Internet Journal of Allied Health Sciences and Practice

Volume 21 | Number 1 Article 1

January 2023

Computer-Based Assessment of Non-Cognitive Attributes of Occupational Therapy Students: A Pilot Evaluation

Sara Maher
Wayne State University, Sara.maher@wayne.edu

Regina N. Parnell
Wayne State University, ad9049@wayne.edu

E. Whitney G. Moore
Wayne State University, whitneymoore@wayne.edu

Peter Frade
Wayne State University, peter.frade@wayne.edu

Barry Markman
Wayne State University, barry.markman@wayne.edu

Recommended Citation

Maher S, Parnell RN, Moore EG, Frade P, Markman B. Computer-Based Assessment of Non-Cognitive Attributes of Occupational Therapy Students: A Pilot Evaluation. The Internet Journal of Allied Health Sciences and Practice. 2023 Jan 04;21(1), Article 1.

Computer-Based Assessment of Non-Cognitive Attributes of Occupational Therapy Students: A Pilot Evaluation

Abstract

Purpose: Holistic admissions tools have been considered for use to diversify student admissions. The ability to screen non-cognitive attributes of applicants is an important element of holistic admissions. The objective of this study was to establish instrument reliability, inter-rater reliability, validity, item discrimination, item difficulty, and bias of the Computerized Assessment of Non-Cognitive Attributes of Health Care Professionals (CANA-HP). Methods: A novel methodology was developed to screen non-cognitive attributes of health profession applicants. Using a cross-sectional design, a convenience sample of students invited for interviews into a Mid-western occupational therapy program were recruited for participation. The 37 participants who consented, completed a demographic survey followed by the 12 question CANA-HP. Results: Open-ended questions had adequate internal reliability, discrimination, and difficulty. Multiple choice questions had low reliability and discrimination. No correlation was found between the CANA-HP and standardized cognitive assessments, except non-science GPA which was significantly correlated to the total open-ended (essay) scores and total overall score. Conclusions and Recommendations: The CANA-HP was not biased toward individuals from varied demographic backgrounds. Predictive validity of this tool and non-cognitive measurements of success are still needed. Occupational therapists in educational settings could consider adding open-ended ethical questions to the application process when interviews are not feasible. Classes in the non-sciences and other non-cognitive markers of success may also help identify students with 'grit' and critical reasoning which are important to practicing clinicians

Author Bio(s)

Sara F. Maher, PT, PhD, DScPT, is an Associate Professor and Interim Associate Dean of Health Sciences at Wayne State University. Dr. Maher's scholarship focuses on identifying tools and interventions which best enhance learning and identifying methodologies which target equal opportunities for learners of all genders, ethnicities and socio-economic status.

Shlomo Sawilowsky, PhD, was a Distinguished Faculty Fellow and Professor of Educational Evaluation & Research in the College of Education at Wayne State University (deceased). His areas of expertise included expert testimony, statistics, psychometry, research design, and program evaluation.

Regina Parnell, PhD, OTR/L, is an Assistant Professor and Admissions Coordinator for the Occupational Therapy Program at Wayne State University. Her primary scholarship is focused on Health Symptoms and Single African American Mothers and Health Inequalities: Hurried Women African American.

E. Whitney G. Moore, PhD, CSCS*D, is an Associate Professor in the College of Education, teaching Exercise & Sport Sciences and Educational Evaluation & Research programs at Wayne State University. She researches the implementation of planned missing data designs; assisting researchers with high quality data while decreasing study participant burden.

Peter D. Frade, PhD, is an Emeritus Professor who served as the Associate Dean of Health Sciences and Associate Professor in the Eugene Applebaum College of Pharmacy and Health Sciences at Wayne State University. His primary area of specialization surrounds hyphenated instrumental techniques as applied to clinical and forensic protocols.

Barry S. Markman, PhD, is a Professor of Educational Psychology, and Professor and Program Coordinator for Education Evaluation and Research at Wayne State University (retired). He is a Licensed Psychologist, with research interests in research design and statistical applications, and adolescent and adult ADHD and

applications of cognitive behavioral therapy.

Acknowledgements

This research was funded by a Faculty Research Grant Program through the College of Pharmacy and Health Sciences at Wayne State University. We publish in memory of Dr. Shlomo Sawilowsky, who passed away on January 11, 2021, a great mentor, teacher and friend.

The Internet Journal of Allied Health Sciences and Practice
Dedicated to allied health professional practice and education
Vol. 20 No. 4 ISSN 1540-580X

Computer-Based Assessment of Non-Cognitive Attributes of Occupational Therapy Students: A Pilot Evaluation

Sara Maher
Regina N. Parnell
E. Whitney G. Moore
Peter Frade
Barry Markman

Wayne State University

United States

ABSTRACT

Purpose: Holistic admissions tools have been considered for use to diversify student admissions. The ability to screen non-cognitive attributes of applicants is an important element of holistic admissions. The objective of this study was to establish instrument reliability, inter-rater reliability, validity, item discrimination, item difficulty, and bias of the Computerized Assessment of Non-Cognitive Attributes of Health Care Professionals (CANA-HP). **Methods:** A novel methodology was developed to screen non-cognitive attributes of health profession applicants. Using a cross-sectional design, a convenience sample of students invited for interviews into a Mid-western occupational therapy program were recruited for participation. The 37 participants who consented, completed a demographic survey followed by the 12 question CANA-HP. **Results:** Open-ended questions had adequate internal reliability, discrimination, and difficulty. Multiple choice questions had low reliability and discrimination. No correlation was found between the CANA-HP and standardized cognitive assessments, except non-science GPA which was significantly correlated to the total open-ended (essay) scores and total overall score. **Conclusions and Recommendations:** The CANA-HP was not biased toward individuals from varied demographic backgrounds. Predictive validity of this tool and non-cognitive measurements of success are still needed. Occupational therapists in educational settings could consider adding open-ended ethical questions to the application process when interviews are not feasible. Classes in the non-sciences and other non-cognitive markers of success may also help identify students with "grit" and critical reasoning which are important to practicing clinicians.

Keywords: holistic admission, non-cognitive attributes, occupational therapy students

INTRODUCTION

Racial and ethnic congruity between health care provider and patient has been associated with better health outcomes.[1] For minority clients, matches are limited due to the underrepresentation of minorities in health care. By the year 2030, it is projected children of racial/ethnic minorities will account for over 50% of the population under the age of 18 in the United States.[2] This will further exacerbate health care disparities unless efforts are made to increase the number of minority health care practitioners. A Healthy People agenda, developed by the U.S. Department of Health and Human Services in 2000, focused on reducing health inequity.[2] To achieve health equity, health care professionals should possess background and skills to enable effective treatment of patients from diverse backgrounds.[2,3]

Despite the recognized benefits of a diverse workforce, the proportion of under-represented applicants into health professions remains low. In the 2017-2018 Occupational Therapy Annual Data Report, disparities existed in the percentage of accepted students.[4] The following ethnic / racial distributions were observed: 80-85% Caucasian, 3-5% African American, 7% Asian, < 1% American Indian, < 1% Pacific Islander, and 5-8% unspecified. Similar trends were observed in physical therapy, nursing, and physician assistants where Caucasian student admissions ranged between 73.0% – 89.6% of candidates.[5-7] In medicine, initiatives to increase racial and ethnic diversity have had little impact: The composition of the medical profession has remained unchanged for the past 20 years.[8]

An important factor contributing to the lack of diversity has been identified as the 'leaking pipeline' or departure of students from their pathway into a health profession.[8,9] Barriers contributing to the 'leaking pipeline' among students who were underrepresented in medicine (URiM) included 1.) inadequate institutional support and resources leading to poor pre-college academic preparation (less rigorous high schools, low grade point averages); 2.) limited personal resources and social / family conflict (supporting themselves or family members, family pressure); 3.) lack of access to information, mentoring and advising (lack of connections to college such as not participating in extracurricular activities); and 4.) social barriers (work-life balance, uncertain job market). [8,9] According to Upshur et al, in consideration of 'pipeline' barriers, institutions should consider the approach used to construct programs and how program measure students' success.[9] Examining the admission process and how student success is measured during admissions may a first step in addressing the leaking pipeline.

One strategy to change student admission processes and measurements of student success may be holistic review; a flexible, individualized method of assessing an applicant's attributes, experiences, and academic metrics to determine the individual's potential contributions to the profession.[3, 10] Holistic admission is considered broad-based because it looks beyond an applicant's academic preparation, one of the identified barriers to URiM students, instead focusing on candidates' non-cognitive attributes, traits not measured by standardized tests (e.g., Graduate Record Examination (GRE).[11] Non-cognitive variables "provided viable alternatives in assessing the abilities of people of color, women, international students, older students, students with disabilities, LGBTQ students, or others with experiences which are different from those of young, White, heterosexual, able-bodies, Eurocentric males in the United States."[12]

Although a number of tools have been developed for cognitive performance, a common tool has yet to be developed to screen non-cognitive attributes of applicants into health profession fields. The purpose of this study was to examine the psychometric properties of a novel methodology called the Computer-based Assessment of Non-cognitive Attributes of Health Professionals (CANA-HP), an assessment tool developed by one of the authors. Specifically, internal consistency (reliability), inter-rater reliability, construct validity, item difficulty / discrimination, bias, and predictive validity of the instrument were examined.

Development of CANA-HP

Sedlacek & Brooks identified eight constructs thought to be important to the success of minority undergraduates, which eventually formed the Non-Cognitive Questionnaire.[13,14] To ensure applicability to a variety of health care professions, the author compared these eight constructs to professional attributes for four health professions found at the authors' university. In occupational therapy, seven core values were identified.[15] Similar core values were identified for three other health professions which included eight professional attributes in nurse anesthesia, six professional competencies in physician assistants, and seven core values in physical therapy.[16-18]

The author assessed commonalities in key words and descriptors for Sedlacek's constructs, and definitions of each professional attribute or core value. When substantial similarities between the non-cognitive constructs and professional attributes was assessed, the construct was matched with the professional concept. No matches were made for constructs and attributes assessed as dissimlar.

Table 1. Comparison of Non-cognitive Attributes with Profession Specific Attributes

Sedlacek & Brooks Non-Cognitive Variables	Nurse Anesthesia	Occupational Therapy	Physician Assistant	Physical Therapist
Positive self-concept	Well (healthy choices)	Freedom (autonomy)	Professionalism (healthy behaviors)	Excellence (skilled service)
Realistic self-appraisal	Situationally aware	Truth	Practice-based learning & improvement	Accountability
Ability to navigate systems & culture	Culturally competent	Dignity / Equality / Justice	Systems-based practice	Social responsibility / Integrity
Leadership	Leader / Teacher	Prudence	Patient Care	Professional Duty
Community service	Professionally engaged	Altruism		Altruism
*Communication & interpersonal**	Collaborative		Interpersonal & communication skills	Compassion / Caring
Delayed gratification**				
Strong support system**				
Knowledge of field**	Evidence based practice		Medical Knowledge	

*Not part of Sedlacek's original eight non-cognitive attributes.
**Excluded non-cognitive attributes of Sedlacek's original eight

The overlap between Sedlacek's eight constructs and the professional attributes of the four health professions are depicted in Table 1. A total of six factors were applicable to all four programs: (1) positive self-concept; (2) realistic self-appraisal; (3) ability to navigate systems and cultures; (4) leadership; (5) community service; and (6) communication and interpersonal skills. Although communication and interpersonal skills were not part of Tracey and Sedlacek's constructs, three programs identifed these as critical and were included in this study.[13, 15-17] Three of Sedlacek's constructs; delayed gratification, strong support system and knowledge field, were not included, as they were not identified as a core value of the professions or were considered profession dependent.[12]

The CANA-HP was designed as a computer assessment consisting of situational judgement tests. Situational judgements tests, which are short scenarios about a work problem or situation, have been used to examine non-cognitive traits in either open-ended /essay or multiple-choice format.[19, 20] For each non-cognitive variable in Table 1, two situational judgement tests were developed; one open-ended question (essay) and one multiple choice. Both sets of questions were developed by the primary author in on-going conjunction with expert feedback from health care professionals. Further, preliminary questions were tested on n = 9 randomly selected students from four health professional programs at the authors' university. The students provided quantitative and qualitative information on the pilot test. Following student input and professional review, the questions were amended and finalized for this study.

METHODS

Approval was obtained from the Institutional Review Board at the university where the study was conducted, prior to recruitment.

Participants

This study only focused on one healthcare profession, occupational therapy, because the admission procedures vary significantly across the healthcare science programs at the university. Informed consent was obtained for all participants.

Procedures

A convenience sample of applicants invited for in-person individual and small group interviews in the occupational therapy program, were invited to participate in the study. All applicants were interviewed but not all applicants consented to take the study's computerized assessment. Applicants were advised their occupational therapy application would not be impacted by not

participating in the study, and they could withdraw at any time. Applicants were excluded from the study if they did not sign an informed consent, or if the applicant was not at least 18 years of age.

Prior to interview, the Admission Chair of the Occupational Therapy Program sent applicants an electronic invitation to participate in the study and an informed consent. On the day of the interview, the primary investigator met with applicants in a computer lab, described the study, and answered any questions the applicants had. The assessment was administered via Qualtrics software (Provo, Utah).

The first survey question asked for informed consent. If consent was confirmed, the applicant was directed to the demographic section of the survey. Following the last demographic question, participants were asked if the research team could access the admission application to retrieve verified graduate record exam (GRE) and grade point average (GPA) scores. If the applicant selected yes, he/she created a unique identifier which was added to the Qualtrics Survey. The unique identifier and applicant name was written on a 3x5 card and given to primary researcher. Next the applicant moved directly into the 12 situational judgment test items.

For the multiple-choice items, each participant was asked to select the three BEST answers from a total of seven choices. Scores ranged from –6 to 9 for each question; three correct answers were worth 3 points each, one neutral answer was worth 0 points, and the remaining three answers had increasing negative value (-1, -2, and -3). The answers were self-graded by Qualtrics software. Two reviewers independently scored the open-ended items on seven, 5-point Likert scale rubrics (range of 1-5), for a total score from 7-35 for each question. The reviewers consisted of one research assistant and the primary investigator.

Data Analysis

All data were analyzed using IBM SPSS Statistics for Windows, version 26.0 (IBM Corp., Armonk, N.Y., USA) or Iteman v. 4.3 (ASC, 2013).[21] Descriptive statistics of the population were conducted, and significance was set at $p \leq .05$.

Cronbach's coefficient alpha was conducted on the twelve items, using $\alpha \geq 0.7$ as evidence of reliability.[22] Interclass correlation coefficients (ICC) estimates were calculated for interrater reliability using average rating (k = 2), consistency-agreement, and 2-way random-effects model. Interrater reliability was classified as poor (< 0.50), moderate (0.50 - 0.75), good (0.75 - 0.90) or excellent (> 0.90).[23]

Pearson's correlation coefficients were calculated between CANA-HP scores and GRE and GPA scores at the time of program admission. To control for Type I error, Bonferroni corrections were applied to results of the multiple comparisons.

The CANA-HP used partial scoring for all items, and the scales were considered polytomous in nature. For item difficulty and discrimination, scores for both question types were adjusted. Iteman software limits each variable to a maximum of 15 options (possible scores) and does not recognize negative values. Therefore, for the six multiple-choice items, each participant's initial total score ranged from –6 to 9. The scores were adjusted so each negative score was converted to a zero. All other scores remained the same. For the open-ended items, each construct was rated on 7 characteristics using a Likert scale (1-5), with a final score from 7-35. Because of the 15-option limit, each final score was divided by 7 (the number of characteristics) to give an average score (1-5) for the question.

Item difficulty was analyzed using mean average (P), and test discrimination was conducted with Pearson point-biserial correlation (Rpbis). The P value was the average of item responses. A good rating scale had a mean close to 50% of the maximum score for the item.[21] The Rpbis value ranged from -0 to 1.0 with a minimal acceptable range starting between 0.10 – 0.20. A negative point-biserial indicated a very poor item, and a score of 0.0 indicated no differentiation.[21]

Fisher's exact tests, with Bonferroni adjustments, were conducted to determine if the CANA-HP scenarios were biased for minorities, Pell-grant status, socio-economic status, or family college history. Due to the small sample size, all categories were collapsed to increase the number of individuals in each category. Binary categories were created for race (minority or Caucasian), Pell-Grant status (recipient or non-recipient), and family college attendance (first generation or not first generation). Income and the three scores on the items (final written, final open-ended, and total overall score) were collapsed into quartiles.

Simple linear regressions were conducted to assess if CANA-HP scores predicted performance in two OT classes; OT 5310 (Movement Assessment and Intervention) offered in the fall and OT 6100 (Occupational Therapy Assessments and Interventions I) offered in the winter. OT 5310 focused on theory and motor skill learning related to movement assessment including range of motion, strength, sensation, and coordination. OT 6100 focused on musculoskeletal and orthopedic diagnoses, assessments and

interventions, and impact on life occupations. Binary logistic regressions were calculated to determine if CANA-HP scores predicted a participant assuming a leadership role within the program.

RESULTS

Table 2. Participant Demographics

Variable	Category	Frequency	Percent
Gender	Female	32	86.5%
	Male	5	13.5%
Race/Ethnicity	Black	3	8.1%
	Hispanic	2	5.4%
	Multi-racial	1	10.8%
	Middle-eastern	4	2.7%
	White / Caucasian	27	73.0%
Pell Grant Recipient	Yes	9	24.3%
	No	27	73%
	Missing	1	2.7%
First Generation College	Yes	8	21.6%
	No	29	78%

Variable	N	Mean	SD	Range
Age	37	23.00 yrs.	(± 3.76)	20-43 yrs.
Family Income	30	$84,313	(± $51,586)	$15,000 – $200,000
Science GPA	37	3.46	(± 0.36)	2.68 - 4.0
Non-science GPA	37	3.61	(± 0.31)	2.91 - 4.0
Undergraduate GPA	37	3.51	(± 0.34)	2.55 - 4.0
Verbal GRE	11	145.4	(± 5.0)	136 - 151
Quantitative GRE	11	145.6	(± 6.0)	132 - 154
Analytic GRE	9	3.7	(± 0.8)	2.0 - 4.5

Thirty-eight applicants were interviewed and 37 (97.4%) agreed to participate in the study. Demographics for these participants are compiled in Table 2. They were primarily female (86.5%), Caucasian (73%), with a mean age of 23.0 (± 3.76).

Reliability for all CANA-HP items was low (α = 0.492). Although rules of thumb abound, this magnitude did not meet the modest criterion for evidence of reliability set at α ≥ 0.7.[22] The decision was made to conduct separate analyses for both question types. Reliability for the six multiple-choice items was α = 0.091, with minimal variation (α = .064 – 0.171) between items if any item was deleted. Reliability for the open-ended items was found to be minimally adequate (α = 0.7060), and the six items had minimal variation if deleted, (α = 0.582 to 0.717). All open-ended items appeared to strengthen overall reliability. Because the number of items for both items was small, a Spearman Brown correction was computed. To achieve a minimally adequate reliability for the entire CANA-HP, the number of items would need to be tripled (α = 0.744).

Results of interrater reliability can be seen in Table 3. Scores between the two raters ranged from 'moderate' for positive self-concept (0.67) and realistic self-appraisal (0.67) to 'excellent' for navigate systems / culture (0.91). On average, the six open-ended items took a total of 6 minutes and 29 seconds to grade with a range from 50 seconds to 1 minute and 22 seconds per question.

Table 3. Intraclass Correlations Coefficients for Average Measures for the Open-ended Items

	95% Confidence Interval			F Test with True Value 0			
Construct	Intraclass Correlation[b]	Lower Bound	Upper Bound	Value	df1	df2	Sig.
Positive self-concept	.672	.364	.831	3.052	36	36	.001
Realistic self-appraisal	.674	.367	.832	3.068	36	36	.001
Navigate systems / culture	.908	.822	.953	10.892	36	36	.000
Leadership	.891	.788	.944	9.168	36	36	.000
Community service	.827	.665	.911	5.789	36	36	.000
Communication / interpersonal	.817	.645	.906	5.468	36	36	.000

Note: Two-way random effects model where both people effects and measures effects are random.
a. The estimator is the same, whether the interaction effect is present or not.
b. Type C intraclass correlation coefficients using a consistency definition. The between-measure variance is excluded from the denominator variance.

Three GPA scores were significantly correlated to each other ($p \leq .001$) (Table 4). The total score on the multiple-choice items was not correlated to any of the GPA scores (p = .084 - .699). However, non-science GPA was significantly correlated to the total score on the open-ended items (p = .002) and total overall score (p = .008). GRE scores were not required for admission into the occupational therapy program. Therefore, results of the correlation analysis for GRE and the question score were based on the nine individuals, or approximately 24.3% of the sample population. None of the GRE scores was correlated to any other measure examined in this study (p = .052 - .999).

Table 4. Correlations between Three Types of GPA, Three Types of GRE, and Question Scores

Variable		Science GPA	Non-science GPA	Undergrad GPA	Total Multiple Choice	Total Open-ended	Total Overall Score
Science GPA	Pearson Correlation	1	.678**	.657**	-.066	.269	.241
	Sig. (2-tailed)		.000	.000	.699	.107	.150
	N	37	37	37	37	37	37
Non-science GPA	Pearson Correlation	.678**	1	.867**	-.163	.496**	.429**
	Sig. (2-tailed)	.000		.000	.335	.002	.008
	N	37	37	37	37	37	37
Undergrad GPA	Pearson Correlation	.657**	.867**	1	-.288	.326	.210
	Sig. (2-tailed)	.000	.000		.084	.049	.211
	N	37	37	37	37	37	37

Variable		Verbal GRE	Quant-itative GRE	Analytic GRE	Total Multiple Choice	Total Open-ended	Total Overall Score
Verbal GRE	Pearson Correlation	1	.648	-.020	.662	.000	.276
	Sig. (2-tailed)		.059	.959	.052	.999	.472
	N	9	9	9	9	9	9
Quantitative GRE	Pearson Correlation	.648	1	-.411	.155	.008	.073
	Sig. (2-tailed)	.059		.272	.691	.983	.852
	N	9	9	9	9	9	9
Analytic GRE	Pearson Correlation	-.020	-.411	1	-.204	.298	.221
	Sig. (2-tailed)	.959	.272		.598	.435	.567
	N	9	9	9	9	9	9

** Correlation is significant with a Bonferroni adjustment (p = .05/5 = .01).
GPA = grade point average
GRE – graduate record examination

Reliability between item types was poor, therefore the decision was made to separately analyze item difficulty and discrimination for the two question types. For the multiple-choice items, the maximum score was 9, thus a mean (P) of 4.5 was considered 'good'. Three of the multiple-choice items had good item difficulty; positive self-concept (P = 4.81), navigate systems / culture (P = 5.08)

and community / interpersonal (P = 4.43). Community service was too difficult (P = 2.68), and realistic self-appraisal (P = 7.00) and leadership (P = 7.14) were too easy. Only one item from the multiple-choice items appeared to discriminate between test takers, realistic self-appraisal (Rpbis = 0.26). Positive self-concept (Rbpis = 0.04), communication / interpersonal (Rpbis = 0.02), and navigate systems / culture (Rpbis = 0) had low or no discrimination. Leadership (Rpbis = -0.05) and community service (Rpbis = -0.01) may be considered poor items due to their negative Pearson point-serial correlation values.

For the open-ended items, the maximum score was 5 for each item, making a mean (P) of 2.5 as a 'good' rating for item difficulty. The items about communication / interpersonal (P = 2.54), community service (P = 2.74), and leadership (P = 3.14) were considered to have appropriate difficulty. The remaining three items might be considered easy (P = 3.51 – 3.70). All six items had appropriate discrimination (Rbpis = 0.15 – 0.56).

No statistically significant differences were found between scores on the three outcomes measures based upon sex (p = .394 - .925), ethnicity (p = .029 – 1.00), Pell grant status (p = .394 - .694), family college history (p = .124 - .948), or income level (p = .070 - .477).

Due to poor reliability of the multiple-choice items, regression analyses were conducted using only written scores. CANA-HP written scores did not significantly predict final grade scores in OT 5310 ($F(1,26) = .436$, $p = .515$) or OT 6100 ($F(1,25) = .038$, $p = .847$). CANA-HP scores also did not predict if a student assumed a leadership role in the program; written total (Wald = .023, df = 1, p = .880); written leadership (Wald = .067, df = 1, p = .796).

DISCUSSION

The aim of this study was to examine the internal consistency (reliability), inter-rater reliability, construct validity, predictive validity, item difficulty and discrimination, and bias of the CANA-HP, a measurement tool designed to look beyond traditional cognitive assessment tools (e.g., GPA) which were identified as a barrier for under-represented student in healthcare pipelines.

Reliability

Reliability for the six open-ended items ($\alpha = 0.71$) was similar to reliability for the typed CASPeR ($\alpha = 0.72$).[24] Items with strong internal consistency should show moderate correlations (0.70 - 0.90 for a large number of items). When correlation gets too high, there is a concern the items being measured may be redundant with a potential for limitations in the content validity.[25] Hence, the open-ended items of the CANA-HP had a minimally acceptable level of internal consistency, without being redundant.

Analysis of the multiple-choice items of the CANA-HP contained questions in which the reader chose three BEST options from a list of seven. Patterson et al. used multiple choice questions with medical school applicants and reported internal consistency from $\alpha = 0.43 – 0.94$, compared with $\alpha = 0.091$ for this study.[20] Although multiple-choice format is cost-effective and efficient, these results did not support internal consistency of these questions. To improve internal consistency, we postulate increasing the number of items, weighting scores from 1-7 based on appropriateness (e.g., Autonomy & Control Orientations) or perhaps changing the instructions to select a single best answer instead of the BEST three options.[26] A second consideration is to run the sample with a larger number of students including applicants from other health care professions. Patterson et al. for example, used medical school applicants and the differences in the two populations may have impacted these results.[20]

Interrater Reliability

The CANA-HP was found to have a range from $r = 0.67 – 0.91$ for the two raters, compared to a general interrater reliability of $r = 0.81$ for CASPeR.[24] Scoring of the CANA-HP took only 6 minutes and 29 seconds for all six items (roughly one minute, five seconds per question), compared with 24 minutes to score the 12 items of the CASPeR tool. The CANA-HP is reliable and less time intensive based on these results

Construct Validity

Because the CANA-HP was designed to measure non-cognitive attributes, the hypothesis was this methodology would not be significantly correlated to multiple-choice measures of cognitive abilities, thus verifying the tools construct validity. The final score on the multiple-choice items was not significantly correlated to GPA scores, supporting the hypothesis. However, the total score on the open-ended items and total overall scores were significantly correlated to non-science GPA, representing a medium effect size ($r = 0.496$ open-ended items, $r = 0.429$ total overall score).[27] Non-science GPA is comprised of a wide variety of classes a student takes, which may contain a mixture of both cognitive and experiential learning, a subset of intelligence where individuals learn to interpret novel tasks.[28] Kalsbeek reported individuals from non-traditional backgrounds may use this type of intelligence during initial exposure to new subject matter.[9] Thus, non-science GPA may represent another non-cognitive outcome.

GRE scores were not correlated to any of the total scores on the CANA-HP items, supporting the hypothesis of this study and again supporting construct validity of the assessment. These results are similar to those found with CASPeR where no correlation was reported between the cognitive portion of the MCAT test and CASPeR outcomes.[19]

Predictive Validity

Based on the results of this study, predictive validity for two classes (cognitive) and leadership (non-cognitive) could not be determined. Other than performance on standardized tests, predictive validity of non-cognitive tools has not been reported. Future studies could examine predictive validity in CANA in non-cognitive attributes desired in health care providers such as performance on practical examinations, objective structured clinical examinations, and clinical experiences. It is important to determine how non-cognitive attributes impact patient care.

Item Difficulty and Discrimination

The open-ended items had a minimally acceptable level of reliability, appropriate discrimination and three of the six items had appropriate difficulty. Like the open-ended items, three multiple-choice items had appropriate difficulty. However, the reliability for these items was low with poor discrimination. Item discrimination and difficulty cannot be compared to other tools as it has not been reported previously.

Bias

There was no statistically significant difference on the CANA-HP's three outcome measures for five demographic variables *(typical indicators for minority populations*). Hence, the CANA-HP is not biased toward these variables. These results should be interpreted with caution as sample sizes for each variable was small and had to be collapsed for analysis.

Limitations

The participants in this study were a small sample of convenience, limited to applicants of one health profession program at the university where the study was conducted. Therefore, results are not generalizable beyond these subjects. Multiple choice questions may be difficult to design without significant expertise. Although the consultants and primary researcher in this study had a long history in their health professions (minimum of 20 years in the profession), they had no previous knowledge writing questions in the format of situational judgment tests. This may have impacted results. The constructs used in this study may have been difficult to measure as none had a criterion reference. In addition, each construct had only one question for both the multiple-choice and open-ended items. Domain reliability could not be obtained, and more questions could have resulted in higher reliability scores. The open-ended questions had a higher weighted value which might have influenced the total overall scores for the CANA-HP. Future studies should increase the weight given to the multiple-choice items.

Implications for Occupational Therapy Practice

These findings have implications for occupational therapy practice: The benefits of matching health care providers to clients with similar cultural attributes has been established, as well the persistent limited diversity among occupational therapy practitioners.[1,4] The use of holistic admission strategies and non-cognitive measures of success represents one attempt to change the admission process and perhaps retain more under-represented applicants who start a pipeline into a health care profession.

The use of non-science markers of success is an important topic for consideration. While didactic measurements predict didactic performance (e.g., exams), non-cognitive measures better predict experiential performance, and both are needed among clinicians.[29] There has been increasing recognition in medicine of the need to assess skills in the arts and humanities which may help form deeper relations with patients, maintain joy and meaning in treatment, and develop empathy and resilience.[30, 31] Resilience or "grit" (consistency in performance of whatever type) may be one of the best ways to predict later success as a health care provider,[29] and yet markers of grit may not exist in most admission processes. It is also accepted success as a clinical practitioner requires solid skills in critical reasoning, and yet critical reasoning is also rarely assessed by most admission committees.[29] Students accepted into medical school who took fewer classes in traditional sciences (organic chemistry, physics, calculus) performed at levels equivalent to those on traditional pathways.[30] In addition, MCAT total scores were not found to be related to undergraduate minor.[32] These results suggest courses and tools which identify non-cognitive measures of student success are as equally important as traditional cognitive markers and do not impact student performance in health programs.

CONCLUSION

These results support the CANA-HP has no correlation with standardized cognitive assessments (GRE and GPA scores), highlighting its potential value in assessing important non-cognitive attributes of applicants. The one exception was non-science GPA which should be further examined. Classes in the non-sciences and non-cognitive markers of success may help identify students with 'grit' and critical reasoning which are important to practicing clinicians. Results from this study support the use of

open-ended (essay) questions as a way to assess non-cognitive attributes of applicants into for occupational therapy programs. Multiple choice questions were not supported in their current form; further research is needed if multiple choice questions are to be used in the future. Predictive validity of this methodology and other measurements of non-cognitive skills needs to be further assessed and developed.

References

1. Traylor, A. H., Schmittdiel, J. A., Uratsu, C. S., Mangione, C. M., & Subramanian, U. (2010). The predictors of patient-physician race and ethnic concordance: a medical facility fixed-effects approach. *Health Services Research*, 45(3), 792–805. https://doi.org/10.1111/j.1475-6773.2010.01086.x
2. Meadows, J. (2014). Blueprint for teaching cultural competence in physical therapy education. Retrieved October 27, 2018, from https://www.apta.org/Educators /Curriculum/APTA/CulturalCompetence/
3. Artinian, N., Drees, B., Glazer, G., Harris, K., Kaufman, L., Lopez, N., & Michaels, J. (2017). Holistic admissions in the health professions: Strategies for Leaders. College and University: *The Journal of the American Association of Collegiate Registrars*, 92(2), 65–68.
4. Harvison, N. (2017-2018). Academic programs annual data report. Retrieved October 27, 2018, from https://www.aota.org/~/media/Corporate/Files/EducationCareers/ Educators/2017-2018-Annual-Data-Report.pdf
5. Chana, T. (2017-2018). Aggregate program data: 2017-2018 physical therapist education programs fact sheets. Retrieved October 27, 2018, from http://www.capteonline.org/uploadedFiles/CAPTEorg/About_CAPTE/Resources /Aggregate_Program_Data/AggregateProgramData_PTPrograms.pdf
6. DiBaise, M., Salisbury, H., Hertelendy, A., & Muma, R. (2015). Strategies and perceived barriers to recruitment of underrepresented minority students in physician assistant programs. *Journal of Physician Assistant Education*, 26(1), 19-27. doi:10.1097/jpa.0000000000000005
7. Gould, W. (2014, July 1). Improving diversity in graduate nurse anesthesia programs. Retrieved October 16, 2018, from https://minoritynurse.com/improving-diversity-in-graduate-nurse-anesthesia-programs/
8. Freeman, BK., Landry A., Trevino, R., Grande, D., & Shea, JA. (2016). Understanding the leaky pipeline: Perceived barriers to pursuing a career in medicine or dentistry amon underrepresented-in-medicine undergraduate students. *Academic Medicine*, 91(5), 987-993. doi: 10.1097/ACM.0000000000001020
9. Upsur, CC., et al. (2018). The health equity scholars program: Innovation in the leaky pipeline. *Journal of Racial and Ethnical Health Disparities*, 5(2), 342-350. doi: 10.1007/x40615-017-0376-7.
10. Witzburg, R., & Sondheimer, H. (2013). Holistic review: Shaping the medical profession one applicant at a time. *New England Journal of Medicine*, 368(17), 1565-1567.
11. Kalsbeek, D. (2013). Employing noncognitive variables to improve admissions, and increase student diversity and retention. *Strategic Enrollment Management Quarterly*, 1(2), 132-152. https://doi.org/10.1002/sem3.20016
12. Sedlacek, W. (2017). Measuring noncognitive variables: Improving admissions, success, and retention for underrepresented students (1 ed.). Stylus Publishing, LLC.
13. Sedlacek, W., & Brooks, G. (1976). Racism in American education: a model for change. Nelson-Hall.
14. Tracey, T., & Sedlacek, W. (1989). Factor structure of the non-cognitive questionnaire-revised across samples of black and white college students. *Educational and Psychological Measurement*, 49, 637-648. https://doi.org/10.1177/001316448904900316
15. Kanny, E. (1993). Core values and attitudes of occupational therapy practice. *American Journal of Occupational Therapy*, 47(12), 1085-1086.
16. American Association of Nurse Anesthetists (AANA). (2016). Professional attributes of the nurse anesthetist: Practice considerations. Retrieved from https://www.aana.com/docs/ default-source/practice-aana-com-web-documents-(all)/professional-attributes-of-the-nurse-anesthetist.pdf
17. American Academy of Physician Assistants (AAPA). (2012). Competencies for the Physician Assistant Profession. Retrieved from https://www.aapa.org/wp-content/uploads/2017/02/PA-Competencies-updated.pdf
18. American Physical Therapy Association (APTA). (2010). Core values for the physical therapist and physical therapist assistant. Retrieved from https://www.apta.org/siteassets/pdfs/policies/ core-values-endorsement.pdf
19. Dore, K., Reiter, H., Kreguer, S., & Norman, G. (2017). CASPeR, an online pre-interview sceen for personal / professional characteristics: prediction of national licensing exam scores. *Advances in Health Science Education*, 22(2), 327-336. doi:10.1007/s10459-016-9739-9
20. Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory, ad practice: AMME Guide No. 100. *Medical Teacher*, 38(1), 3-17. doi:10.3109/0142159X.2015.1072619

21. Guyer, R., & Thompson, N.A., (2013). User's manual for Iteman 4.3. Woodbury, MN: Assessment Systems Corporation.
22. Fraenkel, J., Wallen, N., & Hyun, H. (2014). How to design and evaluate research in education (9th ed.). McGraw-Hill Education.
23. Koo, TK., Li, MY. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*. 15(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012
24. Dore, K., Reiter, H., Eva, K., Krueger, S., Scriven, E., Siu, E., & Norman, G. (2009). Extending the interview to all medical school candidates - The computer-based multiple sample evaluation of non-cognitive skills (CMSENS). *Academic Medicine*, 84(10), S9-12. doi: 10.1097/ACM.0b013e3181b3705a
25. Portney, LG., Watkins, MP. (2020). Foundations of clinical research: Applications to evidence-based practice (4th ed.). F.A. Davis.
26. Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, 73(5), 642–650. doi: 10.1037/0022-0663.73.5.642
27. Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed.). SAGE Publications, Inc.
28. Sternberg, R. (1985). Beyond IQ: A Triarchic Theory of Human Intelligence (1st ed.). Cambridge University Press.
29. Wilcox, RE., & Lawson, KA. (2017). Predicting performance in health professions education programs from admission information – comparison of other health professions with pharmacy. *Currents in Pharmacy Teaching and Learning*, 10, 529-541. https://doi.org/10.1016/j.cptl.2017.12.004
30. Muller D., & Kase, N. (2010). Challenging traditional premedical requirements as predictors of success in medical school: The Mount Sinai school of medicine humanities and medicine program. *Academic Medicine*, 85(8), 1378-1383. Doi: 10.1097/ACM.0b013e3183dbf22a
31. Wald, HS., McFarland, J., & Markovina, I. (2019). Medical humanities in medical education and practice. *Medical Teacher*, 41(5), 492-496. https://doi.org/10.1080/0142159x.2018.1497151
32. Ganjoo, R., Schwartz, L., Boss, M., McHarg, M., & Dobrydneva, Y. (2020). Predictors of success on the MCAT among post-baccalaureate pre-medicine students. Heliyon, 6, e03778. https://doi.org/10.1016/j.heliyon.2020.e03778.