COMPUTER-SCORED PSYCHOLOGICAL ASSESSMENT OF CANDIDATES FOR ADMISSION TO MEDICAL SCHOOL.

M. Anbar and M. L. Raulin.

Dept.of Biophysical Sciences, School of Medicine & Biomedical
Sciences, and Psychological Service Center, Faculty of Social
Sciences, State Univ. of New York, Buffalo, NY.

AN OVERVIEW

The evaluation of candidates for medical school currently focuses on appraisal of scholastic ability, leaving the assessment of personality traits to a brief personal interview plus "reading between the lines" of letters of recommendation. While scholastic performance is critical in medical school, especially in the preclinical sciences, lacking analytical thinking ability or possessing antisocial personality traits, such as combativeness, excessive selfishness or ineffective communication skills, predict poor professional performance in clinical practice. Extensive psychological tests, including interviews with trained psychologists, that could detect such undesirable traits, are too expansive to be applied to the hundreds of candidates selected for personal interview by each medical school. Personality assessment during brief personal interviews, generally by an untrained interviewer, suffers from severe shortcomings. In addition to the relative ease of deceiving untrained interviewers, these shortcomings include subjectivity biased by the personality match of the interviewer and interviewee, lack of standardization because of the parallel participation of many interviewers, each with a different experience and personality, and finally the lack of opportunity to assess the interviewee's behavior in a real-life like environment, when the interviewed person is relatively free of stress and less self-conscious.

One way to improve the quality of personality assessment of candidates, and minimize the acceptance of misfits, is the use of computerized tests to supplement personal interviews. Computerized tests cannot evaluate poise, demeanor, composure and savoir-faire, which can be assessed only in a person-to-person interaction, but computers can evaluate many personality traits that are hard to analyze in a casual interview.

For several years, we have been developing computerized psychological assessment tools based on role-playing in a simulated environment.[1] Such role-playing involves handling of simulated persons who challenge the testee in different scenarios. The testee plays different roles, ranging from a figure of authority, such as a student counselor, to a citizen intimidated by a rude law enforcement officer. These computerized simulations involve an unrestricted natural language interaction, achieved by using CASIP[2] as the authoring tool. CASIP parses the testee's input and recognizes answers by the presence of keywords, or their contextual synonyms, in specified positions in the input. The computer's response depends on all the interactions that took place from the start of the session, giving the testee the impression of a dialog with a live person. CASIP-authored programs yield a verbatim record of the man-machine dialog, including several measures of conduct in giving each answer.

CASIP also allows the scoring of each of the testee's responses on up to eight different dimensions. The total, mean, minimum and maximum scores on each of these dimensions, as well the number of answers above or below the midpoint of each scoring scale, are part of the computer output. These scoring parameters provide a combination of scores that characterizes the testee by a multidimensional behavioral profile. CASIP's automatic scoring thus overcomes the major shortcomings to personality assessment by interviews, namely subjectivity, lack of normativeness and excessive level of effort.

Ten scenarios were developed and tested on approximately one hundred medical school applicants. From those we selected five scenarios and after additional improvements in answer recognition and scoring, we tested them on over fifty undergraduate psychology students. We now plan to administer them to the entire freshmen medical class. Each scenario is stopped after 15 dialog cycles or after 8 minutes, whichever comes first; they may also terminate by certain extreme statements or actions of the testee. To implement automatic scoring, we carefully analyzed the records of the testees to determine the attributes that can be derived from each scenario.

We developed a scoring manual that scores each potential answer of the testee, often in conjunction with any of the answers previously given. Certain scoring dimensions are averaged, while in others we look for extremes, since a single critical answer, or a sequence of two answers, can unambiguously reveal a certain attribute of the testee's personality. If an undesirable attribute is flagged by more than one scenario, we feel that the test has achieved its goal of detecting a risky candidate. To validate these computerized tests, we intend to test the entire freshman class several years in a row, follow up the testees through residency and find to what extent these tests predict professional behavior. This vast effort is worthwhile if it significantly improves screening of candidates for medical school.

THE TESTING SCENARIOS

The battery of 5 testing scenarios, that takes up to 40 minutes to complete, starts with a scenario where the testee is at a somewhat higher social status than the computer-emulated person. Here is the introduction to that scenario: *"You are a student peer counselor. Your role is to advise students on subjects concerning their health and welfare. John, a student, is sent to you by the resident advisor because several students have been complaining about his smoking. What will you say to John?"* Testees have followed several different strategies, ranging from authoritative to friendly, and from trying to convince John to quit smoking, which is not called for by this scenario, to discussing the rights of others to avoid his secondhand smoke. John has an inconsistent personality that oscillates between militancy and compliance, making his handling intentionally difficult and frustrating. There is a twist when John mentions, if appropriately interrogated, that the complaint against him has an ulterior motive - he thinks he is hated because he is a better student than his peers.

The second scenario puts the testee in a confrontational position with a petty bureaucrat on campus. This is how it opens: *"You try to check out from the library a book that is essential for a term paper due tomorrow, and the librarian will not let you. She insists that you have an overdue fine of sixteen dollars and twenty-five cents because a previous book was returned late. You know that you returned that book on time, and that the library is at fault.*

This library does not issue receipts when books are returned and paying the fine is regarded as a final settlement; usually there is no way to recover a fine once paid. Books in this library are not stamped with dates of check out and check in. What will you say to the librarian at this point?" She is consistently authoritative and non-yielding to a level that evokes frustration. This scenario probes self-confidence and persistence without over-combativeness. Testees handle this situation in different ways, ranging from paying the fine right away, pleading for understanding, suggesting different possibilities that might have led to the unjust fine, inventing fake witnesses and receipts, to becoming abusively combative.

The third scenario puts the testee in an equal status to the emulated person under conditions that may call for suspicion and aggression: *"You are living in a dormitory. You just noticed that your wallet containing your monthly allowance, credit card and driver's license is missing. There is a young man named Bob in the room, he is a high-school classmate of Pat, your roommate. Bob arrived just yesterday. Pat is going to be in class for the rest of the afternoon. There is a phone on the desk, so you may call Campus Security. What will you say now to Bob?"* This scenario ends with a twist:*"It is 7PM. Pat returns at last and says, while in the doorway: "I found your wallet in the hallway near the elevator. Since I was rushing to class I could not get back and tell you. Here it is. You must have been worried. Weren't you? Well let's go and have dinner. By the way, where is Bob?"* Like in the former scenarios, different testees handle this situation very differently. Their strategies range from immediately accusing Bob, who then leaves indignantly, to "beating around the bush" trying to trap Bob as the culprit, to asking for his help in finding the missing wallet and borrowing money from him.

In the fourth scenario the computer emulates an authoritative bully: *"You buy a pack of pencils in a local drug store. Just as you are leaving the store, a security guard stops you. He accuses you of stealing a pack of gum. The gum was at your feet, just as you were leaving the checkout counter. It must have fallen out of someone else's bag, but to the guard it appears that it fell out of your bag. What will you say now to the guard in your defense?"* Like in the second scenario, the testee is innocent, but this offense is more serious and the consequences of conceding are much more severe. The guard is assertive and stubborn. Strategies used by testees in this situation range from aggressive defiance to submissive denial. This scenario tests the point where the testee breaks down and becomes combative, though alternatives of calling a lawyer or the police do exist.

The fifth and last scenario puts the testee in a situation where the computer emulates an irrational person of equal social status: *"You are at a Delta Tau Delta fraternity party, and one of the brothers, with the smell of beer on his breath, grabs you and says, "Hey...That is my shirt! Give me it! Right now!" He is very insistent that you have his shirt. Surely you know he is wrong. You are separated from your friends and have no extra clothing. How will you talk your way out of such a situation?"* This again is a frustrating situation which has either aggressive or submissive solutions, although certain delay or distraction tactics may also work. As stated, we look for consistencies in the behavior of the same testee in the different scenarios.

THE SCORING MANUAL

The limitations on the length of this paper allow us to present here just one abbreviated example from the scoring manual: The **Social Skills** variable is scored on a seven-point scale. Score = 2: Clearly inappropriate socially. *"Shut up or I'll smack you."* Score = 3: Not a good social response. *"This makes me so angry that I want to light up a cigarette right now, but I won't because it's prohibited here."* Score = 4: A socially responsive but crude response. *"You have the right to kill yourself by smoking, but not others."* Score = 5: An adequate social response, showing some mastery of social skill. *"You may not have trouble now, but you will in the future."* Score = 6: A fair to good response, indicates understanding of the situation. *"You are avoiding the proven link between smoking, cancer and heart attack."* Score 7: A very good response, showing positive interactive skills. *"I cant make you stop; I can only tell you that you would be better off if you stopped smoking."* Score = 8: An unusually good social response, projecting both warmth and understanding, while being responsive to the needs of the situation. *"I can see your problem with the friends of your roommate; let's ask him to help us."* The scores of the different answers of the preliminary stage of this study were used to establish a tentative automatic score. Later we checked the automatic scoring by reevaluating each answer and, when necessary, correcting the scoring accordingly.

SESSION RECORDS

The following is a typical record of the drug store session. The testee's answers are capitalized; punctuation marks were inserted for clarity. Each answer's record includes the time (seconds) it took to think and come up with the answer (Tk), to type it (Typ), and to read it over before pressing "Enter" (Rd); the number of correction backstrokes (Bk); and scores for social skills (SS), frustration (FR), submissiveness (SB), combativeness (CM) and negotiative skills (NG).

Date = 4-25-1991
Time = 13:40:07
Name = John Doe (changed)

SIR, IT WAS ON THE FLOOR.
[Tk:005s|Typ:008s|Rd:000s|Bk:000|SS=5|FR=1|SB=1|CM=6|NG=6]
The gum was at your feet. It fell out of your bag. You better have some explanation then.

I DIDN'T STEAL IT, IF THAT'S WHAT YOU'RE SAYING.
[Tk:004s|Typ:010s|Rd:000s|Bk:000|SS=6|FR=1|SB=1|CM=3|NG=4]
Stop denying something that is an established fact. Both of us know that you're a thief. All I want to know at this point is why did you take it.

I DIDN'T STEAL ANYTHING. IT WAS ON THE FLOOR SO I PICKED IT UP.
[Tk:009s|Typ:016s|Rd:000s|Bk:000|SS=5|FR=1|SB=1|CM=6|NG=6]
Didn't you hear me? The gum was at your feet. It fell out of YOUR bag.

NO, IT DIDN'T. I DIDN'T SEE THE GUM UNTIL I STARTED TO WALK AWAY.
[Tk:004s|Typ:026s|Rd:000s|Bk:006|SS=5|FR=3|SB=1|CM=3|NG=4]

How am I to know? I might just book you.

WHY?
[Tk:004s|Typ:000s|Rd:000s|Bk:000|SS=4|FR=1|SB=1|CM=1|NG=4]
I don't know why! If I were you I wouldn't ask such stupid questions.

GIVE ME A BREAK, I DON'T BELIEVE THIS.
[Tk:004s|Typ:009s|Rd:000s|Bk:000|SS=3|FR=3|SB=1|CM=5|NG=5]
Why don't you believe it? You never thought that you would get caught. Did you?

CAUGHT? I DIDN'T DO ANYTHING.
[Tk:005s|Typ:009s|Rd:000s|Bk:000|SS=5|FR=3|SB=1|CM=3|NG=4]
You heard me. Though I do not know for sure, I'll play safe and book you.

THIS IS RIDICULOUS. I DIDN'T STEAL ANY GUM!
[Tk:014s|Typ:013s|Rd:000s|Bk:000|SS=4|FR=4|SB=1|CM=3|NG=4]
No matter what you say now, I apprehended you trying to steal merchandise.

ASK THE CASHIER WHO CHECKED ME OUT IF SOMEONE JUST BOUGHT GUM.
[Tk:013s|Typ:016s|Rd:010s|Bk:005|SS=5|FR=1|SB=1|CM=1|NG=6]
Nobody else paid for the gum. You are just fishing for excuses.

NO, I'M NOT. I'M TRYING TO SHOW YOU I DIDN'T STEAL THIS GUM!
[Tk:003s|Typ:015s|Rd:000s|Bk:000|SS=5|FR=1|SB=1|CM=3|NG=4]
No one will believe you. You were caught in the act and must be punished.

NO
[Tk:004s|Typ:001s|Rd:001s|Bk:000|SS=3|FR=1|SB=1|CM=1|NG=4]
Don't try to be a 'wise guy' with me, I've put many like you behind bars.

I DIDN'T DO A THING. DID YOU SEE ME TAKE THE GUM?
[Tk:008s|Typ:009s|Rd:001s|Bk:000|SS=5|FR=1|SB=1|CM=3|NG=6]
I didn't see you pick-up the gum, but I know that you dropped it.

IT WAS BY MY FEET AND I PICKED IT UP.
[Tk:006s|Typ:012s|Rd:002s|Bk:000|SS=5|FR=1|SB=1|CM=1|NG=4]
No. You're not telling the truth. It wasn't there before you got to the counter.

IT MUST HAVE BEEN.
[Tk:008s|Typ:004s|Rd:000s|Bk:000|SS=4|FR=1|SB=1|CM=3|NG=3]
Tell me, why didn't you pay for the gum?

WHY ARE YOU ASKING ALL THESE QUESTIONS?
[Tk:027s|Typ:011s|Rd:000s|Bk:001|SS=5|FR=3|SB=1|CM=1|NG=4]
Are you trying to be Sherlock Homes or Doctor Watson? Of course, I asked. Anyway, I am the one asking questions here!