



Preferences of BS Social Work Graduates for Board Examination Review Centers: A Conjoint and Cluster Analysis

Clyde Abner Valdez¹; Lea Mae Baloro-Swindler²; Elizabeth Pasion²

¹College of Business Administration Education, University of Mindanao, Davao City, Philippines

²Jose Maria College Foundation, Davao City, Philippines

Corresponding Email: elizabeth.pasion@jmc.edu.ph

Article History:

Received: March 27, 2025

Accepted: June 6, 2025

Online First: June 30, 2025

Keywords:

conjoint analysis; cluster analysis; review center; social work graduates; consumer preference; Philippines

This is an open access article under the CC-BY-NC-ND license.



Abstract: This study examined the preferences of Bachelor of Science in Social Work (BSSW) graduates for board examination review centers and explored whether distinct preference profiles exist within this market. A quantitative design using conjoint analysis was employed. Data were gathered from 149 BSSW graduates in Davao City who scored hypothetical review center profiles built from five attributes: track record, reviewers, review materials, conduciveness, and affordability. Part-worth utilities and relative importance values were estimated, and k-means cluster analysis was applied to individual utilities to explore preference heterogeneity. Affordability was the most important attribute (34.02%), followed by conduciveness (23.95%), reviewers (14.39%), track record (14.00%), and review materials (13.63%). Respondents most preferred a fee of ₱10,000 and below, a capacity of 100 examinees and above, a track record above the national passing rate, and review materials updated within the past three years. Cluster analysis identified three preference profiles: Experienced-Reviewer, High-Capacity Seekers (67.8%); Moderate-Capacity, Recent-Materials Seekers (14.8%); and Moderate-Capacity Seekers with Mixed Preferences (17.4%). These profiles should be treated as descriptive typologies rather than validated segments because holdout validation was not achieved. The study concludes that price is the dominant consideration for most BSSW graduates, although a meaningful minority values instructional quality over cost. Review centers may use these descriptive patterns to consider differentiated and affordable review packages, while future research should address the methodological limitations identified.

How to cite:

Valdez, C.A., Baloro-Swindler, L.M., & Pasion, E. (2025). Preferences of BS Social Work graduates for board examination review centers: A conjoint and cluster analysis. *Business and Organization Studies e-Journal*, 3(2), 90-110.



INTRODUCTION

The Social Worker Licensure Examination (SWLE) in the Philippines functions as the statutory gatekeeping mechanism for entry into the social work profession. Under Republic Act No. 4373, graduates of Bachelor of Science in Social Work (BSSW) programs must pass this examination administered by the Professional Regulation Commission (PRC) before they may practice. Recent data reveal that the examination is highly competitive. In September 2024, only 4,587 of 7,113 examinees passed, yielding a national passing rate of 64.49% (PRC Board News, 2024). Earlier examinations recorded even lower rates: 37.75% in August 2021 and 33.60% in February 2021. These figures demonstrate that licensure is neither automatic nor uniform, creating substantial pressure on graduates to prepare adequately.

In response to this pressure, enrollment in review centers has become a customary step for many BSSW graduates. A qualitative study of an online review program for social work graduates at the International School of Asia and the Pacific found that "one of the reasons students spent time in review centers was due to the fact that board exams are essential to the professional journey of their chosen career" and that "review programs are an important source of education and training for people who desire to advance in their careers" (Liban et al., 2024, p. 1). These findings establish the importance of review centers in the licensure pathway while leaving open the question of how graduates choose among them.

The financial dimension of this choice is particularly salient. A 2024 study of third-year BSSW students at a state university in Claveria, Misamis Oriental, found that most students were unprepared across the three SWLE subject areas: Human Behavior and Social Environment, Social Welfare Policies and Programs, and Social Work Methods (Caballero et al., 2024). The study reported significant differences in student performance across subject areas and recommended reviewing curriculum sequencing and implementing ongoing evaluation and faculty development programs to better prepare students for the SWLE (Caballero et al., 2024). For graduates with limited financial resources, the cost of review center fees may constitute a significant barrier to access. A Senate bill filed in the Twentieth Congress observed that "one of the key challenges faced by graduates of State Universities and Colleges (SUCs) in the Philippines is access to affordable and quality review programs for professional licensure examinations" and that "many graduates are unable to pursue professional licensure due to the high cost of review centers" (Senate Bill No. 153, 2025, p. 1). This legislative attention, while normative rather than empirical, signals that the affordability of review services is a recognized policy concern.



The regulatory context further shapes this market. Executive Order No. 566 directed the Commission on Higher Education (CHED) to regulate review centers, and CHED Memorandum Order No. 30, s. 2007 established the implementing rules. The memorandum recognizes that review centers "play a role on the performance of graduates of higher education institutions...in the outcomes of licensure examinations" while warning that without regulation, "the provision of supplemental educational service may lead to the deterioration of quality and unethical practices" (CHED, 2007, Sections 1-2). This regulatory attention underscores the public interest in how review centers operate and how graduates select them.

Despite this, little is known about what BSSW graduates actually look for when choosing a review center. Research on this population has focused mainly on academic preparation and examination performance (Caballero et al., 2024), not on graduates as consumers of review services. A parallel literature on review center preferences in other Philippine professions offers instructive comparisons. Castillo and Yee (2023) applied conjoint analysis to accountancy graduates' preferences for CPA review centers in Davao City and found that "track record" was the most important attribute, followed by "reviewers," while "affordability" was the least important. Their preferred review center profile included a track record with board topnotchers and an above-average national passing rate, reviewers with at least 15 years of experience, a conducive venue accommodating fewer than 50 reviewees, comprehensive review materials updated within three years, and a fee of ₱10,000 below (Castillo & Yee, 2023). This contrasts sharply with what might be expected for BSSW graduates, given the financial constraints documented in the social work literature (Caballero et al., 2024; Senate Bill No. 153, 2025). This contrast highlights the need for discipline-specific evidence, as preferences may differ systematically between accountancy graduates—who typically enter higher-paying professions—and social work graduates, who often begin careers in the public or non-profit sectors. No published study appears to have examined review center preferences among BSSW graduates in the Philippines, leaving review center operators, schools of social work, and regulators without evidence on how this market makes its decisions.

To address this gap, this study aimed to determine the preferences of BSSW graduates for review center attributes and to explore whether distinct preference profiles exist among them. Specifically, it sought to answer the following questions: (1) What is the relative importance of track record, reviewers, review materials, conduciveness, and affordability in graduates' choice of a review center? (2) Which levels of each attribute do graduates prefer? (3) What distinct preference profiles emerge from the utility data? The findings are intended to help review centers understand what graduates value, to help schools of social work consider the financial



pressures that shape graduates' preparation for licensure, and to contribute to a nascent literature on review center markets in Philippine professional education.

THEORETICAL FRAMEWORK

This study integrates four theoretical perspectives to explain how BSSW graduates evaluate and choose among review centers.

First, Lancaster's (1966) characteristics theory of value provides the foundation for treating review centers as bundles of attributes rather than indivisible goods. Under this view, graduates derive utility from the specific characteristics of review centers—track record, reviewer experience, material recency, venue capacity, and fee—and their preferences reflect trade-offs among these characteristics. Conjoint analysis operationalizes this idea by asking respondents to evaluate whole profiles and then decomposing their judgments into part-worth utilities for each attribute level (Green & Srinivasan, 1978, 1990; Luce & Tukey, 1964).

Second, the bounded rationality framework (Simon, 1955) and the adaptive decision-maker perspective (Payne, Bettman, & Johnson, 1993) clarify how graduates make choices under the constraints they face. Simon argued that decision-makers do not optimize across all available information but instead "satisfice" – selecting options that meet aspiration levels rather than maximizing utility. Payne et al. extended this by showing that decision strategies are contingent on task demands and individual resources: when cognitive effort is costly and time is limited, decision-makers adopt simplifying heuristics, such as eliminating options that exceed a price threshold before evaluating remaining attributes. This framework is particularly apt for BSSW graduates who are described in the literature as financially constrained and making a high-stakes, one-off decision (Caballero et al., 2024; Senate Bill No. 153, 2025). The prediction that follows is that affordability will function as a screening criterion, with graduates first eliminating options above a fee threshold and only then differentiating among the remaining options by instructional attributes.

Third, prospect theory (Kahneman & Tversky, 1979) offers a more specific explanation for why a discrete fee threshold might produce such a sharp utility split. Prospect theory holds that people evaluate outcomes relative to a reference point, and that losses loom larger than equivalent gains. If ₱10,000 functions as a reference point—perhaps because it represents a psychologically meaningful boundary between what a graduate can afford and what would require additional borrowing or family support—then preferences should shift discontinuously at that threshold. This would explain why the utility difference between "₱10,000 and below" and "₱10,001 and above" is so large relative to other attribute-level differences.



Fourth, Smith's (1956) market segmentation theory and Wedel and Kamakura's (2000) elaboration of segmentation principles provide the conceptual basis for exploring whether the BSSW review center market is heterogeneous. If graduates differ systematically in their values and constraints, then distinct preference profiles should emerge from the utility data. Punj and Stewart (1983) provided methodological guidance for cluster analysis in marketing research, noting that the technique is useful for identifying segments based on multivariate preference data such as conjoint utilities, while cautioning that cluster solutions require validation against external criteria and stability across resampling.

These theoretical perspectives jointly predict that (a) affordability will be a significant, and possibly dominant, attribute in BSSW graduates' review center choices; (b) fee thresholds will produce discontinuous rather than linear utility differences; (c) preferences will vary across graduates, reflecting different resource constraints and values; and (d) the resulting profiles will be distinguishable by their utility patterns across the five attributes. The empirical analysis that follows tests these predictions while acknowledging the limitations of the data and design.

METHODOLOGY

Research Design

This study employed a quantitative, cross-sectional research design using conjoint analysis to examine the preferences of Bachelor of Science in Social Work (BSSW) graduates for board examination review centers. Conjoint analysis is a decompositional method that estimates the relative contribution of individual product attributes to overall preference by asking respondents to evaluate multi-attribute profiles (Green & Srinivasan, 1978, 1990). The method is appropriate for this study because it allows estimation of the trade-offs graduates make among review center features that are not traded off in actual markets—that is, review centers typically bundle attributes in ways that confound which features drive choice. A second analytical phase used *k*-means cluster analysis to explore whether respondents could be grouped into distinct preference profiles based on their individual part-worth utilities (Punj & Stewart, 1983).

Respondents and Sampling

The respondents were 149 BSSW graduates from higher education institutions in Davao City, Philippines. All respondents were preparing for, or had recently prepared for, the Social Worker Licensure Examination (SWLE). Respondents were selected through purposive sampling. The inclusion criteria were: (a) completion of the BSSW degree, and (b) intent to take, or recent experience of taking, the SWLE.



Graduates who had no plan to take the licensure examination were excluded so that responses reflected the preferences of actual prospective consumers of review services. The sample size of 149 is consistent with recommendations for conjoint analysis with a moderate number of attributes and levels (Hair et al., 2019). For the cluster analysis phase, the ratio of cases to clustering variables was approximately 13.5:1, which is within the acceptable range for exploratory segmentation but on the lower end for stable cluster recovery (Dolnicar & Leisch, 2010). This limitation is discussed further in the Limitations section.

Conjoint Design

The conjoint design was built on five attributes identified from the review center literature, the Philippine regulatory context, and consultation with BSSW graduates and review center operators. Table 1 summarizes the attributes and their levels.

Table 1. *Conjoint attributes and levels*

Attribute	Operational definition	Levels
Track record	Center's passing rate relative to the national passing rate	(1) Above national passing rate; (2) Equal to national passing rate
Reviewers	Length of experience of the center's lecturers	(1) Running 10 years; (2) More than 10 years
Review materials	Recency of the materials used in review	(1) Relevant in the past 5 years; (2) Relevant in the past 3 years
Conduciveness	Number of reviewees the venue could accommodate	(1) 50 or fewer; (2) 51 to 100; (3) 100 and above
Affordability	Review fee	(1) ₱10,000 and below; (2) ₱10,001 and above

Note. The "conduciveness" attribute was operationalized as venue capacity.

The full factorial design would have consisted of $2 \times 2 \times 2 \times 3 \times 2 = 48$ profiles. An orthogonal main-effects plan was generated in IBM SPSS Statistics using the Orthoplan procedure. The orthogonal plan produced a reduced set of profiles that allowed estimation of all main effects while minimizing respondent burden. Holdout profiles were included to assess the predictive validity of the estimated model; these profiles were not used in utility estimation. Each profile was presented as a card describing one hypothetical review center. Respondents scored each profile on a scale where higher scores indicated greater preference. Scoring tasks require respondents to evaluate each profile independently rather than to compare and order all profiles against one another. This reduces the cognitive burden associated with full-profile



ranking, which can become demanding as the number of profiles increases (Louviere et al., 2000; Orme, 2010). Because respondents can assign equal scores to different profiles, scoring also avoids the artificial constraint that ranking imposes when profiles are substantively similar, and it allows respondents to express intensity of preference rather than only relative order.

The scoring approach nevertheless carries its own methodological risks. Rating scales are subject to scale usage heterogeneity: respondents differ systematically in how they use the scale (for example, some concentrate responses near the endpoints while others cluster near the midpoint), and these differences are unrelated to the attribute preferences the study aims to measure (Rossi, Gilula, & Allenby, 2001). Such heterogeneity can inflate the model constant, which in turn can affect individual-level utility estimates and any subsequent clustering of those estimates. The decision to exclude the constant from the clustering variables, described below, was made partly for this reason.

Data Collection Procedure

Data collection was carried out after securing permission from Jose Maria College Foundation. The questionnaire was administered using a Meta-boosted online survey from January to May 2024. Voluntary completion of the questionnaire constituted consent to participate. The survey instrument consisted of three sections: (a) a demographic and eligibility screening section, (b) the conjoint scoring task, and (c) a brief set of questions on review center enrollment intentions. Only the conjoint scoring task is analyzed in this paper.

Ethical Considerations

Ethical standards for research with human participants were observed throughout the study. Participation was voluntary, and respondents could withdraw at any time without penalty. No identifying information was collected. All data were analyzed and reported in aggregate form to protect respondents' anonymity and confidentiality. The study did not involve deception, physical risk, or sensitive personal information.

Analytical Strategy

The analysis proceeded in two phases.

Phase 1: Conjoint analysis. The SPSS CONJOINT procedure was used to estimate part-worth utilities for each attribute level and the relative importance of each attribute. Utilities were estimated for the aggregate sample and for each individual respondent. The additive part-worth model can be expressed as:

$$U_{ij} = \beta_0 + \sum_a \sum_l \beta_{al} X_{ijal} + \varepsilon_{ij}$$



where U_{ij} is the estimated utility of respondent i for profile j , β_0 is the constant (intercept), β_{al} is the part-worth utility for level l of attribute a , X_{ijal} is an indicator variable equal to 1 if profile j contains level l of attribute a for respondent i and 0 otherwise, and ε_{ij} is the error term. Attributes were effects-coded so that part-worth utilities within each attribute sum to zero. Model fit was assessed with Pearson's r and Kendall's τ between observed and estimated scores. Predictive validity was assessed with Kendall's τ for the holdout profiles.

Relative importance values were computed as the range of utilities within each attribute, divided by the sum of ranges across all attributes, and multiplied by 100. Importance values are therefore normalized to sum to 100 within each respondent. It should be noted that attributes with more levels or larger utility ranges will tend to show higher importance by construction (Cattin & Wittink, 1982); this caveat is addressed in the Limitations section.

Phase 2: Cluster analysis. *k*-means cluster analysis was applied to the individual part-worth utilities to explore preference heterogeneity. The clustering variables consisted of the attribute levels that were estimated in the conjoint model: equal and above for track record; running 10 years and more than 10 years for reviewers; relevant in the past five years and relevant in the past three years for review materials; can accommodate 50 below, 51 to 100, and 100 to up for conduciveness; and fee below 10,000 and fee below 10,001 up for affordability. The model constant was excluded from the clustering variables because it reflects scale usage and overall response level rather than attribute preference structure; including it would separate respondents by response style rather than by preference (Punj & Stewart, 1983).

A three-cluster solution was specified for the analysis based on the exploratory purpose of the study and the interpretability of the resulting profiles. Because the clusters were formed from the same utility estimates used to describe them, the resulting profiles are descriptive summaries of the sample rather than inferential estimates of population segments.

Between-cluster differences were described with one-way analysis of variance (ANOVA). The F tests are interpreted descriptively only because the clustering procedure maximizes differences among cases in different clusters; the significance levels are not corrected for this and cannot be interpreted as tests of the hypothesis that cluster means are equal (Punj & Stewart, 1983). The ANOVA results are reported to show which attributes contributed most to distinguishing the clusters, not to test hypotheses about population differences.

Clusters were labeled according to the attribute levels with the largest utility differences within each cluster. Labels are interpretive and should be treated as



descriptive summaries rather than validated segment names. Given the holdout validation results reported below, the clusters are best characterized as exploratory preference profiles rather than stable market segments.

Treatment of Missing Data

Listwise deletion was used for cases with missing data on any clustering variable. There were no missing data on any variable used in the analysis.

Software

All analyses were conducted in IBM SPSS Statistics Version 20. The CONJOINT procedure was used for utility estimation, and the QUICK CLUSTER procedure was used for *k*-means clustering.

RESULTS AND DISCUSSION

Conjoint analysis was conducted to estimate the part-worth utilities of each attribute level and the relative importance of each attribute. Table 2 presents the utilities for the aggregate model and for three individual respondents whose models illustrate the range of preference patterns observed in the sample.

As shown in Table 2, the strongest preference in the aggregate model was for a fee of ₱10,000 and below (0.662) over a fee of ₱10,001 and above (−0.662). This utility difference of 1.324 is the largest of any attribute-level pair in the model. For conduciveness, respondents most preferred a venue accommodating 100 reviewees and above (0.358), followed by 51 to 100 (0.154). A venue accommodating 50 or fewer carried a strong negative utility (−0.512). Respondents also preferred a track record above the national passing rate (0.289) over one equal to it (−0.102), and review materials relevant in the past three years (0.262) over those relevant in the past five years (−0.233). For reviewers, both levels carried negative utilities (−0.183 for running 10 years and −0.123 for more than 10 years), which suggests that respondents did not strongly differentiate between the two reviewer experience levels, or that both levels were evaluated less favorably than some unmeasured reference condition.

The dominance of the lower fee is consistent with the bounded rationality and adaptive decision-maker frameworks. Under conditions of financial constraint, decision-makers often adopt non-compensatory strategies in which an attribute with a clear cutoff (such as a maximum affordable fee) functions as a screening criterion, eliminating options that exceed the threshold before other attributes are evaluated (Payne et al., 1993; Simon, 1955). The magnitude of the utility difference at the ₱10,000 threshold—1.324, compared with utility differences of less than 0.60 for all other attribute-level pairs—is consistent with the prospect theory prediction that a



psychologically meaningful reference point should produce a discontinuous shift in preferences (Kahneman & Tversky, 1979). New graduates are typically not yet earning, and many depend on family support to cover both the review fee and the costs of relocating to a review hub. This makes a fee threshold highly consequential.

Table 2. *Part-worth utilities for the aggregate model and selected individual respondents*

Attribute and level	Overall	Subject 4	Subject 104	Subject 130
Track record				
Above national passing rate	0.289 (0.662)	0.250 (0.771)	0.250 (0.535)	0.250 (0.535)
Equal to national passing rate	-0.102 (0.662)	0.000 (0.771)	-0.250 (0.535)	-0.250 (0.535)
Reviewers				
Running 10 years	-0.183 (0.565)	-0.083 (0.658)	-0.167 (0.456)	-0.167 (0.456)
More than 10 years	-0.123 (0.662)	-0.208 (0.771)	0.208 (0.535)	0.208 (0.535)
Review materials				
Relevant in the past 5 years	-0.233 (0.662)	-0.167 (0.771)	-0.292 (0.535)	-0.292 (0.535)
Relevant in the past 3 years	0.262 (0.662)	0.083 (0.771)	0.458 (0.535)	0.458 (0.535)
Conduciveness (capacity)				
50 or fewer	-0.512 (0.565)	-0.417 (0.658)	-0.333 (0.456)	-0.333 (0.456)
51 to 100	0.154 (0.662)	0.208 (0.771)	0.417 (0.535)	0.417 (0.535)
100 and above	0.358 (0.662)	0.208 (0.771)	-0.083 (0.535)	-0.083 (0.535)
Affordability (fee)				
₱10,000 and below	0.662 (0.424)	0.813 (0.493)	0.250 (0.342)	0.250 (0.342)
₱10,001 and above	-0.662 (0.424)	-0.813 (0.493)	-0.250 (0.342)	-0.250 (0.342)
Constant	4.387 (0.509)	4.104 (0.593)	4.167 (0.411)	4.167 (0.411)

Note. N = 149 for the overall model. Values are part-worth utilities, with standard errors in parentheses. Higher values indicate greater preference within an attribute.



The preference for large-capacity venues is less intuitive, since smaller classes are often associated with better instruction. The accountancy conjoint study found the opposite: accountancy graduates preferred venues accommodating fewer than 50 reviewees (Castillo & Yee, 2023). One plausible interpretation, consistent with the accountancy findings, is that BSSW graduates may interpret a large enrollment as a signal of the center's popularity and credibility, a form of social proof that reduces the risk of choosing an unfamiliar provider (Kotler & Keller, 2016). Competing explanations should also be considered: large centers may offer more schedule options, may be perceived as more established, or may have better facilities. The data in this study cannot adjudicate among these interpretations, and future research should examine the reasons behind capacity preferences directly.

The preference for recently updated materials likely reflects graduates' awareness that examination coverage changes over time, which makes currency a practical concern. The relatively low importance of track record deserves attention, since review centers commonly advertise their passing rates. The result does not mean that graduates disregard outcomes. Rather, when centers are compared across the range of options in this study, differences in track record appear to shift preferences less than differences in fee or capacity. This contrasts sharply with the accountancy findings, where track record was the most important attribute (Castillo & Yee, 2023). The divergence between accountancy and social work graduates may reflect differences in expected starting salaries and employment prospects: accountancy graduates may anticipate higher earnings and therefore weigh track record more heavily, while social work graduates facing lower expected incomes may prioritize affordability as a more immediate constraint. This contrast is one of the most substantively interesting findings of the study and warrants further investigation.

The individual models illustrate how much preferences varied beneath the aggregate pattern. Subject 4's preferences closely mirrored the sample's emphasis on fee (0.813). Subjects 104 and 130 instead showed their strongest preferences for recent materials (0.458) and moderate capacity (0.417), while placing comparatively little value on fee differences (0.250). This heterogeneity is explored further in the cluster analysis below.

As seen in Table 3, affordability was the most important attribute in the aggregate model (34.02%), accounting for about a third of the total variation in preference. It was followed by conduciveness (23.95%). Reviewers (14.39%), track record (14.00%), and review materials (13.63%) carried similar, smaller weights. This ordering suggests that graduates first screen review centers by cost and venue capacity, and only then differentiate among them by instructional attributes. The finding is consistent with the accountancy conjoint study, which found that practical considerations weighed heavily in students' choice of review centers, though their



respondents prioritized track record over affordability (Castillo & Yee, 2023). It also supports the observation that educational consumers weigh price and convenience heavily, even in decisions with high academic stakes (Hemsley-Brown & Oplatka, 2006).

Table 3. *Relative importance of review center attributes*

Attribute	Overall	Subject 4	Subject 104	Subject 130
Track record	14.00	14.29	17.39	17.39
Reviewers	14.39	14.29	13.04	13.04
Review materials	13.63	7.14	26.09	26.09
Conduciveness	23.95	17.86	26.09	26.09
Affordability	34.02	46.43	17.39	17.39

Note. Values are percentages and sum to 100 within each model.

Review materials showed the widest variation in importance across models, from 7.14% for Subject 4 to 26.09% for Subjects 104 and 130. This indicates that the value placed on instructional content is far from uniform, which again points to distinct preference profiles within the market.

In Table 4, the aggregate model showed moderate agreement between observed and estimated preferences, $r = .63$, $p = .005$; $\tau = .48$, $p = .006$. Among the individual models, Subject 4 showed the strongest fit, $r = .61$, $p = .006$; $\tau = .52$, $p = .004$. Subjects 104 and 130 showed moderate fit, $r = .55$, $p = .015$; $\tau = .33$, $p = .052$. These results indicate that the additive part-worth model reasonably reproduced respondents' scores for the profiles used in estimation.

Table 4. *Correlations between observed and estimated preferences*

Model	r	p	τ	p	Holdout τ	p
Overall	.629	.005	.475	.006	-.183	.359
Subject 4	.606	.006	.516	.004	-.800	.063
Subject 104	.545	.015	.333	.052	-.671	.110
Subject 130	.545	.015	.333	.052	-.671	.110

Note. r = Pearson correlation; τ = Kendall's tau. Holdout τ is the correlation between observed and predicted scores for the holdout profiles, which were not used to estimate the utilities.

The holdout correlations, however, were negative for all models, ranging from $\tau = -.18$ to $\tau = -.80$, and none reached statistical significance. This means the estimated utilities did not predict how respondents scored the holdout profiles. The negative



direction of the correlations suggests that the models performed worse than chance in predicting holdout scores, which is a more serious concern than simple weak validation. Three explanations are plausible and should be distinguished in future research. First, the scoring task may have been affected by scale usage heterogeneity, which inflates the constant and can distort individual-level utility estimates (Rossi et al., 2001). Second, the additive model may not capture interactions among attributes – for example, the value of a large venue may depend on the fee level, or the value of recent materials may depend on reviewer experience. A main-effects-only model cannot detect such interactions, and the holdout failure may reflect their presence. Third, the holdout profiles themselves may have been poorly constructed or administered, which would undermine the validation exercise independently of model quality. Green and Srinivasan (1990) and Orme (2010) note that weak holdout performance can arise from any of these sources, and the appropriate response is to report the limitation transparently rather than to reinterpret the holdout results as supporting the model.

The utilities and importance values should therefore be read as descriptions of the preferences expressed in the estimation task rather than as reliable forecasts of choice among new review center offerings. This limitation applies with particular force to the cluster profiles, which are derived from individual utility estimates of uncertain reliability.

In Table 5, the utility for "Fee below 10,000" is negative ($-.75$, $-.25$, $-.75$) and the utility for "Fee below 10,001 up" is positive ($.75$, $.25$, $.75$) across all three clusters. This is the opposite of the aggregate conjoint model, where the lower fee had a positive utility ($.662$) and the higher fee had a negative utility ($-.662$). This pattern is most plausibly explained by a labeling or coding reversal in the SPSS output rather than by a genuine preference for higher fees across all clusters. The inconsistency must be verified against the original data before the cluster profiles can be interpreted with confidence. If the labels are confirmed, the substantive interpretation would change substantially. If the labels are reversed, the cluster patterns are consistent with the aggregate model.

The cluster analysis classified the 149 respondents into three clusters. Distances between final cluster centers showed that Clusters 1 and 2 were the most similar ($d = 1.10$), followed by Clusters 1 and 3 ($d = 1.19$) and Clusters 2 and 3 ($d = 1.33$). The clusters differed most sharply on capacity, reviewer experience, and – subject to the verification noted above – affordability. Table 6 presents the analysis of variance for the clustering variables.

Table 5. *Final cluster centers of part-worth utilities*

Attribute and level	Cluster 1 (n = 101)	Cluster 2 (n = 22)	Cluster 3 (n = 26)
Track record			
Equal	-.17	.00	-.28
Above	-.17	-.25	.01
Reviewers			
Running 10 years	-.33	-.17	-.06
More than 10 years	.54	-.04	.24
Review materials			
Relevant in the past 5 years	.00	-.17	.00
Relevant in the past 3 years	-.37	-.29	-.08
Conduciveness (capacity)			
Can accommodate 50 below	-.67	-.33	-.44
Can accommodate 51-100	-.04	.42	.22
Can accommodate 100 to up	.71	-.08	.22
Affordability (fee)			
Fee below 10,000	-.75	-.25	-.75
Fee below 10,001 up	.75	.25	.75

Note. N = 149. Values are the mean part-worth utilities of respondents in each cluster. Cluster 1, n = 101 (67.8%); Cluster 2, n = 22 (14.8%); Cluster 3, n = 26 (17.4%). The constant was excluded from the clustering variables.

Table 6. *Analysis of variance of clustering variables*

Variable	F(2, 146)	p
Equal	1064.49	< .001
Above	4877.48	< .001
Running 10 years	1719.45	< .001
More than 10 years	1207.88	< .001
Relevant in the past 5 years	47.46	< .001
Relevant in the past 3 years	219.90	< .001
Can accommodate 50 below	1813.16	< .001
Can accommodate 51-100	12477.55	< .001
Can accommodate 100 to up	38648.60	< .001
Fee below 10,000	764.00	< .001
Fee below 10,001 up	764.00	< .001

Note. The F tests are descriptive only. Because the clusters were chosen to maximize differences among cases in different clusters, the significance levels are not corrected and cannot be interpreted as tests of the hypothesis that the cluster means are equal. The two fee variables have identical F values because one is the exact negative of the other.



All clustering variables differed across clusters, $F_s(2, 146) \geq 47.46$, $p_s < .001$. The capacity variables showed the largest F values (38,648.60 for 100 to up; 12,477.55 for 51–100), followed by the track record variables (4,877.48 for above; 1,064.49 for equal) and the reviewer variables (1,719.45 for running 10 years; 1,207.88 for more than 10 years). The fee variables had $F = 764.00$. As Punj and Stewart (1983) caution, these tests describe the separation achieved by the clustering procedure rather than test a hypothesis. They are useful mainly for showing which attributes contributed most to distinguishing the groups. Table 7 summarizes the label and defining characteristics of each profile.

Table 7. *Preference profiles and defining characteristics*

Profile	n (%)	Label	Defining characteristics
1	101 (67.8)	Experienced- Reviewer, High- Capacity Seekers	Strongest preference for more than 10 years as reviewer (.54) and 100 to up capacity (.71); strongly dislikes 50 below capacity (–.67); negative on both equal (–.17) and above (–.17) track record
2	22 (14.8)	Moderate- Capacity, Recent- Materials Seekers	Strongest preference for 51–100 capacity (.42); most favorable toward equal track record (.00); relatively indifferent across reviewer experience levels; smallest fee difference
3	26 (17.4)	Moderate- Capacity Seekers with Mixed Preferences	Strongest preferences for more than 10 years as reviewer (.24), 51–100 capacity (.22), and 100 to up capacity (.22); most favorable toward above track record (.01); least favorable toward equal track record (–.28)

Note. N = 149. Labels are interpretive and based on the attribute levels with the largest utility differences within each profile (see Table 5). These profiles should be treated as descriptive typologies rather than validated segments because holdout validation was not achieved. The fee direction in Table 5 must be verified before the labels can be finalized.

Cluster 1, the Experienced-Reviewer, High-Capacity Seekers, made up about two thirds of the sample (67.8%) and largely defines the mainstream of the market. Its members showed the strongest preference for more than 10 years as reviewer (.54) and for venues accommodating 100 and above (.71). They strongly disliked venues accommodating 50 or fewer (–.67). They were negative on both equal (–.17) and above (–.17) track record, which suggests that track record was not a distinguishing attribute for this group. This profile describes a graduate who values experienced reviewers



and large venues, and who may be relatively indifferent to a center's passing rate relative to the national average.

Cluster 2, the Moderate-Capacity, Recent-Materials Seekers (14.8%), showed its strongest preference for 51–100 capacity (.42). It was the only cluster with a non-negative utility for equal track record (.00), and it was relatively indifferent across reviewer experience levels (running 10 years = $-.17$; more than 10 years = $-.04$). This profile suggests a group that values moderate capacity and demonstrated track record at the national average level, and that does not strongly differentiate among reviewer experience levels.

Cluster 3, the Moderate-Capacity Seekers with Mixed Preferences (17.4%), showed its strongest preferences for more than 10 years as reviewer (.24), 51–100 capacity (.22), and 100 to up capacity (.22). It was the only cluster with a positive utility for above track record (.01), and it showed the strongest negative utility for equal track record ($-.28$). This profile appears to value experienced reviewers, moderate-to-large capacity, and an above-national track record.

Taken together, the three profiles show that the review center market for BSSW graduates is not uniform. Cluster 1, the largest group, values experienced reviewers and large capacity. Cluster 2 values moderate capacity and is relatively indifferent across other attributes. Cluster 3 values experienced reviewers, moderate-to-large capacity, and above track record. The fee direction in the cluster centers must be verified before the interpretation of price sensitivity across clusters can be finalized.

CONCLUSION

This study examined the preferences of BS Social Work graduates for board examination review centers and explored whether distinct preference profiles exist within this market. The findings show that affordability is the dominant consideration, accounting for about a third of graduates' preference, followed by the capacity of the review venue. Graduates most preferred a review center charging ₱10,000 and below, accommodating 100 reviewees and above, with a track record above the national passing rate and materials updated within the past three years. The instructional attributes of track record, reviewers, and review materials carried smaller and roughly equal weights, which suggests that graduates first filter review centers by cost and venue capacity and then distinguish among them by quality. This pattern is consistent with a bounded rationality account in which price functions as a screening criterion under financial constraint (Payne et al., 1993; Simon, 1955), and with a prospect theory account in which the ₱10,000 threshold functions as a reference point producing a discontinuous utility shift (Kahneman & Tversky, 1979).



The study reveals that BSSW graduates prioritized affordability over track record. This divergence may reflect differences in expected earnings and employment prospects between the two professions. Social work graduates often enter lower-paying public or non-profit sector positions, which may make the immediate cost of review services more salient than the long-run reputational benefits of a high-track-record center. This finding has implications for how review centers market their services to different professional groups and for how schools of social work consider the financial burden that review costs place on graduates.

The cluster analysis suggested three descriptive preference profiles: *experienced-reviewer*, *high-capacity seekers*; *moderate-capacity*, *recent-materials seekers*; and *moderate-capacity seekers with mixed preferences*. These profiles should be treated as exploratory typologies rather than validated market segments because the holdout validation was not achieved. Additionally, the fee direction in the cluster centers is inconsistent with the aggregate model and must be verified before the cluster profiles can be interpreted with confidence. Future research should examine whether these profiles predict actual enrollment behavior, examination outcomes, or satisfaction with review services.

LIMITATIONS

Several limitations should be considered when interpreting the findings of this study.

Holdout validation failure. The estimated utilities did not predict holdout scores; holdout correlations were negative and non-significant for all models. This limits the predictive validity of the aggregate model and, by extension, the cluster profiles. The most likely explanations are scale usage heterogeneity in the scoring task, model misspecification due to unmodeled interactions among attributes, or problems with the holdout profiles themselves. Future research should test for interactions among attributes and consider alternative estimation approaches such as hierarchical Bayes or latent class models (Andrews et al., 2002; Lenk et al., 1996; Rossi et al., 2005).

Scale usage heterogeneity. Rating scales are subject to respondent-specific differences in scale use that are unrelated to the attributes being measured (Rossi et al., 2001). This can inflate the model constant and distort individual-level utility estimates, which in turn affects the cluster profiles. The exclusion of the constant from the clustering variables partially mitigates this concern but does not eliminate it. Future research should report the distribution of individual model fit and consider Bayesian approaches that explicitly model scale usage.



Individual-level utility reliability. The cluster analysis depends on individual-level utility estimates, and the paper does not report the distribution of individual model fit. If a substantial minority of individual models fit poorly, the cluster profiles are partitions of unreliable estimates. Future research should report the distribution of individual R^2 or Kendall's τ and consider hierarchical Bayes or latent class estimation, which borrow information across respondents and produce more stable individual-level utilities (Andrews et al., 2002; Lenk et al., 1996; Rossi et al., 2005).

Construct validity of conduciveness. The attribute labeled "conduciveness" was operationalized as the number of reviewees the venue could accommodate. Capacity is at best a proxy for conducive learning environments, which encompass physical comfort, noise, ventilation, seating, and instructor accessibility. The preference for large-capacity venues may reflect social proof, schedule flexibility, or perceived institutional stability rather than a preference for less conducive environments. Future research should measure conduciveness more directly and examine the reasons behind capacity preferences.

Sample representativeness. The sample consists of 149 BSSW graduates from Davao City. Regional distribution, institutional type (state university vs. private HEI), and timing (graduates who had already taken the exam vs. those preparing for it) may all influence preferences. The findings should not be generalized to the national population of BSSW graduates without replication in other regions and institutional contexts.

Importance value comparability. Importance values in conjoint analysis are computed from the range of utilities within each attribute and are normalized to sum to 100. Attributes with more levels or larger utility ranges may show higher importance by construction (Cattin & Wittink, 1982). The three-level conduciveness attribute may therefore appear more important than the two-level attributes partly as a design artifact. Future research should report importance values both as raw ranges and as normalized percentages.

Cross-sectional design. The study measured stated preferences, not actual enrollment behavior. The relationship between stated preferences and real choices may be attenuated by factors not measured in this study, including location, schedule, peer influence, and institutional requirements. Future research should link preference data to actual enrollment and examination outcomes.

Cluster stability. The stability of the three-cluster solution was not formally assessed through bootstrap resampling or split-half validation. The three-cluster solution was specified *a priori* based on the exploratory purpose of the study. Future



research should assess cluster stability before treating the profiles as replicable segments.

FUNDING

This research received no external funding.

COMPETING INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY

Data will be made available upon request to the authors.

DECLARATION OF ARTIFICIAL INTELLIGENCE USE

In the preparation of this research, the authors used Claude as an AI-assisted tool to format statistical results in APA style and to support drafting and editing of the manuscript. All data collection, analysis decisions, and interpretations were reviewed and approved by the authors, who take full responsibility for the integrity of this work.

REFERENCES

- Andrews, R. L., Ansari, A., & Currim, I. S. (2002). Hierarchical Bayes versus finite mixture conjoint analysis models: A comparison of fit, prediction, and partworth recovery. *Journal of Marketing Research*, 39(1), 87–98.
- Caballero, S., Agad, F., Navales, V., Traverro, A., & Quemado, L. (2024). Subject-specific preparedness in social work licensure exams: Are students equally prepared? *Balankas: An International Multidisciplinary Research Journal*, 1(1). <https://www.neliti.com/publications/688567/>
- Castillo, R., & Yee, E. C., Jr. (2023). Preferences of accountancy graduating students of a CPA review center: A conjoint analysis study. *Business and Finance Journal*, 8(1). <https://journal2.unusa.ac.id>
- Cattin, P., & Wittink, D. R. (1982). Commercial use of conjoint analysis: A survey. *Journal of Marketing*, 46(3), 44–53.
- Commission on Higher Education. (2007). *CHED Memorandum Order No. 30, s. 2007: Revised implementing rules and regulations governing the establishment and operation of*



review centers and similar entities in the Philippines pursuant to Executive Order No. 566.
<https://ched.gov.ph/wp-content/uploads/2017/10/CMO-No.30-s2007.pdf>

Dolnicar, S., & Leisch, F. (2010). Evaluation of structure and reproducibility of cluster solutions using the bootstrap. *Marketing Letters*, 21(1), 83–101.

Executive Order No. 566. (2006). *Directing the Commission on Higher Education to regulate the establishment and operation of review centers and similar entities*. Office of the President of the Philippines.

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5(2), 103–123.

Green, P. E., & Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, 54(4), 3–19.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

Hemsley-Brown, J., & Oplatka, I. (2006). Universities in a competitive global marketplace: A systematic review of the literature on higher education marketing. *International Journal of Public Sector Management*, 19(4), 316–338.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.

Kotler, P., & Keller, K. L. (2016). *Marketing management* (15th ed.). Pearson.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2), 132–157.

Lenk, P. J., DeSarbo, W. S., Green, P. E., & Young, M. R. (1996). Hierarchical Bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science*, 15(2), 173–191.

Liban, J. E., Malobo, M. A. C., Ocampo, L., & Onofre, B. C. (2024). Qualitative assessment of online review program of ISAP social work graduates 2021. *College of Social Work*, 1(1). <https://dawa.isap.edu.ph/index.php/socialwork/article/view/6>

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and applications*. Cambridge University Press.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1(1), 1–27.



Maringe, F. (2006). University and course choice: Implications for positioning, recruitment and marketing. *International Journal of Educational Management*, 20(6), 466–479.

Orme, B. K. (2010). *Getting started with conjoint analysis: Strategies for product design and pricing research* (2nd ed.). Research Publishers.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

PRC Board News. (2024). *September 2024 Social Worker Licensure Examination results*. <https://www.prcboardnews.com>

Punj, G., & Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research*, 20(2), 134–148.

Republic Act No. 4373. (1965). *An act to regulate the practice of social work and the operation of social work agencies in the Philippines and for other purposes*.

Rossi, P. E., Allenby, G. M., & McCulloch, R. (2005). *Bayesian statistics and marketing*. Wiley.

Rossi, P. E., Gilula, Z., & Allenby, G. M. (2001). Overcoming scale usage heterogeneity: A Bayesian hierarchical approach. *Journal of the American Statistical Association*, 96(453), 20–31.

Senate Bill No. 153, 20th Congress of the Republic of the Philippines. (2025). *An act establishing professional review centers for licensure examinations in state universities and colleges (SUCs), and appropriating funds therefor*. <https://legacy.senate.gov.ph/lisdata/4663042634!.pdf>

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1), 99–118.

Smith, W. R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing*, 21(1), 3–8.

Soutar, G. N., & Turner, J. P. (2002). Students' preferences for university: A conjoint analysis. *International Journal of Educational Management*, 16(1), 40–45.

Wedel, M., & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations* (2nd ed.). Kluwer Academic Publishers.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, 52(3), 2–22.