Detailed Description of Development of CHIEF

Using multiple methodologies, pools of qualified persons were identified to participate in four separate advisory panels. Methods for identifying and selecting participants assured that a wide and varied range of abilities, disabilities, attitudes, philosophies, knowledge and skills were reflected in the panel meetings. Each group included a diverse array of 32 participants with expertise in the four areas of disability: mobility, self-care, learning, and communication limitations. Each individual brought his or her personal and professional perspectives and experiences on disability, participation, and the impact of the environment. Each group consisted of professors, researchers and academicians representing the fields of sociology, occupational therapy, economics, public health and philosophy. Universities represented included Boston University, Rutgers University, the University of Denver, Queens University in Ontario, the State University of New York (Buffalo, Plattsburgh), University of Illinois at Chicago, University of California-Berkeley, Ohio State University, and the University of North Carolina.

There were representatives from such advocacy and policy implementation groups as the Institute on Disability and Human Development, the American Foundation for the Blind, the Paralyzed Veterans of America, and the Access Board; while the U.S. government had representatives from the Centers for Disease Control and Prevention and the National Center for Health Statistics. Consumer representatives included Native Americans and individuals with hearing and visual impairments, spinal cord injuries, speech impairments, and cerebral palsy, as well as family members of people with mental retardation and traumatic brain injury. Finally, service providers' input was provided by physicians, occupational and physical therapists, a former independent living center director, a director of a university's disabled student services program, and a vocational rehabilitation counselor. These meetings were very productive resulting in four draft instruments, one from each group. Each draft instrument was designed to be used in a telephone or 'paper pencil' survey that would be appropriate for general population use, as well as applying to the full range of disability categories.

After reviewing and critically assessing the four instruments, project staff decided the best instrument would come from synthesizing the vital elements, conceptualizations, and spirit of each draft into a fifth or "next generation" survey. Advisory panel members continued to be involved, and to participate via mail. Project staff applied advisory group comments and advice to the development of the draft instrument that identified 25 key elements of the environment.

Two forms of the draft instrument were proposed. Both had the same item content, but two different metrics were used to assess environmental impact. In one form, individuals were asked to indicate "how often" a barrier is encountered using response categories of "daily, weekly, monthly, less than monthly, and never". In the other form, individuals were asked to assess the degree to which the environmental element "facilitates or hinders" participation using response categories of "big barrier, little barrier, no impact, little help, and big help." The first form had the advantage of easier response categories, while the second form had the advantage of identifying facilitators as well as barriers.

Initial Pilot Testing

Both forms were tested on a group of 97 people, 50 with disabilities and 47 who indicated they did not have a disability. Results of that pilot testing indicated:

1. The "frequency" response categories were strongly preferred by participants over the "extent of barrier/facilitator" response categories.
2. The "frequency of barrier" response categories better differentiated people with and without disabilities than the "extent of barrier/facilitator" response categories.
3. Correlations between the two response categories, while significant, were relatively low.

Discussions of the results from the comparison of response categories with project staff and representatives who attended advisory panel meetings, yielded a consensus that all 25 items should be retained in the draft instrument, but that a follow-up impact question needed to be added since the correlation between frequency and impact was not particularly high. These discussions led to adding a follow-up question, "When this problem occurs, is it usually a big problem or a little problem?" This question was added after each item where the respondent indicated the frequency of the problem to be anything other than never. The final draft instrument, the Craig Hospital Inventory of Environmental Factors (CHIEF) was distributed to all advisory panel members for review.

Testing of the Environmental Instrument

Instrument Validation - "CHIEF 400 Dataset"

A convenience sample of 409 individuals with disability was recruited for a validation study to test the psychometric properties of the CHIEF. The sample included available people with spinal cord injury and traumatic brain injury who had been treated at Craig Hospital (but not included in prior pilot tests of the instrument). The sample also included individuals recommended for recruitment by advisory panel members, professional colleagues, and acquaintances of other project staff and research participants. In total, the sample included 124 participants with spinal

cord injury, 120 participants with traumatic brain injury and 165 participants with a wide variety of other disabilities. This included 55 persons with multiple sclerosis, 35 persons with amputations, and others with auditory and visual impairments, developmental disabilities, cerebral palsy and some with multiple impairments resulting in disability. While the group of individuals with spinal cord injury was 80% male with an average age of 41 and the group of individuals with traumatic brain injury was 61% male with an average age of 41, the group of individuals with a variety of other impairments were 62% female with an average age of 48.

All 409 study participants were administered the CHIEF. In addition, 103 of the total 409 participants (46 with spinal cord injury, 44 with traumatic brain injury, and 13 with other impairments) were interviewed using CHIEF a second time, approximately two weeks after the first administration in order to assess test-retest reliability. Finally, family members or friends of 125 subjects (46 with spinal cord injury, 54 with traumatic brain injury, and 25 with other

impairments), not included in the test-reliability sub-study, were successfully recruited and asked to complete the CHIEF as a proxy for the subject in order to assess subject-proxy agreement.

This completed dataset is referred to henceforth as the *"CHIEF 400 Dataset"*. Analysis of this data began by defining three methods of scoring each item:

1. A frequency score on a scale of 0-4 indicating the frequency with which barriers were encountered (0=never, 1=less than monthly, 2=monthly, 3=weekly, and 4=daily).
2. A magnitude score on a scale of 0-2 indicating the size of the problem which a barrier typically presented (0=no problem since the barrier was never encountered, 1=a little problem, and 2=a big problem).
3. A frequency-magnitude product score on a scale of 0-8 calculated as the product of the frequency score and the magnitude score, indicating the overall impact of the barrier.

Total scores across the 25 items were calculated as the average frequency score, the average magnitude score, and the average frequency-magnitude product score across all of the non-missing scale items.

Test-Retest Reliability

Test-retest reliability of individual items and the total scale were calculated using both the intra-class correlation coefficient and the percent of cases with exact agreement between both tests. Mean difference scores between the test and retest were also calculated and significance assessed. This process was repeated for frequency scores, magnitude scores, and frequency-magnitude product scores. In general, the product scores showed slightly higher reliability co-efficient and they became the focus of additional psychometric analysis.

Table 1 presents all test-retest comparison data (separately for frequency and magnitude), while *Table 2* presents the test-retest reliabilities for the frequency magnitude product scores. These tables report item and total scale reliability scores. Data are presented separately for spinal cord injury, traumatic brain injury, and "other" impairment groups, as well as total sample reliability

statistics. These data indicate a total scale score ICC reliability of .926, indicating acceptable reliability for the instrument.

Subject-Proxy Agreement

After establishing test-retest reliability for CHIEF, the extent of subject-proxy agreement was analyzed. *Tables 3 and 4* present the results of frequency and magnitude comparisons and frequency magnitude product data respectively in a format identical to *Tables 1 and 2*. Across all disability sub-groups subject proxy interclass correlations ranged from 0.494 to 0.618 with a total scale ICC of 0.618. These data indicate that subject proxy agreement is marginal and result in the recommendation that proxies not be asked to complete CHIEF when subjects are unavailable to do so.

Discriminant Validity

As one method of validating the data collected in CHIEF, differences in response patterns were compared across impairment groupings in an effort to determine if the instrument differentiated among impairment groups in expected ways. *Tables 5, 6 and 7*, report percent frequency distributions of the raw data across the 25 items for spinal cord injury, traumatic brain injury, and other impairment groups respectively. *Table 8* presents the mean frequency-magnitude product scores for persons with spinal cord injury, traumatic brain injury, multiple sclerosis, amputees, and other impairments as well as the total sample mean. Tests of differences among the five groupings were compared using one-way analysis of variance with Bonferroni post hoc comparisons. Significant main effects and significant differences between groups are indicated in the table. It can be seen that the majority of items and sub-scales produce statistically significant differences among the impairment groups. Cases with traumatic brain injury scored dramatically lower on physical barriers than the other groups. These data lend support to the validity of CHIEF by indicating that the tool differentiates scores among different disability groups in ways that are consistent with the unique barriers faced by those groups.

Further Evaluation of the Environmental Instrument

Additional evaluation of the CHIEF was performed to: 1) examine the underlying dimensions that might exist within the context of the 25 items; 2) demonstrates its applicability to large-scale disability surveillance; 3) establish scoring norms; and 4) develop a CHIEF Short Form. This was accomplished by adding the CHIEF to the Behavioral Risk Factor Surveillance System (BRFSS) survey in Colorado. In 1999, a separate population-based sample was drawn and this sample was administered the: 1) BRFSS core survey; 2) BRFSS Quality of Life Module; 3) BRFSS State-added Disability Questions; 4) Craig Handicap Assessment and Reporting Technique Short Form (CHART-SF); and 5) CHIEF. The survey was administered via telephone to 2,259 individuals. This completed dataset will be henceforth referred to as the *"BRFSS Dataset"*.

The BRFSS data was weighted using the standard BRFSS weighting formula. In addition, post-stratification weighting has been applied to account for differences in age and gender between the sample and the population of Colorado. All analyses have been performed using the weighted data; therefore, the results can be generalized to the entire population of Colorado, 18 years or older.

Identification of CHIEF Subscales

Factor analysis was used to identify underlying dimensions, or subscales, within CHIEF. This analysis was performed on the 25 CHIEF items with five factors accounting for 48% of the cumulative variance across the 25 items. After varimax rotation, each item was assigned to the factor with the highest positive loading. This resulted in five factors with 3-7 items included in each factor. Descriptive labels for the factors were assigned including "attitude and support barriers", "services and assistance barriers", "physical and structural barriers", "policy barriers" and "work and school barriers".

The following indicates which items are contained in each subscale of the CHIEF:

Policies Subscale: Policies businesses, policies government, policies employment/education & services community.
Physical/Structural Subscale: Surroundings, natural environment, design home, design community, design work/school, & technology.
Work/School Subscale: Attitudes work/school, help work/school & support work/school.
Attitudes/Support Subscale: Attitudes home, discrimination, support community, attitudes community & support home.
Services/Assistance Subscale: Transportation, medical care, help home, information, education/training, help community & personal equipment.

Scoring Differentiation Between Groups

Across items, subscales and total scores, the CHIEF was able to show differences in reported frequency and magnitude of environmental barriers between groups with a variety of impairments and activity limitations. *Table 9* summarizes the mean and standard deviation for each CHIEF item, subscale and total score using the two datasets (CHIEF 400 and BRFSS) grouped by disability status.

Subjects within the BRFSS Dataset were differentiated by whether or not they had a "disability". This was determined by using a definition where a subject was considered "disabled" if they responded "yes" to any of the following questions: 1) Are you limited in the kind or amount of work you can do because of any impairment or health problem; 2) Because of any impairment or health problem, do you have any trouble learning, remembering or concentrating; 3) Do you use special equipment or help from others to get around; 4) Are you limited in any way in any activities because of any impairment or health problem. Within the "CHIEF 400 Dataset", subjects were differentiated by the same impairment categories as previously described.

Figures A through F provide a graphic summary of the information in *Table 9*. *Figure A* shows the CHIEF subscales and total scale mean scores by disability status. This figure indicates that

both people with and without disabilities experience environmental barriers. However, those with disabilities reported an overall higher frequency and magnitude of environmental barriers. Further, individuals with traumatic brain injury reported greater barriers than those identified as having a disability from the BRFSS data (see definition above), but fewer than individuals with spinal cord injury. Individuals with other types of impairments (i.e., multiple sclerosis, amputees, other auditory, visual and multiple impairments, developmental disabilities, cerebral palsy) reported the greatest barriers.

Figures B through F show the mean scores for each CHIEF subscale and the items on that subscale by disability status. Overall, the same general trend is seen, however some items and subscales do vary by disability status. This analysis confirms that the CHIEF has the ability to differentiate between those with and without disability and between different impairment groups.

Development of the CHIEF Short Form

Several criteria were used to determine which items should be retained for a "short form" version of the CHIEF. In general, these criteria included items which: 1) had the highest frequency of barrier mean scores; 2) had the highest magnitude of barrier mean scores; 3) had the highest item score-subscale score correlations (using the mean frequency-magnitude product score); 4) had the highest item score-total score correlations (using the mean frequency-magnitude product score); 5) were the most frequently reported barriers; and 6) best differentiated between people with and without disability. In addition, taking all of the criteria into consideration, if an item was to be excluded, but it was felt, conceptually should be in the scale, it was retained. The following 12 items within the original five subscales were retained:

Policies Subscale: Policies businesses & policies government
Physical/Structural Subscale: Surroundings & natural environment
Work/School Subscale: Attitudes work/school & help work/school
Attitudes/Support Subscale: Attitudes home & discrimination
Services/Assistance Subscale: Transportation, medical care, help home & information

Table 10 shows the mean and standard deviation for each CHIEF-SF item, subscale and total score using the two datasets (CHIEF 400 and BRFSS) grouped by disability status. *Figures G and H* provide a graphic summary of the information in *Table 10*. *Figure G* shows the CHIEF-SF subscales and total scale mean scores by disability status, and *Figure H* shows the total scale and item mean scores by disability status. These figures further substantiate the findings from the 25-item CHIEF.

For all referenced tables and figures, see CHIEF_Figs.pdf and SF_Figs.pdf.