

Example 6b: Path Analysis for Mediation Predicting Logistic Outcomes
(complete syntax and output available for Mplus and STATA electronically, along with publication)

Figure 1 and Table 1 from: Hoffman, L., & McDowd, J. M. (2010). Simulator driving performance predicts accident reports five years later. *Psychology and Aging*, 25(3), 741-745.

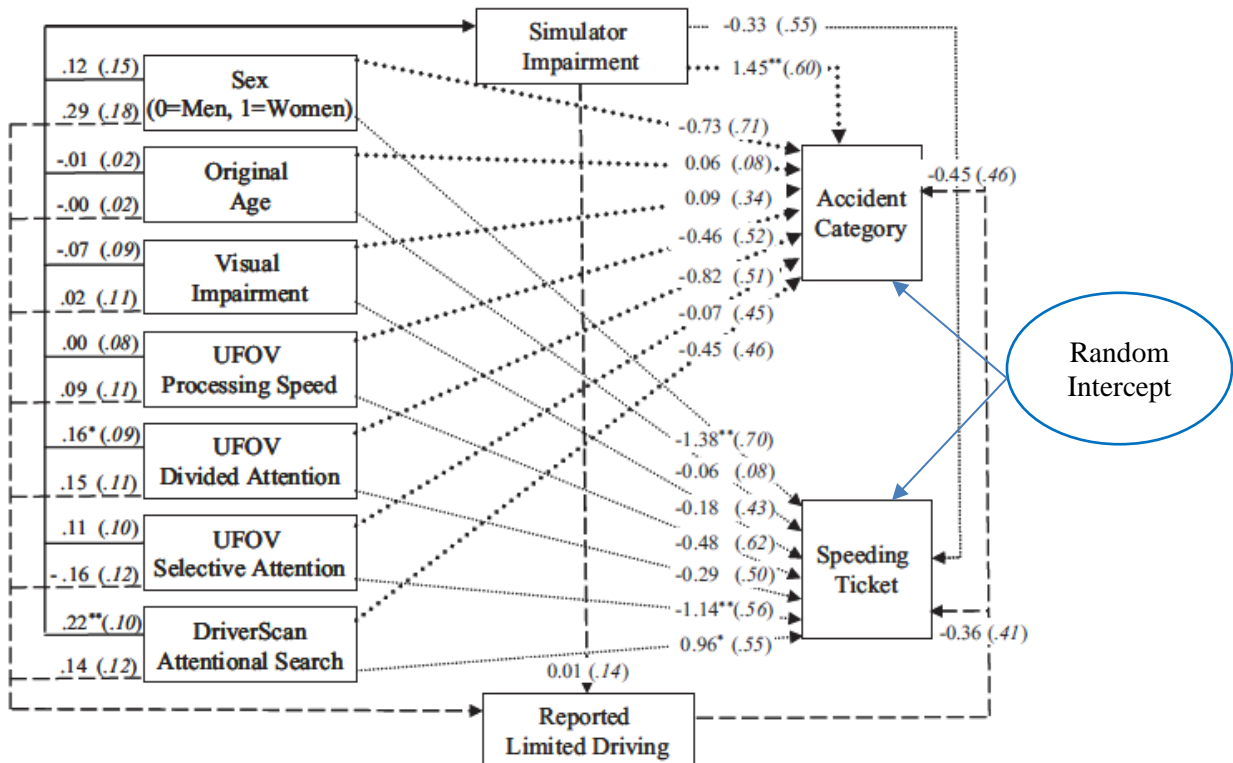
This study reports on follow-up data for 114 of 152 persons originally tested in my dissertation in 2003, which focused on the role of vision and attention in predicting simulator driving impairment. The goal was to see if any of the original study variables (left panel plus simulator impairment) would predict future reports of limited driving, involvement in an at-least-partially-at-fault accident (14/114) or receipt of a speeding ticket (14/144).

For this example, the original model shown in Figure 1 was expanded to include a random intercept latent variable to create covariance between the two binary outcomes (which cannot be added directly when using full-information maximum likelihood estimation). I also added an indirect effect for demonstration purposes. For estimation, I switched to robust ML in Mplus, and invoked robust standard errors (to mimic robust ML) in STATA GSEM. However, because STATA GSEM did equation-wise deletion of missing cases (17 cases), the results do not match those of Mplus.

Table 1
Bivariate Correlations Between Predictor and Outcome Variables

Variable	1	2	3	4	5	6	7	8	9	10	11
Outcomes											
1. Accident report	—										
2. Speeding ticket report	-.14	—									
Predictors											
3. Reported limited driving	-.25	-.22	—								
4. Simulator impairment	.21	-.22	.14	—							
5. Sex (0 = men, 1 = women)	-.23	-.35*	.21*	.13	—						
6. Original age	.04	-.20	.05	.16	-.03	—					
7. Visual impairment	-.00	-.18	.02	.05	-.05	.22*	—				
8. UFOV processing speed	-.15	-.28	.12	.17*	-.10	.12	.13	—			
9. UFOV divided attention	-.24	-.31*	.17*	.40*	.04	.25*	.17	.30*	—		
10. UFOV selective attention	-.08	-.43*	.03	.38*	-.09	.36*	.29*	.29*	.52*	—	
11. DriverScan attentional search	-.16	-.14	.15	.43*	.06	.41*	.20*	.20*	.45*	.60*	—

Note. UFOV = Useful Field of View test.
 * $p < .05$.



Mplus Syntax and Partial Output for Path Model with Random Intercept:

```

TITLE: Example 6b: Mplus Path Analysis for Dissertation Follow-up
DATA: FILE = driver.csv;           ! Can just list file name if in same folder;
        FORMAT = free;               ! FREE (default) or FIXED format;
        TYPE = individual;          ! Individual (default) or matrix data as input;
VARIABLE:
! List of ALL variables in original wide data file, in order;
! Mplus names must use 8 characters or fewer (so rename as needed);
  NAMES = PartID sex age75 cs_1_5 cs_3 cs_6 cs_12 cs_18 far near
        zufov1 zufov2 zufov3 Dscan lane da_task crash stop speed time
        simfac part visfac attfac limit4 ticket2 speed2 follow attr
        nacc2 jacc2 jacc20 acc2;
! List of ALL variables used in model;
  USEVARIABLE = sex age75 visfac zufov1 zufov2 zufov3 Dscan simfac
        limit4 speed2 acc2;
! Missing data identifier;
  MISSING = ALL (-999);
! Select only follow-up cases;
  USEOBS = follow EQ 1;
! Categorical outcomes;
  CATEGORICAL = acc2 speed2;

ANALYSIS: LINK = LOGIT;           ! Link function for categorical outcomes;
        ESTIMATOR = MLR;           ! Robust full-information maximum likelihood;
        INTEGRATION = MONTECARLO(1000); ! Mplus required (#samples);
OUTPUT: CINTERVAL;             ! Print confidence intervals;
        STDYX;                   ! Print fully standardized solution, too;
        SAMPSTAT;               ! Print descriptive statistics;

MODEL: ! * --> Estimated parameter (all listed below for clarity);

! Outcome intercepts (for continuous variables);
[simfac* limit4*];
! Outcome thresholds (for binary variables);
[speed2$1* acc2$1*];

! Regressions: y outcomes ON x predictors (label to do math on later, * implied);
simfac ON sex age75 visfac zufov1 zufov2 zufov3 Dscan (sim1-sim7);
limit4 ON sex age75 visfac zufov1 zufov2 zufov3 Dscan simfac (lim1-lim8);
acc2 ON sex age75 visfac zufov1 zufov2 zufov3 Dscan simfac limit4 (acc1-acc9);
speed2 ON sex age75 visfac zufov1 zufov2 zufov3 Dscan simfac limit4 (spd1-spd9);

! Estimated residual variances for continuous outcomes;
simfac* limit4*;

! Random intercept factor for binary outcome covariance;
RandInt BY speed2@1 acc2@1;
[RandInt@0]; ! Fix fixed intercept to 0;
RandInt*; ! Estimate random intercept variance;

MODEL CONSTRAINT:           ! Like ESTIMATE in SAS or LINCOM in STATA;
  NEW(DStoAcc);             ! List names of estimated effects on NEW;
  DStoAcc = sim7 * acc2;    ! Indirect effect of Dscan --> Sim --> Acc;

```

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	114 → It used all cases for all equations
Number of Free Parameters	40
Loglikelihood	
H0 Value	-318.727
H0 Scaling Correction Factor for MLR	0.9958
Information Criteria	
Akaike (AIC)	717.455
Bayesian (BIC)	826.903
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	700.476

These are all the fit statistics we get—
there is no saturated model or null model
easily possible when not all variables are
conditionally multivariate normal.

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
RANDINT BY				
SPEED2	1.000	0.000	999.000	999.000
ACC2	1.000	0.000	999.000	999.000
SIMFAC ON				
SEX	0.125	0.134	0.931	0.352
AGE75	-0.005	0.017	-0.276	0.782
VISFAC	-0.067	0.085	-0.785	0.433
ZUFOV1	0.002	0.099	0.020	0.984
ZUFOV2	0.166	0.103	1.611	0.107 → Was marginal, now not
ZUFOV3	0.112	0.102	1.095	0.274
DSCAN	0.218	0.098	2.228	0.026 → Also found in dissertation original sample
LIMIT4 ON				
SEX	0.292	0.192	1.520	0.129
AGE75	-0.002	0.023	-0.078	0.938
VISFAC	0.013	0.102	0.127	0.899
ZUFOV1	0.098	0.125	0.782	0.434
ZUFOV2	0.150	0.125	1.204	0.229
ZUFOV3	-0.154	0.134	-1.150	0.250
DSCAN	0.138	0.128	1.076	0.282
SIMFAC	0.018	0.150	0.122	0.903
ACC2 ON				
SEX	-0.761	0.559	-1.361	0.174
AGE75	0.059	0.083	0.711	0.477
VISFAC	0.086	0.310	0.279	0.781
ZUFOV1	-0.450	0.614	-0.733	0.463
ZUFOV2	-0.838	0.384	-2.178	0.029 → New significant result given robust SEs
ZUFOV3	-0.060	0.389	-0.155	0.877
DSCAN	-0.458	0.324	-1.415	0.157
SIMFAC	1.461	0.543	2.692	0.007 → Reason for the publication
LIMIT4	-0.419	0.406	-1.033	0.302
SPEED2 ON				
SEX	-1.388	0.746	-1.862	0.063 → Was p=.048, now NS with robust SEs
AGE75	-0.064	0.052	-1.226	0.220
VISFAC	-0.183	0.350	-0.523	0.601
ZUFOV1	-0.478	0.485	-0.986	0.324
ZUFOV2	-0.286	0.607	-0.471	0.638
ZUFOV3	-1.143	0.375	-3.045	0.002 → Also reported significant in publication
DSCAN	0.964	0.516	1.869	0.062 → Also reported marginal in publication
SIMFAC	-0.346	0.692	-0.500	0.617
LIMIT4	-0.353	0.486	-0.726	0.468
Means				
RANDINT	0.000	0.000	999.000	999.000
Intercepts				
SIMFAC	-0.082	0.103	-0.794	0.427
LIMIT4	0.075	0.150	0.502	0.616
Thresholds				
SPEED2\$1	1.738	0.511	3.399	0.001
ACC2\$1	2.050	0.533	3.843	0.000
Variances				
RANDINT	0.049	0.005	9.158	0.000 → Not included in publication model (should have)
Residual Variances				
SIMFAC	0.465	0.061	7.598	0.000
LIMIT4	0.850	0.114	7.481	0.000
New/Additional Parameters				
DSTOACC	0.319	0.175	1.824	0.068 → Not tested for publication (just demo here)

LOGISTIC REGRESSION ODDS RATIO RESULTS

	Estimate	S.E.	(Est. - 1) / S.E.	Two-Tailed P-Value
ACC2 ON				
SEX	0.467	0.261	-2.040	0.041
AGE75	1.061	0.088	0.690	0.490
VISFAC	1.090	0.337	0.267	0.790
ZUFOV1	0.638	0.391	-0.926	0.354
ZUFOV2	0.433	0.166	-3.409	0.001
ZUFOV3	0.942	0.367	-0.160	0.873
DSCAN	0.633	0.205	-1.795	0.073
SIMFAC	4.309	2.338	1.415	0.157
LIMIT4	0.658	0.267	-1.283	0.199
SPEED2 ON				
SEX	0.250	0.186	-4.032	0.000
AGE75	0.938	0.049	-1.266	0.205
VISFAC	0.833	0.292	-0.574	0.566

ZUFOV1	0.620	0.301	-1.265	0.206
ZUFOV2	0.751	0.456	-0.546	0.585
ZUFOV3	0.319	0.120	-5.692	0.000
DSCAN	2.622	1.352	1.199	0.230
SIMFAC	0.708	0.489	-0.598	0.550
LIMIT4	0.703	0.341	-0.871	0.384

STATA GSEM Syntax and Partial Output for Path Model with Random Intercept:

```

display "STATA Path Model for Example 6b"
display "Results do not match Mplus because of missing data"
gsem
(speed2@1 acc2@1 <-RandInt)
(simfac limit4 speed2 acc2 RandInt@0 <- _cons)
(simfac <- sex age75 visfac zufov1 zufov2 zufov3 dscan)
(limit4 <- sex age75 visfac zufov1 zufov2 zufov3 dscan simfac)
(acc2 <- sex age75 visfac zufov1 zufov2 zufov3 dscan simfac limit4, logit)
(speed2 <- sex age75 visfac zufov1 zufov2 zufov3 dscan simfac limit4, logit),
var(e.simfac e.limit4 e.RandInt)
method(ml) vce(robust)
gsem, coeflegend
nlcom _b[simfac:dscan]*_b[acc2:simfac]
estat eform speed2 acc2

```

```

Generalized structural equation model          Number of obs   =          97 → Number of complete cases

Response      : speed2                      Number of obs   =          94
Family        : Bernoulli
Link          : logit

Response      : acc2                       Number of obs   =          95
Family        : Bernoulli
Link          : logit

Response      : simfac                     Number of obs   =          97
Family        : Gaussian
Link          : identity

Response      : limit4                    Number of obs   =          95
Family        : Gaussian
Link          : identity

```

```

Log pseudolikelihood = -279.9549
( 1) [speed2]RandInt = 1
( 2) [acc2]RandInt = 1

```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	

speed2						
sex	-1.557061	.7879778	-1.98	0.048	-3.101469	-.0126527
age75	-.0546159	.0525276	-1.04	0.298	-.1575682	.0483363
visfac	-.1502138	.3466184	-0.43	0.665	-.8295735	.5291458
zufov1	-.5356348	.4797669	-1.12	0.264	-1.475961	.404691
zufov2	-.3621424	.5854344	-0.62	0.536	-1.509573	.7852879
zufov3	-1.123434	.3804918	-2.95	0.003	-1.869185	-.3776841
dscan	1.072291	.5491541	1.95	0.051	-.0040308	2.148614
simfac	-.2707209	.683736	-0.40	0.692	-1.610819	1.069377
limit4	-.3472816	.4811235	-0.72	0.470	-1.290266	.5957032
RandInt	1	(constrained)				
_cons	-1.441395	.4920703	-2.93	0.003	-2.405836	-.4769553 → intercept, not threshold

acc2						
sex	-.7558821	.68301	-1.11	0.268	-2.094557	.5827929
age75	-.1174961	.0851259	-1.38	0.168	-.2843399	.0493477
visfac	.4131037	.2933547	1.41	0.159	-.1618609	.9880684
zufov1	-1.670182	1.457019	-1.15	0.252	-4.525887	1.185524
zufov2	-.7274938	.3456611	-2.10	0.035	-1.404977	-.0500106
zufov3	-.0820221	.5071776	-0.16	0.872	-1.076072	.9120277
dscan	-.5146463	.3802142	-1.35	0.176	-1.259852	.2305598
simfac	1.795184	.8688706	2.07	0.039	.0922286	3.498139
limit4	-.9255547	.6167973	-1.50	0.133	-2.134455	.2833457
RandInt	1	(constrained)				
_cons	-3.091845	1.068584	-2.89	0.004	-5.18623	-.9974596 → intercept, not threshold

```

simfac |
  sex | .1274706 .137288 0.93 0.353 -.1416089 .3965501
  age75 | -.0081964 .0166646 -0.49 0.623 -.0408585 .0244656
  visfac | -.0621605 .0869903 -0.71 0.475 -.2326583 .1083373
  zufov1 | .0095013 .1016697 0.09 0.926 -.1897675 .2087702
  zufov2 | .180878 .1035175 1.75 0.081 -.0220126 .3837687
  zufov3 | .1043617 .1018883 1.02 0.306 -.0953357 .3040592
  dscan | .2191596 .0996074 2.20 0.028 .0239328 .4143865
  _cons | -.0972922 .1053367 -0.92 0.356 -.3037482 .1091639
-----
limit4 |
  simfac | .0179388 .1506128 0.12 0.905 -.2772568 .3131344
  sex | .1759188 .2057114 0.86 0.392 -.2272682 .5791057
  age75 | -.0034281 .0230333 -0.15 0.882 -.0485726 .0417163
  visfac | .0148541 .1110292 0.13 0.894 -.2027591 .2324674
  zufov1 | .1181027 .131899 0.90 0.371 -.1404145 .3766199
  zufov2 | .1511188 .134478 1.12 0.261 -.1124532 .4146907
  zufov3 | -.1373689 .1355998 -1.01 0.311 -.4031396 .1284018
  dscan | .2287724 .1315016 1.74 0.082 -.0289659 .4865108
  _cons | .1419225 .1618459 0.88 0.381 -.1752896 .4591347
-----
var(e.RandInt) | 3.31e-34 5.77e-33 4.77e-49 2.29e-19
-----
var(e.simfac) | .4618696 .0610618 .3564394 .5984847
var(e.limit4) | .8334947 .1121721 .6402473 1.08507
-----
. nlcom _b[simfac:dscan]*_b[acc2:simfac] // Indirect effect: dscan --> sim --> acc
      _nl_1:  _b[simfac:dscan]*_b[acc2:simfac]
-----
      |      Coef.      Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
      _nl_1 |      .3934318      .2507519      1.57    0.117      -.0980329      .8848964
-----
. estat eform speed2 acc2 // Get odds ratios for binary outcomes
-----
      |      exp(b)      Robust Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
speed2 |
  sex | .2107546 .16607 -1.98 0.048 .0449831 .987427
  age75 | .9468487 .0497357 -1.04 0.298 .8542186 1.049524
  visfac | .8605239 .2982735 -0.43 0.665 .4362353 1.697482
  zufov1 | .5852976 .2808064 -1.12 0.264 .2285591 1.498839
  zufov2 | .6961832 .4075696 -0.62 0.536 .2210044 2.193038
  zufov3 | .3251612 .1237212 -2.95 0.003 .1542494 .685447
  dscan | 2.922067 1.604665 1.95 0.051 .9959773 8.572964
  simfac | .7628294 .5215739 -0.40 0.692 .199724 2.913564
  limit4 | .7066063 .3399649 -0.72 0.470 .2751975 1.814306
  RandInt | 2.718282 (constrained)
  _cons | .2365974 .1164226 -2.93 0.003 .0901901 .6206703
-----
acc2 |
  sex | .4695962 .3207389 -1.11 0.268 .1231248 1.791034
  age75 | .889144 .0756892 -1.38 0.168 .7525109 1.050586
  visfac | 1.511502 .4434062 1.41 0.159 .8505595 2.686041
  zufov1 | .1882128 .2742297 -1.15 0.252 .0108251 3.2724
  zufov2 | .4831182 .1669952 -2.10 0.035 .2453727 .9512193
  zufov3 | .9212516 .4672382 -0.16 0.872 .3409321 2.489365
  dscan | .5977119 .2272586 -1.35 0.176 .2836959 1.259305
  simfac | 6.020581 5.231106 2.07 0.039 1.096616 33.05388
  limit4 | .3963115 .2444439 -1.50 0.133 .118309 1.327564
  RandInt | 2.718282 (constrained)
  _cons | .0454181 .048533 -2.89 0.004 .0055931 .3688152
-----

```

For a sample results section, please see the original manuscript, with the following addition:

A covariance between the two binary outcomes was created using a random intercept latent factor, in which the factor loadings to each binary outcome were fixed to 1 for identification, the latent factor mean was fixed to 0 for identification, and the latent factor variance was estimated.