

Baby Books Baseline Group Assessment

Ben Saville and JoAnn Alvarez

March 5, 2009

To assess whether the three study groups are similar at baseline, we examine them with respect to baseline measures of several key variables listed in the Aims document. We used several hypothesis tests that we will outline here. The Pearson's χ^2 test of independence is used to test the null hypothesis that there is no association between two categorical variables. In cases where the contingency tables had low cell counts, we sometimes use Fisher's exact test, which tests the same null hypothesis as the χ^2 test of independence, but uses the exact distribution to calculate p -values. For variables that are ordinal, we used the proportional odds likelihood ratio test, which has more power to detect trends than the χ^2 test because it makes use of the ordering in the ordinal variable. The Kruskal-Wallis test tests the null hypothesis that, for a continuous variable, the central tendency is the same in different groups. It is the nonparametric counterpart to the familiar one-way analysis of variance (ANOVA), meaning that it does not require that the continuous variable be normally distributed.

We did not perform tests on awbq12, the number of jobs the mother held, because there was so little variation, that is, almost everyone had job, that the three treatment groups are automatically similar with respect to this variable. There was an exceptionally low response for awbq16a, awbq16b, and awbq16c, that was in some cases due to skip patterns, that any test results would be unreliable. For the public assistance variables, we did not test the variables with very low variability. Information about these variables can be found in Tables 1 and 2.

The Pearson's χ^2 test statistic for the variable awbq06, which is current marital status, is undefined because some of the cell counts are zero. The p -value for the Fisher's exact test is 0.26. The variable called "marital" is a collapsed version of awbq06 that appears after awbq06 in Table 4.

The table of frequencies for race and study group had small expected counts, so that the χ^2 test of independence may not be appropriate (although this is debated in the literature). For this reason, we present the Fisher's exact test, which had a p -value of 0.425.

All of the tests we conducted did not reject the null hypothesis at the 0.05 level, with the exception of income, whose p -value was 0.027. This low p -value could only be due to the large number of tests. (For a type-I error rate of 0.05, one would expect about one in every twenty tests to reject the null hypothesis when it is really true.) The income differences among the three groups was mainly due to the No Book group having fewer people in the lowest income category and more people in the highest income category than the VU Book group. Since this difference does not favor the experimental book group, one can be confident in attributing potential outcome differences to the experimental treatment. Overall, the three groups are similar with respect to the variables we tested. One purpose of randomization is to achieve group balance for all confounding variables, even those that are unknown to influence the outcomes.

Table 1: Employment and Income Variables

4 Variables		198 Observations	
awbq12 : How many paying jobs currently			
n	missing	unique	
93	105	2	
1 1 job (87, 94%), 2 2 jobs (6, 6%)			
awbq16a : Do anything else 4 pay (babysittg etc.)?			
n	missing	unique	
26	172	2	
0 No (21, 81%), 1 Yes (5, 19%)			
awbq16b : Hours of alt. work (16a) in past 7 days			
n	missing	unique	Mean
5	193	4	13.8
0 (2, 40%), 5.5 (1, 20%), 27.5 (1, 20%), 36 (1, 20%)			
awbq16c : Hrly rate: alt. work (16a)- past 7 days			
n	missing	unique	Mean
5	193	4	3.672
0 (2, 40%), 2.35999989509583 (1, 20%), 6 (1, 20%), 10 (1, 20%)			

Table 2: Public Assistance Variables

4 Variables		198 Observations	
awbq18c : Do you receive SSI?			
n	missing	unique	
197	1	2	
0 No (195, 99%), 1 Yes (2, 1%)			
awbq18d : Do you receive Veteran benefits?			
n	missing	unique	
197	1	2	
0 No (195, 99%), 1 Yes (2, 1%)			
awbq18f : Do you receive Disability Income (SDI)?			
n	missing	unique	
197	1	2	
0 No (196, 99%), 1 Yes (1, 1%)			
awbq18gs : Do you receive Other; response specified			
n	missing	unique	
29	169	5	
AmeriChoice (4, 14%), TennCare (21, 72%), TennCare-AmeriChoice (1, 3%) unemployment (2, 7%), unemployment, TennCare (1, 3%)			

Table 3: Descriptive Statistics by Study group (experimental condition)

	N	1 VUBook	2 CommercialBook	3 NoBookGiven	Test Statistic
		<i>N</i> = 62	<i>N</i> = 66	<i>N</i> = 70	
college : At least some college	198	50% (31)	42% (28)	41% (29)	$\chi_2^2 = 1.14, P = 0.565^1$
No college		50% (31)	58% (38)	59% (41)	
awbq04ord : 1 Some high school	198	24% (15)	33% (22)	21% (15)	$\chi_2^2 = 1.99, P = 0.369^2$
2 Completed high school or GED		26% (16)	24% (16)	37% (26)	
3 Some college, but no degree		37% (23)	32% (21)	19% (13)	
4 Associate degree		6% (4)	8% (5)	7% (5)	
5 Bachelor's degree		5% (3)	2% (1)	4% (3)	
6 Some graduate school		0% (0)	0% (0)	1% (1)	
7 Graduate degree		2% (1)	2% (1)	10% (7)	
Mom's age at this ivw	198	19.500 21.700 24.750	19.325 20.550 23.475	20.100 22.100 24.775	$F_{2,195} = 1.9, P = 0.152^3$
race : Asian	198	0% (0)	2% (1)	1% (1)	$\chi_{10}^2 = 8.86, P = 0.545^1$
black		66% (41)	64% (42)	51% (36)	
Latino		2% (1)	5% (3)	3% (2)	
multiracial		6% (4)	2% (1)	7% (5)	
other		2% (1)	2% (1)	0% (0)	
white		24% (15)	27% (18)	37% (26)	
White : non-white	198	74% (46)	68% (45)	60% (42)	$\chi_2^2 = 3.05, P = 0.218^1$
white		26% (16)	32% (21)	40% (28)	
income : 1 Less than \$8,000	120	29% (13)	29% (10)	15% (6)	$\chi_2^2 = 7.2, P = 0.027^2$
2 \$8,000 - \$12,000		22% (10)	15% (5)	20% (8)	
3 \$12,001 - \$16,000		11% (5)	18% (6)	2% (1)	
4 \$16,001 - \$21,000		4% (2)	15% (5)	12% (5)	
5 \$21,001 - \$26,000		13% (6)	6% (2)	12% (5)	
6 \$26,001 - \$30,000		4% (2)	3% (1)	5% (2)	
7 \$30,001 - \$40,000		4% (2)	9% (3)	10% (4)	
8 \$40,001 - \$50,000		7% (3)	0% (0)	7% (3)	
over \$50,000		4% (2)	6% (2)	17% (7)	

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.

N is the number of non-missing values.

Numbers after percents are frequencies.

Tests used: ¹Pearson test; ²Proportional odds likelihood ratio test;

³Kruskal-Wallis test

Table 4: Descriptive Statistics by Study group (experimental condition)

	N	1 VUBook N = 62	2 CommercialBook N = 66	3 NoBookGiven N = 70	Test Statistic
marital : Single	197	79% (49)	86% (57)	70% (48)	$\chi_4^2 = 6.38, P = 0.173^1$
Married		18% (11)	9% (6)	23% (16)	
Other		3% (2)	5% (3)	7% (5)	
employment : Not working	197	54% (33)	47% (31)	46% (32)	$\chi_2^2 = 1.47, P = 0.48^2$
Part time		18% (11)	23% (15)	16% (11)	
Full time		28% (17)	30% (20)	39% (27)	
Main job: wks worked for pay in last mo	170	0 2 4	0 3 4	0 2 4	$F_{2,167} = 0.03, P = 0.972^3$
Main job: mos worked for pay in last 12	187	5.0000 7.0000 11.0625	4.3750 8.0000 12.0000	4.5000 10.0000 12.0000	$F_{2,184} = 0.98, P = 0.379^3$
Hours worked in past 7 days	93	15.0 30.0 40.0	14.5 20.0 40.0	8.0 30.0 40.0	$F_{2,90} = 0.08, P = 0.925^3$
Hourly rate of pay in past 7 days	88	7.100 7.500 10.000	7.525 9.000 10.185	7.550 9.000 12.750	$F_{2,85} = 0.62, P = 0.543^3$
WP total score; range [0,55]	194	7.00 14.00 20.25	9.00 11.50 18.50	12.00 15.50 18.00	$F_{2,191} = 2.01, P = 0.137^3$

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.

N is the number of non-missing values.

Numbers after percents are frequencies.

Tests used: ¹Pearson test; ²Proportional odds likelihood ratio test;

³Kruskal-Wallis test

Table 5: Descriptive Statistics by Study group (experimental condition)

	N	1 VUBook <i>N</i> = 62	2 CommercialBook <i>N</i> = 66	3 NoBookGiven <i>N</i> = 70	Test Statistic
Do you receive WIC? : 0 No	197	32% (20)	41% (27)	48% (33)	$\chi_2^2 = 3.29, P = 0.193$
1 Yes		68% (42)	59% (39)	52% (36)	
Do you receive Food stamps? : 0 No	197	52% (32)	56% (37)	67% (46)	$\chi_2^2 = 3.26, P = 0.195$
1 Yes		48% (30)	44% (29)	33% (23)	
Do you receive Family First? : 0 No	197	89% (55)	89% (59)	91% (63)	$\chi_2^2 = 0.26, P = 0.877$
1 Yes		11% (7)	11% (7)	9% (6)	
Do you receive Other? : 0 No	197	87% (54)	86% (57)	83% (57)	$\chi_2^2 = 0.62, P = 0.735$
1 Yes		13% (8)	14% (9)	17% (12)	

N is the number of non-missing values.

Numbers after percents are frequencies.

Test used: Pearson test

Table 6: Descriptive Statistics by Study group (experimental condition)

	N	1 VUBook <i>N</i> = 62			2 CommercialBook <i>N</i> = 66			3 NoBookGiven <i>N</i> = 70			Test Statistic
O1B total # correct	198	10.25	12.00	14.00	10.00	12.00	14.75	11.00	13.00	15.00	$F_{2,195} = 1.51, P = 0.223^1$
O1B % correct, based on # answered	198	33.8675	40.0450	45.1600	32.2600	38.7100	47.5825	35.7775	41.9400	48.3900	$F_{2,195} = 1.45, P = 0.238^1$
RSP summary score	193	14	15	16	13	15	16	13	15	16	$F_{2,190} = 0.81, P = 0.446^1$
PRP total score	198	80.00	87.00	92.00	83.25	88.00	92.00	84.00	90.00	92.00	$F_{2,195} = 1.27, P = 0.283^1$
PRP Teaching Efficacy Subscore	198	28.00	29.00	31.00	28.25	30.00	31.00	29.00	30.00	31.00	$F_{2,195} = 1.84, P = 0.162^1$
PRP positive Affect Subscore	196	16.75	18.00	19.00	17.00	18.00	19.00	17.00	18.00	19.00	$F_{2,193} = 0.23, P = 0.794^1$
PRP Verbal Participation Subscore	196	19.75	22.00	24.00	20.00	22.50	24.00	21.00	22.00	24.00	$F_{2,193} = 0.25, P = 0.783^1$
PRP Reading Instruction Subscore	198	6	6	8	6	6	8	6	7	8	$F_{2,195} = 0.77, P = 0.463^1$
PRP Resource Subscore	197	13	15	16	14	16	16	15	16	16	$F_{2,194} = 1.96, P = 0.143^1$
PES Uplift Frequency	198	9	10	10	9	10	10	8	10	10	$F_{2,195} = 0.15, P = 0.863^1$
PES Hassle Frequency	198	6.00	7.00	8.75	6.00	8.00	9.00	5.25	7.00	9.00	$F_{2,195} = 0.71, P = 0.495^1$
CES summary score: [0-30]	198	6.25	9.00	12.00	7.00	11.00	12.00	6.00	9.00	12.00	$F_{2,195} = 0.85, P = 0.429^1$
Do you plan on breastfeeding your baby? : 1 Yes	198	74%	(46)		64%	(42)		76%	(53)		$\chi_4^2 = 5.49, P = 0.241^2$
2 No		15%	(9)		29%	(19)		19%	(13)		
3 Haven't decided yet		11%	(7)		8%	(5)		6%	(4)		
Total # correct, AWBI05-10	198	5	5	6	4	5	6	5	5	6	$F_{2,195} = 0.5, P = 0.609^1$

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.

N is the number of non-missing values.

Numbers after percents are frequencies.

Tests used: ¹Kruskal-Wallis test; ²Pearson test