A realist framework investigating the effectiveness of Data-driven Learning for Secondary School EFL students in Southern China

Xin Xu
Supervisor: Pascual Pérez-Paredes
RSLE Presentation, 06 June 2018
Faculty of Education, University of Cambridge

EFL learners from a low-income family in a rural area, **however able**, are **least likely** to succeed in second-language acquisition.

(*Xinhua Journal,* 2016; Li & Yuan, 2013)

My research focus:

In Need of innovative pedagogy

Secondary school EFL learners

Data-driven learning

In need of new knowledge gain
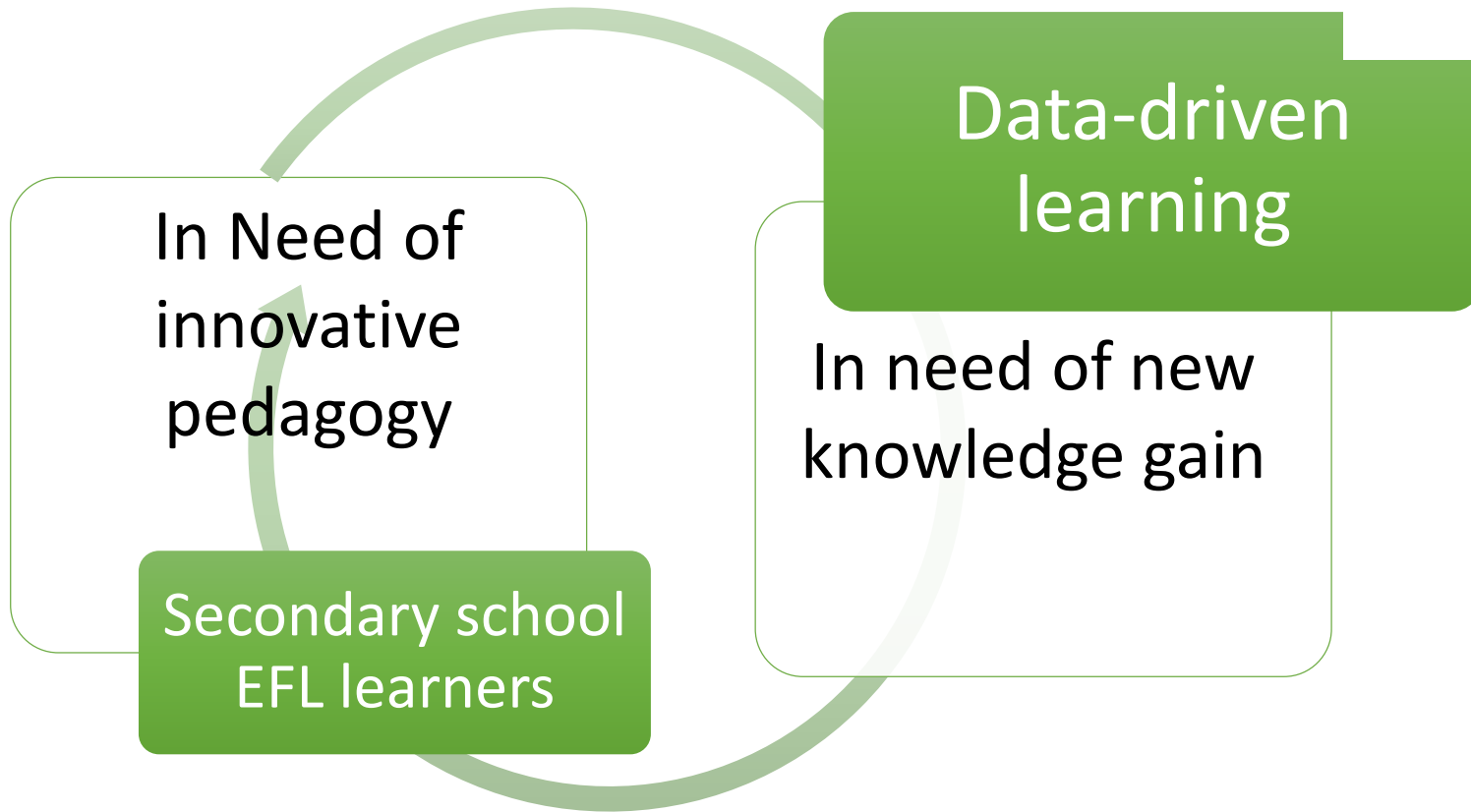
Paper-based DDL

Data-Driven Learning(DDL)

- Corpus linguistic
- Linguistic data (Johns, 1991)
- Paper-based DDL (Boulton, 2010)
- Inductive instruction

英语课前随堂测试系列 2

姓名＿＿＿＿＿＿＿＿   分数＿＿＿＿＿＿＿＿

1. This point is especially important when ＿＿＿＿＿＿＿＿(consider) technical literature.

2. ＿＿＿＿＿＿＿＿(judge) by the number of messages, everyone on the new list is having a great time.

3. Promoting those soldiers who commit to family and serve our country will encourage them to be ＿＿＿＿＿＿＿＿(rely).

4. it is not always the case that they had enough accurate information to come up with the same ＿＿＿＿＿＿＿＿(judge) as those with better information sources.

5. She should be more ＿＿＿＿＿＿＿＿(consider) of her mother and yet the work must be done or there would be no income.

6. The＿＿＿＿＿＿＿＿(origin) design has been changed three times since 1901.

7. I was ＿＿＿＿＿＿＿＿(anxiety) to start experiencing normal life again.

8. In the absence of a policewoman, the female field-worker was＿＿＿(rely) upon for comfort by female members of the public.

9. This procedure usually is ＿＿＿＿＿＿＿＿(consider) minor surgery.

10. The ＿＿＿＿＿＿＿＿(origin) of boxed construction is unknown.

1. Anxiety—anxious—anxiously

Mcintosh even claiming that "the level of **anxiety** diminished as local communities became
Taming of the Shrew' demonstrating widespread **anxiety** about controlling women; David E.Underdown, '
a few hours and therefore helps minimise the **anxiety** associated with awaiting a diagnosis (32). </p><p> In
are especially vulnerable to depression and **anxiety** compared to their older counterparts and
in a daily exercise routine. His reported **anxiety** and low mood during the past 6 weeks obviously
poverty exists, rates of social stress and **anxiety** tend to be higher 7. This is due to "psychosocial
he watches the child picked up by her mother. His **anxiety** is reinforced by the cognac which he purchases

from being in situations that they fear or feel **anxious** about. It also arms patients with the knowledge
play a big part in the patient's fear and feel **anxious** about or anxiety and by excluding such
not be affected, as the participants would be **anxious** about achieving an impressive final score in
much of an importance, 'Every one of them was **anxious** to carry away something memorial...pieces of
of extra - Europeans, as the Europeans were **anxious** to express the fact that they shared some basic
they were already familiar. They were also **anxious** to convey that there was a common humanity and
for reasons of self-advancement. They were **anxious** to depict the natives as savages in order to
spirit also allows youngest siblings to be less **anxious** and reserved, thus promoting an extroverted

cry of one who foresees ill for the sleepers and **anxiously** bids them awake.' <p> The mood continues in the same
She has very little to do on stage other than wait **anxiously** for Corder and follow the instructions he gives
a particular type, either securely attached, **anxiously** resistant and anxious avoidantly attached.
of transcending the boundaries, the more **anxiously** the narrow community watched over the deeds,

2. Original—Origin—Originally

the responses given. In terms of errors made the **original** responses showed that participants during the
levels at university. The word "others" in the **original** questionnaire was changed to "my new friends"
questionnaires were scored according to the **original** criteria. Second marking only occurred on the
" or "new friends" were included, 22 out of the **original** 40 questions. This was to see whether self
is only six nanometres (nm) away from the **original** red photopsin (Winderickx, Lindsey, Sanocki,

through varying perceptions of madness and its **origins** . Whilst there were significant implications
and twentieth centuries. The primary **origins** of women's madness were considered a result of
the role of psychology in determining the **origins** of social behaviour (Simmel 1971a). Simmel was
operations'. It is important to note, that the **origins** of this alienation were, to a great extent,
2: A suggested evolutionary pathway for the **origin** of mitochondria 9 <p> Figure 2 shows the possible
endogenous and exogenous hypotheses for the **origin** of undulipodia. Supporters of the endogenous

this suggestion does seem more practical than I **originally** thought. However, as a realistic emancipation
group only, for example, question two **originally** presents "I feel others do things much better
(Eich et al, 2000). Lloyd and Lishman (1975) **originally** demonstrated mood congruent retrieval,
Keane, 2000). Bower, Monteiro and Gilligan **originally** investigated this, using hypnosis to induce
because the emotions causally belonged **originally** to the events being recalled, and the

Main research question

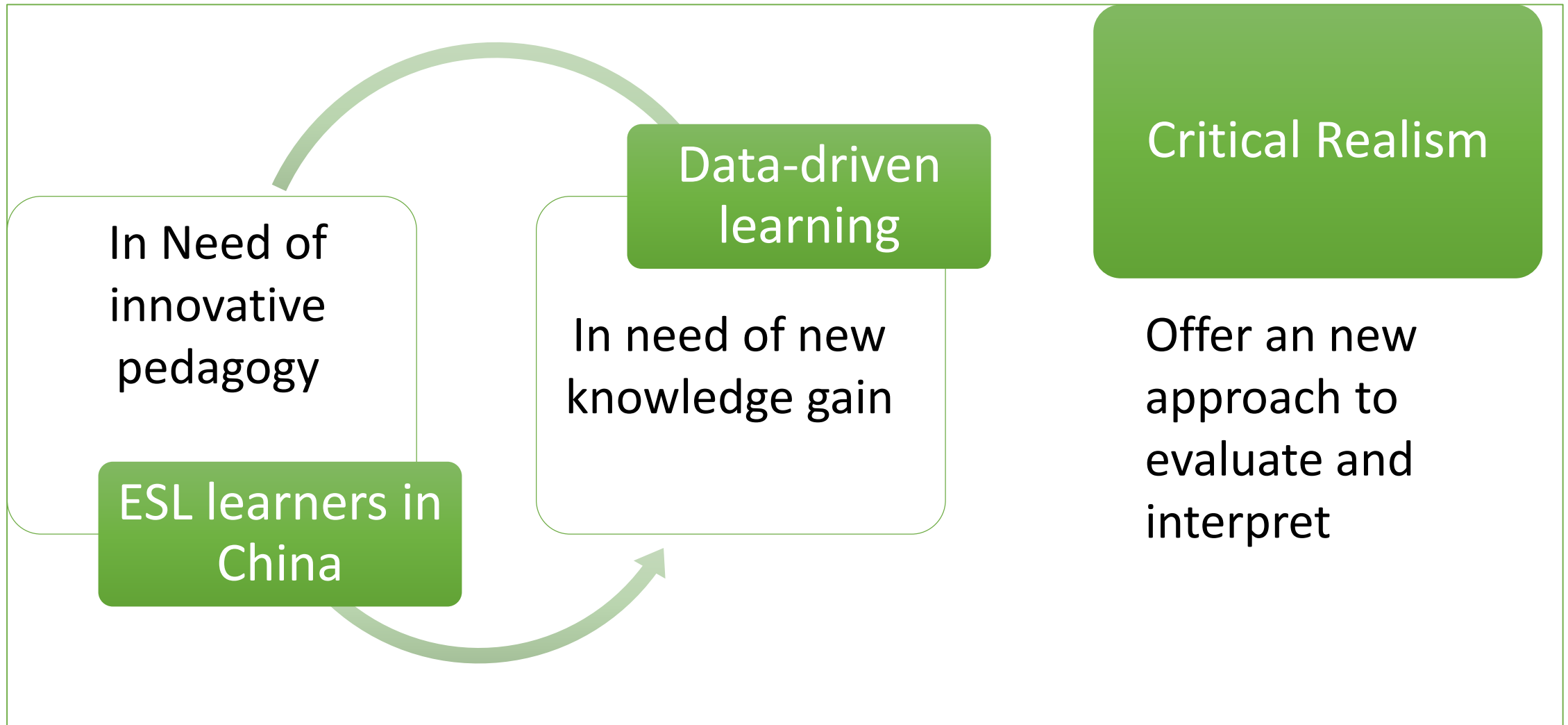1. To what extent can paper-based DDL materials impact learner's language gain?

Empirical Level
-experienced and observed events
-events understood through human interpretation

Actual Level
-events occur, whether observed or not

Real Level
-causal mechanisms within objects or structures cause events at empirical level to occur
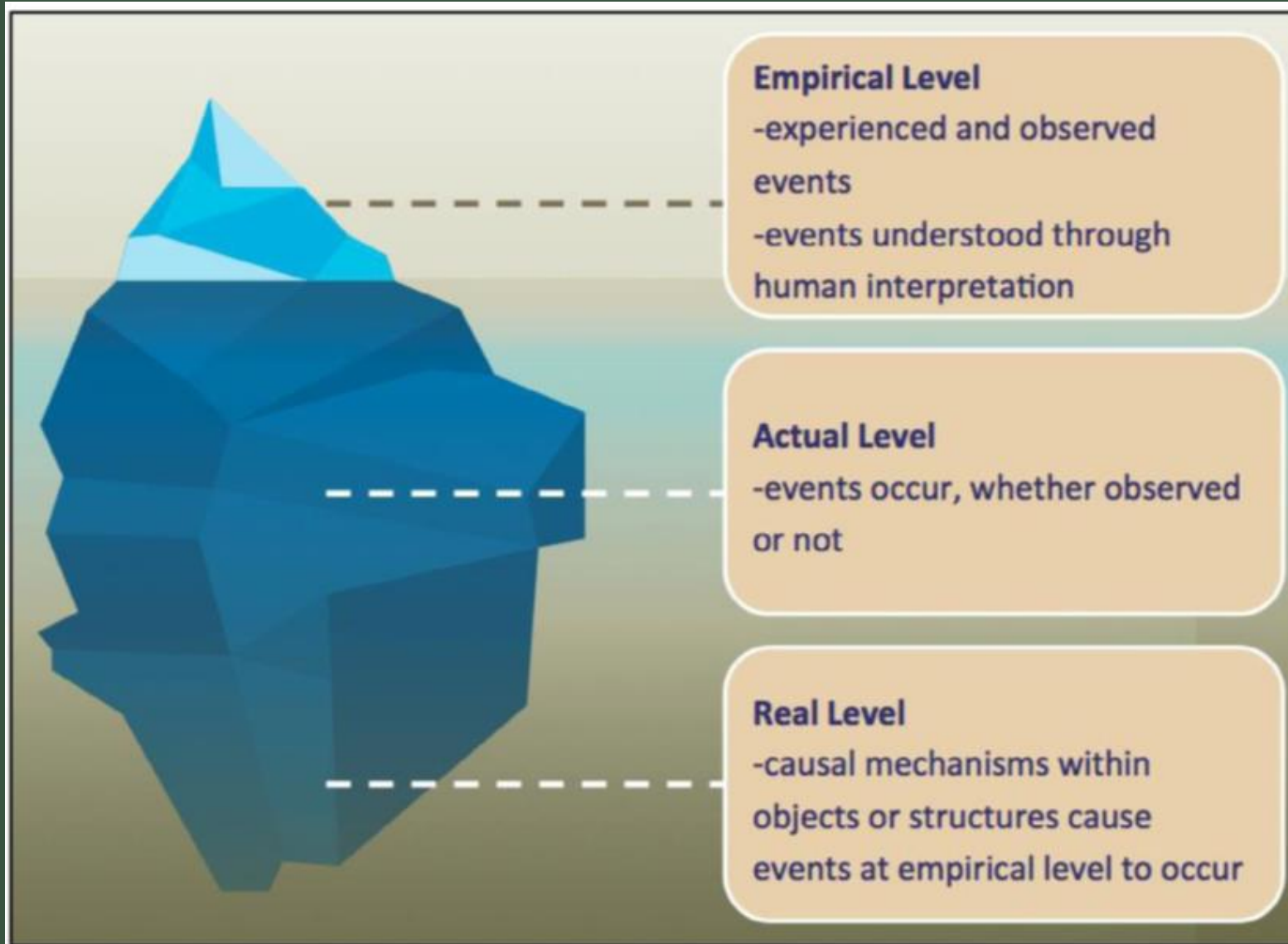
(Fletcher, 2017)

- Realist Ontology, Constructivist Epistemology

- stratified reality

- CMO:
  - Regularity/outcome = mechanism + context (Pawson & Tilley, 1997)

Context + **Mechanism** = **Outcome**

Context	+	New mechanism	=	Outcome pattern
Small isolated area, with local burglars	+	Familiarization with new measures and recognition of unchanged risks	=	Gradual return to previous burglary rates
Small isolated area, with local burglars	+	Refreshment of perceptions of scheme efficacy via publicity	=	Renewed fall in burglary rate

Context + ???
= outcome

Main research question

1. To what extent can paper-based DDL materials impact learner's language gain?

2. How do the underlying mechanisms in using paper-based DDL materials give rise to such language gain?

Site and participants

A classroom of 58 students

18-19 years old

Han ethnicity

in their final year

secondary school in Hunan Province, China

22 female, 36 male

L1 Mandarin

Study English for 8-9 years

Low-level EFL learners

A mixed-methods approach

Research Question		Data collection Method		N.	analysis
DDL's impact on Language gain (Empirical Level)	vocabulary gain	Student achievement data: pre-test, post-test and delayed post-test	One round of TAs	40	Two-way ANOVA
			Three rounds of Intervention	40	
	Improved Language sensitivity Pattern noticing	Student interview	After three rounds of intervention; one-to-one	12	Tape-recorded and transcribe; coding
		Observation	During 4 lessons, unstrutured	4	
Mechanism (Real level)		Student interview	After three rounds of intervention	12	same

Data Collection Timeline

context

I &M

X X1 X2 Y Y1 Y2 Z Z1 Z2 A A1 A2

TAs   Intervention 1   Intervention 2   Intervention 3
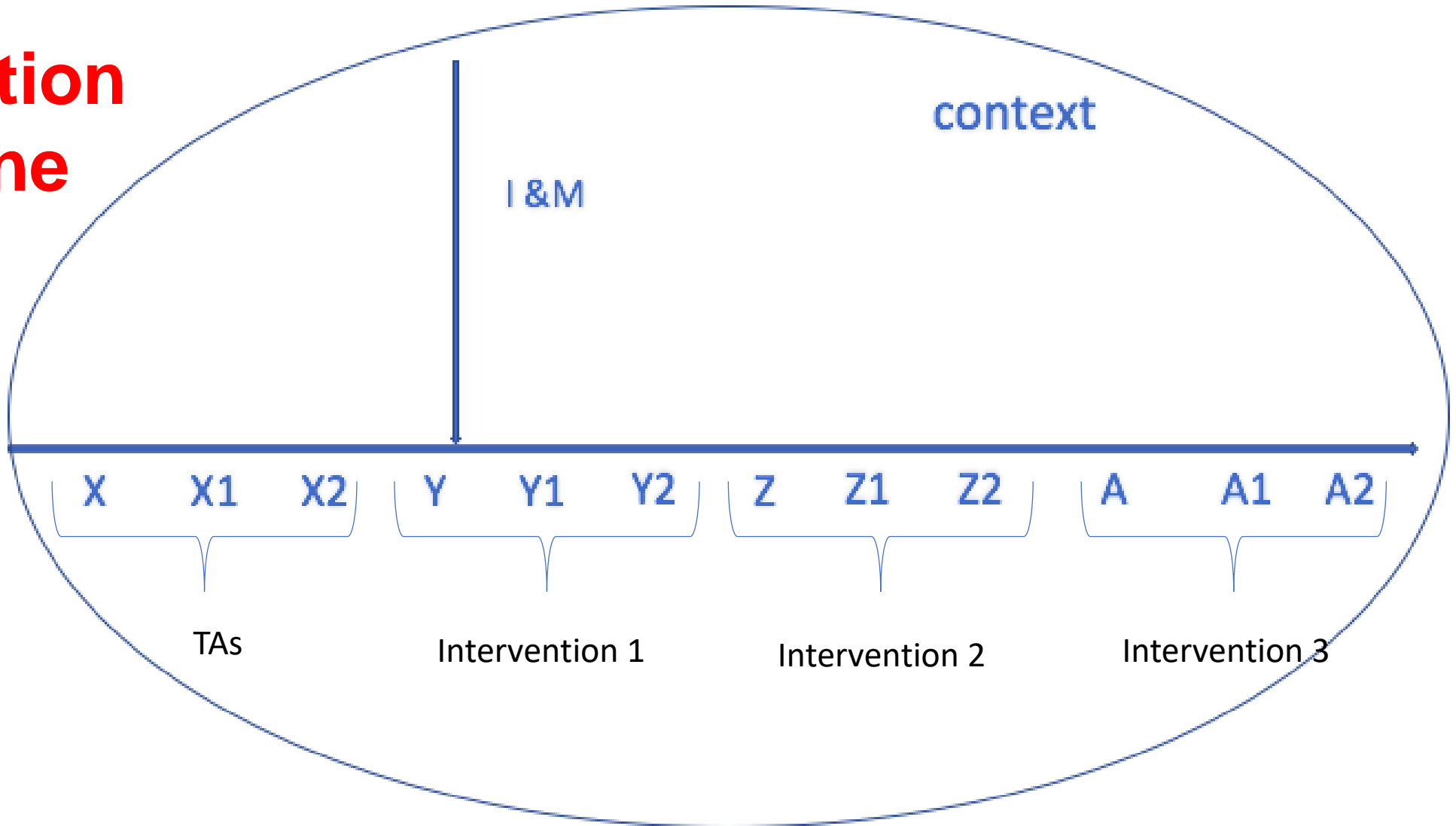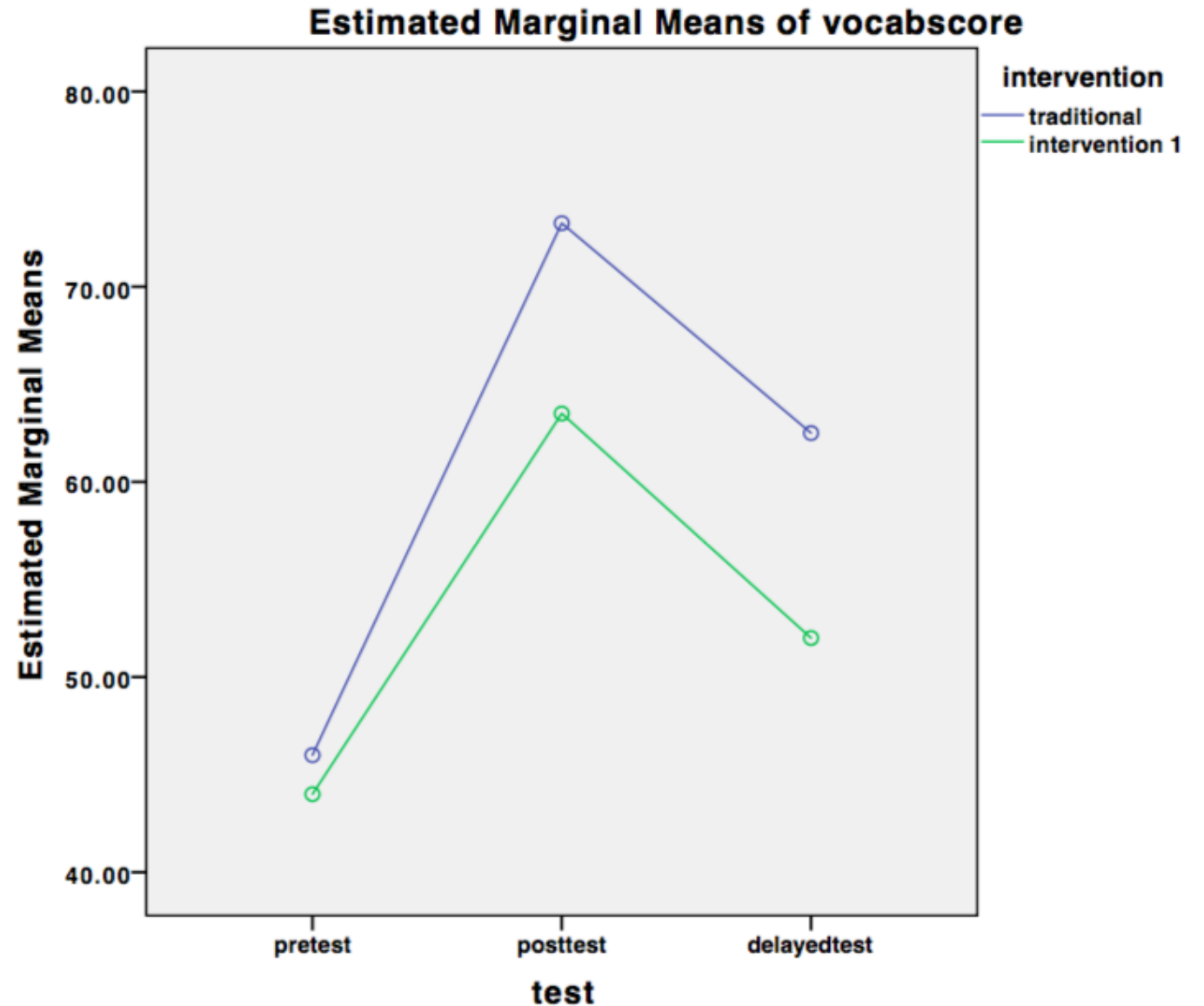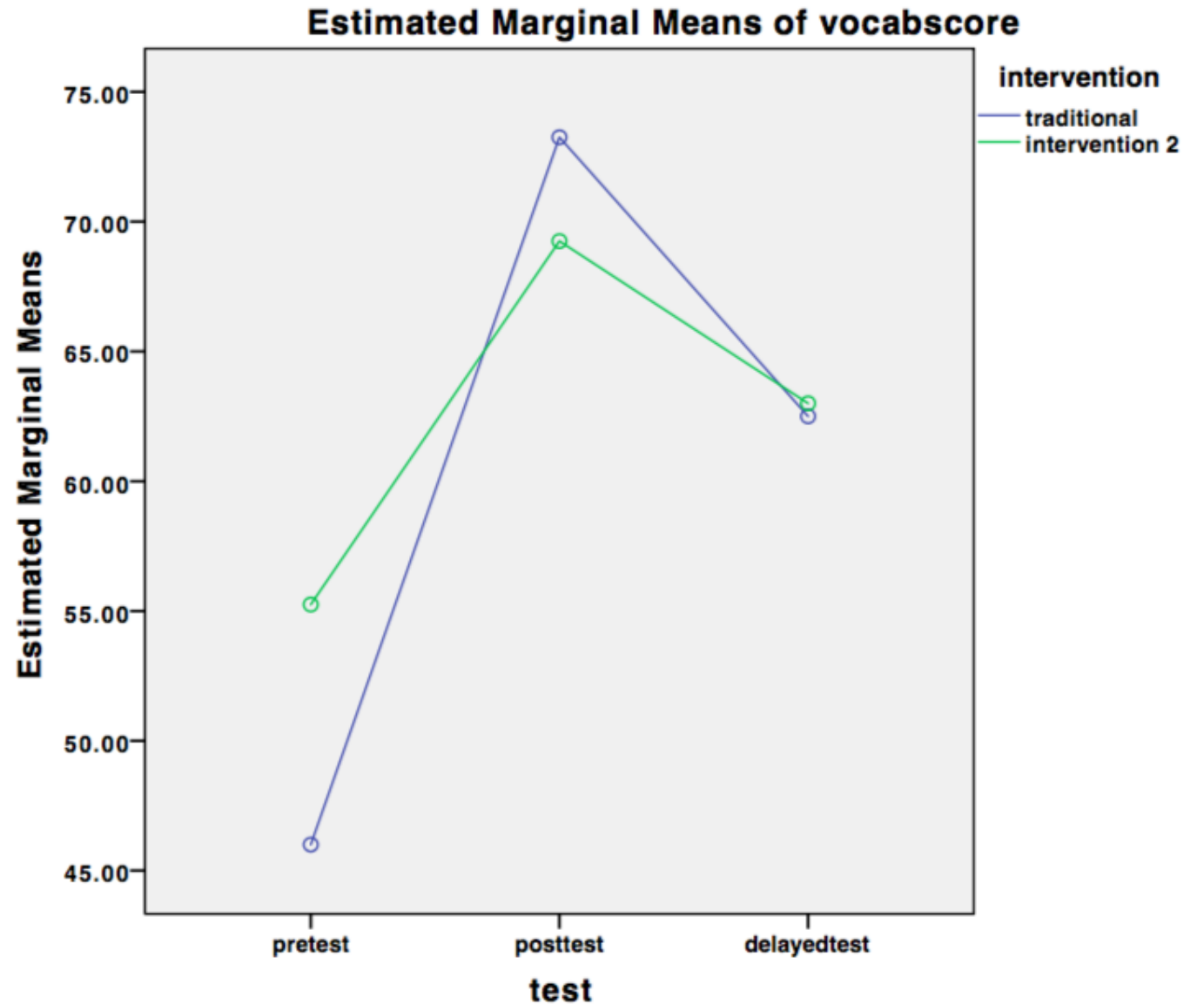
Quantitative Data

1. Histogram: **Normality Assumption** met
2. Levene's test: assumption of sphericity/ **homogeneity of varia**nce is not violated, the P value is smaller that 0.05 (all variance are equal)
3. **Two-way ANOVA** with interaction
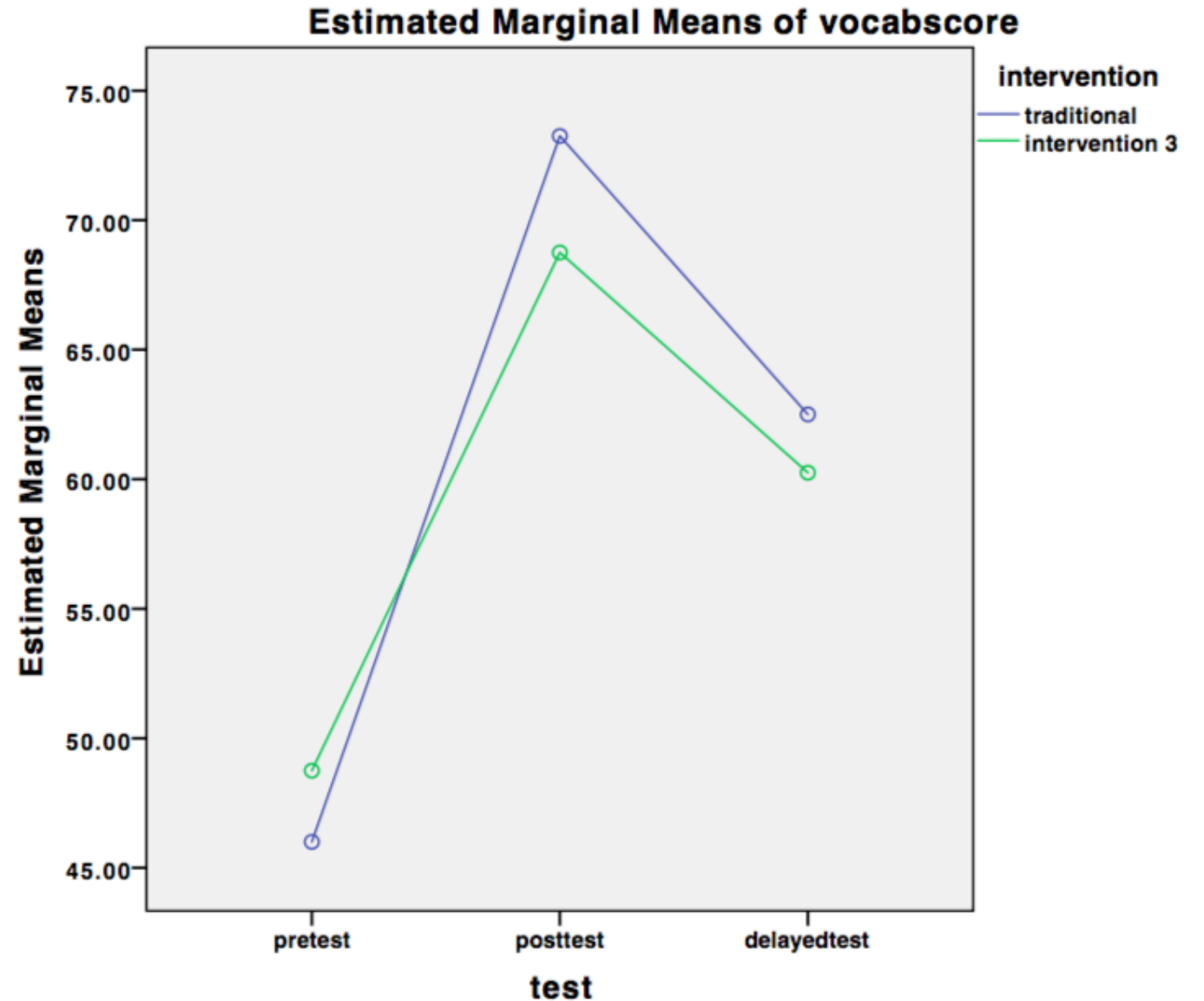4. **Post hoc** test: to see where the significance lie

Two-way ANOVA 2

Estimated Marginal Means of vocabscore

Two-way ANOVA 3

Estimated Marginal Means of vocabscore

Post hoc test

Multiple Comparisons TRADITIONAL

Dependent Variable: vocabscore

Bonferroni

(I) test	(J) test	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
pretest	posttest	-27.25000*	5.16346	.000	-39.7915	-14.7085
	delayedtest	-16.50000*	5.16346	.005	-29.0415	-3.9585
posttest	pretest	27.25000*	5.16346	.000	14.7085	39.7915
	delayedtest	10.75000	5.16346	.119	-1.7915	23.2915
delayedtest	pretest	16.50000*	5.16346	.005	3.9585	29.0415
	posttest	-10.75000	5.16346	.119	-23.2915	1.7915

*. The mean difference is significant at the 0.05 level.

Post hoc test

Multiple Comparisons INTERVENTION 1

Dependent Variable: vocabscore

Bonferroni

(I) test	(J) test	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
pretest	posttest	-19.50000*	4.24214	.000	-29.8037	-9.1963
	delayedtest	-8.00000	4.24214	.185	-18.3037	2.3037
posttest	pretest	19.50000*	4.24214	.000	9.1963	29.8037
	delayedtest	11.50000*	4.24214	.023	1.1963	21.8037
delayedtest	pretest	8.00000	4.24214	.185	-2.3037	18.3037
	posttest	-11.50000*	4.24214	.023	-21.8037	-1.1963

*. The mean difference is significant at the 0.05 level.

Context + Mechanism= Outcome

Not at work?

Compared to TAs
1: significantly lower
2:no significance
3:no significance

Within Intervention:
Always positively significant result btw Pre- and Post- test; progress
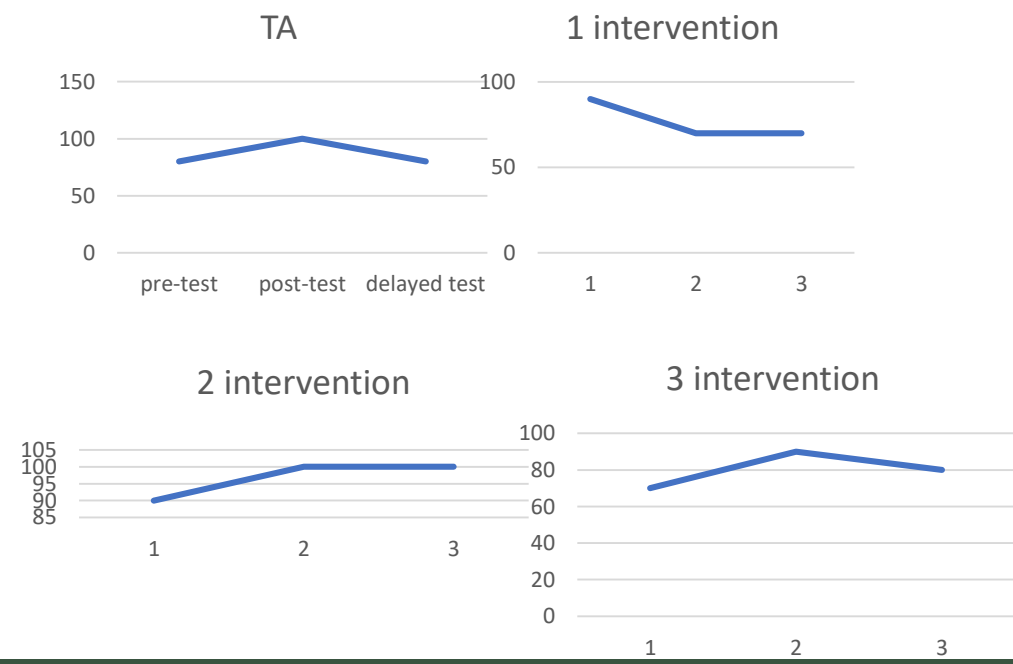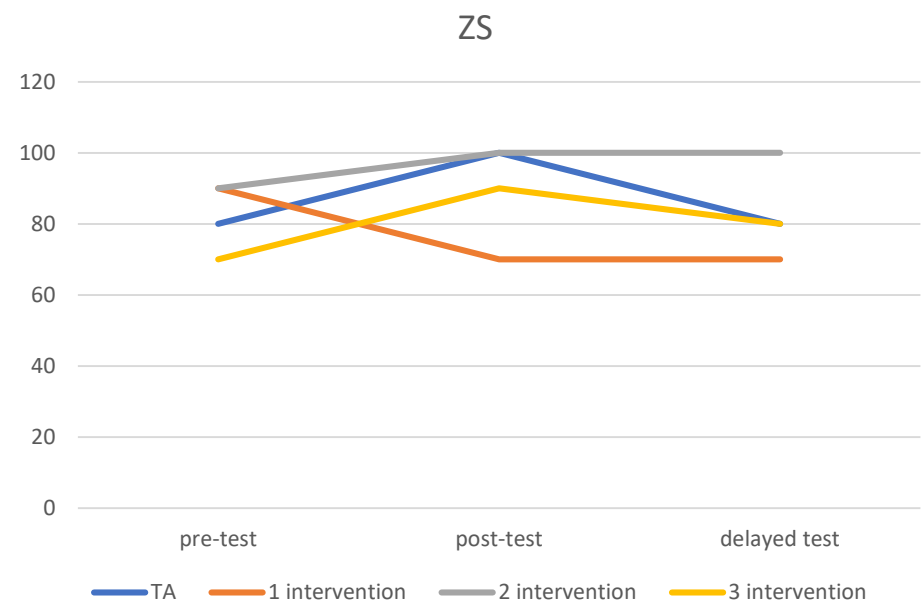
Qualitative

Observation Notes

Interview Transcripts

XX：How long did teacher usually spend on teaching vocabulary?

ZS：It's longer when we were first year students, but it is only around 10-20 minutes now.

XX：Wow, can you comment on its comparison to DDL, which takes 1-2 lessons to cover half the vocabulary in one unit?

ZS：DDL session takes longer, but worth spending the time as we can learn it better. It is one thing to memorise the spelling and Chinese translation of a word, it is another thing to know whether it is a noun or verb and understand its daily usage. Now I realize that tit is important to do the latter and incorporate that into learning vocabulary, but not yet reached the stage of fluently applying it.

ZS：It is easy to do exam question when you not knowing this, so that you do it according to 'how you feel'. I hesitated more now as I haven't fully understood its usage and application. I know that adverb goes with verb, but I have difficulty identifying the adverb or verb in the exam paper.

XX: Why is that?

ZS：Because when we learn vocabulary, we do not pay attention to that. It would be much better if we can be exposed to DDL earlier, like in our first year in secondary school.

ZS — line chart with axis 0 to 120, x-axis: pre-test, post-test, delayed test. Series: TA, 1 intervention, 2 intervention, 3 intervention.

TA — line chart, axis 0 to 150, x-axis: pre-test, post-test, delayed test.

1 intervention — line chart, axis 0 to 100, x-axis: 1, 2, 3.

2 intervention — line chart, axis 85 to 105, x-axis: 1, 2, 3.

3 intervention — line chart, axis 0 to 100, x-axis: 1, 2, 3.

Context + Mechanism= Outcome

Fixed **curriculum goals** (Smith, 2014)—time constraint;
58 students VS 1 English teacher–scaffolding and facilitation;
Exam-oriented; explicit **rule-based** language teaching and learning

Increased linguistics sensitivity;
pattern noticing and induction ability

Context + Mechanism= Outcome

Fixed **curriculum goals** (Smith, 2014)—time constraint;
58 students VS 1 English teacher– scaffolding and facilitation;
Exam-oriented; explicit **rule-based** language teaching and learning;
DDL lesson last shorter

Pre-existing: Culture and expectation of Language teaching and learning

DDL triggering: inductive learning, pattern noticing and internalizing, TBC

Qualitative Data: Increased linguistics sensitivity; pattern noticing and induction ability

Quantitative Data
Compared to TAs
1: significantly lower
2:no significance
3:no significance

Within Intervention:
Always positively significant result btw Pre- and Post- test; progress

Discussion

Deny DDL?
Or just the learning curve?

Wang, 2013:

Increased vocabulary gain qualitatively, but it takes longer than TAs, which make DDL only complementary

Sweller, Ayers, & Kalyuga, 2011:

pattern induction is a natural process, it **reduces the cognitive load** of processing

Barrett, Dunbar, & Lycett, 2002:

rules are an artificial intellectual abstraction, whereas the human brain is programmed to detect patterns in the world around us (e.g.,)

Aston, 1998, :
This allows the learner to proceed toward the target norm by **progressive approximations** DDL arguably promotes such skills, which should **be transferable to new contexts** and thus produce **better learning outside the classroom**, increasing learner autonomy and lifelong learning for them.

TBC

Limitations

- Translation issues

- Limited Intervention

- Quantitative data: Missing students; Sleepy student; Only one round of Test; Spelling mistakes; difficulty of the Vocabulary

- Qualitative data

References

- Boulton, A. (2010). Data-Driven Learning: Taking the Computer Out of the Equation. *Language Learning, 60*(3), 534–572.

- Easton, G. (2010). Critical realism in case study research. *Industrial Marketing Management, 37*, 118–128.

- Edwards, P. K., O'Mahoney, J., & Vincent, S. (Eds.). (2014). *Studying organizations using critical realism: A practical guide*. Oxford: Oxford University Press.

- Fletcher, A. J. (2017). Applying critical realism in qualitative research: methodology meets method: International Journal of Social Research Methodology: Vol 20, No 2. *International Journal of Social Research Methodology, 20*(2), 181–194.

- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: SAGE Publications Ltd. Retrieved from

Parr, S. (2013). Integrating critical realist and feminist methodologies: Ethical and analytical dilemmas. *International Journal of Social Research Methodology, 18*, 193–207.

Johns, T. (1991). "Should you be persuaded": Two samples of data-driven learning materials. In T. Johns and P. King (Eds.),*Classroom Concordancing ELR Journal, 4* (pp. 1–16).

Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100–115).

Sripicharn, P. (2000). Data-driven learning materials as a way to teach lexis in context. In C. Heffer, H. Sauntson & G. Fox (Eds.), *Words in context: A tribute to John Sinclair on his retirement* (pp. 169–178).

Many thanks to:

Supervisor:
Pascual Pérez-Paredes

I welcome questions or comments at
xx802@cam.ac.uk
Thanks!
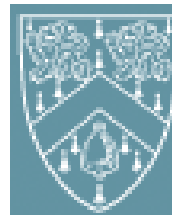
Supplementary Materials
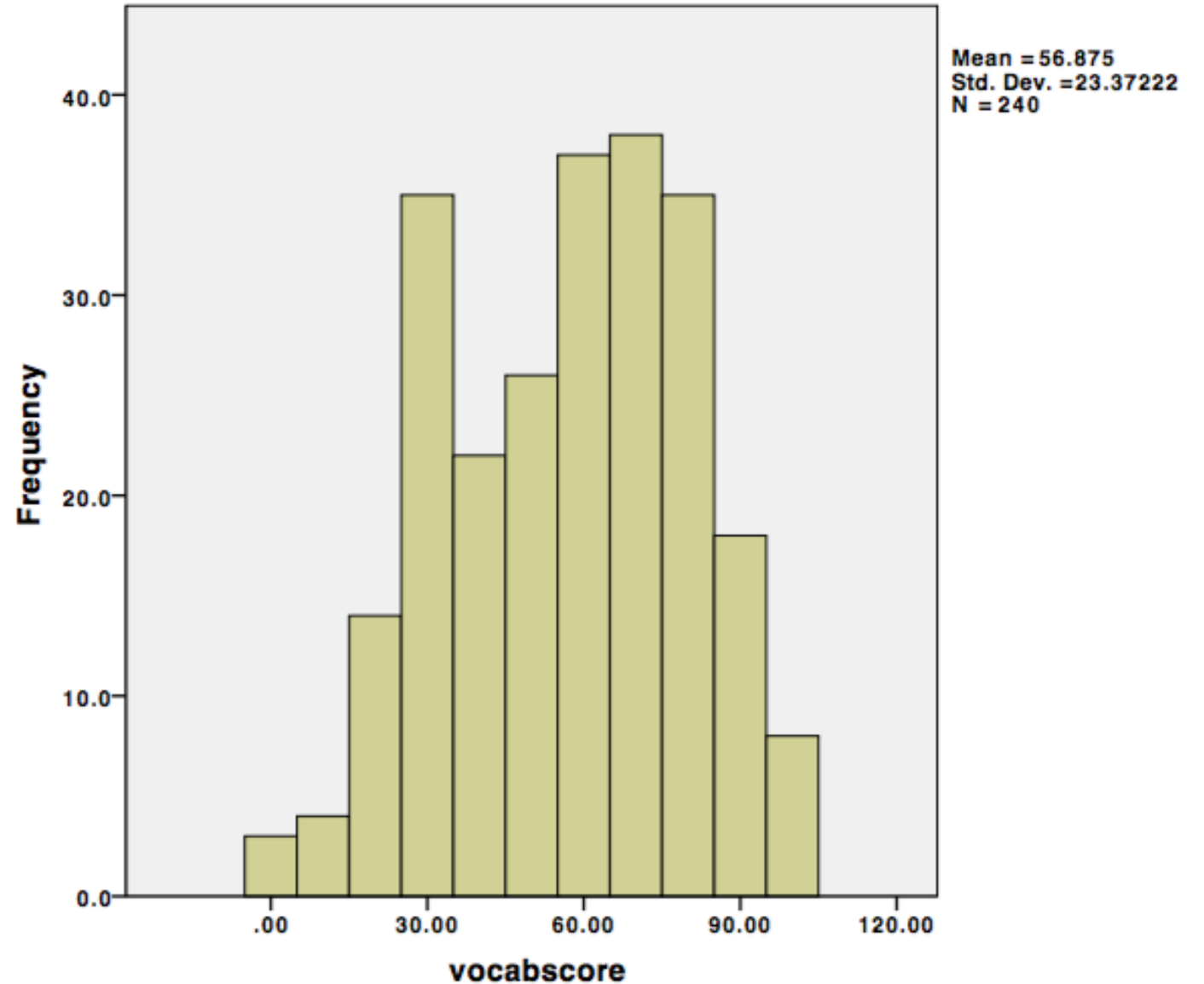
Assumption Test

Levene's Test of Equality of Error Variances[a]

Dependent Variable: vocabscore

F	df1	df2	Sig.
1.096	5	234	.363

Mean = 56.875
Std. Dev. = 23.37222
N = 240

- Intervention material

- Material design; British Academic Written English Corpus (BAWE)

- Test questions: Sketch Engine for Language Learning (https://skell.sketchengine.co.uk/run.cgi/skell)

- usage-based perspective: "frequency of exposure" and "salience of the feature"

Paper-based DDL

- intervention

Main materials:

c) The following concordances all come from the same novel.
   Can you identify the pattern of usage in each group?

1. Harsnet (typed Goldberg), does she entice him there or does he **want** to come? Does she entice him or does he appear by chance? Ovid
2. wrote Harsnet, it is this which excites me, this which makes me **want** to go on. As if my whole life, he wrote, had been spent
3. , it had all become too complicated, too extreme, I did not **want** to know any of it until it was all over, until I had made
4. What about the right hand side? I said. That's what I **want** to know. But he couldn't help. Only kept repeating it was quite
5. , he wrote, this is a message from the past. I just **want** to tell you. Goldberg, pushing aside pad and pen, drew the little
6. and less and less well of course, but the truth was I did not **want** to wake up. How the days, instead of each being distinct from each
7. story, you can have one. If not, not. If you **want** to walk round it, you can do so. If you want to get
8. you with it, he said. My best picture. You didn't **want** to win their lousy prize, did you? I said. You did too
9. a room. And so on. Otherwise perhaps too dangerous. I don't **want** anyone to get lost in it, I said. To be unable to get
10. now? I want it to be as though I had never been. I **want** it to be as though I had never taken that turning. But that can not
11. difficult to know if idea is really valid. Will really yield what I **want** it to. You can dream and dream but only what happens in this room
12. Goldberg too: Help me, I'm so unhappy. What do they **want** me to do? Blow their noses for them? And why me? Why
13. ? Go where? wrote Harsnet (typed Goldberg). Where did they **want** me to go? Mushrooms grow in the dark, I said to Hilda.
14. to stifle doubts by crushing you with sheer bulk, he wrote. I **want** my doubts to play and dance. And Goldberg, drawing his pad
15. glass to be seen, I want it placed in a morgue and I **want** people to come in and see it, pay money and come in and see
16. to be done. I myself am guilty, he wrote, in that I **want** the glass to be seen, I want it placed in a morgue and I

want

vt
   (=desire, wish for) **vouloir**
   (=need) **avoir besoin de**
   if you want **si vous voulez**
   what do you want? **Qu'est-ce que vous voulez?**
   (=should) sb wants to do sth * qn devrait faire qch

n
   (=poverty) **besoin** *m*
   (=lack) **manque** *m* **wants**

npl
   (=requirements) **besoins** *mpl*

want (wants *plural & 3rd person present*) (**wanting** *present participle*) (**wanted** *past tense & past participle*)
   1 **verb** If you **want** something, you feel a desire or a need for it. I want a drink... *V n*
   2 **verb** You can say that you **want** *to* say something to indicate that you are about to say it. Look, I wanted to apologize for today. I think I was a little hard on you. *V to-inf*

Paper-based DDL teaching materials

Traditional inductive teaching materials