Examining admission and formation outcomes for Catholic clergy applicants with the MMPI-2-RF: A Prospective Study

Paul B. Ingram
Texas Tech University
Dwight D. Eisenhower Veteran Affairs

Anthony Isacco
Chatham University

Nicholas C. Borgogna
Texas Tech University

Psychological evaluations of clergy applicants to the Catholic Church are an important gatekeeping mechanism during the admission process. However, limited research exists on the validity of assessments for this evaluative purpose and none have examined the predictive utility of the MMPI-2-RF to determine if an applicant will be accepted to formation, or if they ultimately ordained. This study fills that gap in the literature by investigating the predictive validity of MMPI-2-RF scales in 147 male applicants evaluated as part of their application for seminary or diaconate formation programs in a mid-sized Catholic diocese in the United States. Group analyses (e.g., Kruskal-Wallis tests) with admission status as the independent variable and MMPI-2-RF scales as the dependent variables yielded significant differences, most notably, those participants not admitted had higher mean scores on F-r, Fp-r, EID, RC7, THD, RC8, RC1, MLS, NUC, and JCP than the other three groups. Relative risk ratios were estimated for all MMPI-2-RF scales, indicating that higher scores are generally associated with a lower likelihood of admission and, ultimately, ordination. Limitations and future directions of research are also discussed.

Practical significance: This study evaluates the predictive utility of the MMPI-2-RF for use in clergy evaluations and supports its use in this capacity.

Keywords: MMPI-2-RF, Catholic Clergy applicants, priests, deacons, psychological assessment

The Catholic Church is the largest single religious institution in the United States, counting approximately 51 million adults and over 17,000 parishes (Pew Research Center, 2018). Parishes are led by ordained clergy (priests and deacons) that minister to the people in a variety of organizational, pastoral, and interpersonal ways. The psychological health of applicants to the priesthood and diaconate is a point of emphasis in the Catholic Church. Accordingly, a comprehensive psychological evaluation is a required component of the admission process for clergy formation programs. As such, effective psychological assessments are critical for identifying and screening out candidates that may be inappropriate for these vocations, based on psychopathology, addictive behavior, emotional immaturity, personality characteristics incongruent with effective ministry, or deviant sexual interests and behaviors (USCCB, 2015). Despite the importance placed on accepting psychologically healthy candidates for clergy roles, little is known about specific measurement indicators that might predict relevant admission outcomes such as acceptance or rejection into the formation program, retention or drop-out, and formation completion that is capped by ordination to a clergy role. There are standardized guidelines that describe a comprehensive,

The views, opinions, and/or findings contained in this article are those of the authors and should not be construed as an official Department of Veterans Affairs or U.S. Government position, policy or decision unless so designated by other official documentation. Paul Ingram receives research funding from the University of Minnesota Press, publisher of the MMPI-2-RF. Correspondence concerning this paper should be addressed to Paul B. Ingram, Department of Psychological Sciences, Texas Tech University, Lubbock, TX, 79409. Email: pbingram@gmail.com.

multi-modal psychological evaluation of clergy applicants (USCCB, 2015). The use of broadband personality measures is well suited to the evaluative needs outlined within these guidelines, given that they assess an array of psychological concerns as well as relevant test-taker response styles.

The Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) is a widely validated and frequently utilized broadband instrument that is particularly well suited to assessments of clergy applicants. While the MMPI-2 is the most widely used test in clergy applicant evaluations (McGlone et al., 2010), the scale revisions of the MMPI-2-RF and their associated psychometric improvements (see Tellegen & Ben-Porath, 2008/2011) provide sufficient reason to believe the MMPI-2-RF may be an effective tool. The MMPI-2-RF's 9 validity and 42 substantive scales assess a range of psychological concerns that are relevant to clergy applicant evaluations. For instance, scales on the internalizing and interpersonal domains are likely to predict clergy candidate retention given the role of social (Gautier, 2018; Ferrari, Reed, & Guerrero, 2017; Wong, 2014) and emotional functioning and maturity factors (USCCB, 2015; Sunardi, 2014) in these outcomes. Externalizing scales provide coverage of similarly important considerations, such as those related to addictive behaviors and problems with authority figures (USCCB, 2015). There is some limited research supporting the MMPI-2-RF's utility within this population and evidence suggests that the general trends observed on the MMPI-2-RF are conceptually consistent with the more established literature on the MMPI-2 (see Plante & Lackey, 2007; Plante et al., 2005). Specifically, clergy applicants on the MMPI-2-RF have a low general level of symptom endorsement that meets recommended cut scores and a higher frequency of under-reporting scale elevation (Isacco et al., 2020b). This pattern is not only similar to results with clergy applicants on the MMPI-2, but also with other populations undergoing public service employment evaluations where emotional and interpersonal scales are critical for employment determinations (e.g., Corey & Ben-Porath, 2018).

Despite this early evidence supporting the MMPI-2-RF for use in this population, there is generally a dearth of research on its predictive utility. A robust examination of the MMPI-2-RF's predictive utility would offer substantial benefit to psychologists conducting admission evaluations of clergy applicants. Demonstrating predictive validity is important for establishing test validation (Cronbach & Meehl, 1955), and identifying specific MMPI-2-RF scales that are associated with admission, retention, and ordination as that knowledge would improve clinical decision-making. Therefore, the purpose of this study is to use psychological data from the MMPI-2-RF obtained from clergy applicant evaluations to examine predictors of admission and retention outcomes. We posed two research questions. First, what are the differences in MMPI-2-RF scales between three distinct groups of clergy applicants: those who were (i) not admitted (ii) admitted but left formation and (iii) those whose who were admitted and completed formation (i.e., became an ordained clergy member in the Catholic Church)? Second, based on criteria set forth by the Catholic Church regarding candidacy expectations (USCCB, 2015) that emphasize social-emotional functioning, behavioral concerns, and openness to authority figures, what are meaningful differences between the three groups on the internalizing, interpersonal, externalizing, and validity scales?.

Methods

Participants

This study included 147 male participants who completed a psychological evaluation as part of their application to seminary or diaconate formation. Participants in this study include applicants to the seminary (a formation program for men that are discerning the priesthood) as well as to applicants to the diaconate (a formation program to become a deacon). A priest is ordained to administer the sacraments and be a spiritual leader in a faith community; priests cannot marry. A deacon is ordained to assist the priest in ministerial activities; deacons can marry and have children. Thus, priests and deacons are similar but distinct categories of ordained clergy in the Catholic Church.

Table 1
Participant Demographics

	Not Admitted n = 35	Admitted, Still in n = 107	Admitted, Ordained n = 43	Admitted, Left n = 15
Age M(SD)	41.6 (14.6)	39.8 (14.2)	42.9 (14.3)	36.7 (14.1)
Education, M(SD)	16.9 (2.5)	17.7 (3.5)	17.7 (2.8)	16.5 (2.2)
Race				
White	94.3%	98.1%	97.7%	80.0%
Hispanic	2.9%	-	-	-
Multiracial	2.9%	1.8%	2.3%	6.7%
African American	-	0.9%	-	6.7%
Asian	-	1.80%	-	6.7%
% with Children	45.7%	51.9%	39.5%	60.0%
Marital Status				
Single	42.9%	48.1%	58.1%	40.0%
Married	57.1%	49.1%	41.9%	60.0%

In our sample, seminary applicants (n = 67) were not married and did not have children, mean age was 25.8 (range 18-57 years-old), mean years of education was 15.94 (range 11 – 24 years in school), 94% were White or European American, 1 African-American or Black, 2 Asian-American, and 1 Multiracial. Diaconate applicants' (n = 80) mean age was 51.24 (SD = 2.7; range 34 – 60), mean education years was 18.18 (SD = 2.6; range 16 – 27), 94% were married, 87% were biological fathers, 96% were White/European-American, 2 identified as multiracial and 1 identified as Hispanic. Participants were descriptively similar across each of the outcome groups in this study (see Table 1).

Procedures

Participants were applicants to the seminary or diaconate formation programs in a mid-sized Catholic diocese in the mid-Atlantic region of the United States. All participants took part in a standardized psychological evaluation consisting of a clinical interview, objective and projective tests, and a feedback session between 2013 - 2020. The report from the psychological evaluations became part of the applicant's admission file in the diocese. All participants evaluated between 2013-2016 completed the MMPI-2 or the MMPI-2-RF between 2017 - 2020. All MMPI data was computer scored and the T-Scores were inputted into SPSS v.26. Consistent with previous research using similar clinical databases, MMPI-2 item responses were converted to MMPI-2-RF scale scores for analysis (Tarescavage et al., 2015). Demographic data were obtained on a basic intake form that participants completed prior to the psychological evaluation and included questions about the applicant's age, marital status, parental status, employment, race/ethnicity, and obtained education. Permission to use archival test data from the diocesan admission files was granted by the diocese to the second author. The study was approved by the Chatham University's Institutional Review Board.

Measures

Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF): The MMPI-2-RF is a widely utilized self-report assessment of personality and psychopathology (Ben-Porath, 2012). The test consists of 338 true-false items and includes 9 validity scales which assess profile interpretability as well as the hierarchically organized substantive, clinical scales. These 42 substantive scales include 3 Higher Order construct scales, 9 clinical scales referred to as the Restructured Clinical (RC) scales, 23 specific problem scales (examining specific somatic/cognitive, internalizing, externalizing, and interpersonal problems), 2 interest scales, and the scales of the Personality Psychopathology 5 (PSY-5). The MMPI-2-RF is extensively validated and includes reliability (test-retest, internal consistency, and standard error of measurement) and extra-test validity data (diagnostic formulations, intake demographics, record review forms, etc.) for each scale in the Technical Manual (Tellegen & Ben-Porath, 2008/2011).

Data Analysis Plan

First, we conducted a series of Kruskal-Wallis tests with admission status (Not Admitted, Admitted Left, and Admitted Ordained) entered as an independent variable with all MMPI-2-RF scales entered as dependent variables. Kruskal-Wallis tests were utilized rather than univariate Analysis of Variances (ANOVA) because Kruskal-Wallis does not assume data normality, which was violated across several MMPI-2-RF scales. The Kruskal-Wallis test also helps to address small and unequal sample sizes, which exist in our study sample amongst outcome groups. To determine significance, we utilized a family-wise Bonferroni corrected significance to adjust for multiple comparisons, as is common in the MMPI-2-RF literature (for instance, see Ingram et al., 2020). Specifically, our adjusted p-values were .05/9 = .006 for the Validity scales; .05/3 = .017 for the H-O scales; .05/9 = .006 for the RC scales, .05/5 = .01 for the Somatic/Cognitive scales, .05/9 = .006 for the Internalizing scales, .05/4 = .013 for the Externalizing Scales, .05/5 = .01 for the Interpersonal scales, and .05/5 = .01 for the PSY-5 scales. We then conducted a series of Mann-Whitney U post-hoc tests on scales with a significant overall omnibus statistic to determine which groups differed meaningfully and calculated effect sizes for those differences with Hedges' g. Effect sizes were interpreted as clinically meaningful if they demonstrated a Hedge's g value greater than .4 (Ferguson, 2009). Lastly, we calculated Relative Risk Ratios for scale scores identified as significant on the Kruskal-Wallis test at cut-score values of T55 and T60. Scores with a confidence interval that crosses 1 (e.g., .50 - 1.50) are considered to not have a reliably different rate of risk between groups

Results

Between group analyses identified several scales that differed meaningfully between groups with known admission outcomes (not admitted, admitted left, and admitted ordained). These differences were evident across some of the validity scales as well as all four of the five MMPI-2-RF scale domains, including (1) Internalization, (2) Thought Dysfunction, (3) Externalization, and (4) Somatic/Cognitive Complaints (Table 2). Specifically, this includes F-r (H [2] 13.21, $p < .001$), Fp-r (H [2] = 10.39, p = .006), EID (H [2] = 8.59, p = .014), RC7 (H [2] = 13.54, $p < .001$), THD (H [2] = 9.27, p = .010), RC8 (H [2] = 11.38, p = .003), JCP (H [2] = 9.72, p = .008), RC1 (H [2] = 18.15, $p < .001$), MLS (H [2] = 10.70, p = .005), and NUC (H [2] = 9.76, p = .008) differed meaningfully. Only the Interpersonal Functioning domain did not have any scales which demonstrated statistical significance.

Post hoc testing using Mann-Whitney U tests and Hedges' g effect size estimates demonstrated notable magnitudes of effect comprising differences on the scales identified above (Tables 2 and 3). In general, effect sizes were within the moderate-to-large ranges and in each case exceeded necessary levels to indicate clinically meaningful differences (Ferguson, 2009). The not admitted group had significantly higher means on F-r, Fp-r, EID, RC7, THD, RC8, RC1, MLS, NUC, and JCP scales compared to the admitted and left group (100% of identified scales) and on F-r, EID, RC7, THD, RC8, RC1, and NUC scales compared to the admitted and ordained group (70% of identified scales). Those who were admitted and left differed from those who were admitted and ordained on 40% of the identified scales, with those who were admitted left having higher means on the RC7 and NUC and those who were admitted and ordained group had higher means on the Fp-r and MLS scales (Tables 2 and 3).

To expand the interpretive utility of the MMPI-2-RF substantive scales, Relative Risk Ratios (RRR) were estimated using T55 and T60 cut scores across outcome groups (Tables 4, 5, and 6). In general, RRR analyses reveal that individuals with highest MMPI-2-RF scale scores are those with the highest likelihood of not being admitted or, ultimately, leaving the ordination program once admitted. The scales assessing internalizing and somatic/cognitive concerns were those which demonstrate the most frequent and consistent utility across scale type (i.e., Higher-Order, Restructured Clinical, and Specific Problem) and cut values. Likewise, those with scale scores below T55 regularly had the highest likelihood of positive outcomes. Specific problem scales assessing externalization and behavioral dysfunction were less likely to relate to different outcome risk at T55 levels; however, when an applicant's score was above T60 on an externalizing scale this almost always indicated comparable risk to other MMPI-2-RF scales.

Table 2

Kruskal-Wallis tests on MMPI-2-RF Validity, Higher-Order, and Restructured Clinical scales across admission group

	Admitted Ordained[1]		Admitted Left[2]		Not Admitted[3]			Hedges' g Effect Size		
	n = 15		*n* = 43		*n* = 35					
Domain/Scale	*M*	*SD*	*M*	*SD*	*M*	*SD*	*H* (2)	1 v 2	1 v 3	2 v 3
Validity										
VRIN-r	41.3	3.3	40.8	7.6	45.9	9.9	*6.99*	*.07*	*.54*	*.59*
TRIN-r	53.6	4.0	52.4	3.6	55.2	5.0	4.63	*.32*	*.36*	*.65*
F-r	**42.7**	**1.8**	**43.5**	**3.4**	**48.9** [ab]	**11.0**	**13.2***	*.26*	**.67**	**.69**
Fp-r	**46.1** [c]	**5.6**	**43.7**	**4.8**	**48.3** [b]	**8.3**	***10.4****	***.48***	*.29*	***.70***
Fs-r	42.5	2.1	45.5	6.1	49.0	15.1	*5.27*	*.56*	*.51*	*.32*
FBS-r	50.7	6.7	48.4	9.3	52.5	8.0	*3.95*	*.26*	*.24*	*.47*
RBS-r	50.7	7.6	47.0	8.4	51.9	9.6	*3.89*	*.45*	*.13*	*.55*
L-r	59.4	12.7	58.9	14.5	58.7	12.9	*0.03*	*.04*	*.06*	*.02*
K-r	60.6	6.5	58.4	13.1	57.0	10.9	*1.26*	*.19*	*.37*	*.12*
H-O										
EID	**38.6**	**6.7**	**39.0**	**8.0**	**44.1** [ab]	**9.0**	***8.6****	*.05*	**.69**	***.60***
THD	**42.6**	**5.6**	**44.4**	**6.6**	**49.9** [ab]	**10.2**	***9.3****	*.28*	**.93**	***.66***
BXD	42.3	6.7	41.3	8.4	45.5	8.4	*6.06*	*.13*	*.40*	*.50*
RC										
RCd	42.9	5.5	43.3	7.8	46.4	9.1	*3.09*	*.06*	*.43*	*.37*
RC1	**38.6**	**4.4**	**41.9**	**8.4**	**49.0** [ab]	**11.7**	***18.2****	*.43*	***1.03***	***.71***
RC2	44.3	8.2	42.0	7.3	45.9	8.2	*5.46*	*.31*	*.20*	*.51*
RC3	40.9	4.3	41.6	6.3	44.8	8.6	*3.93*	*.12*	*.69*	*.43*
RC4	42.2	6.7	42.1	8.6	47.0	7.7	*9.31*	*.01*	*.65*	*.60*
RC6	51.7	6.3	49.3	7.4	51.9	10.2	*1.55*	*.34*	*.02*	*.30*
RC7	**36.8**	**4.7**	**40.4** [d]	**5.9**	**44.6** [ab]	**8.7**	***13.5****	***.65***	***1.01***	***.57***
RC8	**43.1**	**5.7**	**45.1**	**6.7**	**51.3** [ab]	**10.4**	***11.4****	*.31*	***.88***	***.72***
RC9	40.8	6.3	41.5	8.9	44.0	9.1	*1.57*	*.08*	*.38*	*.28*

Notes: All means are standardized T-scores. *Signifies statistical significance at familywise-corrected levels. a = Mann-Whitney U Post-hoc tests indicated participants in the Not Admitted group had a significantly higher mean rank than those in the Admitted Ordained group at the $p < .05$ level. b = Mann-Whitney U Post-hoc tests indicated participants in the Not Admitted group had a significantly higher mean rank than those in the Admitted Left group at the $p < .05$ level. d = Mann-Whitney U Post-hoc tests indicated participants in the Admitted Left group had a significantly higher mean rank than those in the Admitted Ordained group at the $p < .05$ level than those in the Admitted Left group at the $p < .05$ level. c = Mann-Whitney U Post-hoc tests indicated participants in the Admitted Ordained group had a significantly higher mean rank than those in the Admitted Left group at the $p < .05$ level.

Discussion

Psychological evaluations of clergy applicants are considered a gatekeeping mechanism during the admission process. However, research has not previously identified psychological factors that may be detected during the admission evaluation process that predicts if an applicant is rejected, accepted, leaves formation, or is ultimately ordained. Such research is important to psychologists conducting evaluations of clergy applicants as it provides clear guidance about specific factors associated with admission outcomes. This study contributed to filling that noticeable gap in the extant literature by investigating the predictive validity of MMPI-2-RF scales in applicants evaluated as part of their application for seminary or diaconate formation programs in a Catholic diocese.

In general, there were meaningful MMPI-2-RF scale score differences between individuals who were accepted for those formation programs and those who were not, as well as scale score differences between those who successfully completed training programs

(e.g., became ordained clergy) and those who did not. These differences also extend from this study's subsamples and the comparison group of male clergy provided within the MMPI-2-RF Technical Manual (Tellegen & Ben-Porath, 2008/2011). Relative to the other outcome groups, the Admitted, ordained group frequently had scales with means differing more than a medium effect from the comparison group (e.g., 5T; Rosnow et al., 2000). These scores were lower than the Technical manual's comparison group as well as from those in this study who were admitted but not yet ordained, and from those who were not admitted. These findings lend some support for the vetting and formation process that ordained men have lower scores, which is a sign of achieving the stated goal of the church of ordaining men to clergy roles in good psychological health. Conversely, differences between those in our study who were not admitted, those who were admitted and left, and the Technical manual are less pronounced relative to those in our sample who are ordained. Differences across clergy applicant groups relative to one, and to the comparison group provided in the Technical Manual, emphasize the importance of considering outcome subgroups (e.g., those not admitted or those ordained) rather than applicants as a whole.

Table 3

Kruskal-Wallis tests on MMPI-2-RF Specific Problem, PSY-5, and Interest scales across

	Admitted Ordained[1]		Admitted Left[2]		Not Admitted[3]			Hedges' *g* Effect Size		
	n = 15		*n* = 43		*n* = 35					
Domain/Scale	*M*	*SD*	*M*	*SD*	*M*	*SD*	*H* (2)	1 v 2	1 v 3	2 v 3
Somatic/Cognitive										
MLS	**46.5** [c]	**6.8**	**42.7**	**7.7**	**47.4** [b]	**7.5**	***10.7****	***.50***	***.12***	***.62***
GIC	46.0	0	47.0	4.8	48.7	8.1	*2.7*	*.24*	*.39*	*.26*
HPC	45.0	6.7	45.0	6.6	49.1	8.7	*6.9*	*.00*	*.52*	*.54*
NUC	**41.8**	**3.1**	**46.2** [d]	**7.5**	**50.9** [a]	**11.8**	***9.8****	***.66***	***.90***	*.49*
COG	42.7	4.6	44.8	6.3	50.4	11.8	*8.8*	*.36*	*.75*	*.61*
Internalizing										
SUI	45.0	0	45.0	0.0	47.2	7.4	*5.1*	*.00*	*.35*	*.45*
HLP	43.2	5.5	41.7	4.8	43.9	7.5	*1.7*	*.30*	*.10*	*.36*
SFD	43.3	3.5	46.3	8.7	49.1	9.2	*6.5*	*.39*	*.73*	*.33*
NFC	43.7	5.1	43.4	4.4	48.1	9.1	*7.4*	*.06*	*.54*	*.68*
STW	42.3	7.0	42.5	7.8	47.1	9.7	*6.1*	*.03*	*.53*	*.53*
AXY	44.0	0	45.4	4.4	47.8	9.1	*3.6*	*.37*	*.51*	*.35*
ANP	41.4	4.2	42.9	5.7	46.1	7.6	*6.9*	*.28*	*.73*	*.48*
BRF	44.3	5.2	43.9	3.4	47.2	7.7	*6.5*	*.10*	*.41*	*.58*
MSF	42.4	7.2	42.7	6.4	44.4	6.5	*1.7*	*.05*	*.30*	*.26*
Externalizing										
JCP	**44.9**	**6.6**	**44.1**	**8.6**	**48.7** [b]	**8.3**	***9.7****	*.10*	*.49*	***.56***
SUB	46.2	6.0	45.7	6.1	47.3	8.2	*0.8*	*.08*	*.14*	*.23*
AGG	44.0	5.3	42.3	7.1	45.0	8.2	*2.9*	*.25*	*.13*	*.36*
ACT	41.7	6.2	45.3	9.4	45.5	9.0	*1.8*	*.41*	*.46*	*.02*
Interpersonal										
FML	42.5	5.9	43.1	7.0	44.8	7.6	*1.6*	*.09*	*.32*	*.23*
IPP	49.9	5.8	45.9	6.5	49.9	10.8	*4.4*	*.63*	*.00*	*.46*
SAV	49.1	5.5	46.6	7.2	50.2	11.2	*1.8*	*.37*	*.11*	*.39*
SHY	41.7	5.7	41.7	7.0	43.8	8.9	*0.9*	*.00*	*.26*	*.27*
DSF	45.9	4.9	45.0	3.6	48.2	9.8	*3.2*	*.23*	*.27*	*.45*
PSY-5 Scales										
AGGR	46.9	6.1	48.6	8.9	49.3	8.5	*1.1*	*.21*	*.31*	*.08*
PSYC	46.8	6.3	45.9	6.9	51.8	10.4	*7.7*	*.13*	*.53*	*.68*
DISC	44.6	6.4	44.2	9.0	48.2	7.8	*6.4*	*.05*	*.49*	*.47*
NEGE	39.0	4.9	41.4	9.0	46.1	10.9	*7.3*	*.29*	*.74*	*.48*
NTR	51.3	8.2	47.9	9.3	50.3	9.6	*1.9*	*.38*	*.11*	*.25*
Interest										
AES	42.4	7.5	44.8	10.1	48.5	9.2	*5.5*	*.25*	*.70*	*.38*
MEC	54.1	8.6	53.4	9.3	54.2	9.5	*0.1*	*.08*	*.01*	*.09*

Notes: All means are standardized T-scores. *Signifies statistical significance at familywise-corrected levels. a = Mann-Whitney U Post-hoc tests indicated participants in the Not Admitted group had a significantly higher mean rank than those in the Admitted Ordained group at the $p < .05$ level. b = Mann-Whitney U Post-hoc tests indicated participants in the Not Admitted group had a significantly higher mean rank thd = Mann-Whitney U Post-hoc tests indicated participants in the Admitted Left group had a significantly higher mean rank than those in the Admitted Ordained group at the $p < .05$ level.an those in the Admitted Left group at the $p < .05$ level. c = Mann-Whitney U Post-hoc tests indicated participants in the Admitted Ordained group had a significantly higher mean rank than those in the Admitted Left group at the $p < .05$ level.

Table 4.

Relative Risk Ratios of Higher-Order and Restructured Clinical Scales

		Non-Admission			Leave Early		
			95% CI			95% CI	
Scale	Cut-Score	RRR	Lower	Upper	RRR	Lower	Upper
Higher Order							
EID	T55	**1.61**	**1.41**	**2.01**	**1.27**	**1.16**	**1.40**
	T60	**1.70**	**1.56**	**1.96**	-	-	-
THD	T55	**1.61**	**1.45**	**1.91**	**1.28**	**1.17**	**1.41**
	T60	**1.72**	**1.60**	**1.92**	-	-	-
BXD	T55	**1.57**	**1.36**	**1.96**	1.08	.64	1.63
	T60	1.46	1.10	2.57	1.21	.57	4.60
Restructured Clinical							
RCd	**T55**	**1.65**	**1.49**	**1.94**	**1.27**	**1.16**	**1.40**
	T60	1.58	1.34	2.16	-	-	-
RC1	T55	**1.53**	**1.30**	**1.98**	**1.28**	**1.17**	**1.41**
	T60	**1.67**	**1.53**	**1.91**	**1.27**	**1.16**	**1.40**
RC2	T55	1.49	1.24	1.99	1.01	.57	1.80
	T60	1.35	.90	2.82	**1.48**	**1.08**	**3.19**
RC3	T55	**1.62**	**1.41**	**2.06**	**1.26**	**1.15**	**1.38**
	T60	**1.70**	**1.56**	**1.97**	**1.26**	**1.15**	**1.38**
RC4	T55	**1.61**	**1.43**	**1.95**	1.08	.74	1.63
	T60	-	-	-	-	-	-
RC6	T55	1.28	.96	1.85	**1.50**	**1.20**	**2.28**
	T60	1.00	.33	3.03	**1.17**	**1.04**	**1.32**
RC7	T55	**1.72**	**1.61**	**1.91**	**1.26**	**1.15**	**1.38**
	T60	**1.72**	**1.59**	**1.97**	-	-	-
RC8	T55	**1.55**	**1.36**	**1.89**	**1.25**	**1.10**	**1.70**
	T60	**1.77**	**1.68**	**1.93**	-	-	-
RC9	T55	**1.61**	**1.42**	**1.98**	**1.27**	**1.16**	**1.40**
	T60	**1.67**	**1.48**	**2.11**	-	-	-

Note. RRR = Relative Risk Ratio. CI = Confidence Interval. When a RRR has a confidence interval that crosses 1, this indicates no reliable difference in risk. Bolded values indicate those values which differ meaningfully in risk between the groups beyond the 95% CI.EID = Emotional/Internalizing Dysfunction; THD = Thought Dysfunction; BXD = Behavioral/Externalizing Dysfunction; RC = Restructured Clinical scales; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC3 = Cynicism; RC4 = Antisocial Behavior; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation.

Below we summarize specific differences between the outcome groups (i.e., admitted and not admitted; admitted-ordained and admitted-left). Scales associated with a variety of different symptom pathologies were evident across these differences with magnitudes of difference typically classified as medium to large in effect for those scales with significant differences. Two specific scales are particularly relevant to the psychological evaluations of clergy applicants: Emotional/Internalizing Dysfunction (EID) and Dysfunctional Negative Emotions (RC7). The USCCB guidelines (2015) for psychological evaluations uses the term "affective maturity" to describe a set of social and emotional characteristics ideal for a clergy applicant to gain admittance. Psychologists are guided to identify emotional deficits that are contraindications of the clergy role in the Catholic Church, which would help rule out applicants. These findings contribute to a more sensitive cut-point for clinical determinations about emotional deficits (i.e., affective immaturity) using the MMPI-2-RF that are predictive of non-admittance. Relatedly, those applicants that were admitted but left formation scored higher on RC7, which taps into anger, fear, and anxiety; emotions that would likely impede open and emotionally grounded discernment of a clergy vocation. Future research would benefit from exploring reasons for those emotions that were associated with leaving formation.

Table 5.
Relative Risk Ratios for the Somatic/Cognitive and Internalizing Scales

		Non-Admission			Leave Early		
			95% CI			95% CI	
Scale	Cut-Score	RRR	Lower	Upper	RRR		
Somatic/Cognitive							
MLS	T55	**1.56**	**1.34**	**2.00**	1.01	.57	1.80
	T60	**1.71**	**1.57**	**1.98**	**1.26**	**1.15**	**1.39**
GIC	T55	**1.55**	**1.31**	**2.05**	**1.27**	**1.17**	**1.40**
	T60	**1.55**	**1.31**	**2.05**	**1.27**	**1.17**	**1.40**
HPC	T55	**1.62**	**1.46**	**1.93**	1.01	.57	1.80
	T60	**1.66**	**1.48**	**2.07**	**1.48**	**1.08**	**3.19**
NUC	T55	**1.53**	**1.31**	**1.93**	**1.29**	**1.18**	**1.43**
	T60	**1.62**	**1.44**	**1.96**	**1.29**	**1.17**	**1.33**
COG	T55	**1.67**	**1.53**	**1.91**	**1.27**	**1.16**	**1.40**
	T60	**1.80**	**1.74**	**1.88**	-	-	-
Internalizing							
SUI	T65	**1.55**	**1.29**	**2.17**	-	-	-
HLP	T55	**1.56**	**1.34**	**2.00**	**1.26**	**1.15**	**1.38**
	T60	**1.56**	**1.34**	**2.00**	**1.26**	**1.15**	**1.38**
SFD	T55	**1.50**	**1.27**	**1.94**	**1.29**	**1.18**	**1.43**
	T60	**1.62**	**1.43**	**1.99**	**1.29**	**1.18**	**1.43**
NFC	T55	**1.53**	**1.30**	**1.98**	**1.27**	**1.16**	**1.40**
	T60	**1.71**	**1.57**	**1.97**	**1.27**	**1.16**	**1.40**
STW	T55	**1.54**	**1.32**	**1.94**	1.08	.74	1.63
	T60	**1.68**	**1.53**	**1.97**	**1.26**	**1.15**	**1.39**
AXY	T55	**1.53**	**1.30**	**1.98**	**1.28**	**1.17**	**1.41**
	T60	**1.71**	**1.57**	**1.97**	-	-	-
ANP	T55	1.05	.27	5.43	**1.26**	**1.15**	**1.38**
	T60	**1.52**	**1.16**	**3.00**	-	-	-
BRF	T55	**1.57**	**1.38**	**1.91**	1.01	.57	1.80
	T60	**1.73**	**1.60**	**1.99**		-	-
MSF	T55	1.29	.72	4.19	.50	.00	5.05
	T60	-	-	-	-	-	-

Note . RRR= Relative Risk Ratio. CI = Confidence Interval. When a RRR has a confidence interval that crosses 1, this indicates no reliable difference in risk. Bolded values indicate those values which differ meaningfully in risk between the groups beyond the 95% CI. SUI scores are presented only at T65 since a single item endorsement results in this elevation. A dash indicates that RRR was unable to be calculated due to a lack of elevations at the given T-score level for the outcome group.MLS = Malaise; GIC = Gastrointestinal Complaints; HPC = Head Pain Complaints; NUC = Neurological Complaints; COG = Cognitive Complaints; SUI = Suicide/Death Ideation; HLP = Helplessness/Hopelessness; SFD = Self-Doubt; NFC = Inefficacy; STW = Stress/Worry; AXY = Anxiety; ANP = Anger-Proneness; BRF = Behavior-Restricting Fears; MSF = Multiple Specific Fears.

Table 6.
Relative Risk Ratios for the Externalizing, Interpersonal, and PSY-5 Scales

		Non-Admission			Leave Early		
			95% CI			95% CI	
Scale	Cut-Score	RRR	Lower	Upper	RRR	Lower	Upper
Externalizing							
JCP	T55	**1.66**	**1.52**	**1.91**	1.11	.47	3.37
	T60	1.34	.87	2.93	**1.25**	**1.14**	**1.38**
SUB	T55	**1.44**	**1.17**	**1.97**	**1.54**	**1.26**	**2.26**
	T60	**1.71**	**1.57**	**1.98**	**1.23**	**1.12**	**1.35**
AGG	T55	**1.43**	**1.11**	**2.14**	1.08	.74	1.63
	T60	**1.62**	**1.41**	**2.06**	**1.26**	**1.15**	**1.38**
ACT	T55	.88	.15	2.96	-	-	-
	T60	**1.52**	**1.16**	**3.03**	**1.29**	**1.17**	**1.42**
Interpersonal							
FML	T55	1.30	.86	2.44	1.01	.57	1.80
	T60	**1.54**	**1.25**	**2.37**	**1.26**	**1.15**	**1.39**
IPP	T55	1.20	.82	1.94	1.36	.95	2.50
	T60	**1.42**	**1.14**	**1.98**	**1.52**	**1.15**	**3.13**
SAV	T55	.81	.17	2.21	1.28	.84	2.42
	T60	**1.66**	**1.52**	**1.93**	-	-	-
SHY	T55	**1.55**	**1.31**	**2.05**	**1.26**	**1.15**	**1.38**
	T60	**1.44**	**1.06**	**2.58**	**1.26**	**1.15**	**1.38**
DSF	T55	1.30	.95	2.01	1.39	.97	2.76
	T60	**1.70**	**1.56**	**1.97**	-	-	-
PSY-5							
AGGR-r	T55	1.29	.95	1.92	1.16	.94	1.45
	T60	**1.49**	**1.24**	**2.01**	1.15	.87	1.58
PSYC-R	T55	**1.51**	**1.30**	**1.90**	1.11	.47	3.37
	T60	**1.73**	**1.62**	**1.92**	-	-	-
DISC-R	T55	**1.61**	**1.43**	**1.93**	1.08	.74	1.63
	T60	1.37	.93	2.78	**1.26**	**1.15**	**1.39**
NEGE-R	T55	**1.51**	**1.26**	**2.01**	**1.28**	**1.17**	**1.41**
	T60	**1.71**	**1.57**	**1.98**	**1.28**	**1.17**	**1.41**
INTR-R	T55	**1.03**	**.56**	**1.95**	1.28	.84	2.42
	T60	.72	-.02	2.46	.95	.60	1.46

Note . RRR= Relative Risk Ratio. CI = Confidence Interval. Bolded values indicate those values which differ meaningfully in risk between the groups. JCP = Juvenile Conduct Problems; SUB = Substance Abuse; AGG = Aggression; ACT = Activation; FML = Family Problems; IPP = Interpersonal Passivity; SAV = Social Avoidance; SHY = Shyness; DSF = Disaffiliativeness; AGGr-r = Aggressiveness–Revised; PSYC-r = Psychoticism–Revised; DISC-r = Disconstraint–Revised; NEGE-r = Negative Emotionality/Neuroticism–Revised; INTR-r = Introversion/Low Positive Emotions.

For example, candidates that were anxious about classwork, angry with authority figures in formation, or fearful of speaking up about sexual harassment may have all left formation after admittance due to emotional dysfunction, but for very different contexts of their emotions.

In using the MMPI-2-RF with clergy applicants, one challenge frequently encountered by psychologists is that substantive scale scores are often below traditional recommended cut values and have elevated K-r and L-r scales (T < 65) (see Isacco et al., 2020b). Indeed, respondents in this sample frequently scored around 10 points lower than the normative sample with a standard deviation approximately half that of the normative sample. Regardless of outcome group, means observed across each group regularly reflected this pattern of below normative sample performance. This pattern of low scores suggests range restriction for this population and more reliance on subjective data to base clinical recommendations. Calculating relative risk ratios at T65 cut-scores was, therefore, limited because of the restricted endorsement on MMPI-2-RF items. Despite lower scale scores and probable range restriction, a number of the MMPI-2-RF scales retain their capacity to meaningfully differentiate between outcome groups (admitted-ordained, admitted-left, and not-admitted), with medium to large effects.

In the case that an individual was to exceed the recommended cut score for a MMPI-2-RF substantive scale (see Tellegen & Ben-Porath, 2008/2011), such performance would be highly atypical and indicative of substantial impairment in functioning for the associated domain. However, this is not likely to be a regular occurrence. As a result, psychologists will be forced to make determinations about clinical implications of scale scores that fall below recommended cut values. Scales on which means differences emerged in this study are likely the best indicators for use in psychological assessments of clergy applicants given their demonstrated utility. At present, there is no research on relationships between MMPI-2-RF scale scores and clergy specific performance metrics. Thus, it would benefit psychologists if research could target expanded interpretation of MMPI-2-RF scales with criteria that are of interest and relevance to seminary or diaconate applicants (e.g., Isacco et al., 2020a). In this research, it will be important for researchers to account for attenuated relationships due to restricted ranges that emerge, at least in part, due to the admission context in which applicants complete the MMPI-2-RF.

Related to the probable restriction of range in this population, clergy applicants may be motivated to engage in positive impression management that contributes to deflation of observed scores. This pattern of lower problem endorsement is common in, and has

been observed for, samples in similar admission contexts such as public service professions (Sellbom et al., 2007; Tarescavage et al., 2015). Indeed, frequent low scores and the associated concerns (discussed above) have led to adapted uses of test scores such as adjusting interpretive scale cutoffs. To facilitate decision making capacity across alternative scale cut scores, the estimated relative risk ratios in this study for each of the MMPI-2-RF substantive scales provides an expanded context for interpretation. In general, scales have small increases in risk for non-desired outcomes (i.e., non-admission or leaving after admission) across evaluated alternative cut-score. Elevations on the MMPI-2-RF substantive scales at lowered thresholds (e.g., T55 and T60) remain associated with less desirable outcomes (i.e., admission or completion of training), particularly as it relates to constructs assessing affective maturity. For instance, relative risk ratios calculated across the various Somatic/Cognitive scales consistently suggested that those who scored above the alternative cut-scores (i.e., T55 and T60) were approximately two third more likely to not be admitted whereas on the same scales they were often one thirds less likely to become ordained if admitted. The most consistent scales for predicting risk include those assessing Somatic/Cognitive concerns as well as Internalizing pathology (e.g., EID, RC7, SFD), although there are some notable exceptions to the later (e.g., RC2). As one might expect, proximal outcomes (admission) have larger risk estimates than distant outcomes, which are measured several years after the psychological assessment is conducted (ordainment).

Our findings are important to consider in the conceptualization of clergy applicants. Clergy acting out behavior (e.g., sexual offenses) receives most of the attention in the popular media and psychological literature. It is rare for clergy applicants to endorse externalizing behavior such as aggression, juvenile conduct problems, and substance use on the MMPI-2-RF despite the high associated impact. Yet, the internalizing concerns are more measurable at the admission stage and are of comparative importance as the Catholic Church has emphasized the need for clergy to be grounded in their emotions, capable of managing the stresses of ministry, and secure in their psychological identities. Psychologists should not discount elevated internalizing scales for potential risk of acting out as emergent findings have pointed to a relationship between emotional deficits and risk factors of sexual offending, as measured on the MMPI-2-RF (Isacco et al., 2020a). In general, the proportion of individuals with positive outcomes declines as scores increase, supporting the utility of many MMPI-2-RF scales when using alternative cut-values. Thus, our data provides support for psychologists to use lower threshold cut scores than the typical T65 with clergy applicants in the same vein as public servant candidates. Psychologists are advised to pay particular attention to even mild evaluations in substantive scales of the MMPI-2-RF among clergy applicants. For example, RC3 (Cynicism) emerged as a significantly impactful scale of negative outcomes in this study. Cynical thinking that views human nature in a negative light, distrusts others, and has manipulative and exploitative qualities are antithetical to the desired characteristics of Catholic clergy. It is suggested that psychologists emphasize the preference for very low scores on RC3 among admitted clergy applicants. Given the lower thresholds and range restriction, the use of additional assessment tools (e.g., a clinical interview) and psychological tests may also be helpful to further identify and clarify the presence of a psychological concern with clergy applicants that may impact admission decisions (Isacco et al., 2020b). A multi-modal approach to psychological evaluations is consistent with Church guidelines (USCCB, 2015) and clinical guidance for robust psychological assessments (Wright, 2011).

This study also has some limitations that warrant discussion. First, our study includes both seminary and diaconate applicants. We did not have a sufficient sample size to compare these groups even though they have some notable distinctions between them, despite their similarities. Second, this study was unable to examine differential relationships between the MMPI-2-RF for individuals that were asked to leave formation from those who left on their own accord. This limitation reflects our not having criterion information available about applicants from their time in formation, other than program admission and completion status. Thus, studies examining criterion information about behavior during formation, or extended data about reason(s) for leaving, would be useful in catering recommendations during applicant evaluation.

We also acknowledge that there may be some concern about criterion contamination in the event that the MMPI test data was used in non-admission decisions. However, it is important to contextualize our study within the admission process of clergy applicants. The MMPI data is one piece of a comprehensive psychological evaluation and the psychological evaluation is one component of an extensive admission process that in-

volves other data sources (e.g., letters of recommendation, background checks). The final admission decision is rendered by the Bishop, informed by all of the data sources and the diocesan admission committee recommendation. The evaluating psychologist is not part of the admission committee and only supplies the psychological report for their consideration. Thus, criterion contamination is unlikely to invalidate our findings. We acknowledge that the clinical utility of this study, with MMPI-2-RF data, may be lessened by the newly released MMPI-3. However, practitioners are known to be slow in the transition to new tests (see McGlone et al., 2010) and it is equally likely that use of the MMPI-2-RF will continue until the published evidence on the MMPI-3 allows for broader utilization in psychological evaluations. Future research with clergy applicants with the MMPI-3 is encouraged. Lastly, our sample size was relatively small for those individuals who were admitted and became ordained, which limited our capacity to detect meaningful relationships. In order to address this concern, we utilized Hedge's g as an effect estimate because of its sample size correction that helps account for small and unequal effects, effects which provide clear indication of meaningful differences between these groups with consistent medium to large effects. However, further study and replication of our findings are warranted to better identify predictive factors of ordination, which is the ultimate outcome of the application and formation process

The above limitations notwithstanding, this study is the first to examine the MMPI-2-RF in predicting prospective, admission-based outcomes for Catholic clergy applicants. Our results provide support for using the MMPI-2-RF within these evaluations. Our results also provide support for the MMPI-2-RF substantive scales. We offer a needed contrast between outcomes of clergy applicants that is currently missing from the existing literature and provide guidance on scale interpretation of outcome risk. Lastly, our research highlights the importance of further study on this population and to provide expanded interpretation of population relevant criterion. For example, expected differences in interpersonal scales were not found in this study but a clergy role is inherently social. Thus, identification of interpersonal indicators of admission rejection or ordination would be valuable information to the extant literature.

References

Ben-Porath, Y. S. (2012). Interpreting the MMPI-2-RF. Minneapolis: Minnesota Press.

Corey, D. M., & Ben-Porath, Y. S. (2018). Assessing police and other public safety personnel using the MMPI-2-RF: A practical guide. University of Minnesota Press.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. https:// https://doi.org/10.1037/h0040957

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40, 532–538. https://doi.org/10.1037/a0015808

Ferrari, J.R. (2017). Called and formed: Personality dimensions and leadership styles among Catholic deacons and men in formation. Pastoral Psychology, 66, 225-237. DOI: 10.1007/s11089-016-0741

Gautier, M. L. (2018). The Class of 2018 Survey of ordinands to the priesthood: A Report to the secretariat of clergy, consecrated life & vocations for the United States Conference of Catholic Bishops. CARA. Washington, D.C.: Georgetown University Press.

Ingram, P. B., Tarescavage, A. M., Ben-Porath, Y. S., Oehlert, M. E., & Bergquist, B. K. (2020). External correlates of the MMPI-2-Restructured Form across a national sample of veterans. *Journal of Personality Assessment.* doi: 10.1080/00223891.2020.1732995

Isacco, A., Ingram, P., Finn, K., Dimoff, J., & Gebler, B. (2020). A novel approach to examining personality risk factors of sexual offending in clergy applicants. Spirituality in Clinical Practice, 7(4), 246-261. https://doi.org/10.1037/scp0000224

Isacco, A., Finn, K., Tirabassi, D., & Meade, K., & Plante, T. (2020). Are Catholic clergy applicants psychologically healthy? *Spirituality in Clinical Practice*, 7(4), 230-245. https://doi.org/10.1037/scp0000229

John Jay College of Criminal Justice. (2011). The causes and con text of sexual abuse of minors by Catholic priests in the United States, 1950–2010. Washington, DC: United States Conference of Catholic Bishops.

John Jay College Research Team (2004). The nature and scope of sexual abuse of minors by Catholic priests and deacons in the United States, 1950-2002. Washington, DC: United States Conference of Catholic Bishops (USCCB).

McGlone, G. J., Ortiz, F. A., & Karney, R. J. (2010). A survey study of psychological assessment practices in the screening and ad mission process of candidates to the priesthood in the U.S. Catholic Church. *Professional Psychology: Research and Practice*, 41(6), 526–532. https://doi.org/10.1037/a0021546

Pew Research Center, (2018, October). 7 Facts about US Catholics. Retrieved from https://www.pewresearch.org/fact-tank/2018/10/10/7-facts-about-american-catholics/

Plante, T. G., & Lackey, K. (2007). Are successful applicants to the Roman Catholic deaconate psychologically healthy? *Pastoral Psychology*, 55(6), 789–795. https://doi.org/10.1007/s11089-007-0081-2

Plante, T. G., Lackey, K., & Louie, C. (2005). Are successful appli cants to the Roman Catholic deaconate psychologically healthy? Pastoral Psychology, 55(6), 789–795. https://doi.org/10.1007/s11089-007-0081-2

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological Science*, 11, 446–453.

Sellbom, M., Fischler, G., & Ben-Porath, Y. (2007). Identifying MMPI-2 predictors of police officer integrity and misconduct. *Criminal Justice and Behavior*, 34, 985-1004. 10.1177/0093854807301224.

Sunardi, Y. (2014). Predictive factors for commitment to the priestly vocation: A study of priests and seminarians. [Un published doctoral dissertation]. Marquette University, Milwaukee, Wisconsin.

Tarescavage, A.M., Corey, D.M., & Ben-Porath, Y.S. (2015). Minnesota Multiphasic Personality Inventory–2–Restructured Form (MMPI-2-RF) Predictors of Police Officer Problem Behavior. Assessment, 22(1), 116-132.

Tarescavage, A. M., Alosco, M. L., Ben-Porath, Y. S., Wood, A., & Luna-Jones, L. (2015). Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) Scores generated from the MMPI-2 and MMPI-2-RF test booklets: Internal structure comparability in a sample of criminal defendants. *Assessment*, 22(2), 188–197. https://doi.org/10.1177/1073191114537347

Tellegen, A., & Ben-Porath, Y. S. (2008/2011). The Minnesota Multiphasic Personality Inventory-2 Restructured Form: Technical manual. Minneapolis, MN: University of Minnesota Press.

USCCB (2015). Guidelines for the use of psychology in seminary admissions. Washington, DC: United States Catholic Conference of Bishops.

Wong, A.C.K. (2014) What factors help seminary students continue in their academic programs? *Journal of Adult Theological Education*, 11(2), 150- 164,

Wright, A.J. (2011). Conducting psychological assessment: A guide for practitioners. New