          OT
```


Planned Comparison: A1

```
Response: OT
Term A: GroupAgg
```

```
Alpha=0.004  Error Term=S(AB)  DF=94  MSE=3.016186
```

```
Comparison Value=-0.4385714   T-Value=1.3481   Prob>|T|=0.180854   Decision(0.004)=Do Not Reject
Comparison Std Error = 0.3253169   Comparison Confidence Interval = -1.398542 to 0.5213993
```

	Comparison		
Group	Coefficient	Count	Mean
01_pre	-1	57	1.210357
02_post	1	57	0.7717857

Notes:
This section presents the results of a standard set of planned comparisons. The first comparison
compares the first group with those below it (alphabetically). The second comparison compares the
second group with those below it (alphabetically) ignoring the first. This continues to the next
to the last group.

Bonferroni (All-Pairwise) Multiple Comparison Test

```
Response: OT
Term B: Subject
```

```
Alpha=0.004  Error Term=S(AB)  DF=94  MSE=3.016186 Critical Value=4.0964
```

			Different From
Group	Count	Mean	Groups
Sub01	4	0.75	
Sub03	16	1.125	
Sub04	10	0.9	
Sub05	10	0.4	
Sub06	14	0.7142857	
Sub07	20	1.45	
Sub08	4	0.5	
Sub09	14	1.357143	
Sub10	8	1.5	
Sub12	14	1.214286	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   OT

Bonferroni (With Control) Multiple-Comparison Test

Response: OT
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=3.6408

If Control Group Is	Count	Mean	Different From Treatment Groups
Sub01	4	0.75	
Sub03	16	1.125	
Sub04	10	0.9	
Sub05	10	0.4	
Sub06	14	0.7142857	
Sub07	20	1.45	
Sub08	4	0.5	
Sub09	14	1.357143	
Sub10	8	1.5	
Sub12	14	1.214286	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Tukey-Kramer Multiple-Comparison Test

Response: OT
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=5.6741

Group	Count	Mean	Different From Groups
Sub01	4	0.75	
Sub03	16	1.125	
Sub04	10	0.9	
Sub05	10	0.4	
Sub06	14	0.7142857	
Sub07	20	1.45	
Sub08	4	0.5	
Sub09	14	1.357143	
Sub10	8	1.5	
Sub12	14	1.214286	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   OT

Bonferroni (All-Pairwise) Multiple Comparison Test

Response: OT
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=4.4794

			Different From
Group	**Count**	**Mean**	**Groups**
01_pre,Sub01	2	1	
01_pre,Sub03	8	0.875	
01_pre,Sub04	5	0.6	
01_pre,Sub05	5	0.8	
01_pre,Sub06	7	1.285714	
01_pre,Sub07	10	1.4	
01_pre,Sub08	2	1	
01_pre,Sub09	7	2.142857	
01_pre,Sub10	4	1	
01_pre,Sub12	7	2	
02_post,Sub01	2	0.5	
02_post,Sub03	8	1.375	
02_post,Sub04	5	1.2	
02_post,Sub05	5	5.551115E-17	
02_post,Sub06	7	0.1428571	
02_post,Sub07	10	1.5	
02_post,Sub08	2	4.551914E-15	
02_post,Sub09	7	0.5714286	
02_post,Sub10	4	2	
02_post,Sub12	7	0.4285714	

Notes:
This section presents the results of all paired comparisons among the means.
Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Tukey-Kramer's method.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   OT

Bonferroni (With Control) Multiple-Comparison Test

Response: OT
Term AB: GroupAgg,Subject

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=3.016186 Critical Value=3.8565

If Control Group Is	Count	Mean	Different From Treatment Groups
01_pre,Sub01	2	1	
01_pre,Sub03	8	0.875	
01_pre,Sub04	5	0.6	
01_pre,Sub05	5	0.8	
01_pre,Sub06	7	1.285714	
01_pre,Sub07	10	1.4	
01_pre,Sub08	2	1	
01_pre,Sub09	7	2.142857	
01_pre,Sub10	4	1	
01_pre,Sub12	7	2	
02_post,Sub01	2	0.5	
02_post,Sub03	8	1.375	
02_post,Sub04	5	1.2	
02_post,Sub05	5	5.551115E-17	
02_post,Sub06	7	0.1428571	
02_post,Sub07	10	1.5	
02_post,Sub08	2	4.551914E-15	
02_post,Sub09	7	0.5714286	
02_post,Sub10	4	2	
02_post,Sub12	7	0.4285714	

Notes:
This section presents the results of comparisons of each treatment group with the control
group. Since this procedure uses the Bonferroni inequality, it is not as accurate as the
Dunnett's method.
Since the actual control group is not specified, a separate report line is generated assuming
that each group is the control group. Only use the line of the report that uses the actual control
group.

Analysis of Variance Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Response   OT

Tukey-Kramer Multiple-Comparison Test

Response: OT
Term AB: GroupAgg,Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=3.016186 Critical Value=6.1848

Group	Count	Mean	Different From Groups
01_pre,Sub01	2	1	
01_pre,Sub03	8	0.875	
01_pre,Sub04	5	0.6	
01_pre,Sub05	5	0.8	
01_pre,Sub06	7	1.285714	
01_pre,Sub07	10	1.4	
01_pre,Sub08	2	1	
01_pre,Sub09	7	2.142857	
01_pre,Sub10	4	1	
01_pre,Sub12	7	2	
02_post,Sub01	2	0.5	
02_post,Sub03	8	1.375	
02_post,Sub04	5	1.2	
02_post,Sub05	5	5.551115E-17	
02_post,Sub06	7	0.1428571	
02_post,Sub07	10	1.5	
02_post,Sub08	2	4.551914E-15	
02_post,Sub09	7	0.5714286	
02_post,Sub10	4	2	
02_post,Sub12	7	0.4285714	

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Frequency Table Reports

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   PtID_YN

Frequency Distribution of GroupAgg

GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	54	54	48.65%	48.65%	||||||||||||||||||||
02_post	57	111	51.35%	100.00%	|||||||||||||||||||||

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	54	55.50	48.6%	50.0%	0.0405
02_post	57	55.50	51.4%	50.0%	0.0405

Chi-Square = 0.0811 with df = 1   Probability Level = 0.775838

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	4	4	3.60%	3.60%		
Sub03	16	20	14.41%	18.02%	|||||
Sub04	10	30	9.01%	27.03%	|||
Sub05	9	39	8.11%	35.14%	|||
Sub06	14	53	12.61%	47.75%	|||||
Sub07	20	73	18.02%	65.77%	|||||||
Sub08	3	76	2.70%	68.47%		
Sub09	13	89	11.71%	80.18%	||||
Sub10	8	97	7.21%	87.39%	||
Sub12	14	111	12.61%	100.00%	|||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	4	11.10	3.6%	10.0%	4.5414
Sub03	16	11.10	14.4%	10.0%	2.1631
Sub04	10	11.10	9.0%	10.0%	0.1090
Sub05	9	11.10	8.1%	10.0%	0.3973
Sub06	14	11.10	12.6%	10.0%	0.7577
Sub07	20	11.10	18.0%	10.0%	7.1360
Sub08	3	11.10	2.7%	10.0%	5.9108
Sub09	13	11.10	11.7%	10.0%	0.3252
Sub10	8	11.10	7.2%	10.0%	0.8658
Sub12	14	11.10	12.6%	10.0%	0.7577

Chi-Square = 22.9640 with df = 9   Probability Level = 0.006278

Counts Report

Variables	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
GroupAgg	54	57	0	0	0	0	0	111
Subject	0	0	4	16	10	9	14	111
Total	54	57	4	16	10	9	14	222

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   PtID_YN

Counts Report (Continued)

			Values			
Variables						
	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	0	0	0	0	0	111
Subject	20	3	13	8	14	111
Total	20	3	13	8	14	222

Combined Report

			Values					
Variables								
		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	54	57	0	0	0	0	111
Subject	Count	0	0	4	16	10	9	111
Total	Count	54	57	4	16	10	9	222

		Values					
Variables							
		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	111
Subject	Count	14	20	3	13	8	14	111
Total	Count	14	20	3	13	8	14	222

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   PtID_YN

Plots Section

Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable
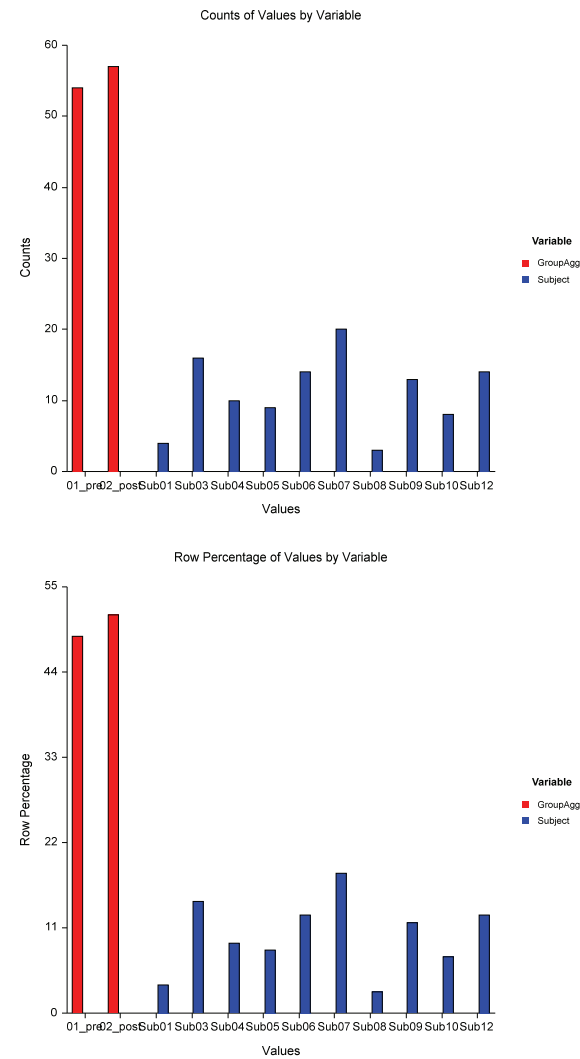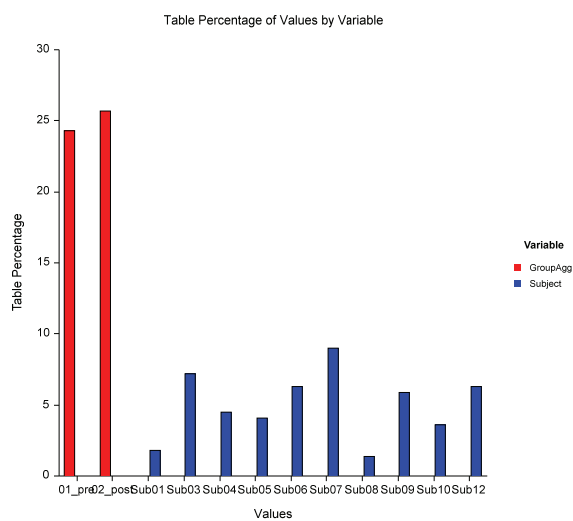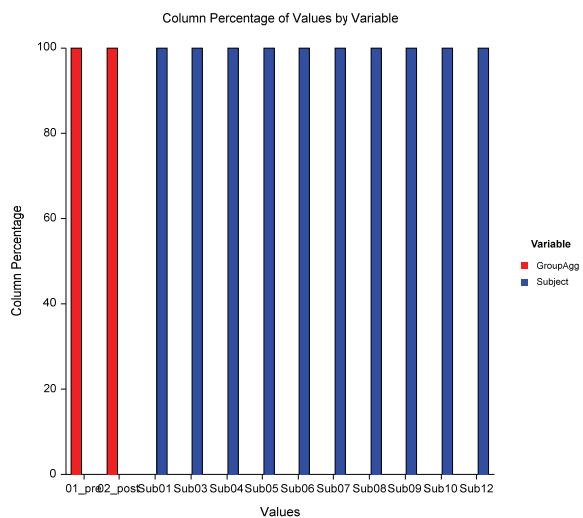
Frequency Table Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   BigS_YN

Frequency Distribution of GroupAgg

GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	43	43	43.88%	43.88%	|||||||||||||||||||
02_post	55	98	56.12%	100.00%	|||||||||||||||||||||||

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	43	49.00	43.9%	50.0%	0.7347
02_post	55	49.00	56.1%	50.0%	0.7347

Chi-Square = 1.4694 with df = 1   Probability Level = 0.225442

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	3	3	3.06%	3.06%		
Sub03	12	15	12.24%	15.31%	||||
Sub04	9	24	9.18%	24.49%	|||
Sub05	8	32	8.16%	32.65%	|||
Sub06	12	44	12.24%	44.90%	||||
Sub07	20	64	20.41%	65.31%	|||||||||
Sub08	4	68	4.08%	69.39%		
Sub09	12	80	12.24%	81.63%	||||
Sub10	5	85	5.10%	86.73%	||
Sub12	13	98	13.27%	100.00%	|||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	3	9.80	3.1%	10.0%	4.7184
Sub03	12	9.80	12.2%	10.0%	0.4939
Sub04	9	9.80	9.2%	10.0%	0.0653
Sub05	8	9.80	8.2%	10.0%	0.3306
Sub06	12	9.80	12.2%	10.0%	0.4939
Sub07	20	9.80	20.4%	10.0%	10.6163
Sub08	4	9.80	4.1%	10.0%	3.4327
Sub09	12	9.80	12.2%	10.0%	0.4939
Sub10	5	9.80	5.1%	10.0%	2.3510
Sub12	13	9.80	13.3%	10.0%	1.0449

Chi-Square = 24.0408 with df = 9   Probability Level = 0.004237

Counts Report

Variables	Values							
	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
GroupAgg	43	55	0	0	0	0	0	98
Subject	0	0	3	12	9	8	12	98
Total	43	55	3	12	9	8	12	196

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   BigS_YN

Counts Report (Continued)

Variables	Values					
	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	0	0	0	0	0	98
Subject	20	4	12	5	13	98
Total	20	4	12	5	13	196

Combined Report

Variables		Values						
		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	43	55	0	0	0	0	98
Subject	Count	0	0	3	12	9	8	98
Total	Count	43	55	3	12	9	8	196

Variables		Values						
		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	98
Subject	Count	12	20	4	12	5	13	98
Total	Count	12	20	4	12	5	13	196

Frequency Table Report

Dataset	C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency	BigS_YN

Plots Section
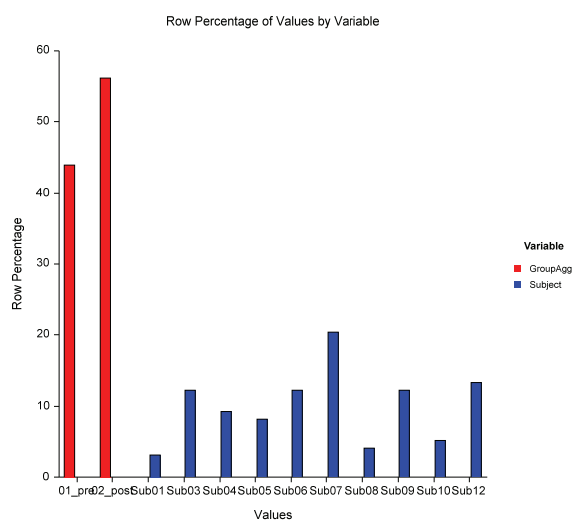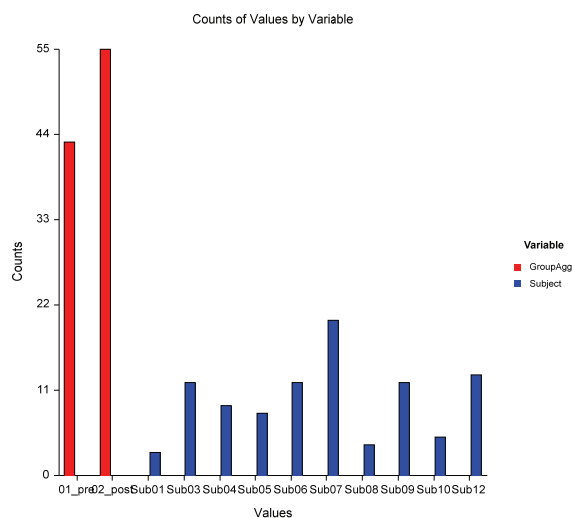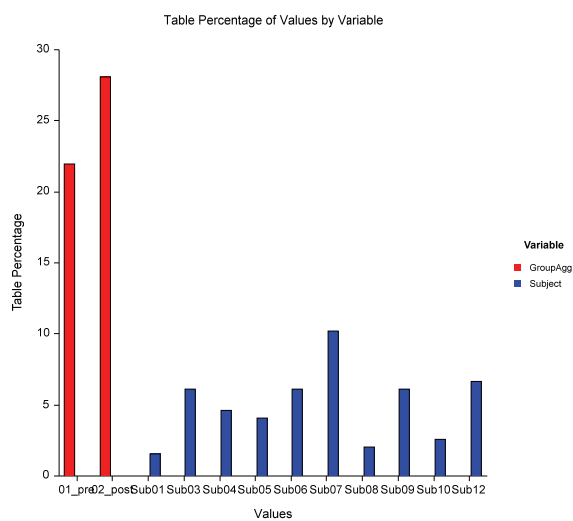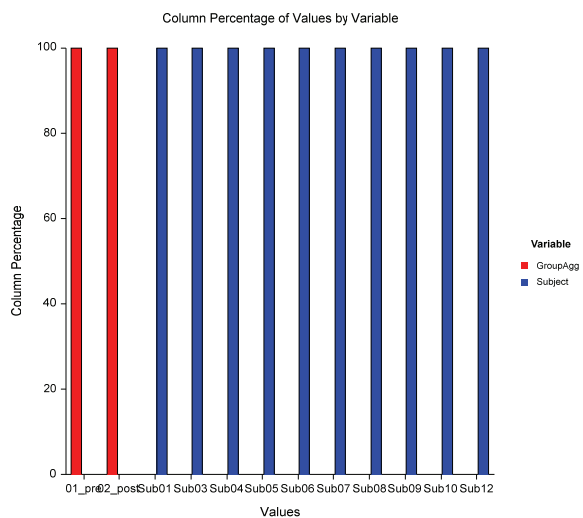
Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable

Frequency Table Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   CurrCond_YN

Frequency Distribution of GroupAgg

GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	48	48	46.60%	46.60%	|||||||||||||||||
02_post	55	103	53.40%	100.00%	||||||||||||||||||||

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	48	51.50	46.6%	50.0%	0.2379
02_post	55	51.50	53.4%	50.0%	0.2379

Chi-Square = 0.4757 with df = 1   Probability Level = 0.490364

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	3	3	2.91%	2.91%	|
Sub03	15	18	14.56%	17.48%	|||||
Sub04	8	26	7.77%	25.24%	|||
Sub05	9	35	8.74%	33.98%	|||
Sub06	14	49	13.59%	47.57%	|||||
Sub07	19	68	18.45%	66.02%	|||||||
Sub08	4	72	3.88%	69.90%	|
Sub09	13	85	12.62%	82.52%	|||||
Sub10	5	90	4.85%	87.38%	|
Sub12	13	103	12.62%	100.00%	|||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	3	10.30	2.9%	10.0%	5.1738
Sub03	15	10.30	14.6%	10.0%	2.1447
Sub04	8	10.30	7.8%	10.0%	0.5136
Sub05	9	10.30	8.7%	10.0%	0.1641
Sub06	14	10.30	13.6%	10.0%	1.3291
Sub07	19	10.30	18.4%	10.0%	7.3485
Sub08	4	10.30	3.9%	10.0%	3.8534
Sub09	13	10.30	12.6%	10.0%	0.7078
Sub10	5	10.30	4.9%	10.0%	2.7272
Sub12	13	10.30	12.6%	10.0%	0.7078

Chi-Square = 24.6699 with df = 9   Probability Level = 0.003359

Counts Report

Variables	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
				Values				
GroupAgg	48	55	0	0	0	0	0	103
Subject	0	0	3	15	8	9	14	103
Total	48	55	3	15	8	9	14	206

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   CurrCond_YN

Counts Report (Continued)

		Sub07	Sub08	Sub09	Sub10	Sub12	Total
Variables							
GroupAgg		0	0	0	0	0	103
Subject		19	4	13	5	13	103
Total		19	4	13	5	13	206

Values

Combined Report

Values

Variables		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	48	55	0	0	0	0	103
Subject	Count	0	0	3	15	8	9	103
Total	Count	48	55	3	15	8	9	206

Values

Variables		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	103
Subject	Count	14	19	4	13	5	13	103
Total	Count	14	19	4	13	5	13	206

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   CurrCond_YN

Plots Section
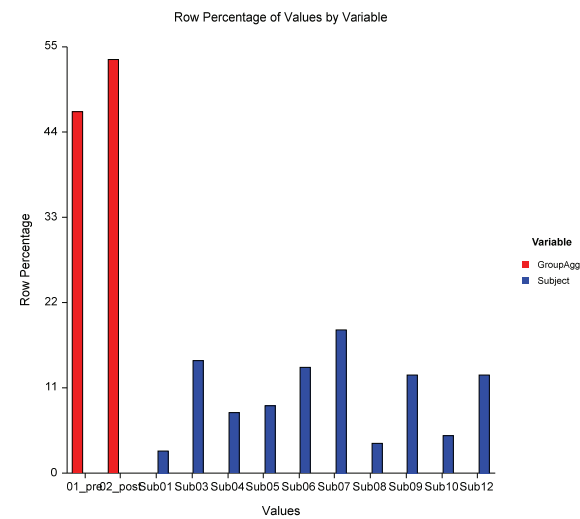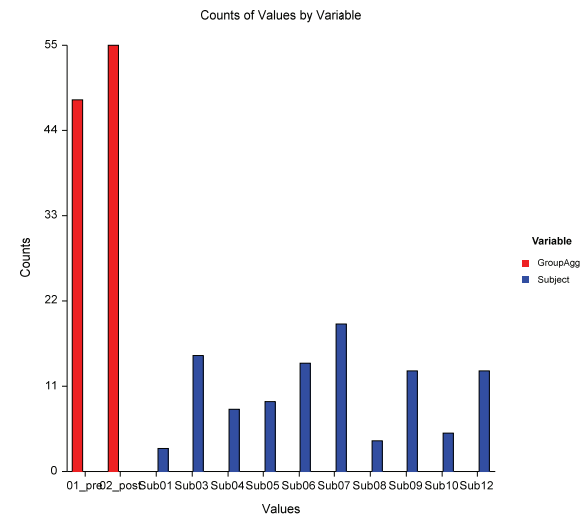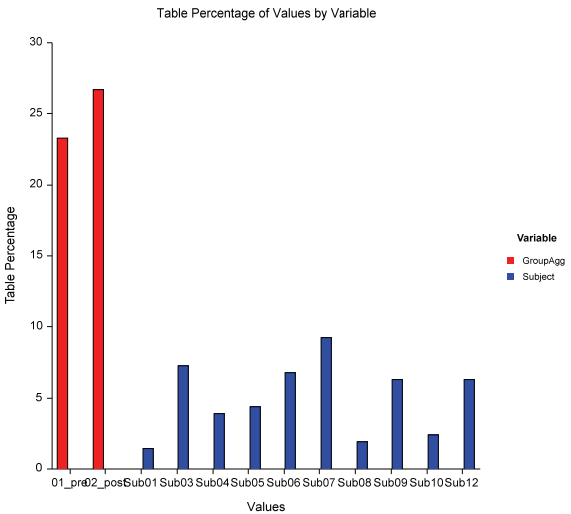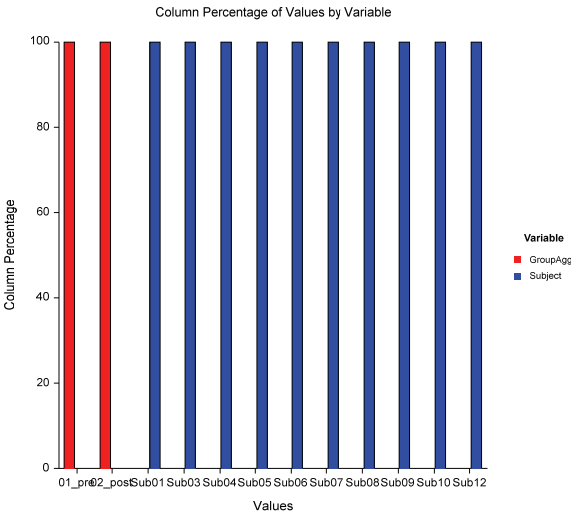
Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   PofC_YN

Frequency Distribution of GroupAgg

GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	51	51	47.22%	47.22%	||||||||||||||||||
02_post	57	108	52.78%	100.00%	||||||||||||||||||||

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	51	54.00	47.2%	50.0%	0.1667
02_post	57	54.00	52.8%	50.0%	0.1667

Chi-Square = 0.3333 with df = 1   Probability Level = 0.563703

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	4	4	3.70%	3.70%	|
Sub03	14	18	12.96%	16.67%	|||||
Sub04	9	27	8.33%	25.00%	|||
Sub05	8	35	7.41%	32.41%	||
Sub06	14	49	12.96%	45.37%	|||||
Sub07	20	69	18.52%	63.89%	|||||||
Sub08	4	73	3.70%	67.59%	|
Sub09	13	86	12.04%	79.63%	||||
Sub10	8	94	7.41%	87.04%	||
Sub12	14	108	12.96%	100.00%	|||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	4	10.80	3.7%	10.0%	4.2815
Sub03	14	10.80	13.0%	10.0%	0.9481
Sub04	9	10.80	8.3%	10.0%	0.3000
Sub05	8	10.80	7.4%	10.0%	0.7259
Sub06	14	10.80	13.0%	10.0%	0.9481
Sub07	20	10.80	18.5%	10.0%	7.8370
Sub08	4	10.80	3.7%	10.0%	4.2815
Sub09	13	10.80	12.0%	10.0%	0.4481
Sub10	8	10.80	7.4%	10.0%	0.7259
Sub12	14	10.80	13.0%	10.0%	0.9481

Chi-Square = 21.4444 with df = 9   Probability Level = 0.010817

Counts Report

Variables				Values				
	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
GroupAgg	51	57	0	0	0	0	0	108
Subject	0	0	4	14	9	8	14	108
Total	51	57	4	14	9	8	14	216

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   PofC_YN

Counts Report (Continued)

			Values			
Variables						
	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	0	0	0	0	0	108
Subject	20	4	13	8	14	108
Total	20	4	13	8	14	216

Combined Report

				Values				
Variables								
		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	51	57	0	0	0	0	108
Subject	Count	0	0	4	14	9	8	108
Total	Count	51	57	4	14	9	8	216

				Values				
Variables								
		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	108
Subject	Count	14	20	4	13	8	14	108
Total	Count	14	20	4	13	8	14	216

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   PofC_YN

Plots Section
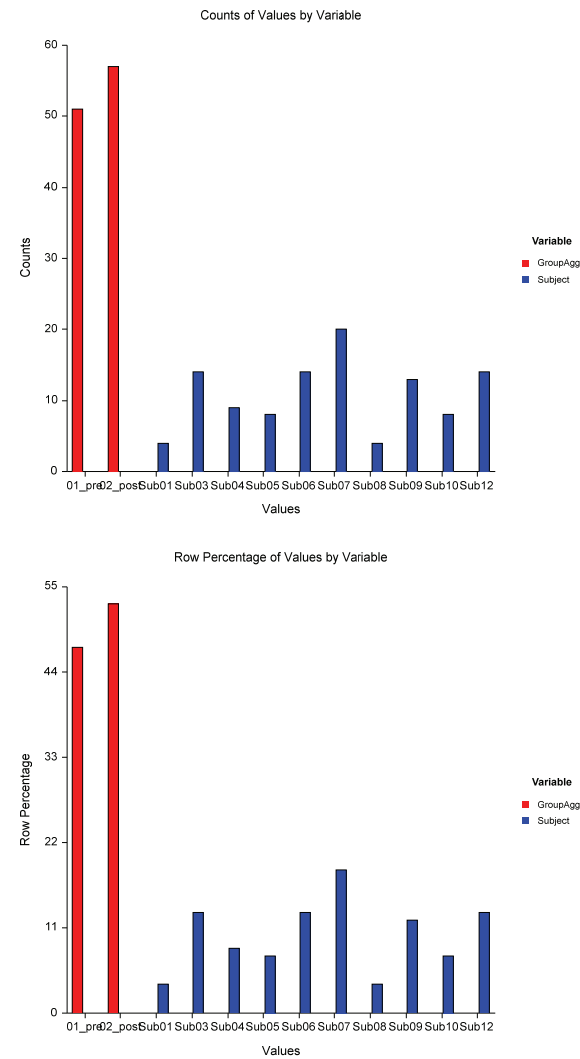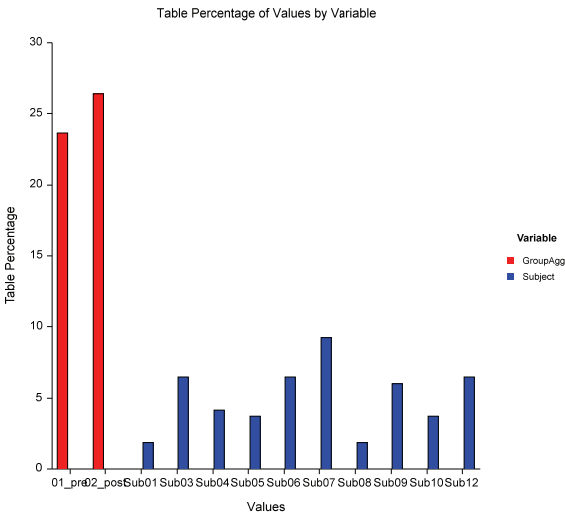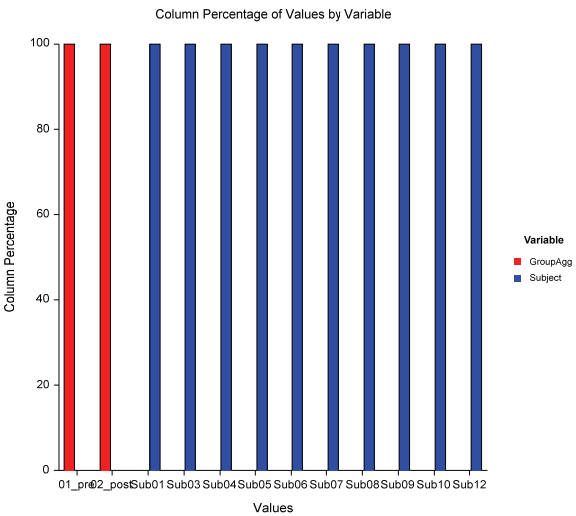
Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable

Frequency Table Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   ContPlan_YN

Frequency Distribution of GroupAgg

GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	28	28	36.36%	36.36%	|||||||||||||
02_post	49	77	63.64%	100.00%	
	||||||||||||||||||||||||				

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	28	38.50	36.4%	50.0%	2.8636
02_post	49	38.50	63.6%	50.0%	2.8636

Chi-Square = 5.7273 with df = 1   Probability Level = 0.016703

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	1	1	1.30%	1.30%	|
Sub03	9	10	11.69%	12.99%	||||
Sub04	7	17	9.09%	22.08%	|||
Sub05	7	24	9.09%	31.17%	|||
Sub06	10	34	12.99%	44.16%	|||||
Sub07	13	47	16.88%	61.04%	||||||
Sub08	3	50	3.90%	64.94%	|
Sub09	12	62	15.58%	80.52%	||||||
Sub10	6	68	7.79%	88.31%	|||
Sub12	9	77	11.69%	100.00%	||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	1	7.70	1.3%	10.0%	5.8299
Sub03	9	7.70	11.7%	10.0%	0.2195
Sub04	7	7.70	9.1%	10.0%	0.0636
Sub05	7	7.70	9.1%	10.0%	0.0636
Sub06	10	7.70	13.0%	10.0%	0.6870
Sub07	13	7.70	16.9%	10.0%	3.6481
Sub08	3	7.70	3.9%	10.0%	2.8688
Sub09	12	7.70	15.6%	10.0%	2.4013
Sub10	6	7.70	7.8%	10.0%	0.3753
Sub12	9	7.70	11.7%	10.0%	0.2195

Chi-Square = 16.3766 with df = 9   Probability Level = 0.059421

Counts Report

Variables				Values				
	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
GroupAgg	28	49	0	0	0	0	0	77
Subject	0	0	1	9	7	7	10	77
Total	28	49	1	9	7	7	10	154

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   ContPlan_YN

Counts Report (Continued)

Variables	Values					
	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	0	0	0	0	0	77
Subject	13	3	12	6	9	77
Total	13	3	12	6	9	154

Combined Report

Variables		Values						
		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	28	49	0	0	0	0	77
Subject	Count	0	0	1	9	7	7	77
Total	Count	28	49	1	9	7	7	154

Variables		Values						
		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	77
Subject	Count	10	13	3	12	6	9	77
Total	Count	10	13	3	12	6	9	154

5/12/2014 10:10:03 PM   3

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   ContPlan_YN

Plots Section
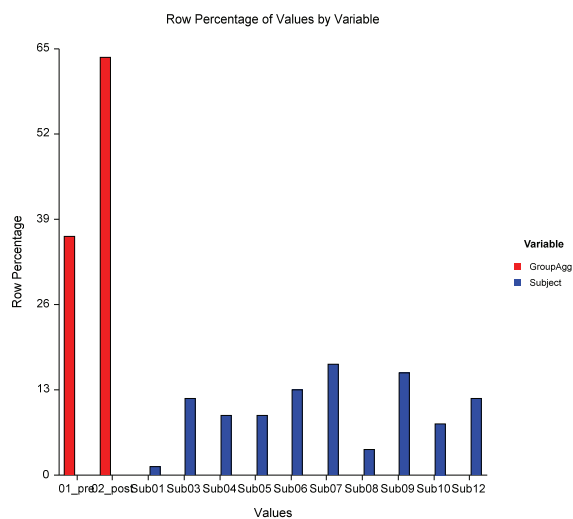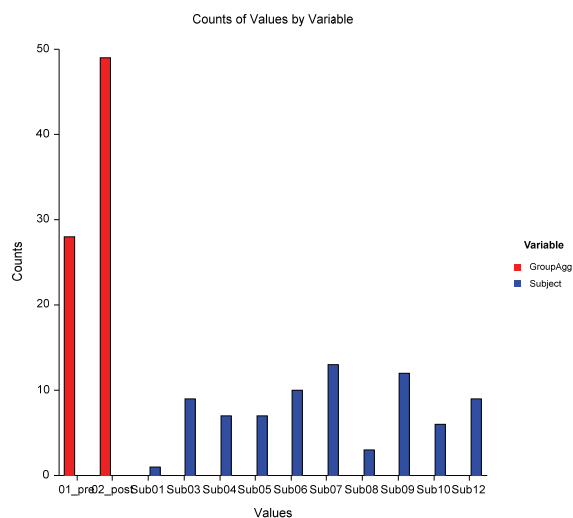
Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable

Frequency Table Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   AllFive_YN

Frequency Distribution of GroupAgg

GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	24	24	32.88%	32.88%	||||||||||||||
02_post	49	73	67.12%	100.00%	

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	24	36.50	32.9%	50.0%	4.2808
02_post	49	36.50	67.1%	50.0%	4.2808

Chi-Square = 8.5616 with df = 1   Probability Level = 0.003433

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	1	1	1.37%	1.37%		
Sub03	8	9	10.96%	12.33%	||||
Sub04	7	16	9.59%	21.92%	|||
Sub05	6	22	8.22%	30.14%	|||
Sub06	9	31	12.33%	42.47%	||||
Sub07	13	44	17.81%	60.27%	|||||||
Sub08	3	47	4.11%	64.38%		
Sub09	12	59	16.44%	80.82%	||||||
Sub10	5	64	6.85%	87.67%	||
Sub12	9	73	12.33%	100.00%	||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	1	7.30	1.4%	10.0%	5.4370
Sub03	8	7.30	11.0%	10.0%	0.0671
Sub04	7	7.30	9.6%	10.0%	0.0123
Sub05	6	7.30	8.2%	10.0%	0.2315
Sub06	9	7.30	12.3%	10.0%	0.3959
Sub07	13	7.30	17.8%	10.0%	4.4507
Sub08	3	7.30	4.1%	10.0%	2.5329
Sub09	12	7.30	16.4%	10.0%	3.0260
Sub10	5	7.30	6.8%	10.0%	0.7247
Sub12	9	7.30	12.3%	10.0%	0.3959

Chi-Square = 17.2740 with df = 9   Probability Level = 0.044595

Counts Report

Variables	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
			Values					
GroupAgg	24	49	0	0	0	0	0	73
Subject	0	0	1	8	7	6	9	73
Total	24	49	1	8	7	6	9	146

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   AllFive_YN

Counts Report (Continued)

		Values				
Variables						
	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	0	0	0	0	0	73
Subject	13	3	12	5	9	73
Total	13	3	12	5	9	146

Combined Report

		Values						
Variables								
		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	24	49	0	0	0	0	73
Subject	Count	0	0	1	8	7	6	73
Total	Count	24	49	1	8	7	6	146

		Values						
Variables								
		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	73
Subject	Count	9	13	3	12	5	9	73
Total	Count	9	13	3	12	5	9	146

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   AllFive_YN

Plots Section
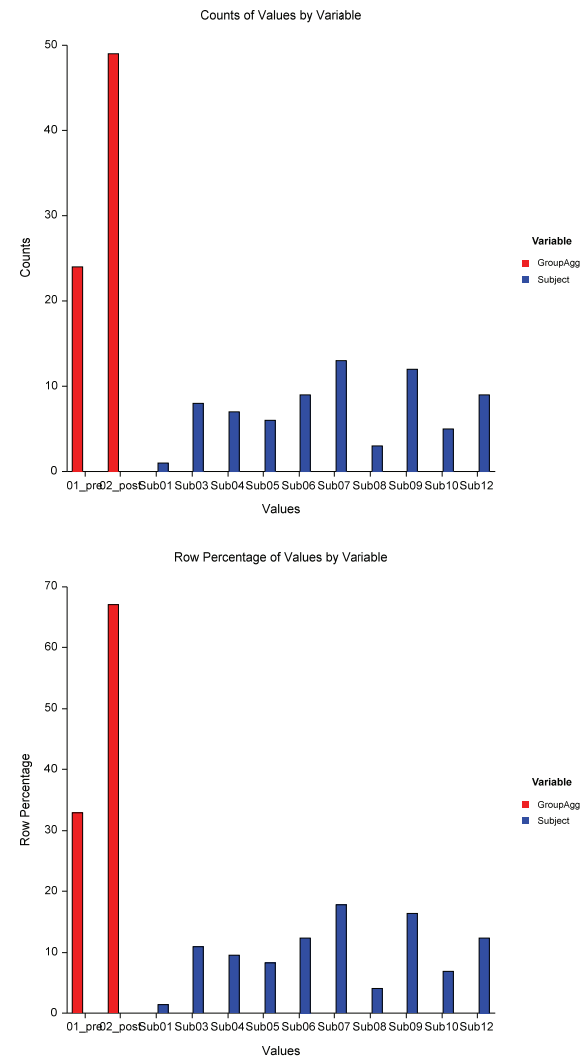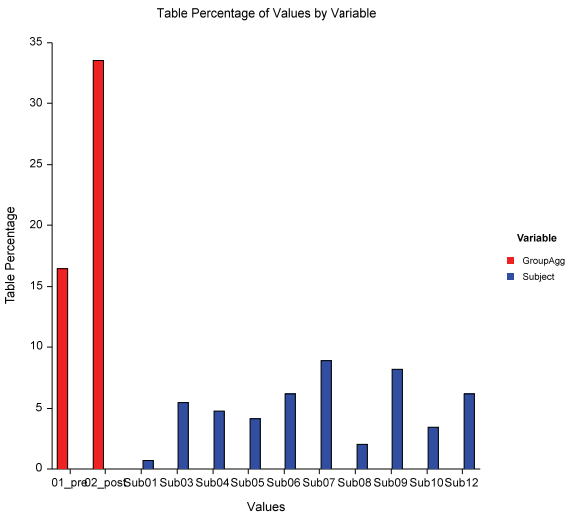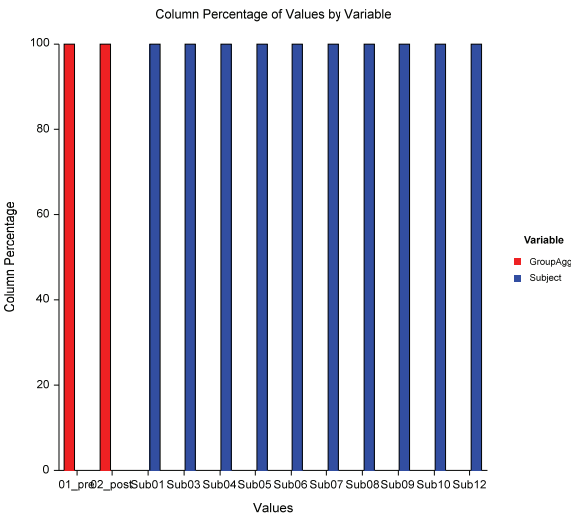
Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable

Frequency Table Report
Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS

Frequency   OT_YN

Frequency Distribution of GroupAgg

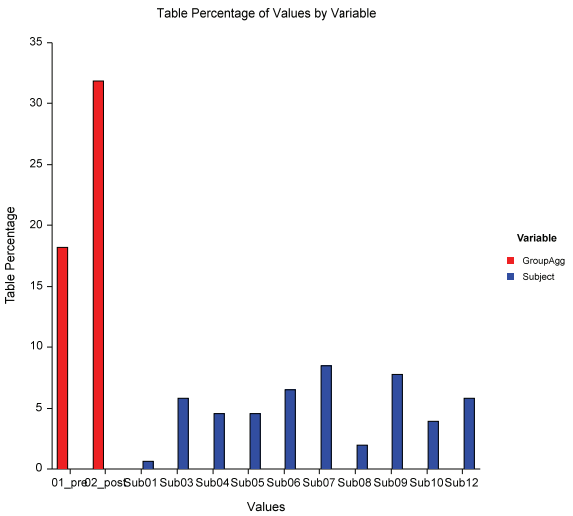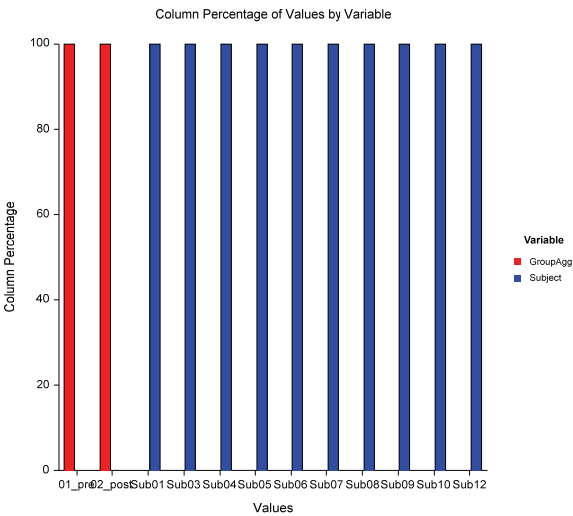| GroupAgg	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
01_pre	32	32	57.14%	57.14%	||||||||||||||||||||||
02_post	24	56	42.86%	100.00%	|||||||||||||||||

Multinomial Test of GroupAgg

GroupAgg	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
01_pre	32	28.00	57.1%	50.0%	0.5714
02_post	24	28.00	42.9%	50.0%	0.5714

Chi-Square = 1.1429 with df = 1   Probability Level = 0.285049

Frequency Distribution of Subject

Subject	Count	Cumulative Count	Percent	Cumulative Percent	Graph of Percent
Sub01	3	3	5.36%	5.36%	||
Sub03	10	13	17.86%	23.21%	|||||||
Sub04	5	18	8.93%	32.14%	|||
Sub05	3	21	5.36%	37.50%	||
Sub06	5	26	8.93%	46.43%	|||
Sub07	9	35	16.07%	62.50%	||||||
Sub08	1	36	1.79%	64.29%	|
Sub09	7	43	12.50%	76.79%	|||||
Sub10	6	49	10.71%	87.50%	||||
Sub12	7	56	12.50%	100.00%	|||||

Multinomial Test of Subject

Subject	Count	Expected Count	Actual Percent	Expected Percent	Chi-Square Amount
Sub01	3	5.60	5.4%	10.0%	1.2071
Sub03	10	5.60	17.9%	10.0%	3.4571
Sub04	5	5.60	8.9%	10.0%	0.0643
Sub05	3	5.60	5.4%	10.0%	1.2071
Sub06	5	5.60	8.9%	10.0%	0.0643
Sub07	9	5.60	16.1%	10.0%	2.0643
Sub08	1	5.60	1.8%	10.0%	3.7786
Sub09	7	5.60	12.5%	10.0%	0.3500
Sub10	6	5.60	10.7%	10.0%	0.0286
Sub12	7	5.60	12.5%	10.0%	0.3500

Chi-Square = 12.5714 with df = 9   Probability Level = 0.182977

Counts Report

Variables				Values				
	01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Sub06	Total
GroupAgg	32	24	0	0	0	0	0	56
Subject	0	0	3	10	5	3	5	56
Total	32	24	3	10	5	3	5	112

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   OT_YN

Counts Report (Continued)

				Values		
Variables						
	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	0	0	0	0	0	56
Subject	9	1	7	6	7	56
Total	9	1	7	6	7	112

Combined Report

				Values				
Variables								
		01_pre	02_post	Sub01	Sub03	Sub04	Sub05	Total
GroupAgg	Count	32	24	0	0	0	0	56
Subject	Count	0	0	3	10	5	3	56
Total	Count	32	24	3	10	5	3	112

				Values				
Variables								
		Sub06	Sub07	Sub08	Sub09	Sub10	Sub12	Total
GroupAgg	Count	0	0	0	0	0	0	56
Subject	Count	5	9	1	7	6	7	56
Total	Count	5	9	1	7	6	7	112

Frequency Table Report

Dataset   C:\...\chapter04_results\PairedStacked_ForANOVA.NCSS
Frequency   OT_YN

Plots Section

Counts of Values by Variable

Row Percentage of Values by Variable

Column Percentage of Values by Variable

Table Percentage of Values by Variable
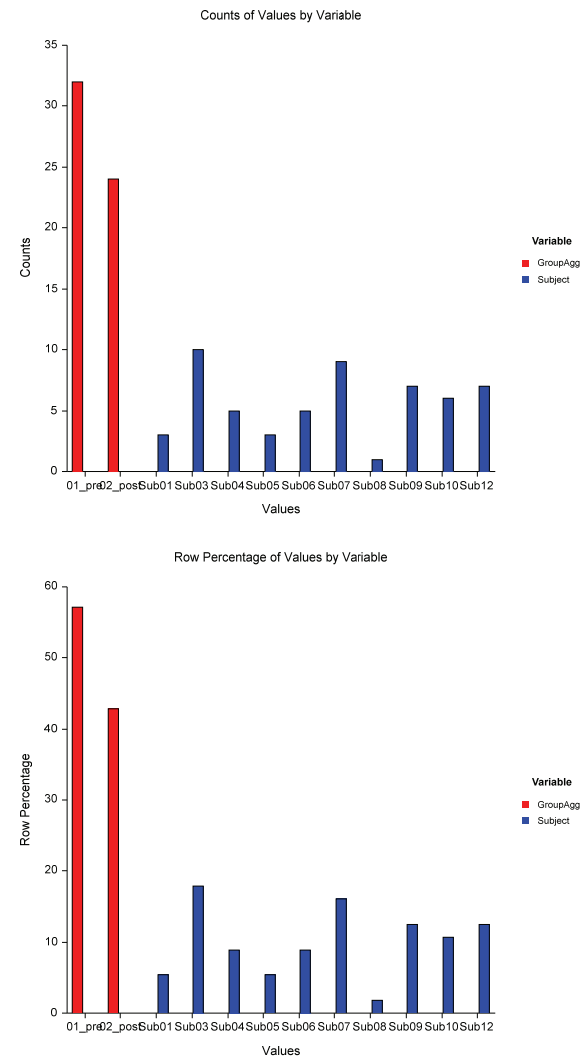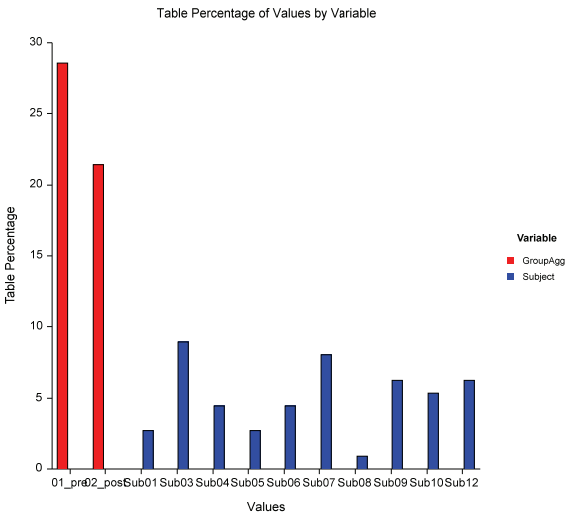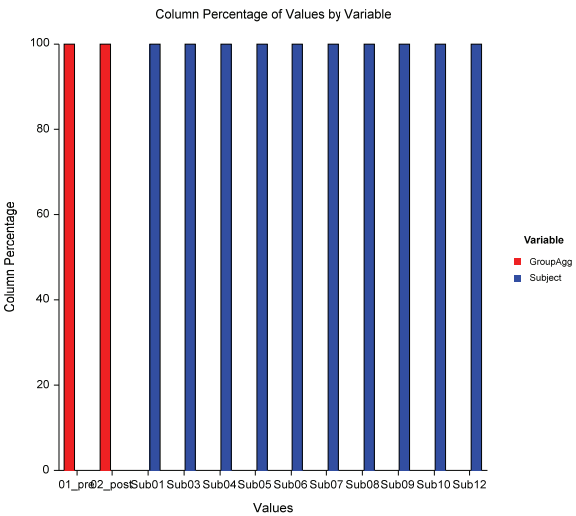
Appendix J. Training data analyses to address internal validity

Pairwise group comparisons to assess training effects on behavior

Baseline vs Pre-training

To strengthen internal validity in the absence of a true control group baseline data were compared to pre-training data. For the mean utterance counts and proportion-based dependent measures, the hypothesis was that significant differences would not exist between the baseline and pre-training groups in order to establish that pre-training handoffs were not significantly different from baseline. In comparing baseline data to the pre-training data, Mann-Whitney U-tests were used to compare mean counts of utterances and Chi-Square (Pearson's) tests of proportions were used to compare dependent measures of proportions.

Pre-training Observation 1 vs. Pre-training Observation 2

To further strengthen internal validity of the study in the absence of a true control group, a second set of analyses compared data from the two sets of pre-training observations: pre-training observation 1 vs. pre-training observation 2. Comparisons between the two pre-training observations serve to address validity issues with respect to history and maturation as well as statistical regression over time. Specifically, to show that the passage of time and practice were not confounding factors and to correct for statistical regression.

The overall hypothesis was that the first and second pre-training observations would not significantly differ across all dependent measures. In comparing pre-training observation 1 with pre-training observation 2, Wilcoxon Rank-Sum tests were conducted for dependent counts of and Chi-Square (Pearson's) tests were used to compare dependent measures of proportions.

Baseline vs Pre-training Results

Median number of utterances

Table 21 summarizes data for the baseline and pre-training group on dependent measures related to median utterances per patient of the five information components as well as for median number of off-task utterances per patient. The third section of Table 21 summarizes the results of Mann-Whitney U Tests comparing dependent measures of median utterance counts between baseline and pre-training. In sum, significant differences were observed between baseline and pre-training in the median number of utterances per patient of the patient ID, current condition, and plan of care. No significant differences were observed in the median utterances of the big sentence between baseline and pre-training ($U = 8495$, $p = 0.538$, $r = 0.04$), median utterances of contingency plans per patient between baseline and pre-training ($U = 8797.5$, $p = 0.886$, $r =$

0.01), or median off-task utterances per patient between baseline and pre-training ($U = 7763.5$, $p = 0.069$, $r = 0.11$)..

Table 21. Median utterances per patient of the five information components and utterances off-task between baseline and pre-training

Counts Per Patient	Baseline			Pre-Training			Statistical Results				
	n	Mdn	Range	n	Mdn	Range	N	U	Z	p‡	r*
Patient ID	202	1	3 [0, 3]	88	2	4 [0, 4]	290	4172	9.70	0.000†	0.57
Big Sentence	202	1	11 [0, 11]	88	2	8 [0, 8]	290	8495	-0.62	0.538	0.04
Current Condition	202	1	14 [0, 14]	88	3	18 [0, 18]	290	6122	4.23	0.000†	0.25
Plan of Care	202	2	9 [0, 9]	88	4	24 [0, 24]	290	4209	7.21	0.000†	0.42
Contingency Plans	202	1	6 [0, 6]	88	1	10 [0, 10]	290	8797.5	-0.14	0.886	0.01
Off-Task	202	1	11 [0, 11]	88	1	9 [0, 9]	290	7763.5	1.82	0.069	0.11

† Significantly lower median number of utterances of the patient ID per patient were observed at baseline ($U = 4172.0$, $p < 0.001$, $r = 0.57$), as well as a significantly lower median number of utterances of the current condition per patient at baseline ($U = 6122$, $p < 0.001$, $r = 0.25$), and a significantly lower median number of utterances of the plan of care per patient at baseline ($U = 4209$, $p < 0.001$, $r = 0.42$).

‡ No significant difference was observed in the median number of utterances of the big sentence between baseline and pre-training ($U = 8495$, $p = 0.538$, $r = 0.04$), contingency plans per patient between baseline and pre-training ($U = 8797.5$, $p = 0.886$, $r = 0.01$), or the median number of off-task utterances per patient between baseline and pre-training ($U = 7763.5$, $p = 0.069$, $r = 0.11$).

* r, the power of the Mann-Whitney U Test, computed by taking $r = \left|\frac{Z}{\sqrt{N}}\right|$.

Proportion of patients

Table 22 summarizes data for the baseline and pre-training groups on dependent measures related to proportion of patients for which each of the five information components were discussed both individually and all at once, as well as for proportion of patients in which off-task discussion occurred.

The third section of Table 22 summarizes the results of Pearson Chi-Square Tests of Proportions comparing each of these dependent measures between the baseline and pre-training groups. In sum, only two significant differences were observed between the two groups, the proportion of patients for which the plan of care was discussed and the proportion of patients for which contingency plans were discussed.

The proportion of patients was not significantly different at baseline versus pre-training for the proportion of patients which residents discussed the patient ID ($\chi^2(1) = 0.531$, $p = 0.466$), the big sentence ($\chi^2(1) = 0.256$, $p = 0.613$), the current condition ($\chi^2(1) = 2.990$, $p = 0.084$), all five components ($\chi^2(1) = 2.237$, $p = 0.135$), or for whom the discussion went off-task ($\chi^2(1) = 1.191$, $p = 0.275$).

Table 22. Proportion of baseline and pre-training patients for which each of the five information components were discussed individually, all-five at once, and discussion went off-task

Proportion of Patients	Baseline			Pre-Training			Chi-Square Results		
	N	n_{true}	Perc	N	n_{true}	Perc	χ^2	df	p
Patient ID	202	198	98.0%	88	85	96.6%	0.531	1	0.466
Big Sentence	202	157	77.7%	88	66	75.0%	0.256	1	0.613
Current Condition	202	146	72.3%	88	72	81.8%	2.990	1	0.084
Plan of Care	202	161	80.0%	88	81	92.1%	6.761	1	0.009[†]
Contingency Plans	202	133	65.8%	88	45	51.1%	5.592	1	0.018[‡]
All-Five Components	202	73	36.1%	88	40	45.5%	2.237	1	0.135
Off-Task Discussion	202	110	54.5%	88	54	61.4%	1.191	1	0.275

[†] The proportion of patients for which the plan of care was discussed was significantly lower at baseline (80.0%) than at pre-training (92.1%), by a Chi-square test of proportions ($\chi^2(1) = 6.761$, $p = 0.009$).

[‡] The proportion of patients for which contingency plans were discussed was significantly higher at baseline (65.8%) than at pre-training (51.1%), by a Chi-square test of proportions ($\chi^2(1) = 5.592$, $p < 0.018$).

Pre-training Observation 1 vs Pre-training Observation 2 Results

Median number of utterances

Table 23 summarizes data for pre-training observations 1 and 2 on dependent measures related to median utterances per patient of the five information components as well as for median number of off-task utterances per patient.

The third section of Table 23 summarizes the results of Wilcoxon Rank-Sum Tests for each of these dependent measures between the two pre-training observations. In short, significant differences were observed between both pre-training observations on the median utterances of the patient ID per patient and median utterances of the big sentence per patient, both significantly higher in pre-training observation 1.

As Table 23 shows, no significant differences were observed in the median utterances of the current condition per patient between the two pre-training observations ($W = 1331.5$, $p = 0.420$, $r = 0.09$), median utterances of the plan of care per patient ($W = 1261$, $p = 0.154$, $r = 0.15$), median utterances of contingency plans per patient ($W = 1313.5$, $p = 0.308$, $r = 0.11$), or median off-task utterances per patient ($W = 1419.5$, $p = 0.971$, $r = 0.00$).

Table 23. Median number of utterances per patient for the five information components and off-task communication and Wilcoxon Rank-Sum Tests comparing pre-training observations 1 vs. 2

Counts Per Patient	Pre-training Obs. 1			Pre-Training Obs. 2			Statistical Results				
	n	Mdn	Range	n	Mdn	Range	N	W	Z	p	r^*
Patient ID	56	2	4 [0, 4]	32	1	1 [0, 1]	88	606.5	-8.21	0.000[†]	0.88

Big Sentence	56	2	8 [0, 8]	32	1	4 [0, 4]	88	1201.5	-2.03	0.043[†] 0.22
Current Condition	56	4	18 [0, 18]	32	3	10 [0, 10]	88	1331.5	-0.81	0.420 0.09
Plan of Care	56	6	24 [0, 24]	32	4	13 [1, 14]	88	1261	-1.42	0.154 0.15
Contingency Plans	56	0.5	10 [0, 10]	32	1	7 [0, 7]	88	1313.5	-1.02	0.308 0.11
Off-Task	56	1	9 [0, 9]	32	1	9 [0, 9]	88	1419.5	-0.04	0.971 0.00

[†] A significantly higher median number of utterances of Patient ID per patient in pre-training observation 1 than in pre-training observation 2 (Wilcoxon $W = 606.5$, $p < 0.001$, $r = 0.88$) as well as a significantly higher median number of utterances of "The Big Sentence" per patient in pre-training observation 1 than in pre-training observation 2 (Wilcoxon $W = 1201.5$, $p = 0.043$, $r = 0.22$).

[*] r, the power of the Wilcoxon Rank-Sum Test, computed by taking $r = \left| \frac{z}{\sqrt{N}} \right|$.

Proportion of patients

Table 24 summarizes data for both pre-training observations 1 and 2 on dependent measures related to proportion of patients for which each of the five information components were discussed both individually and all at once, as well as the proportion of patients for which off-task discussion was observed.

The third section of Table 24 summarizes results of Pearson Chi-Square Tests of Proportions comparing each of these dependent measures between the two pre-training observations. Only one significant difference was observed between the two pre-training observations with respect to measures of proportion. That is, the proportion of patients for which the plan of care was discussed was significantly larger in pre-training observation 2 ($\chi^2(1) = 4.346$, $p = 0.037$). Comparisons across all other dependent measures between the two pre-training observations did not yield significant differences.

Table 24. Proportion of pre-training observation 1 and 2 patients for which each of the five information components were discussed, all-five discussed, and where discussion went off-task

	Pre-training Obs. 1			Pre-Training Obs. 2			Statistical Results		
Proportion of Patients	N	n_{true}	*Perc*	N	n_{true}	*Perc*	χ^2	*df*	*p* ‡
Patient ID	56	54	96.4%	32	31	96.9%	0.012	1	0.912
Big Sentence	56	40	71.4%	32	26	81.3%	1.048	1	0.306
Current Condition	56	44	78.6%	32	28	87.5%	1.091	1	0.296
Plan of Care	56	49	87.5%	32	32	100%	4.346	1	0.037†
Contingency Plans	56	28	50.0%	32	17	53.1%	0.080	1	0.778
All-Five Components	56	25	44.6%	32	15	46.9%	0.041	1	0.840
Off-Task Discussion	56	33	58.9%	32	20	62.5%	0.108	1	0.742

† A significantly higher proportion of patients for which the plan of care was discussed in pre-training observation 2 (100%) as opposed to pre-training observation 1 (87.5%) ($\chi^2(1) = 4.346$, $p = 0.037$).

‡ No significant differences between the pre-training observations found in the proportion of patients for which they discussed Patient ID ($\chi^2(1) = 0.012$, $p = 0.912$), the Big Sentence ($\chi^2(1) = 1.048$, $p = 0.306$), the Current Condition ($\chi^2(1) = 1.091$, $p = 0.296$), the Plan of Care ($\chi^2(1) = 0.080$, $p = 0.778$), all Five Components ($\chi^2(1) = 0.041$, $p = 0.840$). or for whom discussion went off-task ($\chi^2(1) = 0.108$, $p = 0.742$).

Table 25 summarizes data for the pre-training and post-training groups on dependent measures related to median number of utterances per patient for each of the five information components individually as well as for median off-task utterances per patient.

The third section of Table 25 summarizes the results of Wilcoxon Rank-Sum Tests for each of these dependent measures between the pre-training and post-training groups. In sum, the post-training group showed significant improvement across all dependent measures *except* for the median number of utterances of the *current condition* per patient (Wilcoxon W = 5696.5, p = 0.127) and the median number of utterances of the *plan of care* per patient (Wilcoxon W = 5624, p = 0.078).

Table 25. Median utterances per patient of the five information components and off-task discussion with Wilcoxon Rank-Sum Tests results for pre-training vs post-training across the six dependent measures of interest

Counts Per Patient	Pre-Training			Post-Training			Statistical Results				
	n	Mdn	Range	n	Mdn	Range	N	W	Z	p	r*
Patient ID	88	2	4 [0, 4]	75	1	1 [1, 2]	163	4359	-7.14	0.000†	0.56
Big Sentence	88	2	8 [0, 8]	75	1	2 [0, 2]	163	5397	-2.72	0.007†	0.21
Current Condition	88	3	18 [0, 18]	75	2	11 [0, 11]	163	5696.5	-1.52	0.127	0.12
Plan of Care	88	4	24 [0, 24]	75	3	19 [1, 20]	163	5624	-1.76	0.078	0.14
Contingency Plans	88	1	10 [0, 10]	75	1	7 [0, 7]	163	6567.5	2.23	0.026‡	0.17
Off-Task	88	1	9 [0, 9]	75	0	5 [0, 5]	163	5559	-2.09	0.036‡‡	0.16

† A significantly lower median number of utterances of Patient ID per patient was observed post-training (Wilcoxon W = 4359, p < 0.001), as well as a significantly lower median number of utterances of the big sentence per patient post-training (Wilcoxon W = 5397, p = 0.007),

‡ A significant difference was observed in the median number of utterances of Contingency Plans per patient between the pre-training group (Mdn = 1, 95% CI = [0.00, 1.00]) and the post-training group (Mdn = 1, 95% CI = [1.00, 2.00]) by a Wilcoxon Rank-Sum Test (Wilcoxon W = 6567.5, p = 0.026). Note that the 95% CI for the pre-training group was [0, 1] while the 95% CI for the post-training group was [1, 2] indicating the post training group was significantly more likely to discuss contingency plans more often per patient than the pre-training group.

‡‡ A significant difference was also observed in the median number of Off-Task utterances per patient between the pre-training group (Mdn = 1, 95% CI = [0.00, 2.00]) and the post-training group (Mdn = 0, 95% CI = [0.00, 1.00]) by a Wilcoxon Rank-Sum Test (Wilcoxon W = 5559, p = 0.036). Note that the group medians were 1.00 and 0.00 for the pre-training and post-training groups, respectively. The 95% CI for the median of the pre-training group was [0, 2] while the 95% CI for the post-training group was [0, 1] indicating the post training group was significantly less likely to go off-task more than once per patient than the pre-training group.

* r, the power of the Wilcoxon Rank-Sum Test, computed by $r = \left| \frac{Z}{\sqrt{N}} \right|$.

Table 26 summarizes data for the pre-training and post-training groups on dependent measures related to proportion of patients for which each of the five information components were discussed both individually and all at once, as well as the proportion of patients for which off-task discussion was observed.

The third section of Table 26 summarizes results of Pearson Chi-Square Tests of Proportions comparing each of these dependent measures between the pre-training and post-training groups. In sum, only one of the dependent measures of proportions did not differ significantly between the two groups, the proportion of patients for which residents discussed the patient ID ($\chi^2(1) = 2.605$, $p = 0.107$).

For all other dependent measures of proportions (the big sentence, current condition, plan of care, contingency plans, all five components, and off-task discussion), the post-training group showed significant improvement over the pre-training group (Table 26).

Table 26. Proportion of pre-training and post-training patients for which each of the five information components were discussed, all-five per patient, and for off-task discussion

	Pre-Training			Post-Training			Statistical Results		
Proportion of Patients	N	n_{true}	Perc	N	n_{true}	Perc	χ^2	df	p †
Patient ID	88	85	96.6%	75	75	100%	2.605	1	0.107
Big Sentence	88	66	75.0%	75	70	93.3%	9.847	1	0.002
Current Condition	88	72	81.8%	75	72	96.0%	7.908	1	0.005
Plan of Care	88	81	92.1%	75	75	100%	6.234	1	0.013
Contingency Plans	88	45	51.1%	75	65	81.3%	16.236	1	0.000
All-Five Components	88	40	45.5%	75	60	80.0%	20.379	1	0.000
Off-Task Discussion	88	54	61.4%	75	34	45.3%	4.189	1	0.041

† A significant difference was observed in the proportion of patients for which "The Big Sentence" was discussed between the pre-training and post-training groups ($\chi^2(1) = 9.847$, $p = 0.002$), the proportion of patients for which the Current Condition was discussed ($\chi^2(1) = 7.908$, $p = 0.005$), the proportion of patients for which the Plan of Care was discussed ($\chi^2(1) = 6.234$, $p = 0.013$), the proportion of patients for which Contingency Plans were discussed ($\chi^2(1) = 16.236$, $p < 0.001$), the proportion of patients for which All Five Components were discussed ($\chi^2(1) = 20.379$, $p < 0.001$), and the proportion of patients for which the discussion went Off-Task ($\chi^2(1) = 4.189$, $p = 0.041$).

Pre-training post-training knowledge assessments

A two-tailed Chi-Square (Pearson's) test of proportions revealed that the difference in the knowledge assessment scores between pre-training (86.0%) and post-training (96.0%) was not significant ($\chi^2(1) = 3.053$, $p = 0.081$) (see Appendix B, Table 14).

Appendix K. Full statistical output for pre vs. post training comparisons

Dependent measures of proportions

Patient ID

```
Frequency Distribution of GroupAgg
                                     Cumulative              Cumulative    Graph of
GroupAgg                    Count         Count   Percent       Percent    Percent
01_pre                         54            54    48.65%        48.65%    ||||||||||||||||||||
02_post                        57           111    51.35%       100.00%    ||||||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                      Expected      Actual    Expected    Chi-Square
GroupAgg                    Count        Count     Percent     Percent    Amount
01_pre                         54        55.50       48.6%       50.0%     0.0405
02_post                        57        55.50       51.4%       50.0%     0.0405
Chi-Square = 0.0811 with df = 1   Probability Level = 0.775838
```

The Big Sentence

```
Frequency Distribution of GroupAgg
                                     Cumulative              Cumulative    Graph of
GroupAgg                    Count         Count   Percent       Percent    Percent
01_pre                         43            43    43.88%        43.88%    ||||||||||||||||||
02_post                        55            98    56.12%       100.00%    |||||||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                      Expected      Actual    Expected    Chi-Square
GroupAgg                    Count        Count     Percent     Percent    Amount
01_pre                         43        49.00       43.9%       50.0%     0.7347
02_post                        55        49.00       56.1%       50.0%     0.7347
Chi-Square = 1.4694 with df = 1   Probability Level = 0.225442
```

Current Condition

```
Frequency Distribution of GroupAgg
                                     Cumulative              Cumulative    Graph of
GroupAgg                    Count         Count   Percent       Percent    Percent
01_pre                         48            48    46.60%        46.60%    |||||||||||||||||||
02_post                        55           103    53.40%       100.00%    |||||||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                      Expected      Actual    Expected    Chi-Square
GroupAgg                    Count        Count     Percent     Percent    Amount
01_pre                         48        51.50       46.6%       50.0%     0.2379
02_post                        55        51.50       53.4%       50.0%     0.2379
Chi-Square = 0.4757 with df = 1   Probability Level = 0.490364
```

Plan of Care

```
Frequency Distribution of GroupAgg
                                     Cumulative              Cumulative    Graph of
GroupAgg                    Count         Count   Percent       Percent    Percent
01_pre                         51            51    47.22%        47.22%    ||||||||||||||||||
02_post                        57           108    52.78%       100.00%    |||||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                    Expected      Actual      Expected    Chi-Square
GroupAgg                    Count      Count     Percent       Percent       Amount
01_pre                         51      54.00       47.2%         50.0%       0.1667
02_post                        57      54.00       52.8%         50.0%       0.1667
Chi-Square = 0.3333 with df = 1   Probability Level = 0.563703
```

Contingency Plans

```
Frequency Distribution of GroupAgg
                                 Cumulative              Cumulative     Graph of
GroupAgg                    Count      Count     Percent     Percent     Percent
01_pre                         28         28      36.36%      36.36%     ||||||||||||||
02_post                        49         77      63.64%     100.00%
                  |||||||||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                    Expected      Actual      Expected    Chi-Square
GroupAgg                    Count      Count     Percent       Percent       Amount
01_pre                         28      38.50       36.4%         50.0%       2.8636
02_post                        49      38.50       63.6%         50.0%       2.8636
Chi-Square = 5.7273 with df = 1   Probability Level = 0.016703
```

All-Five Information Components

```
Frequency Distribution of GroupAgg
                                 Cumulative              Cumulative     Graph of
GroupAgg                    Count      Count     Percent     Percent     Percent
01_pre                         24         24      32.88%      32.88%     |||||||||||||
02_post                        49         73      67.12%     100.00%
                  |||||||||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                    Expected      Actual      Expected    Chi-Square
GroupAgg                    Count      Count     Percent       Percent       Amount
01_pre                         24      36.50       32.9%         50.0%       4.2808
02_post                        49      36.50       67.1%         50.0%       4.2808
Chi-Square = 8.5616 with df = 1   Probability Level = 0.003433
```

```
McNemar Two-Sided Hypothesis Tests of the Difference (P1 - P2)
H0: P1 = P2 vs. Ha: P1 ≠ P2

              Distribution    Null             Test
Test          of Test         Hypothesis       Statistic      Prob     Conclusion
Method        Statistic       (H0)                 Value      Level     at α = 0.004
McNemar       CS(1)           P1 = P2            21.9298     0.0000     Reject H0
```

Off-Task

```
Frequency Distribution of GroupAgg
                                 Cumulative              Cumulative     Graph of
GroupAgg                    Count      Count     Percent     Percent     Percent
01_pre                         32         32      57.14%      57.14%     ||||||||||||||||||||||
02_post                        24         56      42.86%     100.00%     ||||||||||||||||||
```

```
Multinomial Test of GroupAgg

                                    Expected      Actual      Expected    Chi-Square
GroupAgg                    Count      Count     Percent       Percent       Amount
```

```
01_pre                                 32         28.00       57.1%        50.0%    0.5714
02_post                                24         28.00       42.9%        50.0%    0.5714
Chi-Square = 1.1429 with df = 1    Probability Level = 0.285049
```

Dependent measures of mean utterance counts

Patient ID

```
Analysis of Variance Table
Source                     Sum of       Mean                    Prob       Power
Term              DF       Squares      Square      F-Ratio     Level      (Alpha=0.004)
A: GroupAgg       1        7.44629      7.44629       34.82     0.000000*  0.998081
B: Subject        9        3.250752     0.3611946      1.69     0.102383   0.376931
AB                9        3.561529     0.3957254      1.85     0.069169   0.435324
S                 94       20.1         0.2138298
Total (Adjusted)  113      37.65789
Total             114
* Term significant at alpha = 0.004
```
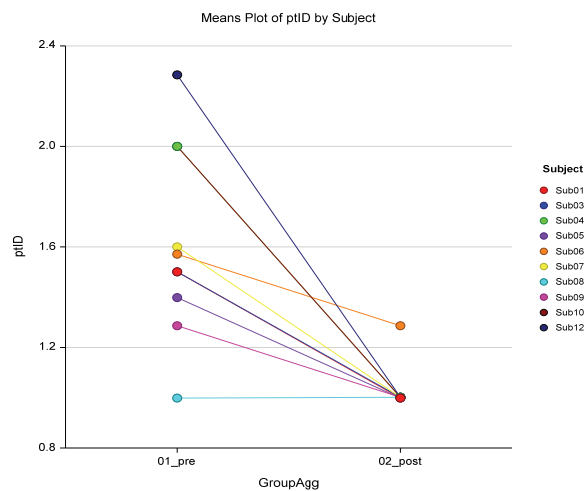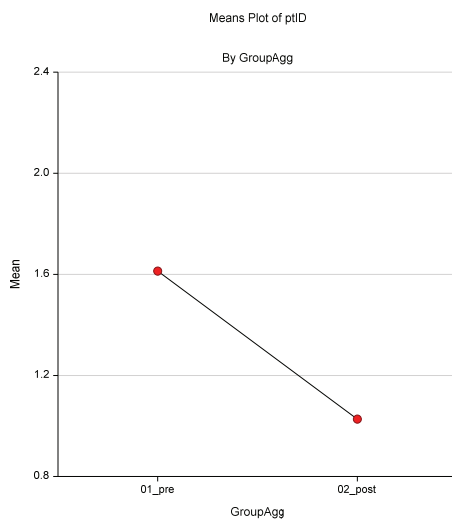
```
Tukey-Kramer Multiple-Comparison Test

Response: ptID
Term A: GroupAgg

Alpha=0.004   Error Term=S(AB)   DF=94   MSE=0.2138298 Critical Value=4.1585


                                                   Different From
Group                  Count        Mean           Groups
01_pre                 57           1.614286        02_post
02_post                57           1.028571        01_pre

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.
```


Means Plot of ptID — By GroupAgg

Means Plot of ptID by Subject

```
Analysis of Variance Table
Source                     Sum of       Mean                    Prob       Power
Term              DF       Squares      Square      F-Ratio     Level      (Alpha=0.004)
A: GroupAgg       1        8.07659      8.07659        9.42     0.002804*  0.549191
B: Subject        9        28.52218     3.169131       3.70     0.000532*  0.898561
```

```
AB                             9       30.15526      3.350585           3.91  0.000302*  0.921444
S                             94       80.58572      0.8572949
Total (Adjusted)             113      145.6579
Total                        114
* Term significant at alpha = 0.004
```

Big Sentence

Tukey-Kramer Multiple-Comparison Test

Response: BigS
Term A: GroupAgg

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.8572949 Critical Value=4.1585

			Different From
Group	Count	Mean	Groups
01_pre	57	1.668571	02_post
02_post	57	1.058571	01_pre

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.

Tukey-Kramer Multiple-Comparison Test

Response: BigS
Term B: Subject

Alpha=0.004 Error Term=S(AB) DF=94 MSE=0.8572949 Critical Value=5.6741

			Different From
Group	Count	Mean	Groups
Sub01	4	0.75	
Sub03	16	0.75	Sub12
Sub04	10	1.6	
Sub05	10	1.1	
Sub06	14	1.071429	
Sub07	20	1.4	
Sub08	4	2.5	
Sub09	14	1.5	
Sub10	8	0.75	
Sub12	14	2.214286	Sub03

Notes:
This report provides multiple comparison tests for all pairwise differences between
the means.