

COMPUTER-ASSISTED MEASUREMENT OF WOUND SIZE
ASSOCIATED WITH SELF-INJURIOUS BEHAVIOR

DAVID M. WILSON

GEORGIAN COURT UNIVERSITY

BRIAN A. IWATA

UNIVERSITY OF FLORIDA

AND

SARAH E. BLOOM

UTAH STATE UNIVERSITY

We evaluated an inexpensive, efficient, and noninvasive technique for measuring tissue damage produced by self-injurious behavior (SIB). The technique involved computerized measurement of wound surface area (WSA) based on digital photographs. In Study 1, we compared photographic measurement to a more commonly used procedure, transparency measurement, in estimating WSA of 20 wound models. Results showed that both methods were reliable and that there was a high degree of correspondence between the 2 sets of measures. In Study 2, we compared photographic WSA measures to direct-observation measures in documenting changes over time in the SIB exhibited by a woman with Prader-Willi syndrome. Results showed that increases and decreases in observed SIB during baseline and treatment conditions corresponded with changes in WSA measures, indicating that the computer-assisted photographic technique may be useful as a corroborative measure or as a primary measure when direct observation of SIB is not feasible.

Key words: differential reinforcement, functional analysis, injury measurement, self-injurious behavior

Measurement of self-injurious behavior (SIB) typically is done through direct observation by counting the number of times that the behavior occurs. Occasionally, however, SIB occurs under conditions that are difficult to observe (when the individual is alone) or in settings where measurement is infrequent (outpatient clinics). In such cases, it may be helpful to estimate changes in SIB through examination of its products, that is, the wounds produced by the behavior.

Product measures of SIB are not preferred as dependent variables for several reasons. First,

This research was supported by a grant from the Florida Agency for Persons with Disabilities. We thank Tiffany Cook, Angela Irwin, Richard Jones II, Paul Niesen, Andreina Santealla, and Renee Saulnier for their help in conducting the study.

Address correspondence to Brian Iwata, Psychology Department, University of Florida, Gainesville, Florida 32611 (e-mail: iwata@ufl.edu).

doi: 10.1901/jaba.45-797

number of wounds may not correlate highly with response frequency because repeated injury to one area of the body may result in only a single wound. Second, the number of wounds may not reflect injury severity because one large wound may be much more serious than several smaller wounds. Third, if wound healing is delayed following decreases in the rate of SIB, conclusions about treatment efficacy based solely on observed physical improvement may be erroneous. Nevertheless, changes in the appearance of wounds may be useful as a dependent variable when direct observation is not feasible. In addition, although response characteristics such as frequency and duration of SIB are important as direct measures of behavior, they are indirect measures of the actual extent of risk posed by the behavior. Thus, product measures also may be helpful as an index of social validity in the evaluation of reductions in responding.

Wound characteristics, such as the number of injuries (St. Lawrence & Drabman, 1983) or the appearance of new injuries (Grace, Thompson, & Fisher, 1996), have been used occasionally as dependent variables in research on the treatment of SIB. A more general measurement procedure that accommodates multiple wound characteristics is the Self-Injury Trauma (SIT) Scale (Iwata, Pace, Kissel, Nau, & Farber, 1990). The SIT Scale classifies SIB according to wound type and provides estimates of both the severity and amount of tissue damage produced by SIB. Iwata *et al.* (1990) compared 50 pairs of independently scored SIT scales and calculated percentage agreement for location, number, and type of injury, as well as overall summary scores. Mean agreement scores for all categories were above 89%, indicating that the SIT Scale was a reliable method for measuring the severity of SIB. One limitation noted by the authors, however, was that trauma indices may be insensitive to small changes in the frequency, duration, or intensity of SIB because they provide little information about any single dimension of SIB as a response. This limitation applies to all research in which wound measures are used as the index of behavior because no attempt has been made to establish correspondence between changes in behavior and changes in wound appearance. As exceptions, two studies included a comparison between self-reported measures of skin picking exhibited by typical adults and either SIT Scale measures (Teng, Woods, & Twohig, 2006) or observers' ratings (Twohig & Woods, 2001), both of which were derived from photographs of wounds taken during pretreatment, posttreatment, and follow-up conditions. Results showed that decreases in self-reported scratching during treatment corresponded to decreases in severity scores, but these data should be interpreted with caution because there was no direct observation of scratching.

Another approach to measurement, common in trauma and wound-care research, is the calculation of actual wound size or wound surface area (WSA). The traditional method for estimating WSA involves placement of a transparency that contains a metric grid over a wound, tracing the wound's perimeter, and counting the number of grid cross points contained in the perimeter (Liskay, Mion, & Davis, 1993). Recent technological advances have made possible the use of more sophisticated, computer-assisted techniques. Thawer, Houghton, Woodbury, Keast, and Campbell (2002) examined the reliability and validity of a computer-assisted, photographic technique for measuring the wounds of 45 human subjects and 38 laboratory mice. Computer-assisted measures were obtained by photographing a wound, digitizing it, and uploading it to a computer. Two assessors then measured each digital image three times using software that calculated WSA (the program was VeV Measurement Documentation). Transparency measures were used as the standard for comparison. An experimenter obtained three tracings of each wound by placing a transparency directly over the wound and tracing the perimeter with an indelible marker. Two assessors then used a planimeter (a mechanical tracing device that estimates the area of a two-dimensional shape) to measure the WSA of each tracing. Results indicated that intrarater and interrater reliability were high for both techniques. Correspondence between the two measures was high for the human wounds but lower for the smaller mice wounds.

Transparency and photographic measurement represent two methods for quantifying wound size that might be useful in corroborating changes in SIB. Both methods have been reported to be reliable (e.g., Griffin, Tolley, Tooms, Reyes, & Clift, 1993; Thawer *et al.*, 2002). However, the transparency method is somewhat invasive and poses risk for contamination. The computer-assisted method, although less invasive, is expensive

(software used by Thawer et al., 2002, cost \$5,000) and has been used exclusively in hospital settings. In addition, because the emphasis in past research has been on measurement of the static (current) condition of wounds, no attempt has been made to correlate changes in wound size with changes in observed behavior. Therefore, we evaluated the utility of an inexpensive and readily available computer-assisted photographic method as a means of measuring injuries produced by SIB. In Study 1, we examined the reliability of and the correspondence between transparency and photographic measures in estimating the size of wound models. In Study 2, we compared photographic WSA measures and direct-observation data during the course of treatment for SIB.

STUDY 1: EVALUATION OF COMPUTER-ASSISTED MEASUREMENT OF WSA

METHOD

Simulated Wounds

We used a sample of simulated wounds in this study for several reasons. First, it would be prohibitively time consuming to obtain samples of actual SIB that show wounds of the full range of sizes and shapes. More significant is the fact that measurement using the standard (transparency) procedure requires repeated placement of material directly on the wound, which we wanted to avoid. Finally, the dynamic nature of an injury is such that it changes over time and requires primary and reliability measures using two different procedures to be obtained within a short time window. To obtain a data set that contained wounds with varied dimensions and whose measurement would pose no risk and would remain constant across repeated observation, we carved 20 wound models out of Styrofoam blocks and colored them with a red marker (see Figure 1). We carved some of the models (9, 10, 12, 13, 14, 16) out of spheres instead of blocks to simulate wounds on curved areas of the body.

The resulting wound models varied in shape, length, area, curvature, and regularity of perimeter.

Procedure

Two experimenters measured all wound models in sequence (1 to 20) using first the transparency method and then the photographic method. Interobserver agreement was calculated separately for each method on a wound-by-wound basis by dividing the smaller WSA by the larger WSA. An index of discrepancy between the two measures was calculated for each wound by subtracting the photograph WSA from the transparency WSA and dividing the difference by the transparency WSA.

Transparency technique. An experimenter placed a mylar transparency of a 0.5-cm grid directly over a wound model and traced the perimeter of the wound with an indelible marker. WSA (in centimeters squared) was estimated by counting the number of grid cross points that were completely within the tracing and then dividing this number by four.

Photographic technique. An experimenter placed a metric ruler adjacent to a wound model and photographed both with a digital camera. No special conditions were arranged during photography; lighting and distance between the camera and model varied unsystematically. Each photograph was uploaded as a JPEG file to a Macintosh computer. Wound measurement was accomplished using ImageJ 1.29x software available from the National Institutes of Health at no cost (<http://rsb.info.nih.gov/ij/>). After opening the photograph in ImageJ (see Figure 2 for a screen shot), the experimenter set the measurement scale by selecting the straight line drawing tool and then drew a 0.5-cm line with the computer's track pad using the ruler in the image as a guide. The experimenter then selected Set Scale from the Analyze drop-down menu and entered the length of the line (0.5 cm) in the Known Distance field of the Set Scale dialogue box.

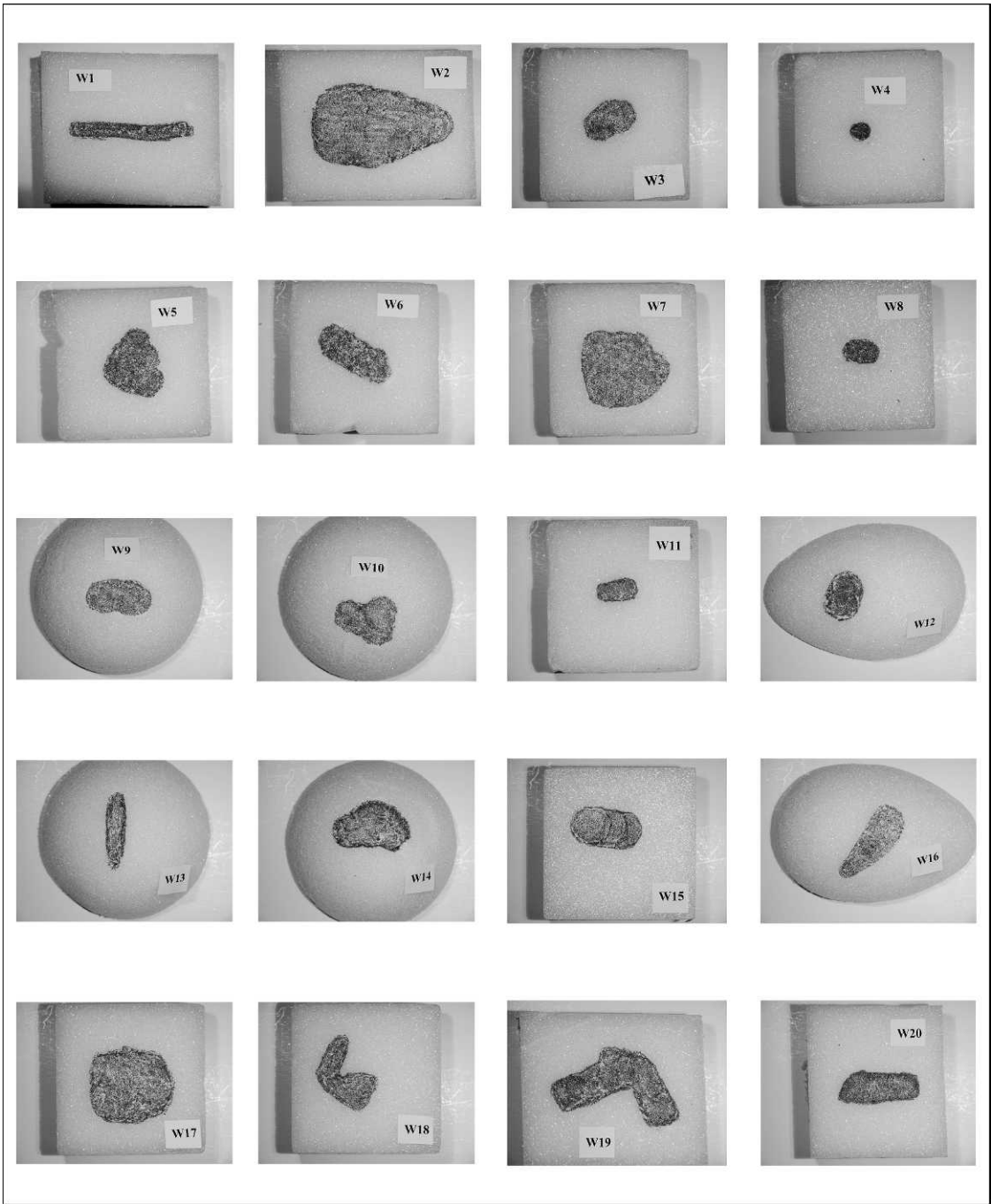


Figure 1. Wound models used in Study 1.

The experimenter then used the freehand drawing tool to trace the perimeter of the wound (the magnifying tool was used occasionally to enlarge the wound image before

tracing the perimeter). WSA (in centimeters squared) was calculated automatically when selecting Measure from the Analyze drop-down menu.

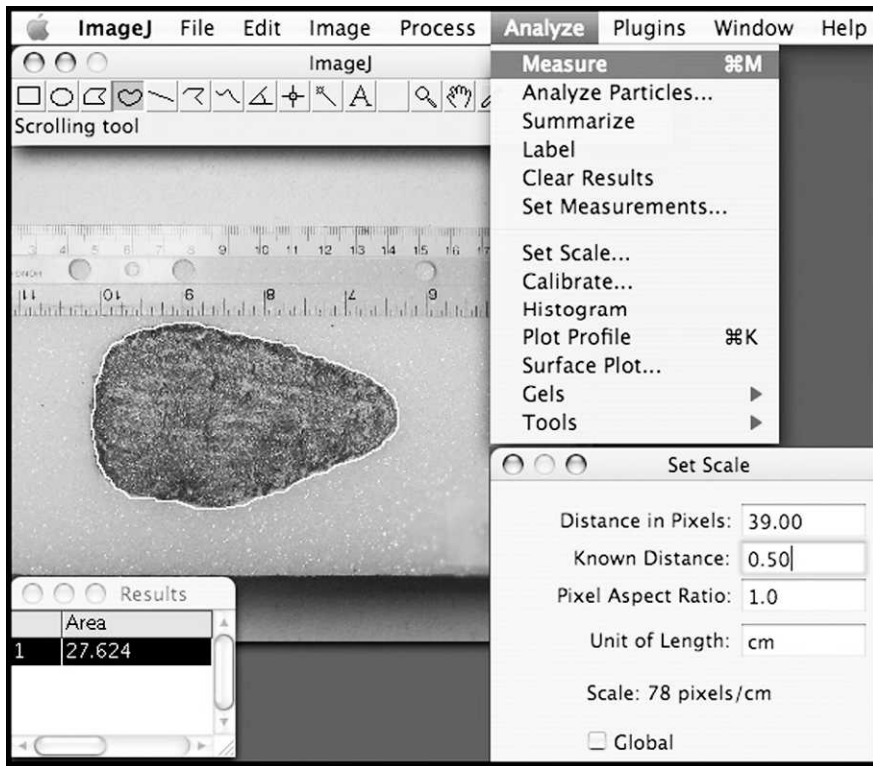


Figure 2. Screen shot of WSA calculation with ImageJ software.

RESULTS

Figure 3 (top) shows separate agreement percentages between the two observers using both measurement procedures for each model. Agreement scores for the transparency and photographic methods were high ($M_s = 95\%$ and 93% , respectively). It is interesting to note that the lowest reliability for both methods was obtained for the smallest wound (4, less than 2 cm^2) because small measurement differences represented larger proportions. This relation held true generally: Wound size was positively correlated with obtained reliability percentage ($r = .56$).

Figure 3 (middle) shows the actual WSA measures derived from the two methods for each wound model and indicates generally high correspondence between the measures. Figure 3 (bottom) shows the discrepancy between the measures obtained on each model.

Photographic WSA measurements were slightly larger than transparency WSA measurements for all but four wound models; this tendency can be attributed to the fact that any part of a wound that extended beyond a grid cross point (but did not reach the next cross point) of a transparency was not counted, whereas the entire (exact) wound perimeter shown in a photograph was included in the computer calculation. The mean discrepancy (calculated by ignoring \pm signs) between the transparency and photographic measures was 8.8% (range, 4.6% to 20%). As was the case with reliability for each of the procedures, WSA discrepancy between the two procedures tended to be correlated with wound size, although in a negative direction ($r = -.58$). That is, smaller wounds yielded larger discrepancies.

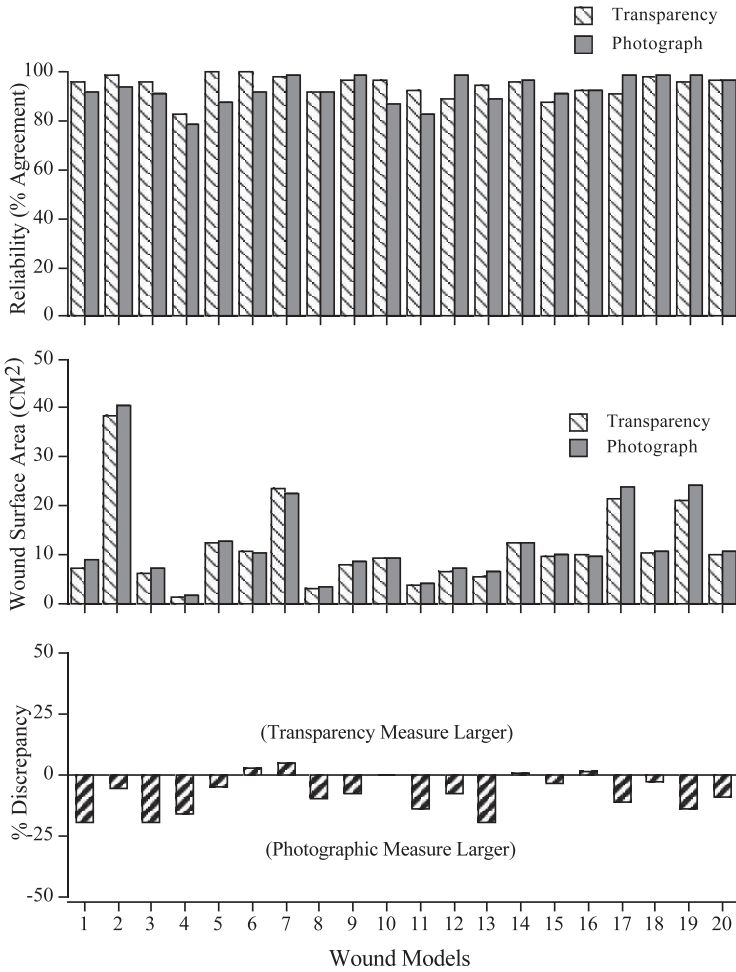


Figure 3. Interobserver agreement for transparency and photographic measures (top), WSA based on the transparency and photographic measures (middle), and discrepancy between the two measures (bottom).

STUDY 2: CORRESPONDENCE
 BETWEEN CHANGES IN OBSERVED
 SIB AND WSA

Although results from Study 1 established the reliability of the photographic method for measuring simulated wounds and its comparability with a more traditional measure, the degree of correspondence between the WSA of actual wounds and the observed frequency of SIB is unknown. Data that show that changes in WSA corresponded with increases and decreases in observed SIB would provide

stronger support for the photographic method as a measure of SIB and would permit greater confidence in the use of WSA as a dependent variable. Therefore, we compared WSA with direct measures of SIB during the course of treatment for an individual who engaged in relatively severe SIB.

PHASE 1: FUNCTIONAL ANALYSIS

Subject and Setting

Denise was a 31-year-old woman who had been diagnosed with Prader-Willi syndrome (PWS) and diabetes and whose SIB consisted of

skin picking, which produced severe wounds on her neck, leg, and buttocks. She lived in a group home, worked at a vocational day program, and required one-to-one staff supervision 24 hr per day.

Response Measurement and Reliability

SIB was defined as the scraping of fingernails, teeth, or an object against the skin, or tampering with a bandage covering an existing wound. Trained observers used laptop computers to record the frequency of SIB during each session. Reliability was assessed by having a second observer independently record data during 39% of the sessions, dividing session time into consecutive 10-s intervals, and comparing observers' records on an interval-by-interval basis. Reliability was calculated by dividing the smaller number of responses in each interval by the larger number of responses and averaging these fractions across the session. Mean agreement between observers was 98% (range, 92% to 100%).

Procedure

We conducted a functional analysis (FA) of Denise's SIB in which four conditions (alone, attention, play, and demand) were arranged in a multielement design (see Iwata, Dorsey, Slifer, Bauman, & Richman, 1982/1994, for procedural details). Because we suspected that Denise's SIB was maintained by automatic reinforcement, we conducted alone sessions to other sessions in a two-to-one ratio. Sessions lasted 10 min and were conducted three to four times daily, 4 to 5 days per week, at Denise's vocational day program.

Results

Figure 4 shows rates of SIB exhibited by Denise during her FA. SIB occurred at low rates initially but increased during all conditions except the attention condition over the course of assessment. Overall, Denise's rates of SIB were highest (although also highly variable) in the alone condition, indicating that her SIB was not likely to be maintained by social contingencies (attention or escape).

PHASE 2: INTERVENTION

Response Measurement and Reliability

Direct observation of SIB. Due to the severity of Denise's SIB, intervention was aimed at reducing its occurrence throughout the day. We trained staff members assigned as Denise's one-on-one therapists to record on a data sheet whether SIB occurred during continuous 15-min intervals, 24 hr per day, 7 days per week across all settings (day program, group home, and community). SIB was defined as either of the following: *Skin picking* was defined as scraping of fingernails, teeth, or an object against the skin that caused bleeding; *blood present* was defined as the presence blood (or a new wound) that had not been observed during a previous interval (this measure was included because it was possible for Denise to scratch while a staff member turned away briefly). Reliability was assessed by having one of the experimenters or a trained research assistant independently collect data with the primary observer during 20% of all intervals, which were scheduled intermittently throughout the day and night in equal proportion during baseline and treatment conditions. The observers' data sheets were compared for overlapping intervals, and reliability was calculated by dividing the number of scoring agreements (on the occurrence or nonoccurrence of skin picking or blood present) by the total number of intervals. Mean reliability was 98% (range, 83% to 100%) for skin picking and 100% for blood present.

Measurement of WSA. Throughout the study, Denise routinely attended wound-care appointments every few days, two to three times per week, for wound cleaning and replacement of dressings. During these appointments, an experimenter photographed each wound. An experimenter subsequently computed the WSA of each wound using the computer-assisted method described in Study 1. Reliability was assessed by having a second observer independently compute 100% of the WSA measures from the same photographs. Reliability was calculated by dividing the smaller WSA by the

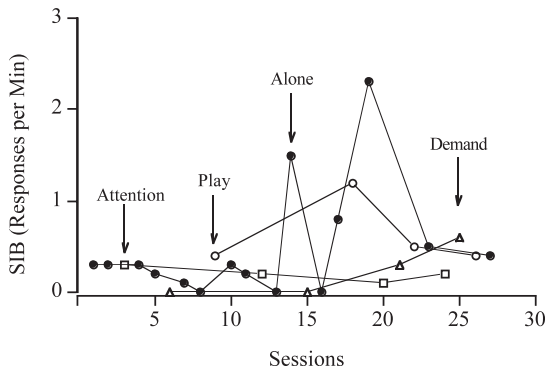


Figure 4. Responses per minute of SIB during Denise's functional analysis.

larger WSA. The mean agreement score for each separate wound measurement was 89% (range, 51% to 100%), and only two scores fell below 80% (51% and 73%).

Design and Procedure

Due to the severity of Denise's SIB, treatment was implemented immediately following the completion of her functional analysis. The effects of intervention were evaluated in a BAB reversal design.

Baseline. SIB and WSA were measured as described above. The occurrence or nonoccurrence of SIB resulted in no programmed consequences. However, an emergency blocking procedure was implemented if Denise (a) removed the bandage from a wound and caused further injury, (b) engaged in 40 or more occurrences of SIB that resulted in bleeding during 1 day, or (c) produced a new wound that required medical attention. All further SIB was blocked, and data collection was terminated for the day. This occurred only once (see below).

Differential reinforcement of other behavior (DRO). Given that Denise's SIB was maintained by automatic reinforcement, and because she already had relatively free access to varied sources of stimulation (leisure activities, social interaction, etc.) throughout the day, we designed an intervention based on a DRO contingency (Cowdery, Iwata, & Pace, 1990).

Access to tangible reinforcers was available according to three schedules across the entire day. First, Denise earned a token for each 15-min interval during which SIB did not occur. The tokens could be exchanged for back-up reinforcers (noncaloric beverages, Magic towels, cosmetics, puzzles, trinkets) four times daily. Second, because Denise was on a calorie-restricted diet as part of a weight management program, a calorie bonus was added to the DRO contingency. At 3-hr intervals during waking hours (five opportunities per day), she received a 20-calorie bonus contingent on the absence of SIB during the previous 3 hr, which she could trade in for food immediately or accumulate and trade in at a later time. Finally, if she engaged in SIB during five or fewer intervals for an entire day, she earned additional reinforcement in the form of an outing to a store to purchase a cup of low-calorie ice cream, a music CD, or a few inexpensive items of her choice.

Results

Figure 5 shows the percentage of intervals with SIB across successive days and the total WSA based on photographs taken during wound-care appointments (total WSA was calculated by adding the values obtained for all wounds during a measurement session). The break in the behavioral data (Days 57 to 62) represents an extended home visit. SIB was low throughout the initial treatment condition and during most of the subsequent baseline condition. During the final day of baseline, however, her SIB increased to above 25% of the intervals, which met the emergency criterion for response blocking. This increase in observed SIB (threefold relative to previous treatment and baseline sessions) seemed sufficient as a reversal and marked the end of baseline. When treatment was reimplemented, SIB decreased and remained low for an additional 3 weeks, after which formal data collection was concluded.

Denise's WSA was 4.98 cm² at the beginning of treatment, which resulted from two injuries,

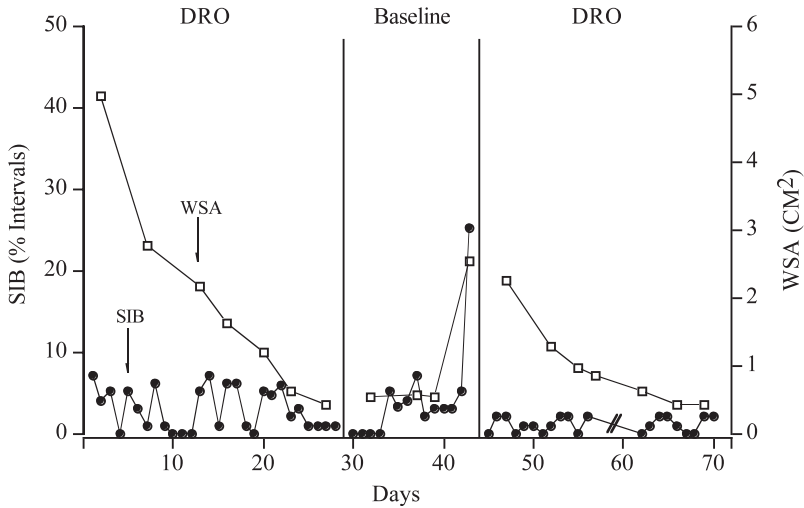


Figure 5. Percentage of intervals during which SIB occurred and WSA measures were taken across days during baseline and treatment conditions.

one on her buttocks (0.48 cm^2) and the other on her neck (4.50 cm^2). Total WSA decreased across treatment sessions and measured 0.44 cm^2 at the end of the phase (buttocks WSA = 0 cm^2 , neck WSA = 0.44 cm^2). The first baseline assessment of WSA also was low (neck wound only, 0.54 cm^2). However, SIB showed a small but abrupt increase on the 5th day of baseline, and a large increase in WSA was observed on the last day (neck wound only, 2.54 cm^2). When treatment was reimplemented, SIB remained low, and WSA gradually decreased again from an initial WSA of 2.25 cm^2 to a final WSA of 0.42 cm^2 (neck wound only). Figure 6 shows photographs of Denise's neck wound taken during the first and last WSA measurements under baseline and treatment conditions.

These results show a general correspondence between directly observed SIB (percentage of intervals) and the product measure of WSA. During the first and second treatment conditions, continued low rates of SIB were associated with gradual decreases in WSA; during baseline, an abrupt increase in observed SIB was associated with an abrupt increase in WSA. Thus, changes in WSA appeared to be

sensitive to sustained decreases in (or continued low levels of) SIB and immediately sensitive to increases in SIB.

GENERAL DISCUSSION

Results of the present studies indicated that measures analogous to WSA based on computer analysis of photographic images of wound models were reliable (Study 1) and that changes in actual WSA measures showed reasonable correspondence with changes in the observed frequency of SIB (Study 2). Thus, the photographic method for measuring wound size may be useful in a number of contexts related to assessment and treatment of SIB.

Results from Study 1 indicated that the photographic method compared very well with the more traditional transparency method and should be considered the preferred method for measuring wound size for several reasons. Perhaps most important is the fact that the photographic method eliminates all contact with the wound, which is a desirable feature when contamination is a concern. Second, the photographic method is inexpensive: Digital cameras are commonplace, and the measurement software is available at no cost. Third,

First Treatment Condition

First WSA measure



Last WSA measure

Baseline Condition

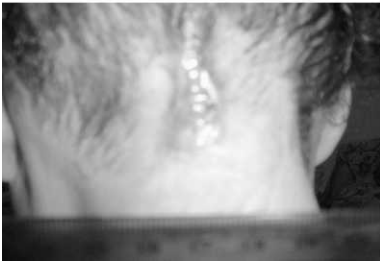
First WSA measure



Last WSA measure

Second Treatment Condition

First WSA measure



Last WSA measure



Figure 6. Photographs taken during the first and last WSA measurements for baseline and treatment conditions.

photographic wound measurement requires little skill: The experimenters who measured wounds in this study had no previous training or experience yet obtained high reliability. Thus, the photographic method seems applicable to a wide range of assessment situations when documentation of wound severity is important.

The comparison of WSA and behavioral data in Study 2 showed that bidirectional changes (increases as well as decreases) in observed SIB corresponded with changes in WSA. The temporal correspondence between these changes could not be determined precisely because it

was not possible to obtain daily WSA measures. However, it appeared that an increase in SIB was reflected immediately in WSA (see Figure 5, the end of baseline), whereas decreases in SIB were followed by decreases in WSA several days later (see Figure 5, both DRO conditions). For this reason, direct measures of responding should always be preferred as evidence of behavioral sensitivity to environmental changes. Typical response measures, however, document changes in rate or duration but do not reflect changes in severity (actual wound appearance), one of the defining features of SIB. WSA measures thus provide a quantitative index for

an important dimension of SIB, one that often is of paramount concern to parents and therapists.

Perhaps the greatest potential application for WSA measures is with SIB that occurs covertly (i.e., only when there are no cues indicative of being observed). Unless behavioral observations can be conducted through one-way windows or with surveillance equipment, SIB might never be seen. WSA measurement represents one way to determine that SIB has or has not occurred, although the detection of decreases in responding may be somewhat delayed. Another application of WSA is in outpatient research or treatment, where individuals typically are seen periodically. Although measures of behavior obtained through self or parental report might be helpful in evaluating the effects of treatment, these measures may be unreliable. Periodic WSA measures could be taken in clinic to corroborate changes in self-reported data. Alternatively, self-reported data could be replaced entirely with photographs taken by clients or parents and e-mailed to researchers or therapists. A similar approach has been shown to be successful in monitoring smoking, another behavior that is difficult to observe directly through measurement of its chief byproduct, carbon monoxide (Dallery & Glenn, 2005).

Several limitations of the photographic method should be noted. First, it may be difficult to determine the exact perimeter of wounds that have highly irregular (discontinuous) patterns or partial scabbing. Second, we measured one type of wound in the present study, lacerations, for which there is a relatively distinct break in the skin. Even so, WSA provides no estimate of wound depth. SIB also produces contusions, which are rendered easily through photographic analysis if the injury has discoloration (bruising), but it is unknown whether the boundaries of swelling can be detected readily in the absence of bruising. Third, wounds that extend around a joint may be difficult to capture in a single photograph,

although it is possible that WSA based on separate photographs could be summed, as was done in this study. Finally, the correspondence between WSA and observed behavior was shown with only one subject. Comparisons of measures for a larger number of individuals would permit greater confidence in the use of WSA as an index of SIB. Still, the present data indicate that computer-assisted photographic analysis of injuries seems to be a reliable and sensitive method for assessing changes in SIB when direct measures of responding are unavailable.

REFERENCES

- Cowdery, G. E., Iwata, B. A., & Pace, G. M. (1990). Effects and side effects of DRO as treatment for self-injurious behavior. *Journal of Applied Behavior Analysis, 23*, 497-506. doi:10.1901/jaba.1990.23-497
- Dallery, J., & Glenn, I. M. (2005). Effects of an Internet-based voucher reinforcement program for smoking abstinence: A feasibility study. *Journal of Applied Behavior Analysis, 38*, 349-357. doi:10.1901/jaba.2005.38-349
- Grace, N. C., Thompson, R., & Fisher, W. W. (1996). The treatment of covert self-injury through contingencies on response products. *Journal of Applied Behavior Analysis, 29*, 239-242. doi:10.1901/jaba.1996.29-239
- Griffin, J. W., Tolley, E. A., Tooms, R. E., Reyes, R. A., & Clift, J. K. (1993). A comparison of photographic and transparency-based methods for measuring wound surface area. *Physical Therapy, 73*, 117-122.
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197-209. doi:10.1901/jaba.1994.27-197 (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2*, 3-20, 1982)
- Iwata, B. A., Pace, G. M., Kissel, R. C., Nau, P. A., & Farber, J. M. (1990). The Self-Injury Trauma (SIT) Scale: A method for quantifying surface tissue damage caused by self-injurious behavior. *Journal of Applied Behavior Analysis, 23*, 99-110. doi:10.1901/jaba.1990.23-99
- Liskay, A. M., Mion, L. C., & Davis, B. R. (1993). Comparison of two devices for wound measurement. *Dermatology Nursing, 5*, 437-441, 434.
- St. Lawrence, J. S., & Drabman, R. S. (1983). Interruption of self-excoriation in a pediatric burn victim. *Journal of Pediatric Psychology, 8*, 155-159.

Teng, E. J., Woods, D. W., & Twohig, M. P. (2006). Habit reversal as a treatment for chronic skin picking: A pilot investigation. *Behavior Modification, 30*, 411–422. doi:10.1177/0145445504265707

Thawer, H. A., Houghton, P. E., Woodbury, M. G., Keast, D., & Campbell, K. (2002). A comparison of computer-assisted and manual wound size measurement. *Ostomy Wound Management, 48*, 46–53.

Twohig, M. P., & Woods, D. W. (2001). Habit reversal as a treatment for chronic skin picking in typically developing adult male siblings. *Journal of Applied Behavior Analysis, 34*, 217–220. doi:10.1901/jaba.2001.34-217

Received September 8, 2006

Final acceptance July 31, 2012

Action Editor, Gregory Hanley