

**UCSF**

**UC San Francisco Electronic Theses and Dissertations**

**Title**

From paper-and-pencil to web: the SF-36 and CRQ

**Permalink**

<https://escholarship.org/uc/item/947838f1>

**Author**

Benjamin, Bonnie

**Publication Date**

2008

Peer reviewed|Thesis/dissertation

**From Paper-and-Pencil to Web: The SF-36 and CRQ**

by

**Bonnie Benjamin**

**THESIS**

**Submitted in partial satisfaction of the requirements for the degree of**

**MASTER OF SCIENCE**

in

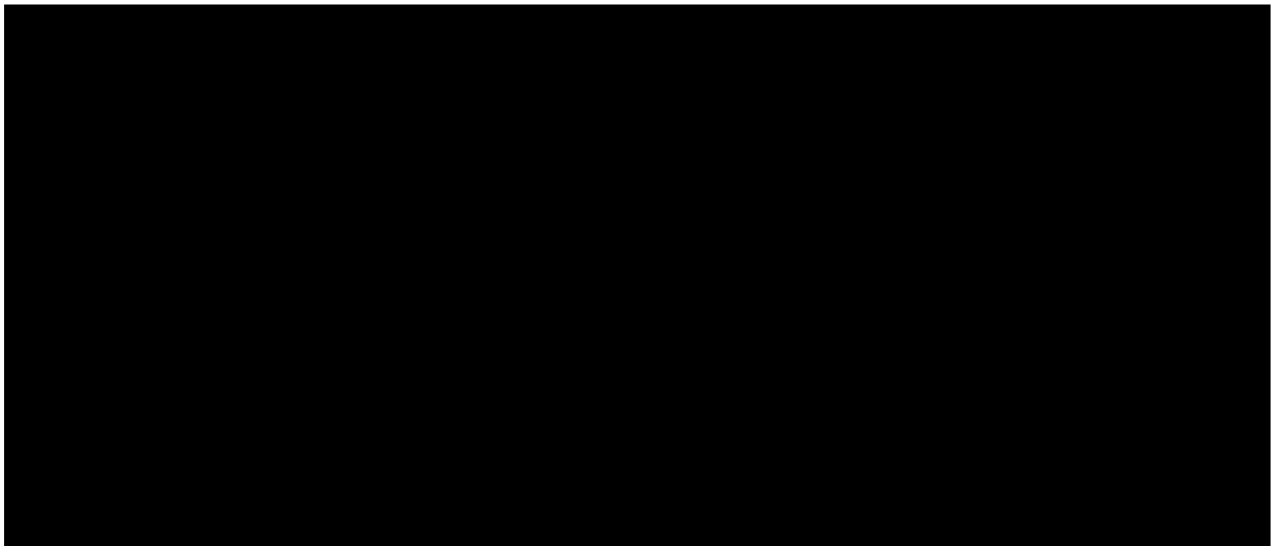
**Nursing**

in the

**GRADUATE DIVISION**

of the

**UNIVERSITY OF CALIFORNIA, SAN FRANCISCO**



Date

University Librarian

Degree Conferred:.....

### Acknowledgements

Thank you to Ginger Carrieri-Kohlman, RN, DNSc, Huong Nguyen RN, PhD and DorAnne Donesky Cuenco, RN, PhD for their guidance.

## Table of Contents

- **Chapter I: Introduction**
- **Chapter II: Review of Literature**
  - **Web versus Paper-and-Pencil**
  - **Influence of Social Desirability Distortion**
  - **Influence of Web Questionnaire Format**
  - **The CRQ**
  - **The SF-36**
- **Chapter III: Methodology**
  - **Study Design**
  - **Sample**
  - **Procedures**
  - **Description of the Web Questionnaires**
  - **Statistical Analysis**
- **Chapter IV: Results**
  - **Participants**
  - **SF-36**
  - **CRQ**
- **Chapter V: Discussion**
  - **Limitations**
  - **Conclusions**
- **References**
- **Tables**

## List of Tables

- **Table 1: Sample Baseline Characteristics**
- **Table 2: SF-36 Results: Mean Differences t-test, Pearson Test-Retest Correlation**
- **Table 3: CRQ Results: Mean Differences t-test, Cronbach's alpha, Pearson Test-Retest Correlation**

## Chapter 1: Introduction

The quality of health care for chronically ill patients is influenced by web-based education and management. This impact is growing rapidly as a global surge in internet use matches a global interest in leveraging the internet in healthcare. Estimates from 2007 indicate that 71% of the population within the USA, and 20% of the global population uses the internet (Miniwatts Marketing Group, 2007). These numbers represent a growth in internet use of 119% in the USA, and 265% globally in the last seven years (Miniwatts Marketing Group). Paralleling this growth is the prioritization of the development of e-health initiatives in the USA and internationally (European Commission, Information Society, 2007; Robert Wood Johnson Foundation, 2008; World Health Organization, 2005). Home telehealth, which includes medical and nursing management of patients via web, is one subset of e-health that is currently increasing. A meta-analysis of telehealth research found that telehealth has a moderate significant positive effect on health across all categories of disease and techniques tested (Dellifraire and Dansky, 2008). However, most of the studies are small, and the efficacy of these studies varies for different types of disease states (Dellifraire and Dansky). Funding and reimbursement for e-health require more conclusive data from large high quality trials that clearly show the benefit and best use of the internet in healthcare management (Agency for Healthcare Research and Quality, 2006; Nguyen, Carrieri-Kohlman, Rankin, Slaughter, Stulbarg, 2004).

Many questionnaires assessing health currently exist in well validated self administered paper-and-pencil format, and currently there is a need for converting these

paper-and-pencil versions to web versions to measure outcomes resulting from health care interventions. Web-based surveys can be beneficial in that they can reach respondents in a wide geographic area, are typically less costly to process than paper or telephone surveys, and provide data quickly in a readily analyzed form (Wyatt, 2000).

The effect of mode of administration is well recognized as a potential influence on data received from questionnaires. Bowling (2005) reviewed 73 papers pertaining to the effect of mode of administration on data quality. Multiple influences on data quality were documented that occur in questionnaire administration, and the difference in influences was concluded to be greatest between modes of administration, rather than within modes. In a discussion of paper versus web questionnaires, Buchanan (2003) states that an online adaptation of an offline test cannot be assumed to retain the psychometric properties of the offline test. Testing is thus necessary to assure integrity of a questionnaire being changed from paper to web format. When web surveys are used to evaluate e-health outcomes, it is important to provide evidence of instrument rigor, as the quality of the research lies in part on the quality of the measurement tools.

Therefore, the major goal of this secondary analysis was to test the equivalency of new web versions of the Short Form-36 (SF-36) and Chronic Respiratory Questionnaire (CRQ) questionnaires as compared to their paper-and-pencil versions. These questionnaires were used for a randomized clinical trial that was comparing an online dyspnea self-management program with an established face to face program in a sample of patients with COPD (Nguyen et al., 2008). Previous studies have been conducted on a web version of the SF-36 (Bell and Kahn, 1996; Soetkino, Mrad, Pao and Lenert, 1997) but additional testing was needed to fully evaluate its psychometric properties with the

format used in the parent study. No studies have been published regarding a web version of the CRQ.



## Chapter 2: Review of Literature

### *Web versus Paper-and-Pencil*

Several research studies have been conducted to compare newly developed web surveys with their paper-and-pencil counterparts. The absence of significant differences in scores between the paper and web versions is revealed in many of the studies. Buchanan and Smith (1999) compared paper (N=224) and web (N=963) versions of a personality assessment survey, the SMS-R, and found them to be similar in mean scores and internal reliability. Pettit (2002) compared web and paper versions of four personality questionnaires in two samples of volunteers (web N=2,649, paper N=458) recruited online and no difference in means and internal reliability were found. Ritter, Lorig, Laurent, and Matthews (2004) compared responses between internet and paper versions of 16 self-report questionnaires useful in evaluation of patient interventions in a randomized convenience sample recruited online (N=397). No significant difference in responses was found between the two versions. The web and paper versions of the Perceived Stress and Center for Epidemiological Studies – Depression Scales were compared by Herrero and Meneses (2006) in a sample of online university students (N=530) and showed similar means and internal consistency between versions. Spek, Nyklicek, Cuijpers, and Pop (2007) compared internal consistency and positive predictive values of an internet administered Edinburgh Depression Scale in a sample of patients with depression (N=407) with previously published results of the scale and found them to be comparable.

Other studies comparing web and paper-and-pencil questionnaires reveal significant differences, though most differences tend to be small. Raat, Mangunkusumo, Landgraf, Kloek and Brug (2007) compared scores of the paper and web version of the CHQ-CF (Child Health Questionnaire Child Form) in a randomized sample of adolescents (N=1071). Significantly higher scores were reported on 4 of 10 scales, but were considered negligible or small. Vallejo, Jordan, Diaz, Comech, and Ortega (2007) compared the results of an web and paper version of the General Health Questionnaire (GHQ) and the Symptom Checklist – 90 (SCL-90) in a sample of university students (N=100), and found that scores on the SCL-90 were significantly higher on paper than online. No significant differences in scores for the GHQ were found. Carlbring et al. (2007) compared scores from internet and paper versions of seven questionnaires used in panic/agoraphobia research. A significant difference in means was noted in five questionnaires. These differences were seen to be clinically significant, but the authors recommended using separate norms for different versions of the questionnaires. Fouladi, McCarthy, and Moller (2002) compared the online (N=234) and paper (N=164) versions of three instruments measuring emotional functioning and attachment in a sample of university students, and some differences were found in means and variance. No notable pattern with differences in means was noted between versions, but difference in variance in the paper version was greater than online.

Investigators have studied the effect the online environment on human behavior when completing questionnaires in order to explain response differences between paper-and-pencil and web versions. In particular the effects of questionnaire design on

behavior and the online environment on behavior have emerged as important factors to consider when implementing and assessing a web questionnaire.

### *Influence of Social Desirability Distortion*

The mode of administration of a questionnaire has long been known to affect the level of self-disclosure (Sudman and Bradburn, 1974). One term for this phenomenon is social desirability distortion, which is defined as "the tendency by respondents, under some conditions and modes of administration, to answer questions in a more socially desirable direction than they would under other conditions or modes of administration" (Richman, Kiesler, Weisband and Drasgow, 1999). When applied to completion of web-based questionnaires, this phenomenon has been called the candor hypothesis, which states that when a subject completes a questionnaire online, the environment leads to increased self disclosure and a decreased effect of social desirability (Joinson, 1999).

The candor hypothesis was supported in a study comparing scores of a web and paper-and-pencil questionnaire evaluating self-consciousness, self-esteem, and social desirability (Joinson, 1999). Lower levels of social desirability and social anxiety were reported for the web questionnaire (Joinson). However, another study evaluating the effect of web and paper-and-pencil administration on scores of three social desirability questionnaires found no significant difference in mean scores (Risko, Ouilty and Oakman, 2006). The effect of social desirability was concluded to be so small such that a very large sample size would be needed to detect a significant difference in scores between modes of administration (Risko et al.). Therefore, it remains unclear whether social desirability distortion varies between paper-and-pencil and web-based modes of questionnaires.

A meta-analysis evaluating the effect of social desirability distortion in 49 studies from 1969-1997 assessing computer and paper-and-pencil administered questionnaires identified factors other than mode of administration that can affect social desirability distortion (Richman, Kiesler, Weisband and Drasgow, 1999). Although this study evaluated computer based questionnaires, not web-based, the findings may be extrapolated to web-based questionnaires due to similarity in format. Results showed that the effect of mode of administration (computer vs. paper-and-pencil) on social desirability distortion was insignificant (Richman et al.). Additional moderators of social desirability were analyzed independently. Significantly less social desirability distortion was noted in computer versus paper-and-pencil questionnaires when respondents were alone, anonymous, and able to backtrack and edit answers (Richman et al.). The effect of social desirability distortion was also less if the publishing date was more recent, suggesting that better computers and software allow for a closer match in format to paper-and-pencil questionnaires, thus producing more similar results (Richman et al.).

Therefore, it has been concluded that the effect of mode, when comparing web-based and paper-and-pencil surveys is weak. The differences in mode itself may not be strong enough to elicit major differences in response, as has been noted earlier (Risko, Ouilty and Oakman, 2006). The other identified factors influencing response, such as anonymity, being alone, and ability to backtrack, may be more important than mode itself, and thus should be considered when administering questionnaires.

#### *Influence of Web Questionnaire Format*

Developing a questionnaire in a web format provides for a multitude of format and design options that are not possible in a paper-and-pencil format. Slight differences

in web format have been shown to influence responses and data quality obtained from questionnaires. A study analyzed the effect of multiple vs. single question screens, radio vs. entry boxes, and long vs. short entry boxes on responses of university students (N=665) to a web questionnaire (Couper, Traugott and Lamias, 2001). Results showed faster completion time and less missing data for multiple item screens vs. single item screens, less missing data for radio buttons vs. entry boxes, and different answer submissions for long vs. short entry boxes (Couper et al.). Another study compared responses of a plain and fancy version of a web questionnaire in a sample of computer users (plain N=9,522, fancy N=2,466) (Dillman, Tortora and Bowker, 1998). The results showed that the fancy web questionnaire that used sophisticated graphics and different background color highlights resulted in lower response and completion rates as well as increased time to complete as compared to the plain version that had no graphics and was black print on white background (Dillman et al.). These results suggest that careful attention needs to be paid to the design details of a web questionnaire.

Guidelines for the development of a web questionnaire advise simplicity. The design of a web questionnaire should be developed to create a user-friendly questionnaire that reduces non-response and measurement error (Dillman, Tortora and Bowker, 1998), otherwise stated as encouraging responses and eliciting accurate and complete responses. Web surveys should focus on articulating questions and directions clearly, keeping advanced graphics to a minimum, formatting questions appropriately for the questionnaire mode, and using design elements with consistency (Dillman and Smyth, 2007). Richman, Kiesler, Weisband and Drasgow (1999) state that, with regards to the effect of the computer format on social desirability distortion, "the more a computer

instrument resembles a traditional instrument, the more the two instruments should produce similar results”.

In summary, the web questionnaire design should be simplistic, consistent, and mimic the format of the paper-and-pencil questionnaire in order to elicit quality data.

### *Questionnaires*

#### *The CRQ*

The CRQ consists of twenty questions that initially were developed and tested as an interview in 1987 as a measure of quality of life for patients with chronic lung disease (Guyatt, Berman, Townsend, Pugsley and Chambers, 1987). The dimensions of quality that are measured include: dyspnea, fatigue, emotional function, and mastery. Responses are scored on a seven point Likert scale. Scores within each domain are added together for a total score, with lower scores indicating a lower quality of life. The questions that measure the dyspnea dimension are individualized in that the patient evaluates dyspnea for five activities identified specifically by that patient. Testing of the CRQ showed that it has appropriate construct validity (Guyatt et al., Wijsktra et al., 1994), sensitivity (Guyatt et al., Harper et al., 1997), reproducibility (Guyatt et al.), and internal reliability (Wijsktra et al., 1994, Harper et al.).

A self administered paper-and-pencil CRQ questionnaire has since been developed to decrease time and cost of administration (Williams, Singh, Sewell, Guyatt and Morgan, 2001). In testing that compared the paper-and-pencil with the interviewer administered versions, no differences were found for mean scores of fatigue and mastery ( $p>0.05$ ). However, there were small but significant differences between the dimensions of emotional function and dyspnea mean scores ( $p<0.05$ ), with higher scores recorded for

both dimensions with the interviewer administered CRQ (Williams et al.). However, these differences were below the described minimum clinically important difference of 0.5 (Redelmeier, Guyatt, and Goldstein, 1996). The paper-and-pencil version also exhibited strong test-retest reliability in the short and long term (Williams et al.). The authors concluded that the paper-and-pencil version could be used in place of the interview version. To date no web versions of the CRQ have been tested to assure quality of the questionnaire in the web mode.

### *The SF-36*

The SF-36 was developed to measure health related quality of life for the general population (Ware and Sherbourne, 1992). The questionnaire evaluates eight health concepts with 36 questions: 1) physical functioning, 2) role limitations because of physical health problems, 3) bodily pain, 4) social functioning, 5) general mental health, 6) role limitations because of emotional problems, 7) vitality, and 8) general health perceptions (Ware and Sherbourne). Responses include yes/no and Likert scales. Each domain is scored separately and transformed to a 0-100 score, with higher scores indicating good health (McHorney, Ware and Raczek, 1993). Typically the SF-36 is self-administered, but can be administered via interview or over the telephone. The SF-36 has been tested to show appropriate construct validity (McHorney et al., 1993), sensitivity (McHorney et al., 1993), and reliability (McHorney, Ware, Lu and Sherbourne, 1994) amongst a diverse patient group (McHorney et al., 1993). The questionnaire has also been shown to be a valid measure of health related quality of life in COPD patients (Mahler and Mackowiak, 1995).

Web and electronic versions of the SF-36 have been created and tested. A web SF-36 was made available anonymously on the internet, and completed by respondents (N=4876) (Bell and Kahn, 1996). The questionnaire was evaluated by comparing the internal reliability from anonymous, web-based results with that obtained from the Medical Outcomes Study (McHorney, Ware and Raczek, 1993; Ware and Sherbourne, 1992). High reliability ( $\alpha=.76-.90$ ) (Bell and Kahn) was found that matched the findings of the Medical Outcomes Study ( $\alpha=.78-.93$ ) (McHorney et al.; Ware and Sherbourne). Another web SF-36 was developed to assess patients with ulcerative colitis (Soetkino, Mrad, Pao and Lenert, 1997). The psychometrics of the web SF-36 were evaluated via internal validity, which was calculated by examining the partial correlations amongst subscales of the web SF-36. Partial correlations ranged from  $r=0.29-0.69$ . This assessment provides little information regarding the validity of the web SF-36, as previous validity testing of the SF-36 showed that it is not expected that the different dimensions measuring physical and mental components will be correlated strongly with each other (McHorney, Ware and Raczek, 1993). Results between an electronic and paper version of the SF-36 in a sample of healthy and chronic pain persons (N=127) were compared (Ryan, Corry, Attewell, and Smithson, 2002). Results showed a significant higher score of social function on the paper version, but the difference was small (<4%).

This study will compare paper-and-pencil and web versions of the SF-36 in the same group of patients, which none of the aforementioned studies did, and provide further details of the assessment of the properties of the web SF-36.



## Chapter 3: Methods

### *Study Design*

A within groups test-retest observational study design was used to compare the paper-and-pencil and the new web versions of the SF-36 and CRQ. Data for the study was collected as part of a randomized repeated measures (0, 3, and 6 months) trial to compare an internet based dyspnea self-management program with a face-to-face program for patients with COPD (Nguyen et al., 2008). The primary study took place at two academic medical centers: the University of California, San Francisco (UCSF), and the University of Washington, Seattle (UW). Approval for the study was given by the institutional review boards from both sites. Data for this secondary analysis was collected from participants at the UW site.

### *Sample*

Participants were recruited from the surrounding area of UW. Recruitment was achieved via online and off-line sources.

The sample inclusion criteria were 1) diagnosis of COPD and clinically stable for at least one month; 2) spirometry results showing at least mild obstructive disease defined as post-bronchodilator forced expiratory volume in one second (FEV1)/forced vital capacity (FVC) ratio <0.70 with FEV1 <80% predicted or FEV1/FVC <0.60 FEV1 >80% predicted; 3) activities of daily living (ADL) limited by dyspnea; 4) actively using a computer and internet with a Windows operating system; 5) oxygen saturation >85% on room air on  $\leq 6$ L/min of nasal oxygen at the end of a six minute walk test. Exclusion criteria were 1) any active symptomatic illness (e.g. cancer, heart failure, ischemic heart

disease with known coronary artery disease with known coronary artery or valvular heart disease, psychiatric illness, and neuromuscular disease; 2) participation in a pulmonary rehabilitation program in the last 12 months; or 3) current participation in >2 days of supervised maintenance exercise.

### *Procedures*

As part of the parent study, participants underwent baseline testing before being randomized to either the internet or face-to-face self management dyspnea program. Baseline assessments included spirometry, six-minute walk tests, and completion of the newly developed web-based SF-36 and CRQ questionnaires. The paper-and-pencil SF-36 and CRQ questionnaires were provided when the participants returned to the UW research lab within a week of their baseline assessment for a two hour exercise consultation with an advanced practice nurse. The participants were instructed to complete the paper-and-pencil questionnaires at home that same day and return them in the mail. Participants were left alone during web questionnaire completion in the research lab, and considered to be alone when at home completing the paper-and-pencil questionnaires. A research staff member was nearby to answer questions for participants in the lab, and available by phone for participants at home. Questionnaires were completed anonymously. Data from the web questionnaires were automatically entered into a database. For the paper-and-pencil versions, data was entered into the database by a research assistant.

### *Description of the Web Questionnaires*

Participants accessed the web surveys by logging on to the study website. For this elderly chronically ill sample, the clearest designs were used to develop the web

questionnaires. Black font was used with a white background. Questions were grouped in variable numbers on several screens, and participants could scroll back and forth between screens. Identical wording and items were used for the web questionnaires as is used in the paper-and-pencil versions. Participants selected answers by clicking on radio buttons associated with the chosen answer. If participants did not complete a question, a prompt was given to complete the question, or press a continue button to leave the question incomplete. On completion of the questionnaire, participants were prompted again to complete incomplete questions. No additional graphics were included in the survey format.

#### *Statistical Analysis*

Statistical analysis was carried out with SPSS version 15.0. Concurrent validity for the SF-36 and CRQ web and paper questionnaire was tested using paired t-test to compare means and Pearson's correlation to measure the strength of association between responses. Internal reliability for the CRQ was tested by calculating a Cronbach's alpha.

## Chapter 4: Results

### *Participants*

Twenty-one participants took both the paper and internet-based surveys. Table 1 shows a comparison between characteristics of three groups: UW participants who took both the internet and paper tests, UW participants who took only the internet test, and UCSF participants. No significant differences in characteristics were identified, except for distance traveled to clinical site for UW participants who took only one test and UW participants who took both tests, and living situation between UCSF participants and UW participants taking both tests.

### *SF-36*

Results for the SF-36 are displayed in Table 2. No significant differences in means between the paper-and-pencil and web version of the SF-36 were found.

There were consistent strong correlations ( $r > .70$ ) for most domains between the paper-and-pencil and web versions. The correlation for the role physical domain between the versions was a moderate to strong correlations ( $r = 0.68$ ), and the domain of role emotional had a weak correlation ( $r = .37$ ,  $p > 0.05$ ).

### *CRQ*

Results for the CRQ are displayed in Table 3. There were no significant differences in the means for all domains between the paper-and-pencil and web versions.

The web and paper-and-pencil CRQ had strong correlations ( $r > .70$ ) for the fatigue, emotional function, and mastery domains. For the dyspnea domain, the correlation was moderate ( $r = 0.62$ ). Internal reliability was strong ( $\alpha > .80$ ) for the

domains of dyspnea, fatigue, and emotion for the paper-and-pencil version, and moderate ( $\alpha=.67$ ) for the mastery domain. Internal reliability was strong ( $\alpha>.80$ ) for the dyspnea, fatigue, and mastery domains, and moderate to strong for the emotional domain ( $\alpha=0.78$ ).

## Chapter 5: Discussion

The results of this secondary analysis demonstrate that the new web versions of the SF-36 and CRQ elicit no significant differences in responses as compared to their paper-and-pencil versions when tested in a small sample of elderly patients with COPD, and can be used when online testing is preferred for clinical or research purposes. The following is a discussion of factors that are likely implicated in these results.

The equivalency of the questionnaires may be attributed to the attention that was given to controlling format and environment. The format of the web questionnaire mimicked that of the paper-and-pencil version, and controlled for differences in format that might influence responses. The effect of social desirability distortion was controlled for in that participants were alone, could backtrack, and completed questionnaires anonymously. With these factors controlled for, the study showed that the effect of mode itself on responses was absent.

Correlations between the corresponding domains for the paper-and-pencil and web versions for the most part were strong. For the SF-36, the outstanding exception was the low correlation for the role emotional domain ( $r=.37$ ). The role emotional domain measures the concept of role limitations due to emotional problems (Ware and Sherbourne, 1992). A review of scoring in the emotional domain reveals large variability for some respondents from one version to another. The reason behind this variability is unclear, though the timing between completion of questionnaires may have confounded results, together with variations in emotions and coping mechanisms that may have occurred in this time period. Though participants were asked to complete and mail back

the paper-and-pencil version on the same day as the web version, there was not strict compliance by all with this request, and some participants may have not completed the paper-and-pencil version on the same day. Emotion and coping may inherently be labile with some people and may have varied between versions, especially if a number of days passed between completing versions.

For the CRQ, only a moderate correlation was seen for the dyspnea score.

However, previous test-retest reliability testing of the CRQ paper-and-pencil questionnaire found the dyspnea dimension to be less reliable than the others (Wijcktra et al., 1994). The dyspnea domain is individualized, as respondents choose from a list of activities to rate dyspnea. Therefore, the domain is best to track changes in one individual, but is difficult to use when comparing results amongst many. A standardized version of the dyspnea domain has now been created for the CRQ (Shunemann et al., 2003), which has been shown to have improved discriminative properties, but reduced responsiveness. For the purpose of comparing scores as a group, the standardized version may have been more appropriate for testing of modes, than the individualized version.

Internal reliability for the CRQ was similar for the paper-and-pencil and web versions with the web version demonstrating overall better internal reliability measures across all domains. The measured internal reliability for both versions is similar to if not better than previous reports of internal reliability for the paper-and-pencil version ( $\alpha=.51-.88$ ) (Wijcktra et al., 1994).

### *Limitations*

Threats to the validity of the study exist. The sample size was very small, and may not have been large enough to detect the effect of mode on the questionnaire results.

Also, this study was confined to COPD patients who were avid users of the computer and internet, and who were primarily Caucasian. Different results may be obtained from respondents with less computer and internet skills, with a different disease, or of differing ethnicity. The environment in which the questionnaires was controlled with regards to being alone, anonymous, and having a staff as a resource. If any changes were to be made with regards to these factors in the administration of the test, results may be influenced.

The effect of format between versions was minimized by following guidelines to develop a simple, easy to complete questionnaire design. Any change to the design used for the survey in this study may affect results obtained from the questionnaire. Moving forward, researchers should continue to be careful of effect of format and consider standardizing format of web surveys so that results between studies can be compared and contrasted.

### *Conclusions*

Web and paper-and-pencil versions of the SF-36 and CRQ were similar with respect to mean scores and correlation in a small group of patients with COPD enrolled in a study to evaluate an internet based self-management dyspnea program. These results support the utilization of these questionnaires in larger scale research to evaluate e-health interventions. When developing a web version of a questionnaire, careful attention should be paid to formatting and the environment of administration. The development of standardized web versions of widely used questionnaires, such as the SF-36 and CRQ should be considered so that comparisons between studies can be more accurate.



## References

- Agency for Healthcare Research and Quality. (2006, February). *Telemedicine for the medicare population – update*. Retrieved May 7, 2008 from <http://www.ahrq.gov/clinic/tp/telemeduftp.htm>
- Bell, D.S. & Kahn, C.E. (1996). Health status assessment via the World Wide Web. *J Am Inform Assoc; Symposium Supplement*, 338-42.
- Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, 27(3), 281-291.
- Buchanan, T. (2003). Internet-based questionnaire assessment: appropriate use in clinical contexts. *Cognitive Behaviour Therapy*, 32(3), 100-109.
- Buchanan, T. & Smith, J. (1999). Using the internet for psychological research: personality testing on the world wide web. *British Journal of Psychology*, 90, 125-144.
- Carlbring, P., Brunt, S., Bohman, S., Austin, D., Richards, J., Ost, L., & Andersson, G. (2005). Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior*, 23, 1421-1434.
- Couper, M., Traugott, M. & Lamias, M. (2001). Web survey design and administration. *Public Opinion Quarterly*, 65, 230-253.
- Dellifraigne, J. & Dansky, K. (2008). Home-based telehealth: a review an meta-analysis. *Journal of Telemedicine and Telecare*, 14(2), 62-66.
- Dillman, D. & Smyth, J. (2007). Design effects in the transition to web-based surveys. *Am J Prev Med*, 32, S90-S96.

- Dillman, D., Tortora, R., & Bowker, D. (1998). Influence of plain vs. fancy design on response rates for web surveys. Retrieved May 7, 2008, from the University of Washington Web site: <http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf>
- European Commission, Information Society. (2007, March). *eHealth ERA report – March 2007: Towards the establishment of a European eHealth research area*. Retrieved May 7, 2008, from [http://ec.europa.eu/information\\_society/newsroom/cf/itemdetail.cfm?item\\_id=334](http://ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=334)
- 6
- Fouladi, R., McCarthy, C. & Moller, N. (2002). Paper-and-pencil or online? Evaluating mode effects on measures of emotional functioning and attachment. *Assessment*, 9(2), 204-215.
- Guyatt, G., Berman, L., Townsend, M., Pugsley, S., & Chambers, L. (1987). A measure of quality of life for clinical trials in chronic lung disease. *Thorax*, 42, 773-778.
- Harper, R., Brazier, J.E., Waterhouse, J.C., Walters, S.J., Jones, N.M., & Howard, P. (1997). Comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax*, 52, 879-887.
- Herrero, J. & Meneses, J. (2006). Short web-based versions of the perceived stress (PSS) and center for epidemiological studies-depression (CESD) scales: A comparison to pencil and paper responses among internet users. *Computers in Human Behavior*, 22, 830-846.
- Joinson, A. (1999). Social desirability, anonymity, and internet-based questionnaires. *Behavior Research Methods, Instruments, and Computers*, 31(3), 433-438.

- Mahler, D. & Mackowiak, J. (1995). Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with COPD. *Chest*, 107(6), 1585-1589.
- McHorney, C., Ware, J., Lu, J. & Sherbourne, C. (1994). The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, 32(1), 40-66.
- McHorney, C., Ware, J. & Raczek, A. (1993). The MOS 36-item short-form health survey (sf-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care*, 31(3), 247-263.
- Miniwatts Marketing Group. (November 2007). *Internet usage and population in North America*. Retrieved March 24, 2008 from <http://internetworldstats.com/stats14.htm>
- Nguyen, H., Carrieri-Kohlman V., Rankin, S., Slaughter, R. & Stulbarg, M. (2004). Internet-based patient education and support interventions: a review of evaluation studies and directions for future research. *Computers in Biology and Medicine*, 34, 95-112.
- Nguyen, H., Donesky-Cuenca, D., Wolphin, S., Reinke, L.F., Benditt, J.O., Paul, S.M., et al. (2008). Randomized controlled trial of an internet-based versus face-to-face dyspnea self-management program for patients with chronic obstructive pulmonary disease: a pilot study. *Journal of Medical Internet Research*, 10(2), e9.

- Raat, H., Mangunkusumo, R., Landgraf, J., Kloek, G. & Brug, J. (2006). Feasibility, reliability, and validity of adolescent health status measurement by the Child Health Questionnaire Child Form (CHQ-CF): internet administration compared with the standard paper version. *Quality of Life Research, 16*, 675-685.
- Redelmeier, D.A., Guyatt, G.H. & Goldstein, R.S. (1996). Assessing the minimum important difference in symptoms: a comparison of two techniques. *Journal of Clinical Epidemiology, 49*, 1215-1219.
- Richman, W., Kiesler, S., Weisband, S. & Drasgow, F. (1999). A meta-analysis study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84*(5), 754-775.
- Risko, E., Quilty, L. & Oakman, J. (2006). Socially desirable responding on the web: investigating the candor hypothesis. *Journal of Personality Assessment, 87*(3), 269-276.
- Ritter, P., Lorig, K., Laurent, D. & Matthews, K. (2004). Internet versus mailed questionnaires: a randomized comparison. *Journal of Internet Medical Research, 6*(3), e29.
- Robert Wood Johnson Foundation. (2008). *Health e-technologies: building the science of ehealth*. Retrieved on May 7, 2008 from [http://www.rwjf.org/applications/solicited/npo.jsp?FUND\\_ID=55054](http://www.rwjf.org/applications/solicited/npo.jsp?FUND_ID=55054).
- Ryan, J., Corry, J., Attewell, R. & Smithson, M. (2002). A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. *Quality of Life Research, 11*, 19-26.

- Shunemann, H., Griffith, L., Jaeschke, R., Goldstein, R., Stubbing, D., Austin, P., et al. (2003). A comparison of the original chronic respiratory questionnaire with a standardized version. *Chest, 124(4)*, 1421-1429.
- Soetkino, R., Mrad, R., Pao, V. & Lenert, L. (1997). Quality of life research on the internet: feasibility and potential biases in patients with ulcerative colitis. *Journal of the American Medical Informatics Association, 4*, 426-435.
- Spek, V., Nyklicek, I., Cuijpers, P. & Pop, V. (2007). Internet administration of the Edinburgh Depression Scale. *Journal of Affective Disorders*, doi: 10.1016/j.jad.2007.07.003.
- Sudman, S. & Bradburn, N. (1974). Response effects in surveys. Chicago: ALDINE Publishing Company.
- Vallejo, M., Jordan, C., Comeche, M. & Ortega, J. (2007). Psychological assessment via the internet: A reliability and validity study of online (vs. paper and pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptom Check-List-90-Revised (SCL-90-R). *Journal of Medical Internet Research, 9(1)*, e2.
- Ware, J. & Sherbourne, C. (1992). The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Medical Care, 30(6)*, 473-483.
- Wijkstra, P., TenVergert, E., Van Altena, R., Otten, V., Postma, D., Kraan, J., et al. (1994). Reliability and validity of the Chronic Respiratory Questionnaire (CRQ). *Thorax, 49*, 465-467.
- Williams, J., Singh, S., Sewell, L., Guyatt, G. & Morgan, L. (2001). Development of a self-reported Chronic Respiratory Questionnaire (CRQ-SR). *Thorax, 56*, 954-959.

World Health Organization. (2005, May). *World health assembly resolution on ehealth*

(*WHA58.28*). Retrieved May 7, 2008 from

[http://www.euro.who.int/telemed/20060713\\_1](http://www.euro.who.int/telemed/20060713_1).

Wyatt, J. (2000). When to use web-based surveys. *Journal of the American Medical*

*Informatics Association*, 7(4), 426-430.

## Tables

Table 1

*Sample Baseline Characteristics*

	UW took both tests (N=21)	UW dropouts (N=14)	UCSF sample (N=15)
Age, years (mean $\pm$ SD)	69.9 $\pm$ 8.6	65.9 $\pm$ 8.9	70.6 $\pm$ 8.6
Female	10 (48%)	6 (43%)	9 (60%)
Caucasian	21 (100%)	13 (93%)	15 (100%)
Education			
High school or some college	8 (38%)	10 (71%)	5 (33%)
College or more	12 (62%)	4 (29%)	10 (67%)
Not currently employed or currently disabled or retired	16 (76%)	8 (57%)	10(67%)
Living situation: with spouse or other	15 (71 %)	11 (79%)	5 (33%) ++
Currently smoking	3 (14.3%)	2 (14.3%)	1 (6.7%)
Distance to clinical site, km (mean $\pm$ SD)	15.8 $\pm$ 11.8	34.4 $\pm$ 36.0 **	27.9 $\pm$ 26.9
BMI, kg/m <sup>2</sup> (mean $\pm$ SD)	27.9 $\pm$ 4.7	28.7 $\pm$ 8.2	27.4 $\pm$ 5.0
FEV1/FVC (mean $\pm$ SD)	0.47 $\pm$ 0.13	0.49 $\pm$ 0.16	0.46 $\pm$ 0.09
FEV1 % predicted (mean $\pm$ SD)	50.5 $\pm$ 19.2	48.0 $\pm$ 19.4	52.3 $\pm$ 11.7
Supplemental oxygen	7 (33%)	3 (21%)	3 (20%)
Comorbidities			
Cardiovascular (HTN and CAD)	13 (62%)	4 (29%)	7 (47%)

	UW took both tests (N=21)	UW dropouts (N=14)	UCSF sample (N=15)
Musculoskeletal (arthritis and other pain)	5 (24%)	5 (36%) check this	2 (13%)
Previous pulmonary rehabilitation	8 (38%)	4 (29%)	6 (40%)
Self-rated computer skills			
Beginner	4 (19%)	1 (7%)	3 (20%)
Intermediate	11 (52.4%)	4 (29%)	11 (73%)
Advanced	6 (29%)	9 (64%)	1 (6.7%)
Self-rated internet skills			
Beginner	5 (24%)	1 (7.1%)	3 (20%)
Intermediate	9 (43%)	4 (29%)	10 (67%)
Advanced	7 (33%)	9 (64%)	2 (13%)

\*\*p<.05 (UW both tests vs. UW internet only)

++ $\Sigma=5.143$ , p<.05 (UW both tests vs. UCSF)



Table 2

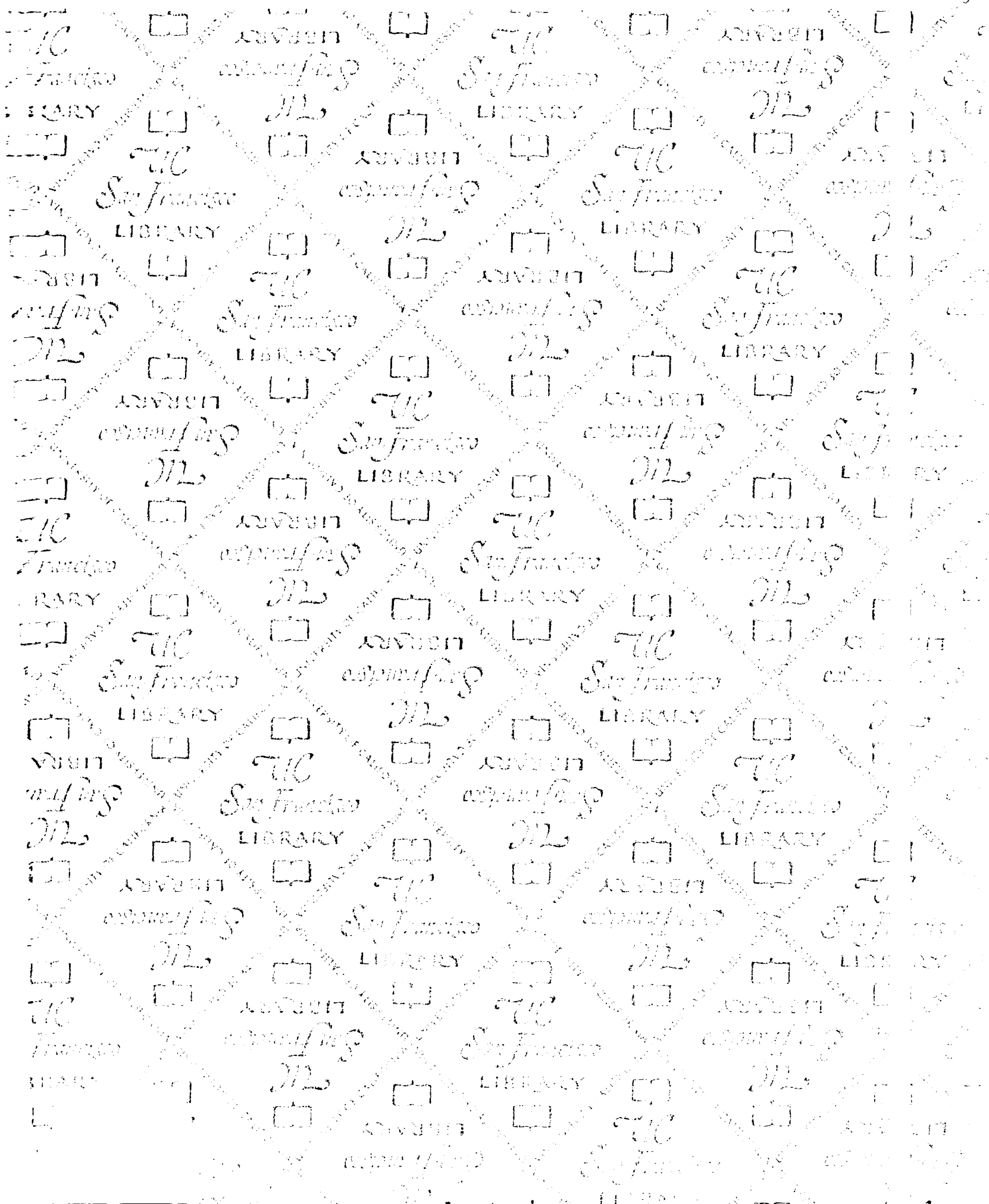
*SF-36 Results: Mean Differences t-test, Pearson Test-Retest Correlation*

	<b>Paper-and-pencil</b>	<b>Online</b>		
<b>SF-36</b>	<b>Mean (SD)</b>	<b>Mean (SD)</b>	<b>t (P value)</b>	<b>r (P value)</b>
Physical function	47.0 (22.8)	47.8 (23.6)	-.370 (.716)	.92 (.000)
Role physical	16.7 (26.6)	25.0 (32.6)	-1.581 (.130)	.68 (.001)
Bodily pain	71.0 (25.8)	69.4 (26.0)	.764 (.454)	.93 (.000)
General health perceptions	40.0 (20.3)	42.2 (18.4)	-1.048 (.307)	.89 (.000)
Vitality	41.0 (26.6)	44.3 (22.7)	-1.362 (.188)	.91 (.000)
Social functioning	71.4 (28.3)	75.6 (20.3)	-.960 (.349)	.71 (.000)
Role emotional	54.0 (41.5)	58.7 (36.4)	-.498 (.624)	.37 (.097)
Mental health	77.3 (15.2)	74.1 (19.5)	1.406 (.175)	.84 (.000)
Reported change	52.4 (23.6)	51.2 (23.0)	.439 (.666)	.86 (.000)
Standardized physical component	32.9 (7.5)	33.7 (8.2)	-.717 (.482)	.80 (.000)
Standardized mental component	49.1 (11.1)	49.6 (10.4)	-.258 (.799)	.70(.001)

Table 3

*CRQ Results: Mean Differences t-test, Cronbach's alpha, Pearson Test-Retest Correlation*

CRQ	Paper-and-pencil		Online		t (P value)	r (P value)
	Mean (SD)	$\alpha$	Mean (SD)	$\alpha$		
Dyspnea	15.1 (5.5)	.89	15.7 (5.6)	.84	-.464 (.649)	.62 (.011)
Fatigue	15.0 (5.8)	.94	15.9 (4.5)	.87	-1.218 (.238)	.79 (.000)
Emotional function	36.2 (7.0)	.86	35.7 (6.0)	.78	.539 (.596)	.82 (.000)
Mastery	20.0 (4.4)	.67	20.1 (4.6)	.86	-1.515 (.145)	.82 (.000)





3 1378 00754 2478

LIBRARY  
USE ONLY

7542478

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

San Francisco  
LIBRARY

