Assessing the Effectiveness of Peer Assisted Study Schemes

DEE 2015, September 2015

Ralf Becker and Maggy Fostier

September 2015

Outline

- **What is PASS**
- Potential Benefits
- Evaluation Issues
- Empirical Results
  - Selection by Unobservables
  - Selection by Observables
- Summary and Outlook

What is PASS

- Formally introduced by the University of Missouri - Kansas City in 1973
- Peer Assisted Study Schemes (aka. Supplementary Instruction)
- Peer support for a course unit or Programme Year
- Higher year students (PASS Leaders, typically 2) meet with lower year students (participants)
- Leaders "*facilitate groups of lower year students to help them deepen their understanding and develop study and learning strategies*"
- No teaching!!!
- weekly sessions, **voluntary**

What is PASS

Peer Assisted Study Schemes are introduced to support:

1. The first year of a degree programme
2. A particular (often difficult) course unit in Year 1 or 2


A PASS Session


Who is the Leader?

Outline

- What is PASS
- **Potential Benefits**
- Evaluation Issues
- Empirical Results
  - Selection by Unobservables
  - Selection by Observables
- Summary and Outlook

Potential Benefits - "certified"

Peer Assisted Study Schemes are hoped to achieve the following (The International Centre for Supplemental Instruction website[4]):

1. Increase retention in course unit
2. Improve student grades in course unit
3. Increase the graduation rates at this institution

These claims received the U.S. Department of Education stamp of approval!
The research had the following basic features:

- Basic comparisons of PASS participants and non-participants.
- Aggregation across schemes and universities
- Control for Self-Selection largely absent

Benefits

The following were reviewed and discussed in Dawson *et al.*, 2014 [2]:

Claim	Evidence
1) improved exam grades	sympathetic yes
2) improved achievement for minority students	no evidence
3) Effectiveness beyond the course unit	no evidence
4) Improved academic skills	perhaps for *information processing* and *motivation levels*
5) Improved satisfaction	no clear evidence, but possibly
6) Enhanced social relationships	no clear evidence, but possibly
7) Improved employability	no clear evidence, but possibly

Benefits for Leaders, School and University[1]

- PASS Leaders
  - Personal development opportunity
  - Skills development - leadership, communication, teamwork etc.
  - Opportunity to reflect, review and re-evaluate
  - Increased academic performance
  - Recognition and Reward
- Discipline Level & University
  - Providing staff with regular & ongoing feedback
  - Highlighted as good practice by QAA and professional bodies
  - Improves student study skills
  - Fostering a spirit of community
  - Widening access to an increasingly diverse student body
  - Reducing student drop out rates
  - Improving the student experience & academic performance

These benefits are mostly based on anecdotal evidence (at most).

[1] from University of Manchester website

Outline

- What is PASS
- Potential Benefits
- **Evaluation Issues**
- Empirical Results
  - Selection by Unobservables
  - Selection by Observables
- Summary and Outlook

Evaluation Issues

Here we focus on grade impact only!

- Attendance is voluntary
- Self selection issues:
  - If better students tend to attend, and they would have done well even without PASS, then the effect of PASS is likely overestimated
  - If worse students tend to attend, and they would have done worse without PASS, then the effect of PASS is likely underestimated

What do we mean by better or worse?

Evaluation Issues

What do we mean by better or worse?

- Pre-requisite knowledge

Evaluation Issues

What do we mean by better or worse?

- Pre-requisite knowledge
- General academic ability

Evaluation Issues

What do we mean by better or worse?

- Pre-requisite knowledge
- General academic ability
- Motivation

Evaluation Issues

What do we mean by better or worse?

- Pre-requisite knowledge
- General academic ability
- Motivation
- Time commitment

Evaluation Issues

What do we mean by better or worse?

- Pre-requisite knowledge
- General academic ability
- Motivation
- Time commitment
- Organisation

Evaluation Issues

What do we mean by better or worse?

- Pre-requisite knowledge
- General academic ability
- Motivation
- Time commitment
- Organisation

If PASS attendance is related to any of these factors, we need to take care of them.
Different methodologies depending on whether we have variables/proxies for these.

The Causal Model

How to think about the problem.

- Consider the ith student
- Assume that you either get the treatment/PASS ($p_i = 1$) or not ($p_i = 0$)
-
- The potential outcomes are

$$y_i = \begin{cases} y_{1i} & , \quad \text{if treatment,} \\ y_{0i} & , \quad \text{if no treatment.} \end{cases} \tag{1}$$

The Causal Model

How to think about the problem.

- Consider the ith student
- Assume that you either get the treatment/PASS ($p_i = 1$) or not ($p_i = 0$)
-
- The potential outcomes are

$$y_i = \begin{cases} y_{1i} & , \quad \text{if treatment,} \\ y_{0i} & , \quad \text{if no treatment.} \end{cases} \tag{1}$$

For each i we only observe one of these!

The Causal Model

How to think about the problem.

- Consider the ith student
- Assume that you either get the treatment/PASS ($p_i = 1$) or not ($p_i = 0$)
- We want to condition the outcomes on a set of covariates q_i
- The potential outcomes are

$$y_i | q_i = \begin{cases} y_{1i} | q_i, & \text{if treatment,} \\ y_{0i} | q_i, & \text{if no treatment.} \end{cases} \quad (2)$$

For each i we only observe one of these!

The treatment effect and the Selection bias

We want to get the treatment effect (integrating over the conditioning variable q_i), the difference between the two potential outcomes (Average Treatment Effect, ATE):

$$E[y_{1i} - y_{0i}] = E[y_{1i}] - E[y_{0i}] \tag{3}$$

If we estimate (naive estimator)

$$E[y_i|p_i = 1] - E[y_i|p_i = 0] \quad = \tag{4}$$

The treatment effect and the Selection bias

We want to get the treatment effect (integrating over the conditioning variable q_i), the difference between the two potential outcomes (Average Treatment Effect, ATE):

$$E[y_{1i} - y_{0i}] = E[y_{1i}] - E[y_{0i}] \tag{3}$$

If we estimate (naive estimator)

$$\begin{aligned} E[y_i|p_i = 1] - E[y_i|p_i = 0] &= E[y_{1i} - y_{0i}] \\ &+ E[y_{0i}|p_i = 1] - E[y_{0i}|p_i = 0] \end{aligned} \tag{4}$$

The treatment effect and the Selection bias

We want to get the treatment effect (integrating over the conditioning variable q_i), the difference between the two potential outcomes (Average Treatment Effect, ATE):

$$E[y_{1i} - y_{0i}] = E[y_{1i}] - E[y_{0i}] \tag{3}$$

If we estimate (naive estimator)

$$\begin{aligned} E[y_i|p_i = 1] - E[y_i|p_i = 0] \quad = \quad & E[y_{1i} - y_{0i}] \\ & + E[y_{0i}|p_i = 1] - E[y_{0i}|p_i = 0] \end{aligned} \tag{4}$$

- If better students attend PASS, then selection bias > 0
- If weaker students attend PASS, then selection bias < 0

Solution Strategies

At the core of the problem is that selection into PASS is non-random

Solution strategies in order of power:

1. Randomised Control Trial (no selection bias)
2. Instrumental Variables Estimation (makes selection bias "irrelevant")
3. Conditioning on variables that control the selection (potentially controls for selection bias due to observed variables)
   3.1 Regression with covariates
   3.2 Matching estimators
4. Panel estimates (can partially control for selection on unobservables)

Outline

- What is PASS
- Potential Benefits
- Evaluation Issues
- **Empirical Results**
  - Selection by Unobservables
  - Selection by Observables
- Summary and Outlook

Data

We look at two datasets

- A Year 1 PASS scheme from the Faculty of Life Sciences
  - About 470 first years in various degree programmes
  - Semester 1 PASS supported Geetics course unit (but with wider brief as well)
  - capacity limited such that some students had access to PASS in Semester 1 (the others in Semester 2)
  - automatic enrolment into PASS group
  - variables: PASS attendance, degree programme and exam results
- 2nd year Econometrics course unit
  - 324 students from various programmes
  - Some students take this course in their 3rd year
  - No binding capacity limit for PASS
  - voluntary PASS sign-up
  - variables: coursework and exam grades, Year 1 grade info (e.g. GPA, statistics), programme, study year, gender, ethnicity

Notation

$$y_i = \text{Exam Grade} \tag{5}$$

$$p_i = \begin{cases} p_i, & \text{enrolled in PASS} \\ pa_i, & \text{no of attended weekly sessions or,} \\ ph_i, & =1 \text{ if } pa_i > 3. \end{cases} \tag{6}$$

$$q_i = \text{Covariates} \tag{7}$$

may distinguish between $q_i^{(o)}$ and $q_i^{(u)}$, observed and unobserved covariates

Randomised Control Trial, RCT

- The capacity constraint (Life Sciences dataset) for PASS delivered an opportunity for a RCT
- But institutional constraints led to the allocation not being totally random!
  Some degree programmes got preferential treatment

Basically one can estimate (educational production function, Todd and Wolpin, 2003, [6])

$$y_i = \alpha + \beta p_i + \gamma q_i + u_i \tag{8}$$

Randomised Control Trial, RCT

- The capacity constraint (Life Sciences dataset) for PASS delivered an opportunity for a RCT
- But institutional constraints led to the allocation not being totally random!
  Some degree programmes got preferential treatment

Basically one can estimate (educational production function, Todd and Wolpin, 2003, [6])

$$y_i = \alpha + \beta p_i + \gamma q_i + u_i \tag{8}$$

In a RCT p_i will be uncorrelated with q_i and u_i

Randomised Control Trial, RCT

- The capacity constraint (Life Sciences dataset) for PASS delivered an opportunity for a RCT
- But institutional constraints led to the allocation not being totally random!
  Some degree programmes got preferential treatment

Basically one can estimate (educational production function, Todd and Wolpin, 2003, [6])

$$y_i = \alpha + \beta p_i + \gamma q_i + u_i \tag{8}$$

In a RCT p_i will be uncorrelated with q_i and u_i

	Est	Sig	Est	Sig
$\hat{\beta}$	5.00	***	1.79	
$\hat{\gamma}$	no		yes	

Randomised Control Trial, RCT

- PASS allocation is clearly related to q_i (containing degree choice) and that is a strong indication that it may also be correlated with further unobserved factors.
- RCT turns out to be a nRCT
  $\Rightarrow$ no reliable result
- Additional issue: PASS enrollment (p_i) is not indicative of engagement

Instrumental variables

From now application to Econometrics dataset
Reconsider:

$$y_i = \alpha + \beta pa_i + \gamma q_i^{(o)} + u_i \tag{9}$$

where $u_i = f(q_i^{(u)})$

- If we could find an instrument z_i that was correlated with pa_i or ph_i but uncorrelated with error term u_i and hence $q_i^{(u)}$, IV estimation could deliver consistent estimate of β

Instrumental variables

From now application to Econometrics dataset
Reconsider:

$$y_i = \alpha + \beta pa_i + \gamma q_i^{(o)} + u_i \tag{9}$$

where $u_i = f(q_i^{(u)})$

- If we could find an instrument z_i that was correlated with pa_i or ph_i but uncorrelated with error term u_i and hence $q_i^{(u)}$, IV estimation could deliver consistent estimate of β
- Gender?, not clear (Ceci, 2014,[1]), it is certainly correlated with attendance (female students are more likely to attend)

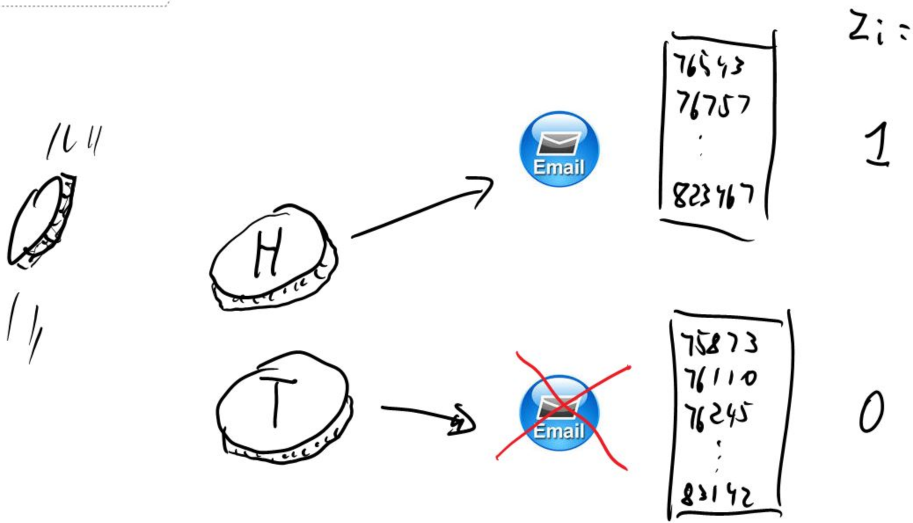
Instrumental variables

From now application to Econometrics dataset
Reconsider:

$$y_i = \alpha + \beta pa_i + \gamma q_i^{(o)} + u_i \tag{9}$$

where $u_i = f(q_i^{(u)})$

- If we could find an instrument z_i that was correlated with pa_i or ph_i but uncorrelated with error term u_i and hence $q_i^{(u)}$, IV estimation could deliver consistent estimate of β
- Gender?, not clear (Ceci, 2014,[1]), it is certainly correlated with attendance (female students are more likely to attend)
- Encouragement?


Zi:
H
T
Email
Email
76543
76757
823467
75873
76110
76245
83142
1
0

Instrumental variables

Reconsider:

$$y_i = \alpha + \beta p_i + \gamma q_i^{(o)} + u_i \tag{10}$$

where $u_i = f(q_i^{(u)})$

- If we could find an instrument z_i that was correlated with pa_i or ph_i but uncorrelated with error term u_i and hence $q_i^{(u)}$, IV estimation could deliver consistent estimate of β
- Gender?, not clear (Ceci, 2014,[1]), it is certainly correlated with attendance (female students are more likely to attend)
- Encouragement? Has no impact on attendance!

Conditioning on observables

Key issue is the self-selection!

$$y_i = \alpha + \beta pa_i + \gamma q_i^{(o)} + u_i \quad (11)$$

where $u_i = f(q_i^{(u)})$

Need to assume that $E[pa_i q_i^{(u)}] = 0$. All systematic selection is on the observables, $q_i^{(o)}$.

Advantage of 2nd year PASS is the conditioning info

- Year 1 GPA
- Year 1 Stats prerequisite
- other personal characteristics (OS, ethnicity, programme dummies, Year 3 student)

Results for conditioning on $q_i^{(o)}$ and IV, pa_i

Mean(y_i) = 56.51; sd(y_i) = 15.86

Method	OLS				IV	
z_i					gen, enc	
pa_i	1.590	***	0.793	***	0.865	
ph_i						
stats			0.118	*	0.118	*
Y1gpa			0.977	***	0.972	***
Y3			2.921	*	2.895	*
O/S			-2.900	**	-2.883	*
P(BSc)			4.067	**	4.125	**
P(IBFE)			3.777		3.806	
P(Other)			3.900		4.077	
P(PPE)			-0.392		-0.336	
R2	0.082		0.495		0.495	
Stage1(F)					5.816	
(p-value)					(0.003)	

Results for conditioning on $q_i^{(o)}$ and IV, ph_i

Mean(y_i) = 56.51; sd(y_i) = 15.86

Method	OLS				IV	
z_i					gen, enc	
pa_i						
ph_i	9.764	***	4.476	**	5.380	
stats			0.117	*	0.116	*
Y1gpa			0.980	***	0.969	***
Y3			3.012	*	2.971	*
O/S			-2.840	*	-2.790	*
P(BSc)			3.804	**	3.881	**
P(IBFE)			3.556		3.575	
P(Other)			3.362		3.649	
P(PPE)			-0.906		-0.887	
R2	0.086		0.493		0.493	
Stage1(F)					4.334	
(p-value)					(0.014)	

Summary so far

- RCT, no results

Summary so far

- RCT, no results
- IV, could control for selection on $q_i^{(u)}$, but at best very weak instruments

Summary so far

- RCT, no results
- IV, could control for selection on $q_i^{(u)}$, but at best very weak instruments
- OLS with conditioning on $q_i^{(o)}$, assuming no selection on $q_i^{(u)}$, better claim for this to be true here as we have richer $q_i^{(o)}$ due to 2nd year scheme.

Summary so far

- RCT, no results
- IV, could control for selection on $q_i^{(u)}$, but at best very weak instruments
- OLS with conditioning on $q_i^{(o)}$, assuming no selection on $q_i^{(u)}$, better claim for this to be true here as we have richer $q_i^{(o)}$ due to 2nd year scheme.
- So far, effect size in the order of $1/3$ of a standard deviation

Matching Estimator

- Matching estimators achieve, conceptually the same as OLS with conditioning:
  Controls for selection on observables
- But is slightly more flexible in that we do not have to assume a linear education production function
- Allows easy calculation of different treatment effect for: all (ATE), treated (ATT), non-treated (ATN)

Matching Estimator

- Matching estimators achieve, conceptually the same as OLS with conditioning:
  Controls for selection on observables
- But is slightly more flexible in that we do not have to assume a linear education production function
- Allows easy calculation of different treatment effect for: all (ATE), treated (ATT), non-treated (ATN)

Consider

$$\delta_{ATT} = E[y_{1i}|q_i, ph_i = 1] - E[y_{0i}|q_i, ph_i = 1] \qquad (12)$$

Matching Estimator

- Matching estimators achieve, conceptually the same as OLS with conditioning:
  Controls for selection on observables
- But is slightly more flexible in that we do not have to assume a linear education production function
- Allows easy calculation of different treatment effect for: all (ATE), treated (ATT), non-treated (ATN)

Consider

$$\delta_{ATT} = E[y_{1i}|q_i, ph_i = 1] - E[y_{0i}|q_i, ph_i = 1] \tag{12}$$

$E[y_{0i}|q_i, ph_i = 1]$ is unobserved/the counterfactual.

Matching Estimators

For ATT the counterfactuals are found from the non-treated by finding (matching) those observations that are most similar in terms of:

- propensity to be a high PASS attender ($ph_i = 1$), *PropMatch*
- similarity of covariates q_i, *CovMatch*

	ATT		ATN		ATE	
	$\hat{\delta}$	sig	$\hat{\delta}$	sig	$\hat{\delta}$	sig
PropMatch	5.258	***	6.327	***	5.956	***
CovMatch	3.486	**	5.010	***	4.481	***

Using Panel features (Econometrics dataset)

Add a time dimension $(t = 1, 2)$ to the educational production function

$$y_{it} = \alpha + \beta pa_{it} + \gamma q_{it}^{(o)} + \delta q_{it}^{(u)} + v_{it} \tag{13}$$

Using Panel features (Econometrics dataset)

Add a time dimension $(t = 1, 2)$ to the educational production function

$$y_{it} = \alpha + \beta pa_{it} + \gamma q_{it}^{(o)} + \delta q_{it}^{(u)} + v_{it} \tag{13}$$

Then subtract the $t = 2$ equation from the $t = 1$ equation, and assuming constant coefficients

Using Panel features (Econometrics dataset)

Add a time dimension $(t = 1, 2)$ to the educational production function

$$y_{it} = \alpha + \beta pa_{it} + \gamma q_{it}^{(o)} + \delta q_{it}^{(u)} + v_{it} \tag{13}$$

Then subtract the $t = 2$ equation from the $t = 1$ equation, and assuming constant coefficients

$$\begin{aligned} y_{i2} - y_{i1} &= \beta(pa_{i2} - pa_{i1}) \\ &\quad + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + \delta(q_{i2}^{(u)} - q_{i1}^{(u)}) \\ &\quad + v_{i2} - v_{i2} \end{aligned} \tag{14}$$

Using Panel features (Econometrics dataset)

Add a time dimension $(t = 1, 2)$ to the educational production function

$$y_{it} = \alpha + \beta pa_{it} + \gamma q_{it}^{(o)} + \delta q_{it}^{(u)} + v_{it} \tag{13}$$

Then subtract the $t = 2$ equation from the $t = 1$ equation, and assuming constant coefficients

$$\begin{aligned} y_{i2} - y_{i1} &= \beta(pa_{i2} - pa_{i1}) \\ &\quad + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + \delta(q_{i2}^{(u)} - q_{i1}^{(u)}) \\ &\quad + v_{i2} - v_{i2} \end{aligned} \tag{14}$$

If δ is constant and $q_{i2}^{(u)} = q_{i1}^{(u)}$, then

Using Panel features (Econometrics dataset)

Add a time dimension $(t = 1, 2)$ to the educational production function

$$y_{it} = \alpha + \beta pa_{it} + \gamma q_{it}^{(o)} + \delta q_{it}^{(u)} + v_{it} \tag{13}$$

Then subtract the $t = 2$ equation from the $t = 1$ equation, and assuming constant coefficients

$$\begin{aligned} y_{i2} - y_{i1} &= \beta(pa_{i2} - pa_{i1}) \\ &\quad + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + \delta(q_{i2}^{(u)} - q_{i1}^{(u)}) \\ &\quad + v_{i2} - v_{i2} \end{aligned} \tag{14}$$

If δ is constant and $q_{i2}^{(u)} = q_{i1}^{(u)}$, then

$$y_{i2} - y_{i1} = \beta(pa_{i2} - pa_{i1}) + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + v_{i2} - v_{i2} \tag{15}$$

Using Panel features (Econometrics dataset)

$$y_{i2} - y_{i1} = \beta(pa_{i2} - pa_{i1}) + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + v_{i2} - v_{i2} \qquad (16)$$

we are basically looking at grade change!

- $pa_{i=1} = 0$ as no PASS scheme in Year 1

Using Panel features (Econometrics dataset)

$$y_{i2} - y_{i1} = \beta(pa_{i2} - pa_{i1}) + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + v_{i2} - v_{i2} \qquad (16)$$

we are basically looking at grade change!

- $pa_{i=1} = 0$ as no PASS scheme in Year 1
- y_{i2} is the Econometrics grade

Using Panel features (Econometrics dataset)

$$y_{i2} - y_{i1} = \beta(pa_{i2} - pa_{i1}) + \gamma(q_{i2}^{(o)} - q_{i1}^{(o)}) + v_{i2} - v_{i2} \qquad (16)$$

we are basically looking at grade change!

- $pa_{i=1} = 0$ as no PASS scheme in Year 1
- y_{i2} is the Econometrics grade
- y_{i1}?
  - Y1 GPA, or
  - Statistics prerequisite course unit

Using Panel features (Econometrics dataset)

$$y_{i2} - y_{i1} = \beta(pa_{i2} - pa_{i1}) + \gamma(q^{(o)}_{i2} - q^{(o)}_{i1}) + v_{i2} - v_{i2} \qquad (16)$$

we are basically looking at grade change!

- $pa_{i=1} = 0$ as no PASS scheme in Year 1
- y_{i2} is the Econometrics grade
- y_{i1}?
  - Y1 GPA, or
  - Statistics prerequisite course unit
- if $q^{(o)}_{i1} = q^{(o)}_{i2}$ then this term disappears

Using Panel features (Econometrics dataset)

Estimation results

Treatment	Y1 grade	$\hat{\beta}$	sig
pa_i	stats	0.631	**
pa_i	Y1 GPA	0.847	**
ph_i	stats	3.410	***
ph_i	Y1 GPA	4.791	***

This model also included constant and qf_i.

Using Panel features (Econometrics dataset)

Estimation results

Treatment	Y1 grade	$\hat{\beta}$	sig
pa_i	stats	0.631	**
pa_i	Y1 GPA	0.847	**
ph_i	stats	3.410	***
ph_i	Y1 GPA	4.791	***

This model also included constant and qf_i.
Recall: sd(y_{i2})=16
Estimates are again in order of 1/3 of a s.d.

Summary and Conclusion

- No watertight evidence on effectiveness of PASS on course unit grade
- Evidence here adds to the existing evidence
- But there exists a potential strategy (better encouragement).
- (Potential) benefits of PASS are multi-facetted.
- Difficult to see research design implemented that will robustly establish these (too many intangibles)

Summary and Conclusion

- No watertight evidence on effectiveness of PASS on course unit grade
- Evidence here adds to the existing evidence
- But there exists a potential strategy (better encouragement).
- (Potential) benefits of PASS are multi-facetted.
- Difficult to see research design implemented that will robustly establish these (too many intangibles)
- I will continue to run the scheme

[Ceci *et al.*, 2014] Ceci, S.J., Ginther, D.K., Khan, S. and Williams, W.M. .
Women in Academic Sciences: A changing landscape,
Psychological Science in the Public Interest, 2014, 15, 75-141.

[Dawson *et al.*, 2014] Dawson, P., van der Meer, J., Skalicky, J. and Cowley, K. .
On the Effectiveness of Supplemental Instruction: A Systematic Review of Supplemental Instruction and Peer-Assisted Study Sessions Literature Between 2001 and 2010,
Review of Educational Research, 2014, 44, 1-16.

[Martin and Arendale, 1992] Supplemental Instruction: Improving First-Year Student Success in High-Risk Courses. The Freshman Year Experience
Monograph Series Number 7,
http://eric.ed.gov/?id=ED354839

[ICSI, 2015] The International Centre for Supplemental Instruction