Statistical Practices of Music Education Researchers: Preliminary Results Regarding Analyses of Variance

Christina Svec

csvec@iastate.edu

Iowa State University – USA

Abstract

The purpose of the current study was to describe the statistical practices of researchers across three prominent music education research journals regarding analyses of variance. The current paper has included information specific to *JRME* and *CRME* from 2000–2015; information specific to *Psychology of Music*, and 2016 data from *JRME* and *CRME* is currently in progress. Research components of ANOVA, MANOVA, and ANCOVA analyses included power analysis, assumptions, descriptive statistics, coefficients, effect sizes, post-hoc analyses, validity, and reliability. Overall, the majority of articles published in *JRME* and *CRME* during the past 15 years have been quantitative in nature; 56% (*JRME*) and 47% (*CRME*) of the quantitative articles used analysis of variance procedures. Preliminary findings indicated research strengths including controlling for validity issues and reporting sample size and selection. Several weaknesses were also highlighted. Implications have been discussed for future research and research pedagogy.

Keywords: research, effect size, assumptions

Statistical Practices of Music Education Researchers: Preliminary Results Regarding Analyses of Variance

Content analyses within music education research journals may be illustrative of previous and current methodological or pedagogical practices. These investigations might also encourage future rigor with regard to analyses, data reporting, and interpretation. Various topics have been investigated across journals including, but not limited to, theoretical frameworks (Miksza & Johnson, 2012), student teaching (Silveira & Diaz, 2014), research paradigms (Jorgensen & Ward-Steinman, 2015), eminence (Standley, 1984; Hamann & Lucas, 1998), qualitative research trends (Lane, 2011), and national music organization conferences (Orman & Price, 2007).

Multiple content analyses have included information specific to quantitative research. Several found that survey research has been a highly represented method across the field (Diaz & Silveira, 2014; Killian, Liu, & Reid, 2012). Killian, Liu, and Reid (2012) found an increase in quantitative research articles from 1991-2011 in the *Journal of Music Teacher Education*; however, research practices were mentioned fewer times than other topics. Silveira and Diaz (2014) labeled the majority of the research as descriptive or experimental. Schmidt and Zdzinski (1993) analyzed *JRME*, *CRME*, and *Psychology of Music* quantitative articles for content and and citation frequencies. Although analyses have discussed quantitative research, very few have discussed research practices.

Studies outside of music education have investigated statistical practices (Armstrong & Henson, 2005; Keselman, Algina, & Kowalchuk, 2001; Keselman et al., 1998). There have been very few articles investigating the pieces of statistical analyses within music education research journals. Given the breadth of current content analyses and the call for increased rigor within music education research journals (Morrison, 2015), a need exists to examine statistical practices within the field.

One article provided the inspiration and framework for which to base the current study. Keselman et al. (1998) investigated statistical practices within three education research journals, specifically regarding "between-subjects univariate designs, between-subjects multivariate designs, repeated measures designs, and covariance designs" (p. 350). Items coded included sample details, methodological details, and the results that were (or were not) reported. Implications included pedagogical suggestions for graduate research courses as well as best practice for in-service researchers.

Therefore, the purpose of the current study was to describe the statistical practices of researchers across three prominent music education research journals regarding analyses of variance. Data collected included all articles published between 2000 and 2016 within the *Journal of Research in Music Education* (*JRME*), the *Bulletin of the Council for Research in Music Education* (*CRME*), and *Psychology of Music*. Research components of ANOVA, MANOVA, and ANCOVA analyses included power analysis, assumptions, descriptive statistics, coefficients, effect sizes, post-hoc analyses, validity, and reliability.

The current paper has included information specific to *JRME* and *CRME* from 2000–2015; information specific to *Psychology of Music*, and 2016 data from *JRME* and *CRME* is currently in progress. *Psychology of Music* data has been coded and is currently being checked for inter-coder reliability.

Preliminary Results

From 2000-2015, analyses of variance have been used in 43% (n = 159) of the total number of research articles published in *JRME* (N = 372) and 54% of the articles that are quantitative (n = 286). In *CRME,* analyses of variance have been used in 27% (n = 100) of the total number of research articles published (N = 369) and 47% of the articles that are quantitative (n = 211). Refer to Table 1 for frequencies of studies and type that used variance analyses by year of both journals. Given the frequency of use, it may be imperative to the integrity of music education research that there is exploration into how the field is using these analyses for the purpose of rigor development and/or maintenance.

Table 1

Overall Frequency of Use from 2000 – 2015: CRME and JRME

Year	Total Articles		Exp/ Descript		Variance Analyses		% Total[1]		% Exp/ Descript[2]	
	JRME	*CRME*	*JRME*	*CRME*	*JRME*	*CRME*	*JRME*	*CRME*	*JRME*	*CRME*
2000	23	44	19	16	13	11	56	25	68	69
2001	24	27	20	19	8	9	33	33	40	47
2002	24	28	19	18	11	10	46	36	58	56
2003	24	23	17	12	8	7	38	30	53	58
2004	22	32	19	18	10	8	45	25	53	44
2005	24	24	21	16	11	10	46	42	52	63
2006	21	21	18	15	11	3	52	14	61	20
2007	22	17	15	12	8	5	36	29	53	42
2008	17	26	12	16	7	8	41	31	58	50
2009	23	22	19	12	12	5	52	23	63	42
2010	22	22	16	11	10	1	45	5	63	9
2011	21	15	17	10	11	5	52	33	65	50
2012	22	19	16	16	7	6	32	32	44	38
2013	27	16	17	7	11	6	41	38	65	86
2014	26	17	20	7	11	3	42	18	55	43
2015	30	16	21	6	9	3	30	19	43	50
Total	372	369	286	211	159	100	43	27	56	47

Note[1] Total percentage of articles that include analysis of variance procedures within the given volume.
Note[2] Percentage of experimental/descriptive articles that include analysis of variance procedures.

Regarding assumptions, 30% (n = 47) of the *JRME* articles reported assumptions. Eighteen (18%) *CRME* articles reported assumptions. Normality, homogeneity of variance, and sphericity were cited most often in both journals across analyses of variance. Refer to Table 2 for frequencies of assumptions by analysis. Refer to Table 2 for frequencies of post-hoc procedures and corrections.

Table 2

Frequency of Analyses and Reported Methodological Components

		n	
Analysis	Type	*JRME*	*CRME*
ANOVA	One-way	39	31
	Factorial	37	17
	Repeated Measures	23	31
	Mixed	57	9
	Not Stated	2	4
	Total	158	92
MANOVA	One-way	2	2
	Factorial	8	4
	Repeated Measures	1	4
	Mixed	18	7
	Not Stated	5	3
	Total	34	22
ANCOVA	One-way	4	-
	Repeated Measures	1	1
	Not Stated	-	5
	Total	5	6
MANCOVA	Factorial	2	-
	Mixed	1	-
	Total	3	-

		n	
Analysis	Assumption	*JRME*	*CRME*
ANOVA	Homogeneity of variance	8	5
	Normality	6	4
	Assumtions met	1	2
MANOVA	Normality	6	2
	Homogeneity of variance	5	1
	Homogeneity of covariance	6	1
	Linearity	3	-
	Homoscedasticity	1	-

ANCOVA	Homogeneity of variance	1	1
	Normality	2	1
Repeated Measures	Normality	4	1
	Homogeneity	7	-
	Sphericity	10	6
	Homogeneity of regression	1	-

	n			*n*	
Procedure	*JRME*	*CRME*	Corrections	*JRME*	*CRME*
Follow-up univariate	17	5	Bonferroni	17	6
t-test	16	3	Greenhouse-Geisser	7	4
Scheffe	15	7	Bonferroni-Holm	1	-
Tukey's HSD	14	11	Data transformation	1	-
Bonferroni	10	5	Harwell	1	-
Fisher's LSD	10	6	Huynh-Feldt	1	1
Pairwise comparisons	8	-	Macmillan and Creelman (1991)	1	-
Follow-up ANOVAs	7	2	Missing variables analysis	1	-
Post-hoc contrasts	4	9	Not labeled	1	2
Bonferroni/Dunn	3	1	Pillai's trace	1	-
Chi-square	3	-	Log 10 transformation	-	1
Tests of simple effects	2	-			
Tukey-Kramer HSD	2	-			
Univariate ANCOVA	2	-			
Dunnet T3	1	-			
Dunnet's C	1	-			
W/in participants' contrasts	1	-			
Polynomial contrasts	1	-			
Friedman ANOVA	1	-			
Newman-Keuls	1	3			
Games-Howell	-	1			
Duncan multiple range	-	1			
z-tests	-	1			
Sidak adjusted paired	-	1			
Total	119	56	Total	32	16

	N		Interpreted	
Effect Size	*JRME*	*CRME*	*JRME*	*CRME*
η^2	28	13	14	6
Partial η^2	43	13	13	5
Cohen's *d*	6	3	3	3
h^2	2	-	2	-
ω^2	1	-	1	-

r^2	-	2	-	1

Notes. Repeat measures were omitted from other groups and grouped together due to similarities across analyses.
n = the number of articles that included each type of analysis. Many articles used more than one type of analysis; therefore, the total is greater than the total number of articles.

Issues concerning validity and reliability included group equality, threats to internal validity, sampling, and tests of reliability. Of the 159 coded *JRME* articles, 43% (*n* = 68) reported equal group sizes throughout the study. Of the 100 coded *CRME* articles, 71% indicated unequal group sizes, and 13 of the 71 indicated group sizes fewer than 20 participants. The most frequently reported control for internal validity across the two journals was selection. The most frequently reported control for external validity across the two journals was controlling for order effects.

Data collected regarding reliability included measure reliability and scoring reliability. Within the 159 *JRME* articles, the most frequently reported reliability procedures included interjudge/inter-rator reliability (*n* = 60) and internal consistency (*n* = 49). Within the 100 *CRME* articles, the most frequently reported reliability procedures included citations for previously used instruments (*n* = 21) and interjudge reliability (*n* = 6). Overall, 34% of *JRME* articles and 35% of *CRME* articles did not report reliability for measures, scoring, or both.

Effect Size Interpretations

Journal of Research in Music Education. Effect sizes were reported in 77 (48%) of the articles and were only interpreted in 33 (21%) of the articles. In other words, group differences were interpreted based on statistical significance in 79% of the articles using analysis of variance procedures. Refer to Table 2 for effect size type and frequency. Examples of interpreted effect sizes included the following: "The effect size for years since attending college was considerably smaller" (Elpus, 2015, p. 328). "The small effect size indicates that the independent variable of facial expression accounted for very little of the total variance in participants' ratings of ensemble expressivity . . . given this small effect size, these results must be approached with caution" (Silvey, 2013, p. 426). "It is also important to note that for all tests in this study resulting in statistically significant differences, partial eta squared ranged from .009 to .016, indicating weak relationships" (Campbell & Thompson, 2007, p. 170).

***Bulletin of the Council for Research in Music Education*.** Effect sizes were reported in 33% and interpreted in 17% of the articles. In other words, overall group differences were interpreted based on statistical significance in 83% analysis of variance procedures. Refer to Table 2 for effect size type and frequency. Examples of interpreted effect sizes included the following:

> "The 3.1 mean difference that separates the attitudinal scores of men and women remains a statistically significant finding in this research. However, the small effect size (*d* = .26) suggests this finding has little practical significance. This prompted a deeper investigation into the data. In doing so, the researchers uncovered that 67.9% of the choral educators participating in the study were female while only 32.1% were male . . . the significance of this finding may ultimately be attributed to teacher specialization rather than gender" (Gerber & Garrity, 2007, p. 83).

"It should be noted that the eta-squared values for significant main effects and interactions were relatively low, varying from .02 to .07, suggesting modest practical significance" (Zdzinski, 2013, p. 81). "There was a significant main effect for order, though the effect size was negligible (partial $\eta^2 = .08$)" (Morrison & Selvey, 2014).

> "While pitch had a significant main effect on the results, the smaller effect size implies that the register difference of a perfect fourth between the F and B-flat excerpts may not be different enough to result in any type of effect due to potential frequency sensitivity for these older adults" (Napoles, 2009, p. 30).

Implications. The majority of articles published in *JRME* and *CRME* during the past 15 years have been quantitative in nature; 56% (*JRME*) and 47% (*CRME*) of the quantitative articles used analysis of variance procedures. Music education research may benefit from more attention to detail and more training regarding analyses of variance reporting practices. Several specific recommendations can be made for future research and researchers:

1. Report (and check) all assumptions appropriate to an analysis in order to decrease sensitivity to Type I and Type II errors
2. Report all post-hoc procedures and corrections

With regard to effect size:

1. Adhere to APA guidelines regarding research rigor. Report and interpret effect sizes.

With regard to research pedagogy:

1. Maximize time with descriptive statistics. Assumptions may be best understood when raw data and descriptive statistics can be observed with understanding.
2. Calculate as much as possible by hand before turning to a computer program. When using a statistics program, train students to use syntax so that they may understand what they are asking of the program.

References

Armstrong, S. A., & Henson, R. K. (2005). Statistical practices of IJPT researchers: A review from 1993-2000. *International Journal of Play Therapy*, *14*(1), 7-26. doi:10.1037/h0088894

Hamann, D. L., & Lucas, K. V. (1998). Establishing journal eminence in music education research. *Journal of Research in Music Education*, *46*, 405-413. doi:10.2307/334552

Jorgensen, E. R., & Ward-Steinman, P. M. (2015). Shifting paradigms in music education research (1953–1978): A theoretical and quantitative reassessment. *Journal of Research in Music Education, 63*, 1-20. doi:10.1177/0022429415601690

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *Mathematical and Statistical Psychology*, *54*, 1-20. doi:10.1348/000711001159357

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B.,…Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research*, *68*, 350-386. doi:10.3102/00346543068003350

Killian, J. N., Liu, J., & Reid, J. F. (2012). The *Journal of Music Teacher Education*: A content analysis of articles 1991-2011. *Journal of Music Teacher Education*, *22*(2), 85-99. doi:10.1177/1057083712467637

Lane, J. (2011). A descriptive analysis of qualitative research published in two eminent music education research journals. *Bulletin of the Council for Research in Music Education*, *188*, 65-76.

Miksza, P., & Johnson, E. (2012). Theoretical frameworks applied in music education research: A content analysis of the Journal of Research in Music Education, 1979 to 2009. *Bulletin of the Council for Research in Music Education*, *193,* 7-30.

Morrison, S. J. (2015). Forum. *Journal of Research in Music Education*, *63*, 143–144. doi:10.1177/0022429415595625

Orman, E. K., & Price, H. E. (2007). Content analysis of four national music organizations' conferences. *Journal of Research in Music Education*, *55*, 148-161. doi:10.1177/002242940705500205

Schmidt, C. P., & Zdzinski, S. F. (1993). Cited quantitative research articles in music education research journals, 1975 – 1990: A content analysis of selected studies. *Journal of Research in Music Education*, *41,* 5-18. doi:10.2307/3345475

Silveira, J. M., & Diaz, F. M. (2014). Student teaching in music: A content analysis of research journals in music education. *Journal of Music Teacher Education*, *23*, 92-104. doi:10.1177/1057083713487215

Standley, J. M. (1984). Productivity and eminence in music research. *Journal of Research in Music Education, 32*, 149-157. doi:10.2307/3344834