Productive Failure

Manu Kapur

Associate Professor of Curriculum, Teaching & Learning
Researcher, Learning Sciences Lab
National Institute of Education, Singapore
manu.kapur@nie.edu.sg

CNIE, 2012

The Role of Failure in Initial Learning

Before learning a concept vs. after learning a concept

Heavily-guided or direct instruction is the most prevailing and intuitive way of designing the initial learning

Examining the Case for Direct Instruction

"Controlled experiments almost uniformly indicate that when dealing with novel information, learners should be explicitly shown what to do and how to do it" (p. 79; Kirschner et al., 2006)…

Cognitive Load: un-guided or minimally-guided instruction increases working memory load that interferes with schema formation (Sweller, 1988)

Substantive empirical support (Sweller & Copper, 1985; Carroll, 1994; Paas, 1992; Klahr & Nigam, 2004)

compared some version(s) of a worked example or strong instructional guidance condition with a pure problem-solving or discovery condition.

Conclusion: there is little efficacy in letting learners solve problems that target novel concepts…

Examining the Case for Direct Instruction

It is not surprising that students in the pure problem solving condition did not learn as much as those in the strongly-guided condition…

But, this do not necessarily imply that there is little efficacy in letting learners solve novel problems on their own

To determine if there is such an efficacy, a stricter comparison is needed:

Direct instruction vs. students first solve novel problems on their own followed by some form of instructional structure or direct instruction

What is Productive Failure?

Understand what students know about a novel concept that they have not been taught yet

Afford opportunities to activate and differentiate prior and intuitive knowledge….to generate, explore, critique, and refine representations and methods for solving complex problems

Invariably, such a process leads to failure to develop the canonical solution(s)

But, this may precisely be the locus of deep learning…provided some form of instructional structure follows subsequently

The case for productive failure…

Past research:

- *A Theory of Constructive Failure (Clifford, 1984)*
- *Desirable difficulties* (Schmidt & Bjork, 1992)
- *Impasse-driven Learning* (VanLehn et al., 2003)
- *Assistance Dilemma* (Koedinger et al., 2008)
- *Preparation for Future Learning* (Schwartz & Bransford, 1999)
- *Inventing to Prepare for Learning* (Schwartz & Martin, 2004)
- *Productive Failure* (Kapur, 2008; Kapur & Kinzer, 2009)

Study 1: Productive Failure

(Kapur, 2008; Kapur & Kinzer, 2009)

Target Concept: Newtonian Kinematics

N = 309, 11th-grade physics students in India

Well-structured Problems (WSP)	Well-structured Problems (WSP)	Ill-structured Problems (ISP)
Ill-structured Problems (ISP)	Well-structured Problems (WSP)	Ill-structured Problems (ISP)

Compared with WSP groups, ISP groups:

1. Generated multiple representations and solution methods (RSMs)
2. Engaged in complex interaction patterns of explanation, critique, elaboration
3. Low convergence in their discussions
4. Poor group performance
5. **BUT**, better individual performance on both well- and ill-structured problems

Study 2
(Kapur, 2012)

Target Concept:
Standard Deviation

Grades: 8/9

Who's the most consistent?

Year	Mike Arwen	Dave Backhand	Ivan Right
1988	14	13	13
1989	9	9	18
1990	14	16	15
1991	10	14	10
1992	15	10	16
1993	11	11	10
1994	15	13	17
1995	11	14	10
1996	16	15	12
1997	12	19	14
1998	16	14	19
1999	12	12	14
2000	17	15	18
2001	13	14	9
2002	17	17	10

Comparing regularity

Mike Arwen : Mean = $\frac{280}{20}$

= 14 goals / year

Mode = 14

Dave Backhand : Mean = $\frac{280}{20}$

= 14 goals / year

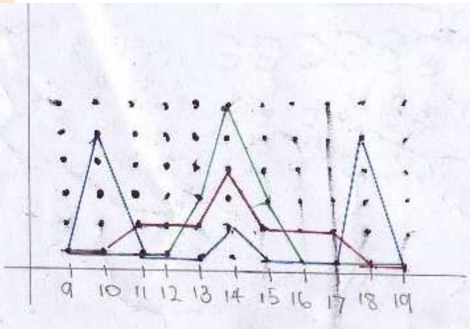
Mode = 14

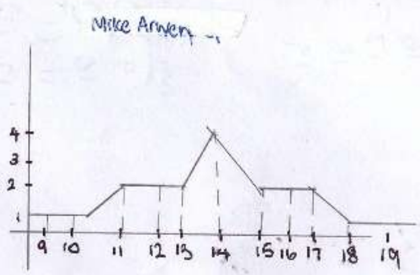
Ivan Right : Mean = $\frac{280}{20}$

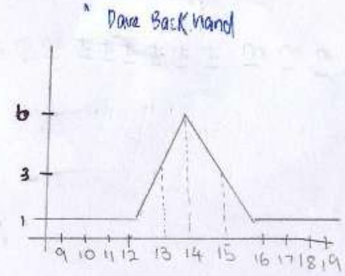
= 14 goals / year

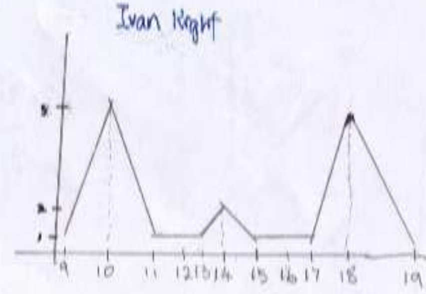
Mode = 18 and 10

	9	10	11	12	13	14	15	16	17	18	19
Mike Arwen	1	1	2	2	2	4	2	2	2	1	1
Dave Backhand	1	1	1	1	3	6	3	1	1	1	1
Ivan Right	1	5	1	1	1	2	1	1	1	5	1

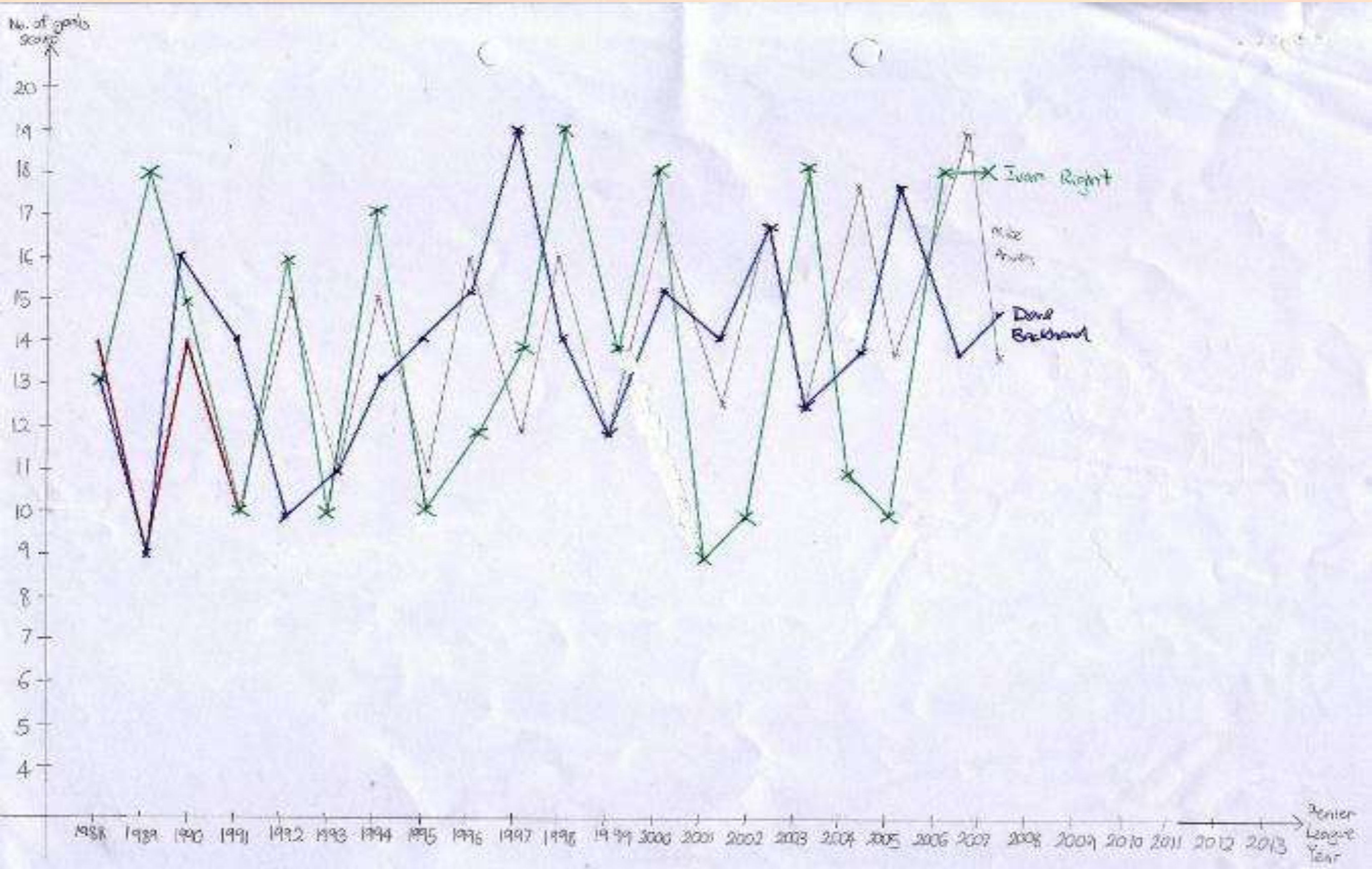








9 10 11 11 12 12 13 13 14 14 14 14 15 15 16 16 17 17 18 19



No. of goals scored
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
Ivan Right
Dave Backhand
Premier League Year
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013

From Question paper: Average = $\frac{280}{20}$

Mike has 8 years < average

4 years = average

8 years > average

Dave has 7 years < average

6 years = average

7 years > average

Ivan has 9 years < average

2 years = average

9 years > average

Frequency of years above, below, and at average

Consistency = years at the mean / years away from the mean

Sum of year-on-year deviation

Mike:
9 − 14 = −5
14 − 9 = 5
10 − 14 = −4
15 − 10 = 5
−4
4
−4
5
−4
4
−4
5
−4
4
−4
5
−4
5
−4
———
0 ✓ mike

Dave:
−4
7
−2
−4
1
2
1
1
4
−5
−2
3
−1
3
−4
1
4
−4
1
———
−2

Ivan:
5
−3
−5
1
−6
7
−7
2
2
5
−5
4
−9
1
8
−7
−1
8
0
———
−5

Range amount for:

Mike Arwen: 9 − 19 = 10
Dave Buckland: 9 − 19 = 10
Ivan Right: 9 − 19 = 10
X

Sum of deviations about the mean

Year	Avg	M.A	D.B	I.R	X		
1988	14	14	13	13	0,	−1,	−1
1989	14	9	4	18	−5,	−5,	4
1990	14	14	16	15	0	+2,	+1
1991	14	10	14	10	−4	0	−4
1992	16	15	10	16	+1	−4	+2
1993	14	11	11	10	−3	−3	−4
1994	14	15	13	17	+1	−1	+3
1995	14	11	14	10	−3	0	−4
1996	14	16	15	12	+2	+1	−2
1997	14	12	19	14	−2	+5	0
1998	14	16	14	19	+2	0	+5
1999	14	12	12	14	−2	−2	0
2000	14	17	15	18	+3	+1	+4
2001	14	13	14	9	−1	0	−5
2002	14	17	17	10	+3	+3	−4
2003	14	13	13	18	−1	−1	+4
2004	14	18	14	11	+4	0	−3
2005	14	14	18	10	0	+4	−4
2006	14	19	14	18	+5	0	+4
2007	14	14	15	18	0	+1	+4
				Total:	[redacted]	[redacted]	[redacted]

Average of year-on-year absolute deviation

MIKE = $\frac{5+5+4+5+4+4+4+5+4+4+4+5+4+4+4+5+4+5+4}{20-1}$

= 84/19 = 4.26

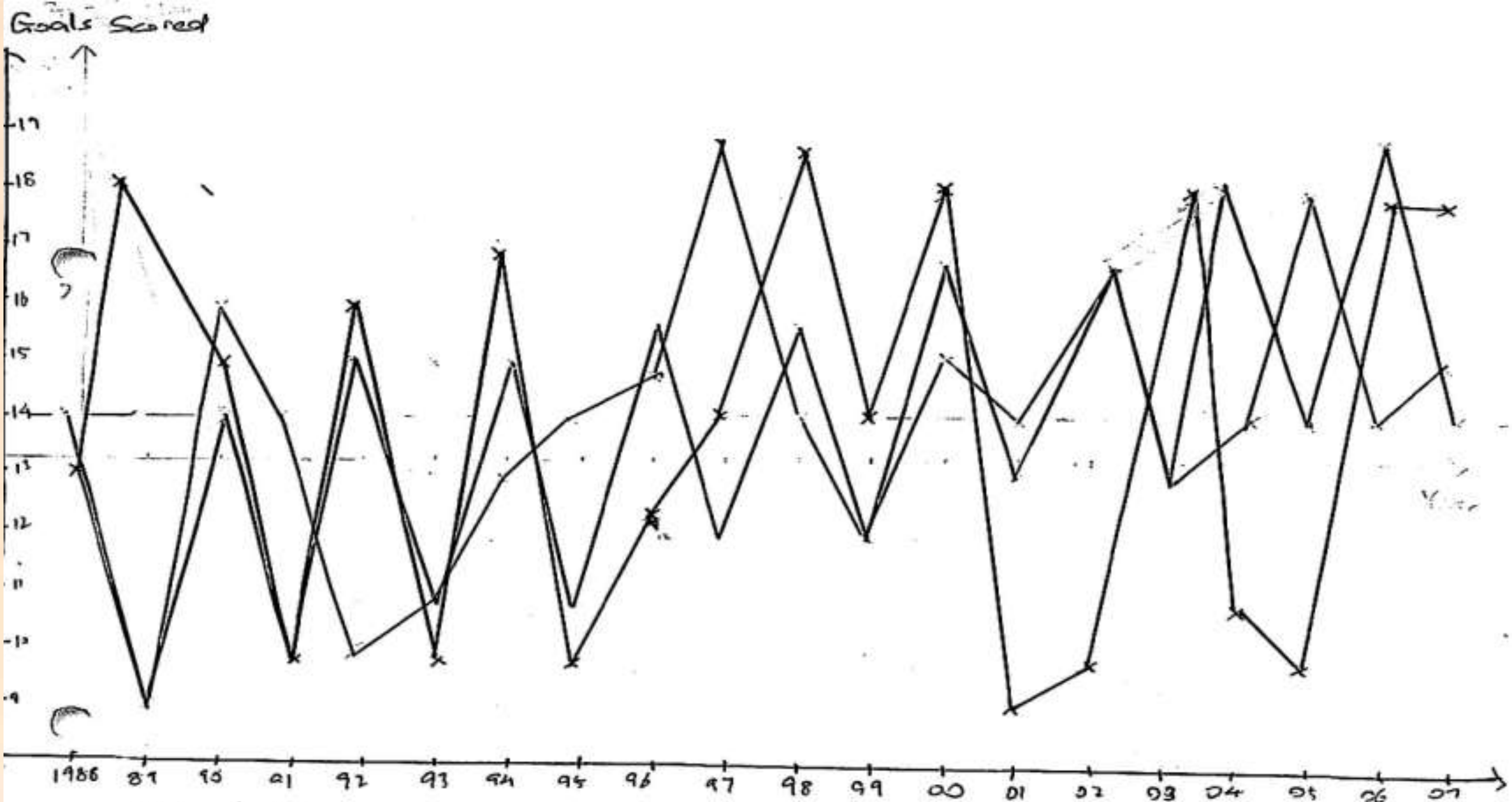
DAVE = $\frac{4+7+2+4+1+2+1+1+4+5+2+3+1+3+4+1+4+4+1}{19}$

= 54/19 = 2.84

DAVE [redacted] is most consistent

IVAN = $\frac{5+3+5+1+6+7+7+2+2+5+5+4+9+1+8+7+1+8+0}{19}$

= 4.79



— Mike Arwen
— Dave Backhand
— Ivan Right

Idea 3 Measure Graph Length

MA $\sqrt{26}+\sqrt{26}+\sqrt{17}+\sqrt{26}+\sqrt{17}+\sqrt{17}+\sqrt{17}+\sqrt{26}+\sqrt{10}+\sqrt{10}+\sqrt{10}+\sqrt{26}+\sqrt{17}+\sqrt{17}+\sqrt{17}+\sqrt{26}+\sqrt{17}+\sqrt{26}+\sqrt{21}=83.26$

DB $\sqrt{17}+\sqrt{50}+\sqrt{5}+\sqrt{17}+\sqrt{3}+\sqrt{5}+\sqrt{2}+\sqrt{2}+\sqrt{17}+\sqrt{26}+\sqrt{5}+\sqrt{10}+\sqrt{2}+\sqrt{5}+\sqrt{10}+\sqrt{2}+\sqrt{17}+\sqrt{17}+\sqrt{2}=56.54$

IR $\sqrt{26}+\sqrt{10}+\sqrt{26}+\sqrt{37}+\sqrt{37}+\sqrt{50}+\sqrt{50}+\sqrt{5}+\sqrt{5}+\sqrt{26}+\sqrt{26}+\sqrt{17}+\sqrt{82}+\sqrt{2}+\sqrt{65}+\sqrt{50}+\sqrt{2}+\sqrt{65}+1=94.54$

∴ Dave Backhand is the most consistent player as he has the shortest 'stretched-out' graph, showing consistency over time.

Azaac Aikeen

3

Core Mechanisms

1. Activation and differentiation of prior knowledge
2. Attention to critical conceptual features
    1. Difference between the mean and consistency
    2. Difference between a qualitative and a quantitative representation
    3. Why must deviations be positive?
    4. Why do we add all the deviations? Why not multiply them?
    5. What is the need for a fixed reference point?
    6. Why is mean a good fixed reference point?
    7. Why must we divide by *n*? …and so on…

Core Mechanisms

1. Activation and differentiation of prior knowledge
2. Attention to critical conceptual features
3. Critiquing, explaining, elaborating upon those features
4. Assembly into canonical solutions

Productive Failure: Students generate Representations & Solution Methods (RSMs), and then they get an opportunity to learn the target concept

Design RSMs	Instructional structure

Study 2: PF vs. DI

Target Concept: Standard Deviation
N = 69, 9th-grade math students from 2 intact classes
Pre-post, quasi-experimental design

Productive Failure Students generate multiple representations and solution methods, followed by DI

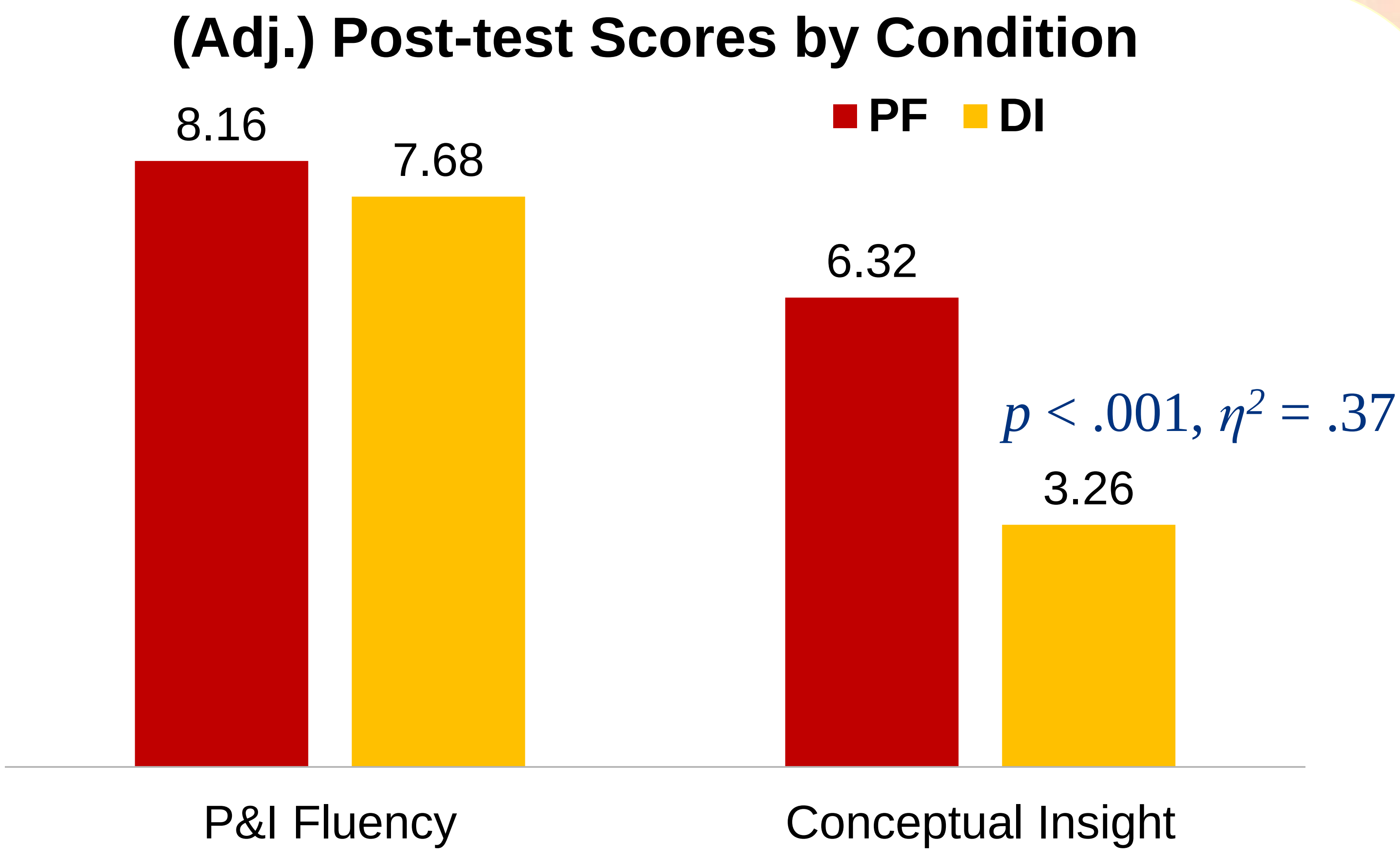
four, 50-min periods

Direct Instruction Teacher explains concept, models problem solving, uses worked-out examples, practice and feedback

four, 50-min periods

DVs:
1) Procedural & Interpretive (P&I) Fluency (3 items)
2) Conceptual Insight (3 items: 2 items on suboptimal solutions, and 1 item on sensitivity to outliers

Study 2: Results



Within PF condition: RSM diversity significantly correlated with post-test scores (Kapur, 2012)

Further Questions

1. Exposure to student-generated RSMs

Core mechanism of activation and differentiation: Is it really necessary for students to generate the RSMs or can these be given to them as worked-out examples to study and evaluate?

In common parlance, must we fail ourselves or can we learn from other people's failures?

2. Attention to Critical Features

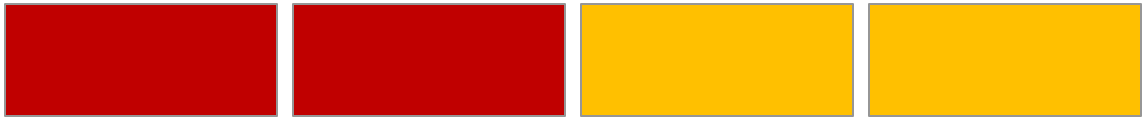
Core mechanism of attention to critical features: Do students really need to generate before receiving the critical features, or would telling the critical features without any generation work just as well?

Study 3: PF vs. Study/Evaluate

Target Concept: Standard Deviation
N = 64, 9th-grade math students from 2 intact classes
Pre-post, quasi-experimental design

Productive Failure Students generate multiple representations and solution methods, followed by DI

four, 50-min periods

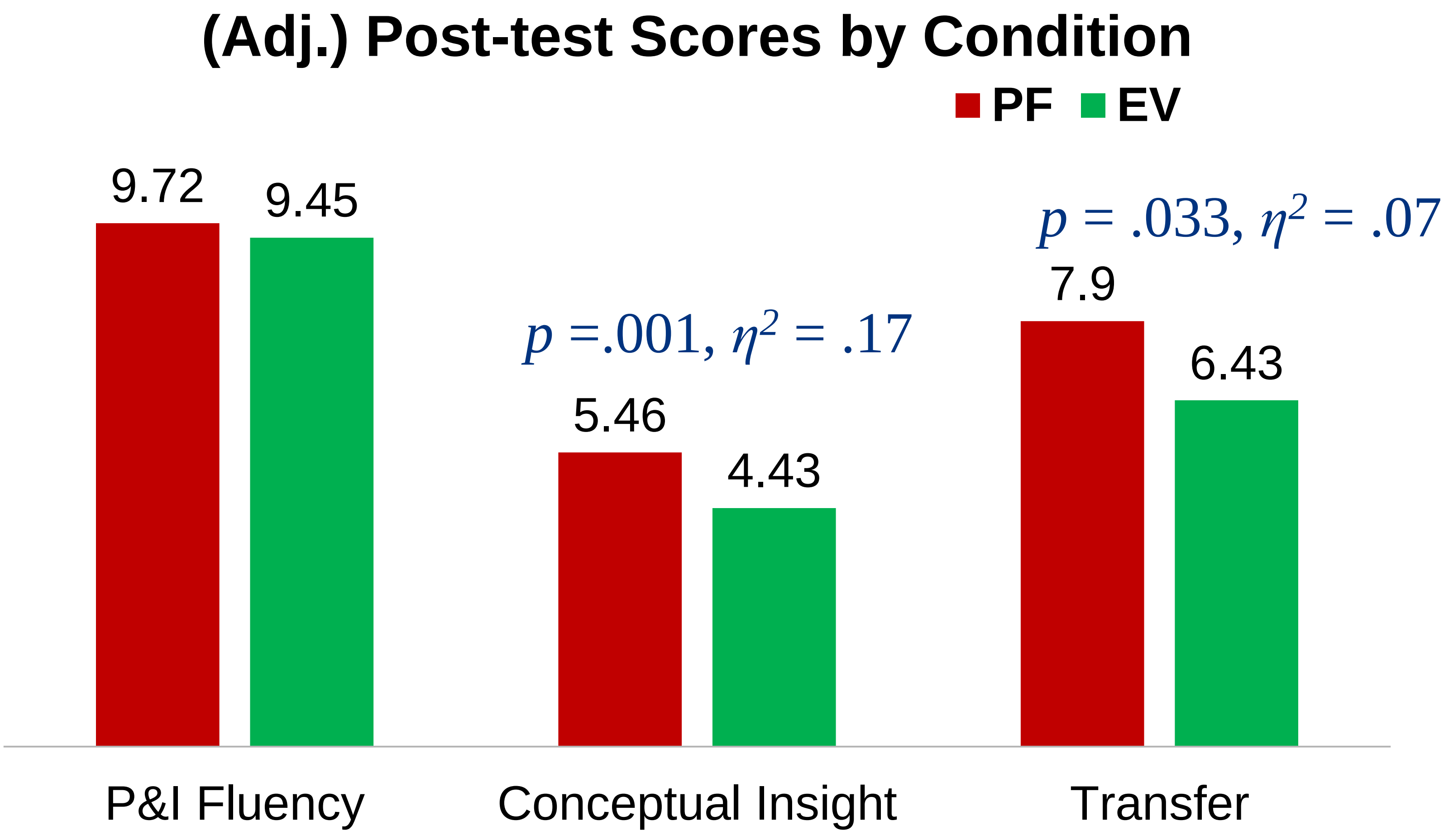


Study & Evaluate Students study and evaluate student-generated RSMs, followed by DI

four, 50-min periods

DVs:
1) Procedural & Interpretive (P&I) Fluency (3 items)
2) Conceptual Insight (3 items)
3) Transfer (3 items: 1 item on transformation, 2 items on normalization)

Study 3: Results

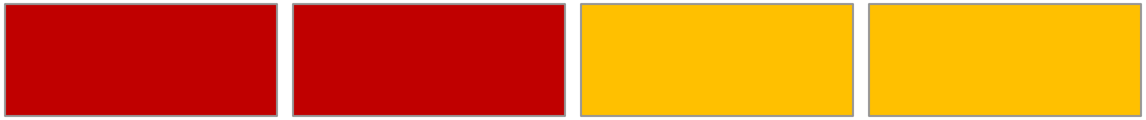


Study 4: PF vs. Strong-DI

Target Concept: Standard Deviation
N = 59, 9th-grade math students from 2 intact classes
Pre-post, quasi-experimental design

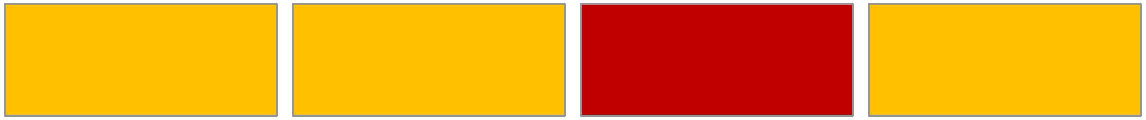
Productive Failure Students generate multiple representations and solution methods, followed by DI

four, 50-min periods



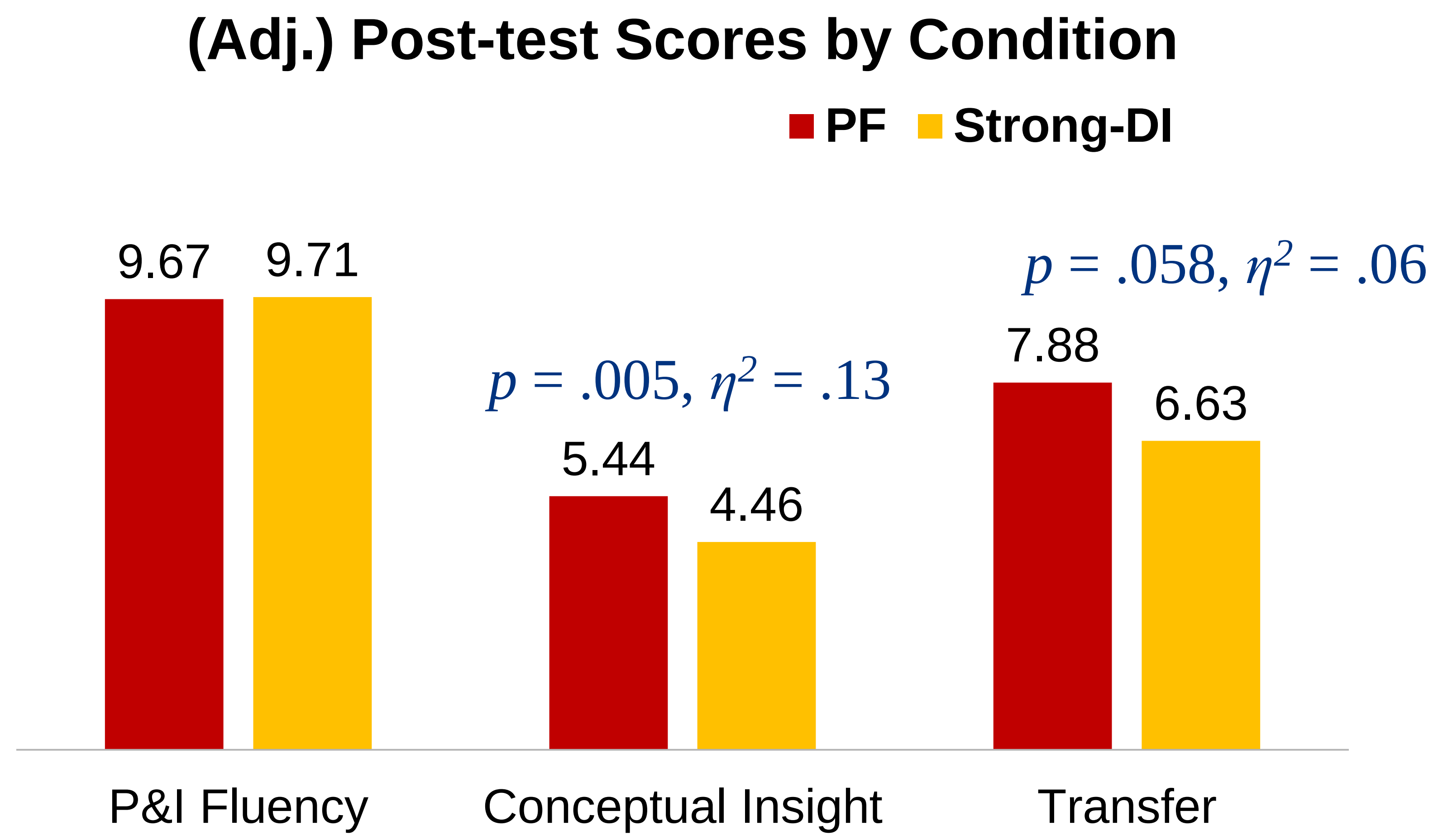
Strong-Direct Instruction Teacher explains concept with explicit discussion of critical features, worked-out examples, practice and feedback

four, 50-min periods



DVs:
1) Procedural & Interpretive (P&I) Fluency (3 items)
2) Conceptual Insight (3 items)
3) Transfer (3 items)

Study 4: Results



Summary

The four studies collectively evidence the core mechanisms embodied in the PF design

Exposure to student-generated RSMs important but evaluation may imposes a heavy load, i.e., we can learn from other people's failures but learning from our own seems better…

Direct instruction with critical features important, and reduced the effect of PF over DI

The core mechanism of activation and differentiation of prior knowledge—formal and intuitive → helps manage cognitive load perhaps?

The core mechanism of knowledge assembly as enacted by teacher consolidation is critical

Summary

- Students that seem strikingly dissimilar on academic competence (PSLE) appear strikingly similar in terms of their design competence
- Design competence correlated with learning gains
- Teachers/experts are generally good estimators of academic competence but not of design competence
- PF teachers consistently report that they are stressed and stretched to work with students' ideas… BUT, they themselves understood the concept of variance deeply

Some final comments on...

- Scaffolding (Woods, Bruner, & Ross, 1976)
- Zone of Proximal Development (Vygotsky, 1978)
- Cognitive Load Theory (Sweller, 2009)

working memory constraints?

"Any instructional theory that ignores the limits of working memory when dealing with novel information or ignores the disappearance of those limits when dealing with familiar information is unlikely to be effective" (p. 77; Kirschner et al., 2006)

Constraints of working memory contingent upon:

- novelty of information / concept being learnt vis-à-vis what students know about the concept
- interaction between working memory (WM) and long term memory (LTM)

Invariant across a constructivist or an information processing perspective: understanding and working with students' priors is critically important

In ending…

Learning vs. performance…

Productive Success	Productive Failure
Unproductive Success	Unproductive Failure

THANK YOU