4.1 Participant Demographics and Trait Distribution

Table 1. Sample Demographics (N=348)

Variable	Category	Count	%
Education			
	Bachelor's degree	106	30.5
	Complete high school	71	20.4
	Graduate degree	66	19.0
	Some college	63	18.1
	Some graduate school	23	6.6
	Incomplete high school	19	5.5
Music Enthusiast			
	Agree	149	42.8
	Strongly agree	121	34.8
	Neutral	60	17.2
	Disagree	13	3.7
	Strongly disagree	5	1.4
Region			
	Southeast	184	52.9
	South	70	20.1
	Northeast	55	15.8
	Center-West	29	8.3
	North	10	2.9
Gender			
	Female	303	87.1
	Male	40	11.5
	Other	5	1.4

Note: Percentages may not sum to 100% due to rounding.

Table 1 presents the demographic characteristics of the study sample. The majority of participants held a Bachelor's degree (30.5%), were female (87.1%), and resided in the Southeast region (52.9%). Additionally, 77.6% of participants identified as music enthusiasts (Agree or Strongly Agree).

Table 2 displays the distribution of participants across low/medium and high levels for each Big Five personality trait. The sample showed notably high levels of Agreeableness (93.1% scoring ≥ 15) and relatively low Extraversion (31.3% scoring ≥ 15). Openness, Conscientiousness, and Neuroticism showed more balanced distributions.

Table 2. Distribution of Personality Traits (N=348)

Trait	Low/Med (< 15)	High (≥ 15)
Openness	107	241
Conscientiousness	130	218
Extraversion	239	109
Agreeableness	24	324
Neuroticism	231	117

Note: Trait scores are categorized as Low/Medium (< 15) and High (≥ 15) on a 20-point scale. Totals per row sum to N=348.

4.2 RQ1: Generating Personality-Aware Explanations at Scale

Table 3 summarizes the best-performing target trait within each explanation condition. Table 4 reports the full set of precision, recall, and F1-scores for all explanation-trait and target-trait combinations.

Explanation trait	**Best target trait (by F1)**	**Precision**	**Recall**	**F1-score**
Agreeableness	Agreeableness	0.156	0.147	0.152
Neuroticism	Neuroticism (Low)	0.237	0.360	0.286
Openness	Openness	0.273	0.265	0.269
Conscientiousness	Conscientiousness	0.143	0.132	0.137
Extraversion	Extraversion	0.095	0.133	0.111

Note: The table reports the highest F1-score observed within each explanation-trait condition. Neuroticism is shown as two target classes (High, Low) because the metrics were computed separately by class.

Table 3. Best-performing target trait within each explanation-trait condition

Target trait	Precision	Recall	F1-score
Explanation trait: Agreeableness			
Openness	0.062	0.059	0.061
Conscientiousness	0.000	0.000	0.000
Extraversion	0.062	0.067	0.065
Agreeableness	0.156	0.147	0.152
Neuroticism (High)	0.031	0.028	0.029
Neuroticism (Low)	0.000	0.000	0.000
Explanation trait: Neuroticism			
Openness	0.000	0.000	0.000
Conscientiousness	0.000	0.000	0.000
Extraversion	0.026	0.033	0.029
Agreeableness	0.079	0.088	0.083
Neuroticism (High)	0.053	0.056	0.054
Neuroticism (Low)	0.237	0.360	0.286
Explanation trait: Openness			
Openness	0.273	0.265	0.269
Conscientiousness	0.000	0.000	0.000
Extraversion	0.000	0.000	0.000
Agreeableness	0.000	0.000	0.000
Neuroticism (High)	0.000	0.000	0.000
Neuroticism (Low)	0.000	0.000	0.000
Explanation trait: Conscientiousness			
Openness	0.086	0.088	0.087
Conscientiousness	0.143	0.132	0.137
Extraversion	0.000	0.000	0.000
Agreeableness	0.000	0.000	0.000
Neuroticism (High)	0.000	0.000	0.000
Neuroticism (Low)	0.057	0.080	0.067
Explanation trait: Extraversion			
Openness	0.000	0.000	0.000
Conscientiousness	0.000	0.000	0.000
Extraversion	0.095	0.133	0.111
Agreeableness	0.048	0.059	0.053
Neuroticism (High)	0.000	0.000	0.000
Neuroticism (Low)	0.000	0.000	0.000

Note: Precision, recall, and F1-score are reported per target trait within each explanation-trait condition. Neuroticism is reported as two target classes (High, Low) because the evaluation outputs were class-specific.

Table 4. Precision, recall, and F1-score by explanation-trait condition and target trait

4.3 RQ2: Personality Traits and Selection Preferences

Table 5 reports the chi-square tests assessing whether trait is associated with explanation type selection, with Holm-adjusted p-values across the five traits. Table 6 presents the Bonferroni-corrected pairwise comparisons between explanation types selection, and Table 7 summarizes which explanation types are most and least selected among participants in the High range based on standardized residuals.

Trait	χ^2	**df**	p	$p_{\mathbf{adj}}$	**Cramér's V**	**Sig.**
Openness	3.134	5	0.67936	1.00000	0.095	ns
Conscientiousness	8.959	5	0.11072	0.55358	0.160	ns
Extraversion	4.318	5	0.50462	1.00000	0.111	ns
Agreeableness	1.535	5	0.90895	1.00000	0.066	ns
Neuroticism	5.619	5	0.34508	1.00000	0.127	ns

Note: Trait levels were defined as Low/Medium (< 15) vs. High (≥ 15) on a 20-point scale. Holm-adjusted p-values are reported across the five omnibus tests. ns = not significant.

Table 5. Chi-square tests of independence between trait and explanation type (Holm-adjusted across five tests)

Pair	χ^2	**Raw** p	**Corrected** p	**Significant**	**Sig.**
Conscientiousness vs Feature-Based	29.96	0.0000	0.0000	Yes	***
Conscientiousness vs Neuroticism	18.23	0.0000	0.0003	Yes	***
Feature-Based vs Openness	12.65	0.0004	0.0056	Yes	**
Extraversion vs Feature-Based	9.21	0.0024	0.0362	Yes	*
Agreeableness vs Conscientiousness	8.71	0.0032	0.0475	Yes	*
Agreeableness vs Feature-Based	6.38	0.0115	0.1728	No	ns
Extraversion vs Conscientiousness	5.97	0.0145	0.2182	No	ns
Neuroticism vs Openness	5.33	0.0210	0.3150	No	ns
Conscientiousness vs Openness	3.66	0.0559	0.8383	No	ns
Extraversion vs Agreeableness	0.17	0.6788	1.0000	No	ns
Extraversion vs Neuroticism	3.16	0.0755	1.0000	No	ns
Extraversion vs Openness	0.19	0.6669	1.0000	No	ns
Agreeableness vs Neuroticism	1.60	0.2061	1.0000	No	ns
Agreeableness vs Openness	0.90	0.3424	1.0000	No	ns
Feature-Based vs Neuroticism	1.38	0.2395	1.0000	No	ns

Note: Bonferroni-corrected p-values are reported. * $p_{\text{corr}} < .05$, ** $p_{\text{corr}} < .01$, *** $p_{\text{corr}} < .001$, ns = not significant.

Table 6. Pairwise chi-square comparisons between explanation types (Bonferroni-corrected)

Trait (High only)	Most preferred	Least preferred
Openness	Feature-based	Conscientiousness
Conscientiousness	Feature-based; Neuroticism	Conscientiousness
Extraversion	Feature-based	Conscientiousness
Agreeableness	Feature-based	Conscientiousness
Neuroticism	Neuroticism	Conscientiousness

Note: "Most preferred" and "Least preferred" indicate explanation types that occurred more or less often than expected among high scorers, based on standardized residuals. The threshold is $|\text{Std Residual}| \geq 2$.

Table 7. High scorers only ($\geq 15/20$): explanation types most and least represented (standardized residuals)