

8
D-A040 598

GEORGIA INST OF TECH ATLANTA COLL OF INDUSTRIAL MAN--ETC F/G 12/2
WHERE HAVE ALL THE MODERATORS GONE: THE PERILS OF TYPE II ERROR--ETC(U)
MAY 77 M R BLOOD, G M MULLET
GIT-M-77-5

N00014-76-C-0828

NL

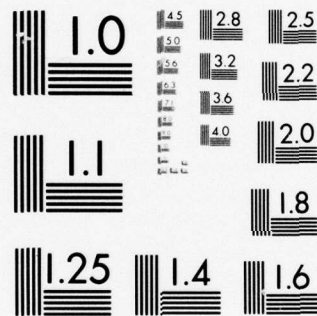
UNCLASSIFIED

1 OF 1
AD A040 598



END

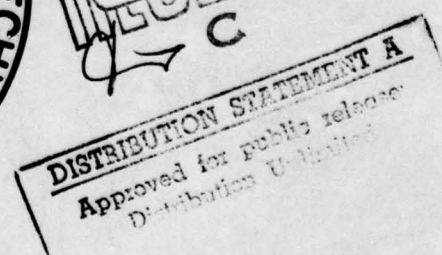
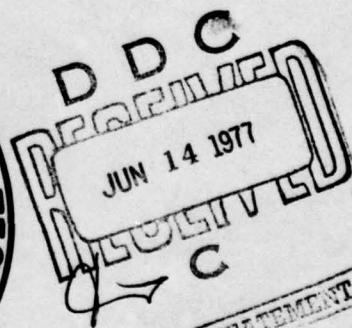
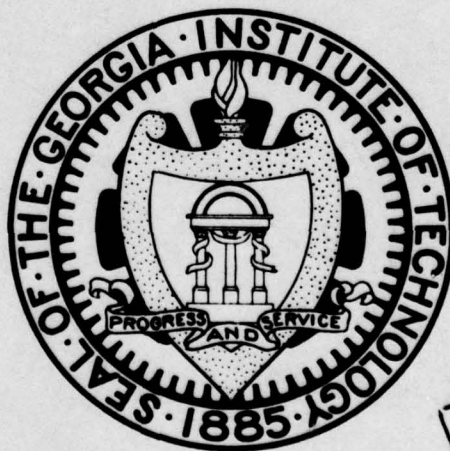
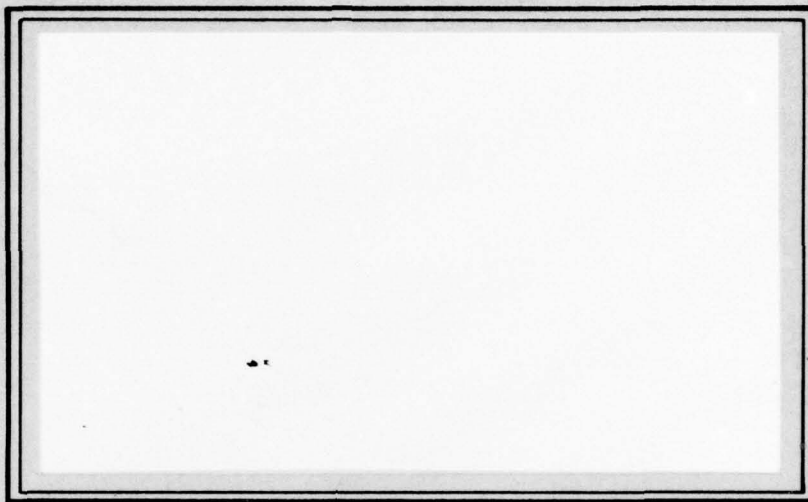
DATE
FILMED
7-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A 040598

68



AD NO. _____
DDC FILE COPY

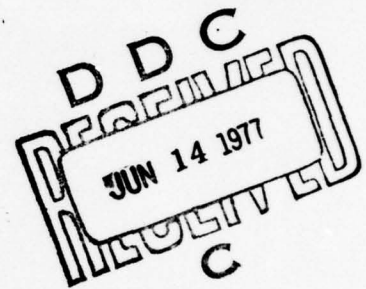
COLLEGE OF INDUSTRIAL MANAGEMENT
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332

Where Have All the Moderators Gone?:

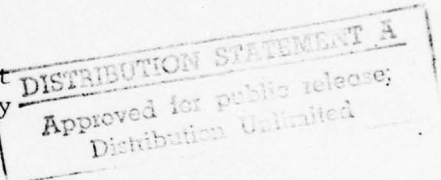
The Perils of Type II Error

Milton R. Blood

Gary M. Mullet



College of Industrial Management
Georgia Institute of Technology



Technical Report #1
M-77-5 (Working Paper Series)

This research was sponsored by the
Organizational Effectiveness Research Program,
Office of Naval Research (Code 452), under
Contract No. N00014-76-C-0828; NR 170-828

Reproduction in whole or in part is
permitted for any purpose of the
United States Government.

Approved for public release; distribution unlimited.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR #1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) WHERE HAVE ALL THE MODERATORS GONE: THE PERILS OF TYPE II ERROR.	5. TYPE OF REPORT & PERIOD COVERED Technical Report.	
7. AUTHOR(s) Milton R. Blood Gary M. Mullet	6. PERFORMING ORG. REPORT NUMBER M-77-5	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Georgia Institute of Technology College of Industrial Management Atlanta, GA 30332	8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0828	
11. CONTROLLING OFFICE NAME AND ADDRESS Organizational Effectiveness Research Programs Office of Naval Research (Code 452) Arlington, VA 22217	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 170-828	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 25p.	12. REPORT DATE May, 1977	
	13. NUMBER OF PAGES 22	
	15. SECURITY CLASS. (of this report) Unclassified	
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE		
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. 14 GIT-M-77-5, TR-1		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Hierarchical regression Multiple correlation Interaction effect Statistical power Linear effect Statistical significance Moderated regression analysis Type II error Moderator effect		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Traditional moderated regression analysis requires a significant and sizable increment in R ² due to the interaction term in order to conclude that there is a moderator effect. A Monte Carlo analysis demonstrates that the traditional procedure will not discover moderators even when they are the only existing effect. A suggestion is made to reverse the usual hierarchical order of the moderated regression analysis. The choice between a linear and an interactive model should be made on the basis of theoretical justification.		

Where Have All the Moderators Gone?:

The Perils of Type II Error

Moderated regression analysis has been used in psychological research since its presentation by Saunders (1956). This technique assesses the influence of a third variable, called a moderator, on the relationship between two other variables. Knowledge of the level of the moderator variable provides information about the strength of the relationship between the other two variables. In his explication of the procedure, Saunders pointed out that a moderator effect will manifest itself as a relationship between the dependent variable and the cross product of the independent and moderator variables. Conceptually, this is equivalent to saying that the relationship between the dependent and independent variables is influenced by the moderator variable.

The relationship thus described allows the postulation of individual differences in the relationships between variables, i.e., one could hypothesize that the relationship between two variables is stronger for individuals who are high on a third characteristic than for individuals who are low on the third characteristic. For example, in the original article Saunders (1956) tested the hypothesis that the relationship between the Engineer Scale of the Strong Interest Blank (independent variable) and the grade point average (dependent variable) of engineering freshman is higher for non-compulsive (moderator variable) students than for compulsive students. This pattern of interrelationships among variables has widespread usefulness in model building. As stated by Saunders (1956), "The class of situations in which the 'moderated multiple regression' might be profitably studied can be made

100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141
142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162
163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183
184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204
205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225
226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246
247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267
268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288
289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309
310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330
331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351
352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372
373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393
394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414
415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435
436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456
457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477
478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498
499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519
520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540
541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561
562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582
583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603
604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624
625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645
646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666
667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687
688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708
709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729
730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750
751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771
772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792
793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813
814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834
835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855
856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876
877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897
898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918
919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939
940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960
961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981
982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002
1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023
1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044
1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065
1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086
1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107
1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128
1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149
1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170
1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191
1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212
1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233
1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254
1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275
1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296
1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317
1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338
1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359
1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380
1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401
1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422
1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443
1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464
1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485
1486	1487	1488	1489	1490	1491	1492	1493													

quite large, and can be seen to include a number of situations of potential practical significance" (p. 210).

The concept of moderators has been utilized in a number of subsequent theories and moderated regression has been empirically investigated in a variety of settings. The steps of the traditional procedure for empirically determining the existence of a moderator effect are presented by Zedeck (1971). "The correlations, slopes, and standard error of estimates should be examined for the following three regression equations:

$$y = a + bx, \quad [1]$$

$$y = a + b_1x + b_2z \quad [2]$$

where z is the potential moderator but is treated as an independent predictor, and

$$y = a + b_1x + b_2z + b_3xz \quad [3]$$

(moderated regression equation). If equations 2 and 3 are significantly different from equation 1, but not from each other, then the variable is an independent predictor and not a moderator variable" (Zedeck, 1971, p. 304). These suggested steps have been followed whenever a moderated regression analysis has been conducted. Cohen (1968) and Cohen and Cohen (1975) show that moderated regression is one conceptualization of an interaction effect in the general linear model. They describe a hierarchical analysis for assessing the existence of an interaction that is equivalent to the procedure outlined above, i.e., one first determines the squared multiple correlation coefficient (R^2) for model 2 and then determines the increment in R^2 caused by adding the interaction term in model 3. The increment¹ is then tested for statistical significance, using model 3 error mean square.

In this hierarchical treatment, the moderator effect must account for dependent variable variance beyond that accounted for by the main effects alone. The main effects are, thus, given a preferential claim to any covariance that is shared by main effects and the moderator effect. Such preference may be unreasonable under the conditions that generally hold when a moderated regression analysis is conducted. The moderated regression analysis is usually done only if there is some theoretical basis to suspect the presence of a moderator effect. If such is the case, the researcher should attempt to use an analysis with reasonable statistical power, i.e., the probability of discovering an existing hypothesized effect. This article will demonstrate that traditional moderated regression analysis procedure, on the contrary, enhances the possibility of type II error, i.e., the failure to recognize the presence of an existing effect.

This subtle but critical flaw in the traditional analysis procedure can explain why the results of moderated regression analyses led Zedeck (1971) to state "In general, moderated regression has not been successful in improving predictions" (p. 302). Competitive demands for journal space and the tendencies of researchers to report only significant finds allows speculation that many instances of failure to find moderated regression have gone unreported. These failures to discover moderator effects and the disappointing characteristics of those effects that have been found result from the features of the traditional analysis procedure that are demonstrated in the data below.

Among the reported studies that have used moderated regression analysis, the results have been quite similar but the conclusions drawn from them have changed from one study to another. Investigators have based their conclusions on two statistical characteristics of the moderated relationship--the statistical

significance of the increment due to the interaction term and the size of the increment.

Requirement of Statistical Significance

As examples of the moderated regression analysis procedure, Saunders (1956) presented three sets of data. In his first example, only one moderator effect was investigated. The increment in R^2 was .013, not significant. A second set of analyses tested the moderator effects of ten potential moderators. In this case, the increments ranged from zero to .025 with three of the ten reaching the .05 level of significance. His third set of demonstration data tested seven potential moderator effects. The increments ranged from zero to .017 with one of the effects significant at the .01 level. He concluded that a variable was operating as a moderator variable whenever the increment reached significance.

Jacobs and Solomon (in press) also concluded that a moderator effect had been shown when the R^2 increment was statistically significant. In their case 11 or 12 moderator effects were significant beyond the .05 level.² The increments in R^2 ranged from .009 to .104. In both the Saunders (1956) and Jacobs and Solomon (in press) studies the criterion for concluding that there is a moderator is the one stated by Zedeck, Cranny, Vale, and Smith (1971), "Presence of a moderator is thus indicated by the form of the equation, specifically by the significance of cross-product terms" (p. 239).

Requirement of a Sizable Increment

Stone (1976) and Stone, Mowday, and Porter (in press) drew a different conclusion from similar results. Stone (1976) showed an increment in R^2 of .016 significant at the .001 level. While acknowledging the statistical

significance of the moderator effect, he went on to conclude from the size of the increment that "the gain in explained...variance is not marked when the... interaction term is introduced into the multiple prediction equation. The practical significance...appears, therefore, to be negligible" (p. 163). In a discussion of a finding of an increment in R^2 of .019, significant at the .01 level, Stone, Mowday, and Porter (in press) state that "the difference between the two multiple correlations while statistically significant,...is of negligible size" (p. 8). In both of these cases the authors have taken the size of the increment in R^2 as an index of the strength of the moderator effect. If a large increment is a legitimate requirement for drawing the conclusion that a moderator should be attended to, then one can wonder if any moderated regression analysis has ever produced a moderator effect that was not "negligible." Further, if the size of the increment in R^2 is taken as an indication of the strength of the moderator effect, the uniformly modest effects that have been reported in the literature would argue that moderator effects are not important tools in scientific endeavors, i.e., conceptual models that include moderator effects could just as well do without them. This article will argue that the size of the increment in R^2 cannot be interpreted in that fashion.

The research reported in this article was conducted to test the efficiency of the traditional moderated regression analysis procedures for discovering moderators. The study considers whether the existence or the strength of moderator effects is accurately reflected in the results from the traditional moderated regression analysis.

Method

Data Generation

In order to simulate data on which to perform the moderated regression

analysis using a Monte Carlo approach, random normal deviates were generated according to the procedures of Box and Muller (1958). In all cases three variables, x , z , and e were constructed for each synthetic "individual." Variable x was treated in subsequent analyses as the independent variable, z was treated as the moderator variable, and e was used as error. The dependent variable, y , was constructed as described below for various cases. Except where it is noted in the definitions of the various cases given below, x and z had means ≈ 10 and standard deviations ≈ 1 ³; e had a mean ≈ 0 and a standard deviation ≈ 1 .

Case One. In the first case the dependent variable was constructed so as to represent a situation with a moderator effect and some error: The generating model was $y = xz + 3$. One hundred samples, each with 100 subjects, were constructed. This case served as the basic moderator effect case, and other cases were generated for comparison purposes. In the additional cases, changes were made in the formula for generating the dependent variable.

Case Two. The second case represents a situation with greater unreliability in the measurement. This effect was accomplished by increasing the variance of the error term. The same model ($y = xz + e$) was used to generate the dependent variable. For this case the standard deviation of the error variable ≈ 5 . Twenty samples, each with 100 subjects, were generated.

Case Three. The third case also used the basic model ($y = xz + e$) to create a dependent variable from a moderator effect. Unlike cases one and two, the independent variable and the moderator variable were constructed to be correlated. This relationship varied because of sample differences between .46 and 1.00 with a median value of .55. Twenty samples, each with 100 subjects, were generated.

Case Four. In the fourth case the dependent variable was built on a linear effects model rather than the interacting, moderator model. The generating model was $y = x + z + e$, thus specifying that y resulted from the additive combination of two unrelated variables plus error. Again, 20 samples of 100 subjects were constructed.

Case Five. The fifth case represents a situation with a dependent variable that is formed from a direct effect of the independent variable plus a moderator effect plus error. The constructive formula was $y = x + xz + e$. There were 20 samples of 100 subjects.

Case Six. In the final case the dependent variable is simply the independent variable plus error, and the formula for generating the dependent variable was $y = x + e$. There were 20 samples of 100 subjects each.

Table 1 summarizes the data generated for the six cases. Two hundred

 Insert Table 1 About Here

samples were generated. The dependent variable, y , was constructed differently in the samples representing the six different cases. Each case represents a different model of underlying effects. In cases one, two, and three the dependent variable results from a moderator effect and random error. The dependent variable is an additive combination of the linear effects of the independent, moderator, and error variables in case four where there is no moderator effect. In case five a combination of a direct linear effect of the independent variable and a moderator effect (plus error) determine the dependent variable. Finally, case six presents a dependent variable that is a sum of the independent variable and error with no moderator effect.

Data Analysis

In each sample the traditional moderated regression analysis was conducted to see if that procedure would "discover" moderator effects where they had been included in the model and would fail to "discover" moderator effects where they were not present.⁴ The increment in R^2 was determined in each sample as it would be in the traditional moderated regression analysis to represent the increase in predictability controlled by the moderator effect. The statistical significance of that increment was then tested using

$$F = \frac{96SS (xz|x,z)}{SS (error)},$$

where the numerator is the error degrees of freedom times the sum of squares of the moderator effect given the linear effects of the independent variable and the moderator variable, and the denominator is the sum of squares for error.⁵

Results

Statistical Significance

As shown in Table 2, the moderator effect was "discovered" by a requirement of statistical significance in Case 1 and Case 3. In both of those models all

Insert Table 2 About Here

of the samples provided an increment in R^2 for the moderator effect that was beyond that expected at the .05 level of significance. Though Case 2 and Case 5 also included a moderator effect in the formula that generated the dependent variable, the increment in R^2 did not reach statistically significant size in any of the samples of either of those cases. Case 2 differed from Case 1

in that the error component was greater in Case 2, i.e., there was more unreliability in the dependent variable. Case 5 differed from Case 1 by the inclusion of a direct relationship between the independent and dependent variables as well as the indirect, or moderated, relationship. The R^2 increment due to the moderator was not significant in any of the samples of the cases where the generating model did not include moderator effects, i.e., Case 4 and Case 6.

Size of the Increment

In all of the 200 samples the maximum increment in R^2 attributable to the moderator effect was .032 in one of the samples of Case 6. And there was not a moderator effect in the generating model for the dependent variable of Case 6! The minimum increment of zero was attained in three of the cases including Case 5 which involved a moderator effect. In general, none of the samples, either with or without actual moderator effects, provided large increments in R^2 .

Success of Moderators in Non-Moderator Models

In the two cases where there is no moderator effect in the actual determination of the dependent variable, Case 4 and Case 6, a comparison was made between prediction by the moderator effect alone versus the true model. For Case 4, prediction of y by x plus z (the true model) provides a range of correlations over the 20 samples from .716 to .875 with a median correlation of .806. If the prediction of y is by the moderator term xz, alone, the correlations range from .708 to .869 with a median of .801. In other words, the moderator model predicts the additive data nearly as well as does the additive model. In two of the samples the multiplicative correlation actually exceeded the additive correlation.

For Case 6 the true model does better than the moderator term at predicting the dependent variable. When y is predicted from x alone (the true model) the correlations range from .640 to .803 with a median of .715. There is no overlap with the correlations using the interaction, xz , as the predictor. The moderator prediction gives a range of correlations from .359 to .619 with a median of .505.

Discussion

Using the traditional moderated regression analysis procedures the influence of a moderator effect was not shown even when the dependent variable was constructed to include a strong moderator effect. Cases 1, 2, and 3 created the dependent variable as a moderator effect plus some random error. Under those conditions Case 2 with a large error component did not demonstrate statistical significance for the moderator effect, and all three of the cases produced such small increments in R^2 attributable to the moderator that the moderator effect could have been called negligible if the size of the increment were interpreted as an indication of the strength of the effect. In Case 5, where the moderator influence is accompanied by both error and by a linear effect of the independent variable, there was neither statistical significance nor a sizable increment. That is, even in situations strongly biased in favor of discovering a moderator effect, i.e., a moderator effect has been used in the creation of the dependent variable, the traditional moderated regression analysis leads to the conclusion of "no moderator."

Why should such a result occur? Why should this analysis be so prone to overlook an effect that is in fact in the data? There are two influences that increase the likelihood of this type II error.

1. The hierarchical nature of the multiple regression procedures robs the moderator term of covariance with the dependent variable.

The traditional moderated regression analysis procedure first regresses the dependent variable on the linear effects of the independent variable and the potential moderator variable. The interaction term is then regressed on only the residual from that regression. The increment in R^2 which is used to judge both the existence and the strength of the moderator effect represents only the relationship with that part of the dependent variable variance that is unrelated to the two linear effects.

Any covariance with the ~~dependent variable~~ that is shared with the linear effects is automatically assigned to the linear effects. This implies a preference for the linear effects that is implicitly a theoretical hypothesis of the generating effects that have produced the dependent variable. This implied priority of the main effects should be supported by the researcher's theoretical understanding of the relationships among the variables; it cannot be accepted simply because the hierarchical analysis has traditionally proceeded in that direction. The ordering of effects in the hierarchical regression analysis is arbitrary. With appropriate theoretical justification the data analyst can, as validly, first assess the variance attributable to the moderator effect, and then determine whether the main effects make a significant and sizable addition to the variance accounted for. The point is, that the decision of the order in which effects are assessed is one that the researcher must support theoretically. Because the order of the removal of variance due to effects is an expression of a theoretical preference, the investigator cannot abnegate the responsibility for this decision on the grounds of tradition. This

decision requires, not just blind application of techniques, but scientific judgment and justification.

2. The linear regression model is a robust method for accounting for variance.

Even if the prediction equation is an imperfect model of the process that generated the data, the linear regression model can often capture a great deal of the variance. The problem of using correlational procedures to assess the appropriateness of an hypothesized model has been discussed in the literature (Alf & Abrahams, 1974; Birnbaum, 1973, 1974; Rorer, 1974). Watson (1972) has demonstrated that this same problem can occur in the analysis of variance. He showed that data generated from an interaction might be accounted for by significant main effects leaving insignificant variance to be accounted for by the interaction.

In a demonstration similar to the one presented in this article Rorer (Note 1) used a linear regression model to replicate the decisions that had been made with several different decision strategies from constructed (synthesized) data. Using 11 different decision strategies that included linear, curvilinear, configural, sequential and categorical processes he was able to closely approximate the decisions of the real functions with a simple linear function. Discussing a hierarchical process for drawing conclusions that is the same as the hierarchical strategy of the traditional moderated regression analysis he states, "These results indicate vividly that it is difficult to infer nonlinearity, even when it exists" (p. 12). "These results show why the search for moderator variables has been so fruitless, just as much as they show why human judges have been unable to improve upon actuarial prediction

by means of their purported ability to recognize and use patterned or unusual relationships. Even if such relationships exist, their incorporation into the decision strategy, be it actuarial or subjective, is not going to result in much improvement over the linear regression function in either case" (p. 13). The linear regression model is robust enough to capture much of the predictable variance of a dependent variable even if that dependent variable sprang from other than linear origins.

The critical distinction for the problem addressed in this paper is the difference between science and application. If a research has only the applied purpose of determining or demonstrating how variance can be accounted for, a statistical data analysis can provide answers. However, if the research has the scientific purpose of discovering or confirming an underlying model or generating process that produced the data, then traditional moderated regression analysis is inadequate for the task. The scientific question may be answered in part by a data analysis if the results demonstrate that a model is incompatible with the data, but the data analysis cannot choose between alternative models that all fit the data. For example, if y is predicted equally well by $x + z$ and by xz , the data analysis is neutral as to choice. The selection of one model over the other rests with the rational scientific judgment and preference of the researcher.

There are viable theoretical contents that may lead a researcher to support a moderator effect as the correct model for his/her data. In his presentation of the moderated regression analysis Saunders (1956) proposed examples where a moderator effect was conceptually supportable. He included such examples as using insight as a moderator of the relationship between self

reports and more objective measures of personality traits and using emotional stability as a moderator of the relationship between academic ability measures and academic success. Another interactive model that has been presented with elaborate theoretical justifications is expectancy theory (Staw, 1976; Vroom, 1964). Researchers in that area are likely to choose the multiplicative expectancy model for explanatory use even when an additive model is equally successful at accounting for variance. Their preference would be based on the rational depiction of the relationship among the variables that is afforded by the expectancy theory. Blood (1977) presents a model of the operation of the cognitive process of self rewarding on work performance. The model includes five variables acting as moderators. The logic of interactive relationships is preferred for the model because of the nature of the variables involved. This theoretical preference is rationally justified.

The issue of theoretical parsimony is not involved in the choice between a linear model and a moderator model. Whether two main effects or one interaction effect is the more parsimonious explanation is a matter of taste. Some people will prefer an additive model and other will elect the multiplicative, but it is not clear that one is simpler than the other. In addition to which, the principle of parsimony is used to choose between theories when other considerations are equal. If other considerations lead to its preference, one may be justified in choosing a less parsimonious model over a more parsimonious one.

This leaves us with the important task of suggesting how a researcher should proceed to assess the strength of a moderator effect. If there are theoretical reasons to support a hypothesis of a moderator effect, the dependent

variable should be directly regressed on the cross-product of the independent variable and the potential moderator,

$$y = xz. \quad [4]$$

Then if desired, the additional contribution of the main effects can be assessed by determining the size and significance of their increment to the squared multiple correlation. That is, the hierarchical regression procedure should proceed in exactly the reverse order from the traditional moderated regression analysis procedure in those cases where a moderator effect is theoretically justified (see Cohen and Cohen, 1975 for analysis procedures).

There may be some cases where theory will suggest that the independent variable will influence the dependent variable partly through a main effect and partly through a moderated effect (as in Case 5). In that event, a researcher may wish to assess the size and significance of the combined main and moderator effects by the regression of

$$y = x + xz. \quad [5]$$

Should this analysis be used, caution should be exercised in interpreting the multiple correlation weights. Because of colinearity influences these will not indicate the relative contributions of the main and moderator effects (see Darlington, 1968; Cohen and Cohen, 1975).

A potential problem with these suggested procedures is that they have the possibility of also encouraging Type II error. If a model is underfit, true relationships can go undetected due to overestimation of the error term (Johnston, 1972). Variance that is uniquely related to the excluded variables is added to the error variance. This means, for example, that if a model of the form of [4 or [5 were used when the true (but unknown) model is [3, then

the real moderator effect could remain statistically insignificant because of this bias. This problem exists whenever a hierarchical regression procedure is used and whenever the model in use may have excluded a variable that controls unique variance. This potential bias did not cause a problem in the present data, and it appears to be a less severe disturbance than the traditional moderated regression analysis procedure.

Whatever the theoretical predispositions of the investigator, an analysis should be chosen such that it is possible to confirm it, i.e., the analysis should not virtually guarantee type II error. The theory should guide the analysis, not the reverse. In the hierarchical analysis traditionally used to discern moderated regression, it is not possible to rule out a moderator as a generating factor for the dependent variable by showing that the main effects alone can account for most of the variance. Thoughtlessly following the traditional analytic procedures cannot replace the exercise of scientific judgment based on carefully drawn theory.

Reference Note

1. Rorer, L. G. A circuitous route to bootstrapping selection procedures
(Research Bulletin Vol. 12, No. 9). Eugene, Ore.: Oregon Research
Institute, 1972.

References

- Alf, E. F., & Abrahams, N. M. Let's give the devil his due: A response to Birnbaum. Psychological Bulletin, 1974, 81, 72-73.
- Birnbaum, M. H. The devil rides again: Correlation as an index of fit. Psychological Bulletin, 1973, 79, 239-242.
- Birnbaum, M. H. Reply to the devil's advocates: Don't confound model testing and measurement. Psychological Bulletin, 1974, 81, 854-859.
- Blood, M. R. Organizational control of performance through self rewarding. In King, B. T., Streufert, S. S., and Fiedler, F. E. (Eds.), Managerial Control and Organizational Democracy. Washington, D. C.: Victor Winston and Sons, 1977.
- Box, G. E. P., & Muller, M. E. A note on the generation of random normal deviates. Annals of Mathematical Statistics, 1958, 29, 610-611.
- Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968, 70, 426-433.
- Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1975.
- Darlington, R. B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.
- Jacobs, R., & Solomon, T. Strategies for the prediction of job performance. Journal of Applied Psychology, in press.
- Johnston, J. Econometric methods, 2nd ed. New York: McGraw-Hill, 1972.
- Rorer, L. G. "What, can the devill speake true?" Psychological Bulletin, 1974, 81, 355-357.
- Saunders, D. R. Moderator variables in prediction. Educational and Psychological Measurement, 1956, 16, 209-222.
- Staw, B. M. Intrinsic and extrinsic motivation. Morristown, N. J.: General Learning Press, 1975.
- Stone, E. F. The moderating effect of work-related values on the job scope-job satisfaction relationship. Organizational Behavior and Human Performance, 1976, 15, 147-167.
- Stone, E. F., Mowday, R. T., & Porter, L. W. Higher-order need strengths as moderators of the job scope-job satisfaction relationship. Journal of Applied Psychology, in press.

References (Con'td.)

Vroom, V. H. Work and motivation. New York: Wiley, 1964.

Watson, J. S. We assess lawfulness, but God knows what the law is. Proceedings of the 80th Annual Convention of the American Psychological Association, 1972, 7, 139-140. (Summary)

Zedeck, S. Problems with the use of "moderator" variables. Psychological Bulletin, 1971, 76, 295-310.

Zedeck, S., Cranny, C. J., Vale, C. A., and Smith, P. C. Comparison of "joint moderators" in three prediction techniques. Journal of Applied Psychology, 1971, 55, 234-240.

Footnotes

¹This increment is called the squared semipartial correlation coefficient by Cohen and Cohen (1975) and the "Usefulness" index by Darlington (1968).

²Because of interdependencies among the variables, these were not 12 independent effects.

³It is important to note that the distributions of independent and moderator variables included only positive numbers. This is usual for social science data. Because of the manner of construction of y , the results would be changed if x and z took on both positive and negative values.

⁴The authors wish to thank Mark Elliott and Julie Bierer for their help with the data generation and analysis for this study.

⁵This is the sum of squares for error of this prediction model, not sum of squares for the "e" term used in generating the dependent variable.

Table 1
Structure of the Simulated Samples

Case	Number of Samples	Subjects per Sample	Generating Formula	Special Features
1	100	100	$y = xz + e$	
2	20	100	$y = xz + e$	$Se = 5Sx = 5Sz$
3	20	100	$y = xz + e$	$mdnr_{xz} = .55$
4	20	100	$y = x + z + e$	
5	20	100	$y = x + xz + e$	
6	20	100	$y = x + e$	

Table 2
Results of Traditional Moderated Regression Analysis

Case	Average Increment in R^2	Minimum Increment in R^2	Maximum Increment in R^2	Increments Significant at .05 level
1	.005	.002	.009	100 of 100
2	.006	.001	.022	0 of 20
3	.006	.003	.010	20 of 20
4	.004	.000	.015	0 of 20
5	.003	.000	.016	0 of 20
6	.008	.000	.032	0 of 20

Distribution List

Office of Naval Research (3 copies)
(Code 452)
800 N. Quincy St.
Arlington, VA 22217

Director
U. S. Naval Research Laboratory (6 copies)
Washington, DC 20390
ATTN: Technical Information Division

Defense Documentation Center (12 copies)
Building 5
Cameron Station
Alexandria, VA 22314

Library, Code 2029 (6 copies)
U. S. Naval Research Laboratory
Washington, DC 20390

Science & Technology Division
Library of Congress
Washington, DC 20540

Director
ONR Branch Office
495 Summer St.
Boston, MA 02210

Psychologist
ONR Branch Office
495 Summer St.
Boston, MA 02210

Director
ONR Branch Office
536 S. Clark St.
Chicago, IL 60605

Director
ONR Branch Office
1030 E. Green St.
Pasadena, CA 91106

Psychologist
ONR Branch Office
1030 E. Green St.
Pasadena, CA 91106

Research Psychologist
ONR Branch Office
536 S. Clark St.
Chicago, IL 60605

Dr. Earl A. Alluisi
Old Dominion University Res. Foundation
Norfolk, VA 23508

Dr. Judith Daly
Decisions & Designs, Inc.
Suite 100
8400 Westpark Dr.
McLean, VA 22101

Dr. James A. Bayton
Department of Psychology
Howard University
Washington, DC 20001

Dr. H. Russell Bernard
Dept. of Sociology & Anthropology
West Virginia University
Morgantown, W. VA 26506

Dr. Arthur Blaiwes
Naval Training Equipment Center
Orlando, FL 32813

Dr. Milton R. Blood
School of Business
Georgia Institute of Technology
Atlanta, GA 30332

Dr. Davis B. Bobrow
University of Maryland
Department of Government & Politics
College Park, MD 20742

Dr. David G. Bowers
Institute for Social Research
University of Michigan
Ann Arbor, MI 48106

Dr. John J. Collins
Vice President
Essex Corporation
6305 Caminito Estrellado
San Diego, CA 92120

Dr. Harry R. Day
University City Science Center
Center for Social Development
3624 Science Center
Philadelphia, PA 19104

Dr. C. Brooklyn Derr
Associate Professor, Code 55
Naval Post Graduate School
Monterey, CA 93940

Distribution List (continued)

Dr. George T. Duncan
Carnegie-Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213

Dr. Samuel L. Gaertner
Department of Psychology
University of Delaware
220 Wolf Hall
Newark, DE 19711

Dr. William E. Gaymon
Suite 200
1055 Thomas Jefferson St., NW
Washington, DC 20007

Dr. Paul S. Goodman
Graduate School of Industrial Admin.
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. J. Richard Hackman
Administrative Sciences
Yale University
56 Hillhouse Ave.
New Haven, CT 06520

Dr. Leo A. Hazlewood
CACI, Inc.
1815 Fort Myer Dr.
Arlington, VA 22209

Dr. Edwin Hollander
Department of Psychology
State University of New York at Buffalo
4230 Ridge Lea Rd.
Buffalo, NY 14226

Mr. Daniel F. Huck
General Research Corp.
Westgate Research Park
McLean, VA 22101

Dr. Charles L. Hulin
Department of Psychology
University of Illinois
Champaign, IL 61820

Dr. Rudi Klauss
Syracuse University
Public Administration Dept.
Maxwell School
Syracuse, NY 13210

Dr. Edward E. Lawler
Battelle Human Affairs Research Centers
4000 N.E. 41st St.
P.O. Box 5395
Seattle, WA 98105

Dr. Arie Y. Lewin
Duke University
Duke Station
Durham, NC 27706

Dr. Morgan W. McCall, Jr.
Center for Creative Leadership
5000 Laurinda Dr.
P.O. Box P-1
Greensboro, NC 27402

Dr. Terence R. Mitchell
School of Business Administration
University of Washington
Seattle, WA 98195

Dr. William H. Mobley
College of Business Administration
University of South Carolina
Columbia, SC 29208

Dr. Thomas D. Morris
The Brookings Institution
1775 Massachusetts Ave., NW
Washington, DC 20036

Dr. James P. Murphy
National Analysts
A Division of Booz-Allen & Hamilton, Inc.
400 Market St.
Philadelphia, PA 19106

Dr. Peter G. Nordlie
Human Sciences Research, Inc.
7710 Old Springhouse Rd.
McLean, VA 22101

Dr. Herbert R. Northrup
Industrial Research Unit
University of Pennsylvania
Philadelphia, PA 19174

Dr. A. F. K. Organski
3068 Institute for Social Research
University of Michigan
Ann Arbor, MI 48104

Distribution List (continued)

Dr. Paul Pedersen
Society for Intercultural Education,
Training and Research
107 MIB, University of Pittsburgh
Pittsburgh, PA 15260

Dr. Manuel Ramirez
Systems and Evaluations
232 Swanton Blvd.
Santa Cruz, CA 95060

Dr. Irwin Sarason
Department of Psychology
University of Washington
Seattle, WA 98195

Dr. S. B. Sells
Texas Christian University
Fort Worth, TX 76129

Dr. H. Wallace Sinaiko
Program Director
Manpower Research & Advisory Services
Smithsonian Institution
801 N. Pitt St. - Suite 120
Alexandria, VA 22314

Mrs. Alice I. Snyder
Mental Health Clinic
Naval Regional Medical Center
Pearl Harbor
FPO San Francisco, CA 96610

Dr. Richard Steers
Graduate School of Management & Business
University of Oregon
Eugene, OR 97403

Dr. Victor H. Vroom
School of Organizational Management
Yale University
56 Hillhouse Ave.
New Haven, CT 06520

Dr. Abraham R. Wagner
Analytical Assessments Corp.
357 South Robertson Blvd.
Beverly Hills, CA 90211

Dr. J. Wilkenfeld
Department of Government & Politics
College Park, MD 20742

AFOSR/NL
Bldg., 410
Bolling AFB
Washington, DC 20332

Military Assistant for Human Resources
OAD (E&LS) ODDR&E
Pentagon 3D129
Washington, DC 20301

HQ, USAF
AFMPC/DPMYP
Randolph AFB, TX 78148

Air University Library/LSE-8110
Maxwell AFB, AL 36112

Office of the Deputy Chief of Staff
for Personnel, Research Office
ATTN: DAPE-PBR
Washington, DC 20310

Chief, Plans & Operations Office
USA Research Institute for the
Behavioral & Social Sciences
Room 278
1300 Wilson Blvd.
Arlington, VA 22209

Army Research Institute (2 cys)
Commonwealth Bldg.
1300 Wilson Blvd.
Rosslyn, VA 22209

ARI Field Unit - Leavenworth
P.O. Box 3122
Fort Leavenworth, KS 66027

Headquarters, Forces Command
AFPE-HR
Ft. McPherson
Atlanta, GA 30330

Dr. A. L. Slafkosky
Code RD-1
HQ US Marine Corps
Washington, DC 20380

Commandant of the Marine Corps.
(Code MPI-20)
Washington, DC 20380

Distribution List (continued)

Chief, Psychological Research Branch
US Coast Guard (G-P-1/62)
400 7th St., SW
Washington, DC 20590

Chief of Naval Personnel
Assistant for Research Liaison (Pers-Or)
Washington, DC 20370

Bureau of Naval Personnel (Pers 6)
Assistant Chief of Naval Personnel for
Human Resource Management
Washington, DC 20370

Bureau of Naval Personnel (Pers-6a3)
Human Resource Management Financial Office
Washington, DC 20370

CDR Paul D. Nelson, MSC, USN
Head, Human Performance Division (Code 44)
Navy Medical R&D Command
Bethesda, MD 20014

Assistant Officer in Charge
Naval Internal Relations Activity
Pentagon, Room 2E329
Washington, DC 20350

Naval Postgraduate School
Monterey, CA 93940
ATTN: Library (Code 2124)

Professor John Senger
Operations Research & Administration
Sciences
Naval Postgraduate School
Monterey, CA 93940

Head, Human Factors Engineering Branch
(1242)
USN Pacific Missile Center
Pt. Mugu, CA 93042

Training Officer
Human Resource Management Center
San Diego, CA 92133 NTC

Scientific Director
Naval Health Research Center
San Diego, CA 92152

Navy Personnel R&D Center (5 cys)
Code 01
San Diego, CA 92152

Commanding Officer
Naval Submarine Medical Research Lab.
Naval Submarine Base New London, Box 900
Groton, CT 06340

Commanding Officer
Naval Training Equipment Center
Technical Library
Orlando, FL 32813

Naval Aerospace Medical Research Lab.
(Code L5)
Naval Aerospace Medical Center
Pensacola, FL 32512

Lt. Rebecca G. Vinson, U.S.N.
Navy Recruiting District, Boston
575 Technology Square
Cambridge, MA 02139

Chief, Naval Technical Training
NAS Memphis (75)
Millington, TN 38054
ATTN: Mr. Tom Warrick, N622

Human Resource Management Center
Box 23
FPO New York 09510

Human Resource Management Center, Norfolk
5621-23 Tidewater Drive
Norfolk, VA 23511

Human Resource Management Center
Bldg., 304
Naval Training Center
San Diego, CA 92133

Office of Naval Research (Code 200)
Arlington, VA 22217

ACOS Research & Program Development
Chief of Naval Education & Training (N-5)
Naval Air Station
Pensacola, FL 32508

Human Resource Management Center
Pearl Harbor
FPO San Francisco, CA 96601

Human Resource Management School
Naval Air Station Memphis (96)
Millington, TN 38054

Distribution List (continued)

Director, Programs Development Division
(Code N-35)
Chief of Naval Education & Training
Naval Air Station
Pensacola, FL 32508

Capt. Charles Baldwin
Bureau of Naval Personnel
Pers 65
Washington, DC 20370

Director, Human Resource Training Department
Naval Amphibious School
Little Creek
Naval Amphibious Base
Norfolk, VA 23521

Navy Material Command
Employee Development Office (Code SA-65)
Room 140 Jefferson Plaza, Bldg., #2
1429 Jeff Davis Highway
Arlington, VA 20360

Human Resource Management Center, Washington
Washington, DC 20370

Dr. Robert B. Tebbs
Franklin University
201 South Grant Ave.
Columbus, OH 43215

HumRRO (ATTN: Library)
300 N. Washington St.
Alexandria, VA 22314

Director, Columbus Office
HumRRO Central Division, Suite 23
2601 Cross Country Dr.
Columbus, GA 31906

Office of the Air Attache
Embassy of Australia
1601 Massachusetts Ave., NW
Washington, DC 20036

Scientific Information Officer
British Embassy
3100 Massachusetts Ave., NW
Washington, DC 20008

Canadian Defense Liaison Staff, Washington
2450 Massachusetts Ave., NW
Washington, DC 20008
ATTN: Chief, Defense Research

Mr. Luigi Petrullo
2431 N. Edgewood St.
Arlington, VA 22207

Dr. John J. Collins
6305 Caminito Estrellado
San Diego, CA 92120