The Use of Interactive Video Client Simulation Scores To Predict Clinical Performance of Occupational Therapy Students

George Tomlin

OBJECTIVE. Student academic performance (grades) has traditionally shown a weak correlation with clinical performance (supervisor ratings). A computer-based interactive video client evaluation program was created to determine whether the decision making of occupational therapy students on a client simulation explained variability in fieldwork supervisor ratings not accounted for by grades.

METHOD. Two student cohorts (N = 43; N = 30) selected evaluation procedures for a client with a cerebrovascular accident from on-screen menus and recorded and interpreted client responses as shown on videodisc. Multiple regressions were calculated in which grades and simulation scores were used to predict fieldwork supervisor ratings.

RESULTS. In cohort 1, grade in a physical disabilities course predicted 31% of the variance in supervisor ratings, whereas the simulation score for “completeness” accounted for an additional 12%, F(2, 39) = 16.75, p < .0001, adjusted \( R^2 = .434 \). Students with better grades received higher ratings, whereas students who evaluated the simulated client less exhaustively received higher ratings. For cohort 2, fieldwork ratings were best predicted by a combination of several simulation efficiency scores (number of correct answers given per unit time), F(4, 21) = 6.17, p < .002, adjusted \( R^2 = .453 \). Mostly, higher efficiency scores were associated with higher supervisor ratings.

CONCLUSION. Video simulations of client encounters may measure decision-making skills not measured by grades, in particular, making correct decisions in a timely manner within a realistic context.


Clinical decision making in the health professions has attracted the attention of practitioners and researchers alike over the past few decades, as theories of complex human performance evolved from those based in behaviorism to those with a cognitivist perspective (Irby, 1990). The inability of behavioral theories to explain or model the performance of human experts, in particular, gave rise to new formulations that highlighted the importance of internal mental processes (pattern recognition, memory structure, metacognition) and that eventually were to uncover the importance of contextual features of the “problem space” in reasoning about practical human problems (Dreyfuss & Dreyfuss, 1986; Schön, 1983). From this work and other research on the contextual features of practical cognition (Rogoff & Lave, 1984) grew the theory of “situated cognition” (Brown, Collins, & Duguid, 1989; Schell & Cervero, 1993). Situated cognition holds that decision making arises inherently from the complex context in which it occurs (Schell & Cervero) (i.e., it cannot be measured devoid of a context). Research in situated cognition was originally done to create a theory of more effective learning environments, but it has evolved into investigating performance assessments in educational settings (Baron, 1991).

Kane (1992) evaluated three performance assessment methods in the health professions (classroom testing, simulation, observation in clinic) for their reliability.
of scoring, internal validity, and external validity (extrapolation to actual practice), and found each method lacking in one respect or another. Simulation, however, was rated by Kane as having high scoring reliability, medium generalizability, and medium extrapolation accuracy. Simulation types in medicine have included paper or computer-based “patient management problems” and using actors as standardized patients (Vu & Barrows, 1994).

Attempts to predict clinical performance using prior measures of competence such as grades have a long history. From seven published studies on occupational therapy students, the highest single correlation found between grades and clinical performance was $r = .39$, and two studies found no correlation at all (Best, 1994; Englehardt, 1957; Ford, 1979; Katz & Mosey, 1980; Kirchner & Holm, 1997; Lind, 1970; Mann & Banasiak, 1985). Several authors recommended examining other predictor variables, such as learning style of student, interpersonal skills, and the pace or workload at individual clinical sites (Best; Mann & Banasiak). None recommended that a performance assessment for clinical skills be administered during the academic program.

Efforts to measure clinical competence in an academic setting in occupational therapy have evolved slowly, beginning with one early reported use of paper patient management problems (Feinstein, Gray, Stern, & Levine, 1982). Neuhaus (1986) described the use of a paper clinical problem-solving exercise, measuring the accuracy, thoroughness, choice value, and treatment priorities of the students’ responses. Neuhaus found great variability in the students’ problem-solving approaches, and concluded that there were no generalized skills of data gathering in the field. More recently occupational therapy educators have sought to design classroom experiences that incorporate those thinking skills necessary in a clinical setting (Babola & Peloquin, 1999; Neistadt, Wight, & Mulligan, 1998). These consisted of paper and videotape classroom exercises and they revealed positive learning outcomes. None examined the correlation of classroom performance with later supervisor ratings of performance during fieldwork.

The first two computerized client encounter simulations in occupational therapy representation (Anson, 1990; Tomlin, 1996) took audio and video approaches, respectively, to the mode of reality. Anson’s simulation was intended to give students practice with the procedures of specific evaluations. Tomlin presented the simulated client via video footage within a computer program, which students could advance only by making clinical decisions and observing and interpreting the client’s responses. To date, no study of the measurement characteristics of student scores on simulated clients in occupational therapy has been published, nor on their correlation with later fieldwork performance.

Interactive video simulations, to the extent that they portray realistic detail about a particular situation with complex, contextual challenges, are capable of fulfilling these conditions for situated cognition (Brown et al., 1989; Rogoff & Lave, 1984). Simulations in the health professions that incorporate interactive video enable the detailed, visual presentation of common clinical diagnoses, make available on a dependable basis the less frequently encountered clinical conditions, and can demonstrate appropriate decision-making approaches to clinical problems (Schwartz, 1989). Over the past 15 years there has been a slow appearance of reports documenting the effects of interactive video experiences among occupational therapy students (Croninger, Tumiel, & Sowa, 1995; Farrow & Sims, 1987; McNurlen, Gilkeson, & Drake, 1996; Tomlin, 1996), but none of the studies included an investigation of the measurement characteristics of the program as a decision-making assessment instrument.

Based on the need of academic programs to promote and assess decision making in students to prepare them for clinical placements, an interactive video client evaluation simulation (Tomlin & Straume, 1994) was created. The purpose of the current exploratory study was to determine if student performance on the video simulation predicted some aspect of clinical performance not predicted by course grades, where the criterion was clinical supervisor ratings of students during fieldwork Level II on the Fieldwork Evaluation (FWE) form.

**Method**

**Research Design**

A multiple regression model using coursework grades and interactive video simulation scores was created to predict the performance of students in their subsequent physical disabilities fieldwork. The regression model was cross-validated on the next cohort of students, in order to assess the stability of the model’s parameters. Reliability and validity of the simulation have been previously examined (Tomlin, 1996). The measurement characteristics of the FWE, although less than excellent, have been found to allow adequate generalizability (Cooper & Crist, 1988).

**Participants**

Students from one occupational therapy educational program who performed the client simulation in 2 successive years (67 students from cohort 1; 50 from cohort 2) were the target population for the current study. At the time of
the simulation students had just completed or were about to complete a neurological treatment course, including lectures and laboratory practice on neurological evaluation theory and procedures including those for clients with stroke. With a few exceptions, students finished their physical disabilities fieldwork within 12 months of completing the simulation.

Instrumentation

Course Grades. The three grades examined were those from functional anatomy, treatment of physical dysfunction, and applied clinical treatment courses, selected on the basis of prior research by Best (1994).

Interactive Video Client Evaluation Simulation. The design of the simulation followed recommendations by Alessi (1988), Hannafin and Hooper (1989), Reigeluth and Schwartz (1989), and others. That is, its fidelity to clinical reality was deliberately moderated for inexperienced students, its program branching was kept simple, student navigation was permitted within small sections but was constrained by an overall linear structure, cued responses were the kind most often required of students, and generally immediate feedback was given. Students were instructed to proceed through the simulation at their own pace.

The scope of the simulation was the evaluation of one client. The case was of a person with a cerebrovascular accident (CVA), judged to be the single most common diagnosis evaluated and treated by occupational therapists (Woodson, 2002). The encounter sought to simulate most of the decisions a practicing therapist might have to make during the evaluation of this client.

A real client with a CVA, undergoing rehabilitation at a local hospital, was videotaped while being reevaluated by a practicing occupational therapist. The client was selected by on-site rehabilitation staff for his presentation of observable movement deficits, ability to understand and follow verbal directions, and his willingness to be videotaped. The evaluation videotaping was guided by a script prepared by four therapists who each had at least 7 years of experience in neurological rehabilitation. The four agreed on the procedures for an ideal evaluation of a client with a CVA, although not on the specific order for these procedures. The videotaping was completed in three sessions, each lasting about 7 days. Three hours of footage of the evaluation was edited and 30 minutes of it pressed onto a random access videodisc, using the following considerations to maximize the value of video information shown per minute of disc space: (a) when the client's movements showed no deficit (as on the unaffected side of the body), a single freeze frame of video was used; (b) movements or behaviors that revealed deficits were shown in full; (c) repeated attempts by the client to perform a task or subtask were edited out; and (d) transitions from place to place (e.g., client rolling the wheelchair from bedroom to clinic) were shown in abbreviated fashion.

From the evaluation script a menu-driven computer program was written by the author and a physical therapy faculty collaborator using Authorware Professional© software (Authorware, 1991). The computer program controlled a videodisc player such that precise video frames were shown at the appropriate moments in the simulation. From on-screen computer menus, students were required to select assessments to perform and answer questions about preparatory procedures. They then observed the evaluation procedure and the client's responses on video, and selected observation and interpretation options shown on the computer. Periodically, they were prompted to give a summary appraisal of the client's performance.

The simulation consisted of evaluation sections—reading the client's chart and interviewing the client (not scored), sensorimotor and functional assessments, and assessing the cognitive and psychosocial functioning of the client (see Appendix). The simulation contained 317 decision points that were used in the current study (for details of a similar simulation program, see Tomlin, 1994). Student responses were automatically saved into individual data files. Completeness and accuracy of responses for each evaluation section and the time taken to finish each section were calculated. Completeness was defined as the number of correct options selected, divided by the total possible number of correct options (C/Total C). Accuracy was defined as the number of correct options selected, divided by this number plus the number of wrong options selected [C/(C+W)]. For example, if a section menu contained 50 options (20 correct and 30 incorrect), and a student selected 15 correct options and 10 incorrect options, the completeness would be 75% (15/20) and the accuracy 60% (15/25). Completeness and accuracy for the total simulation were calculated as well. Time efficiency scores were generated for each section by dividing the number correct in that section by the corresponding time elapsed. Time efficiency for the total simulation was also calculated.

Fieldwork Evaluation Form. The final FWE form completed by the supervising occupational therapist from the physical disabilities placement was the source of data for the dependent variable. The form's three scales (performance, judgment, attitude) have been found to be highly intercorrelated (Best, 1994), so only performance scale scores were used.

Procedure

University human subjects approval was received and consent for use of academic records in a research project was
sought from all potential participants. Students were informed about the simulation as a course requirement at the beginning of the semester in which they completed it. Three weeks were allotted for students to finish the assignment.

After the assignment deadline, user data files were collected. Essential information (number of options selected that were correct and incorrect, time taken) from each section of the simulation was transcribed onto participant data sheets. Course grades and later FWE scores were collected from departmental records.

Data Analysis

Participant data were analyzed using SPSS-X. Normality, variance, and heterogeneity were checked for all variables.

Standard multiple regressions were calculated to discover the amount of variance in fieldwork ratings that could be explained by a combination of course grades and client simulation scores. Because there could be many simulation scores as candidate variables for the multiple regression, the following criteria for inclusion of predictor variables in the model were used: (a) the regression analysis of variance yielded a significant $F$; (b) inclusion of the predictor resulted in an increase in adjusted $R^2$; (c) inclusion of the predictor did not reduce the number of cases in the regression due to missing data; and (d) the total number of predictors was such that there were at least five cases per predictor variable (Tabachnik & Fidell, 1989). Analyses were performed for cohort 1, and the resulting regression model was tested on cohort 2, and revised as indicated.

Results

Participants

The target population was 67 occupational therapy students from cohort 1 and 50 from cohort 2. Due to missing data, delayed fieldwork, or lack of consent, the numbers of final participants were 43 and 30, respectively. Participants were found not to differ significantly from their entire class of students on physical disabilities course grades or fieldwork performance scores, suggesting no differential effects of attrition. After preliminary analysis of the data, one participant from cohort 1 was excluded, due to the appearance of outliers, leaving 72 cases in the analysis.

Means and standard deviations for number correct, number incorrect, time elapsed, and efficiency are shown in Table 1. Students in cohort 2 selected more correct ($t(70) = 2.01, p < .05$), and fewer incorrect options ($t(70) = 6.29, p < .001$), taking nearly the same amount of time, and therefore had higher total completeness, accuracy, and efficiency scores than those in cohort 1 ($t(70) = 3.24, p < .01$).

Table 1. Simulation Total Scores: Completeness, Accuracy, Time, Efficiency

<table>
<thead>
<tr>
<th>Simulation Score</th>
<th>Cohort 1 ($N = 42$)</th>
<th>Cohort 2 ($N = 30$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct (#)</td>
<td>174.0 (22.60)</td>
<td>186.3 (33.51)</td>
</tr>
<tr>
<td>Wrong (#)</td>
<td>91.5 (17.67)</td>
<td>71.2 (16.94)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>68.8 (8.97)</td>
<td>73.7 (13.31)</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>65.6 (4.97)</td>
<td>72.5 (4.33)</td>
</tr>
<tr>
<td>Time (mins)</td>
<td>186.1 (37.35)</td>
<td>188.3 (47.13)</td>
</tr>
<tr>
<td>Efficiency a</td>
<td>0.97 (0.20)</td>
<td>1.10 (0.24)</td>
</tr>
</tbody>
</table>

$a$Completeness = number of correct choices/number of possible correct.

$^b$Accuracy = number correct/(number correct plus number wrong).

$^c$Efficiency = number correct/total time in minutes.

Prediction of Fieldwork Performance

Cohort 1. FWE scores from the 42 students in cohort 1 were first regressed on course grades and simulation scores. The physical disabilities course grade and the component, “completeness,” were significant predictors of FWE, $R^2 = .462$, adjusted $R^2 = .434$, $F(2, 39) = 16.75, p < .0001$. Students with higher grades in the physical disabilities course tended to perform better during fieldwork, whereas students who selected more from among all possible correct options on the simulation tended to be rated lower by supervisors during fieldwork. When the physical disabilities grade was used alone to predict FWE, the adjusted $R^2$ was .311 (31.1%). Adding “completeness” raised adjusted $R^2$ by .123 (12.3%).

Cross Validation of Cohort 1 Model for Cohort 2. Neither independent variable used with cohort 1 was a significant predictor of fieldwork performance for cohort 2. An exploration of other possible significant predictors for cohort 2 students identified several section-efficiency scores. With 4 of 30 cases excluded because of incomplete time data, the efficiency scores of the Transfer, Sensation, Sitting Balance, and Toileting sections combined to predict FWE scores significantly, $R^2 = .540$, adjusted $R^2 = .453$, $F(4, 21) = 6.17, p < .002$.

Discussion

For cohort 1 students the physical disabilities course grade and the simulation score “completeness” together accounted for about 43% of the variability in FWE scores. This amount of prediction power for a course grade (31%) was high in the range of previous studies of occupational therapy students (0% to 39%, median 10%) (Best, 1994; Englehardt, 1957; Ford, 1979; Katz & Mosey, 1980; Kirchner & Holm, 1997; Lind, 1970; Mann & Banasiak, 1985). The particular physical disabilities course that cohort 1 students completed was team-taught by two instructors: one with 15 years of experience in occupational therapy.
education, the other with 10 years’ experience in physical disability settings. Their complementary expertise may account for the unusually strong association in cohort 1 between course grade and later fieldwork performance.

The “completeness” score added about 12% to the variance in FWE scores explained, thus confirming that scores on a video-based client simulation would account for some variability in fieldwork performance among these participants not explained by grades. Surprisingly, students who evaluated more completely were given lower ratings on fieldwork. An ability to set priorities in data gathering during a real evaluation would probably be seen as an asset during fieldwork, where time management with clients and paperwork is crucial. The video-based, time-sensitive context of the simulation seems to have allowed measurement of some aspects of performance under conditions of situated cognition.

The particular cohort 1 predictors were ineffective in explaining cohort 2 FWE scores. The course instructors experimented with a problem-based learning format for cohort 2, which was unfamiliar to the students and may have undermined any correlation of the course grade with later supervisor ratings. Students in cohort 2 did not skip parts of the simulation as often as did some in cohort 1 (i.e., were more complete). A variable with little variance cannot easily be a significant predictor.

The regression model for cohort 2 FWE scores accounted for about the same amount of variability (45%) as cohort 1’s model (43%). The simulation scores that turned out to be significant predictors in cohort 2 were section efficiency variables (number correct in the section divided by time taken in the section). Greater efficiency in the Transfer and Sitting Balance sections was associated with a higher fieldwork rating. Sensation and Toileting section efficiencies, however, were negatively weighted, perhaps acting as “variance suppressor variables” (Tabachnik & Fidell, 1989). Higher efficiency scores would reflect an ability to perform correct actions more promptly—an important skill for clinical practice. Efficiency for the entire simulation was not a predictor for either cohort, probably because the time data did not account for students taking breaks of different lengths during the 2- to 3-hour simulation.

Despite differences, a common pattern emerged from the results of cohorts 1 and 2. For both cohorts some aspect of scoring from the client simulation accounted for variability in fieldwork performance not explained by course grades. In accordance with the theory of situated cognition, the video-simulated evaluation of a client provided students the opportunity to demonstrate some skills similar to those necessary in the complex context of actual student clinical practice—skills that were not entirely or consistently measured by course grades.

In addition, in both cohorts this aspect of predictive simulation scoring had something to do with the ability to make choices while using time efficiently. In cohort 1 it was “completeness” that was a negative predictor of later fieldwork performance. The clinically higher-rated students may have cut short the time they spent with the client simulation, with the realization that excessive thoroughness does not provide additional useful information. In cohort 2, on the other hand, students demonstrated time management skill within sections of the simulation, where efficiency scores were predictors of later clinical performance. The differences in cohort 1 and 2 results may suggest a sensitivity of student decision making to initial expectations for the simulation (how thorough to be, how high a priority is the assignment?), as Elstein, Shulman, and Sprafka (1978) noted with physicians who completed patient management problems. For both cohorts, though, the video-delivered information seems to have played an important role in establishing the realistic context for relevant decision making.

This simulation portrayed only the evaluation phase of a therapist-client interaction, and not any intervention. Other aspects of student preparation (physical skills, communication and social skills, and team skills), which were not measured by this simulation, also influence how well students perform during fieldwork. Students from only one occupational therapy program participated in the study. Another limitation was that 38% (45/117) of the students in the target population were lost from the analysis, for a variety of reasons. The effect of this attrition on the regression outcomes is unknown.

Future study could examine to what extent these results are generalizable beyond the students of one curriculum, or the extent to which they change over time due to intrinsic changes in the way students process video data, changes in the conditions and demands of contemporary fieldwork, or both. Comparison of student and therapist decision making using interactive video client simulations may reveal differences along the beginner-expert continuum in the way clinical evaluation decisions are made.

Conclusions

An interactive video client simulation measured aspects of clinical performance not measured by academic grades, namely, the efficient use of time in gathering evaluation data and the timeliness of making correct decisions during the client encounter. Measuring decision making in a context simulated by video portrayal of a client–therapist interaction, and in accordance with the theory of situated cognition, may allow for more realistic identification of some of the cognitive skills used in clinical practice.
References


### Appendix

#### Client Evaluation Simulation Section Contents

<table>
<thead>
<tr>
<th>Section Name (Client’s Room, in Bed)</th>
<th>Description of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensation</td>
<td>Light Touch, Sharp/Dull, Proprioception</td>
</tr>
<tr>
<td>Motor</td>
<td>Tone, Range of Motion, Active Movement, Synergies for Upper/Lower Extremities</td>
</tr>
<tr>
<td>Reflexes</td>
<td>Clonus, Babinski, Deep Tendon Reflexes</td>
</tr>
<tr>
<td>Bed Mobility</td>
<td>Rolling, Scooting, Bridging, Coming to Sit</td>
</tr>
<tr>
<td>Sitting Balance</td>
<td>Sitting Posture, Static &amp; Dynamic Balance</td>
</tr>
<tr>
<td>Transfer</td>
<td>Client/Wheelchair Preparations, Transfer from Bed to Wheelchair</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section Name (Client’s Room, In Wheelchair)</th>
<th>Description of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toileting</td>
<td>Transfer to/from Wheelchair, Dressing/Undressing</td>
</tr>
<tr>
<td>Grooming</td>
<td>Bathing, Shaving, Tooth Brushing, Hair Combing</td>
</tr>
<tr>
<td>Dressing</td>
<td>Retrieving/Donning Upper/Lower Body Clothes</td>
</tr>
<tr>
<td>Feeding</td>
<td>Containers, Utensils, Finger Foods, Liquids, Solids</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section Name (In Occupational Therapy Clinic)</th>
<th>Description of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of Injury</td>
<td>Sensory Loss, Left Neglect, Impulsivity</td>
</tr>
<tr>
<td>Cognitive</td>
<td>Communication, Memory, Orientation, Insight, Emotional Adjustment</td>
</tr>
</tbody>
</table>

---

**Occupational Therapy and the Older Driver**

Addressing the IADL of Driving and Community Mobility

This 2-day workshop promotes driving as an IADL in all occupational therapy settings and equips therapists to address initial driving and community mobility issues for the older client. The workshop initiates a dialogue between therapists and their clients regarding driving and calls for occupational therapy professionals to collaborate with the programs in their communities to meet the mobility needs of their clients.

Earn 1.3 AOTA CEUs (13 contact hours).

**Save with Early Registration!**

(See Web site for deadlines.)

**$212.50 AOTA Members, $300 Nonmembers**

**Regular Registration:**

**$250 AOTA Members, $355 Nonmembers**

Details at [www.aota.org](http://www.aota.org) (click on Continuing Ed) or Call 800-SAY-AOTA ext. 2837

---

February 26–27  Austin, TX
March 12–13  Washington, DC
April 2–3  Chicago, IL
June 25–26  Providence, RI

---

Development sponsored by the National Highway Traffic Safety Administration