

Are Icelandic boys really better on  
computerized tests than on  
conventional ones?

Gender difference in CBAS

Almar Miðvík Halldórsson, ETI  
PISA 2006 Nordic Conference  
Grand Hótel 18. ágúst 2009

# Topic today

## Test modalities and gender difference

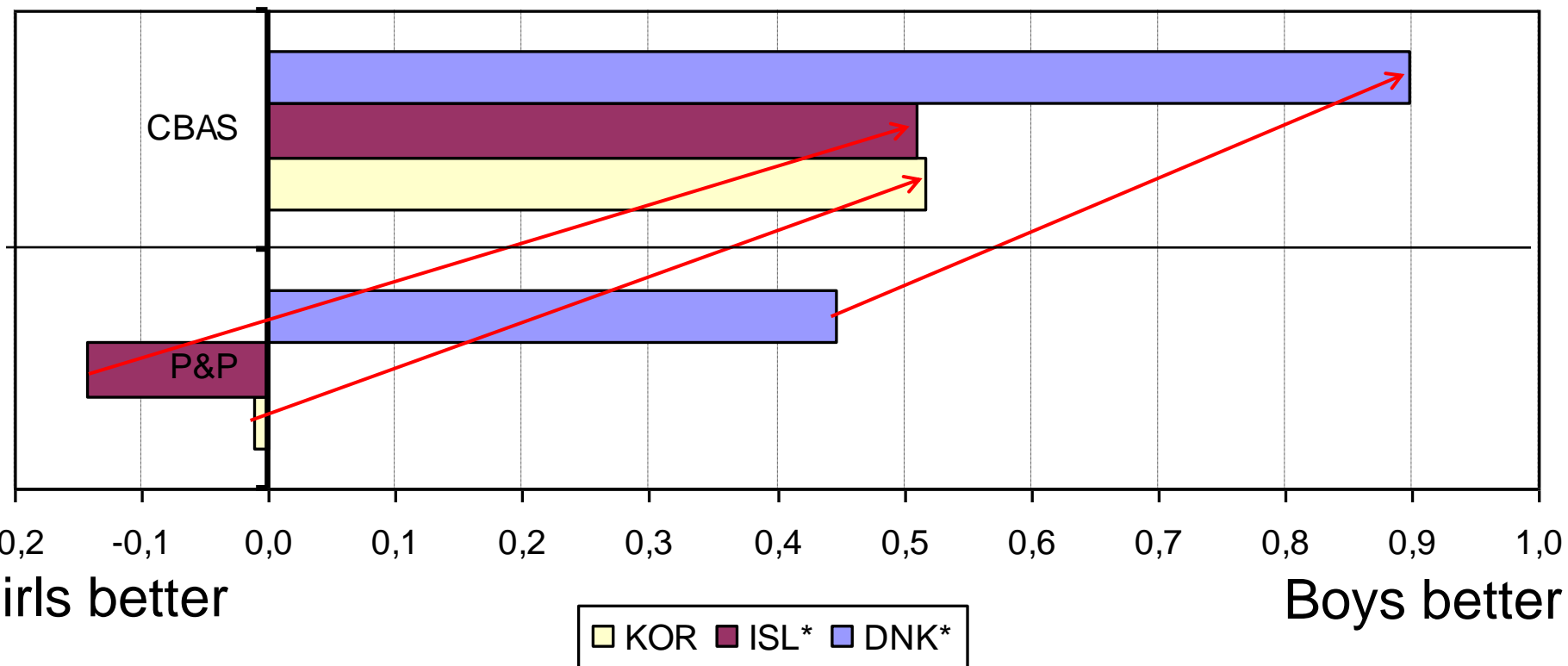
- PISA 2006 results on Science Literacy compared with CBAS 2006 results on Science Literacy
  - Modality for PISA is paper and pencil.
  - Modality for CBAS is a computer.
- Moderator variables: motivation, enjoyment, effort on the test, item difficulty, item interactivity, sub-domains of science and reading load and familiarity with computers

# Gender difference in the three CBAS countries

- PISA p&p:  
In Iceland girls outperform boys on the PISA P&P in Science literacy whereas in Denmark there was a gender difference favoring boys.
- CBAS:  
Boys outperform girls in all three countries.
- In Iceland and Korea, boys' performance increased on CBAS while girls' performance decreased.

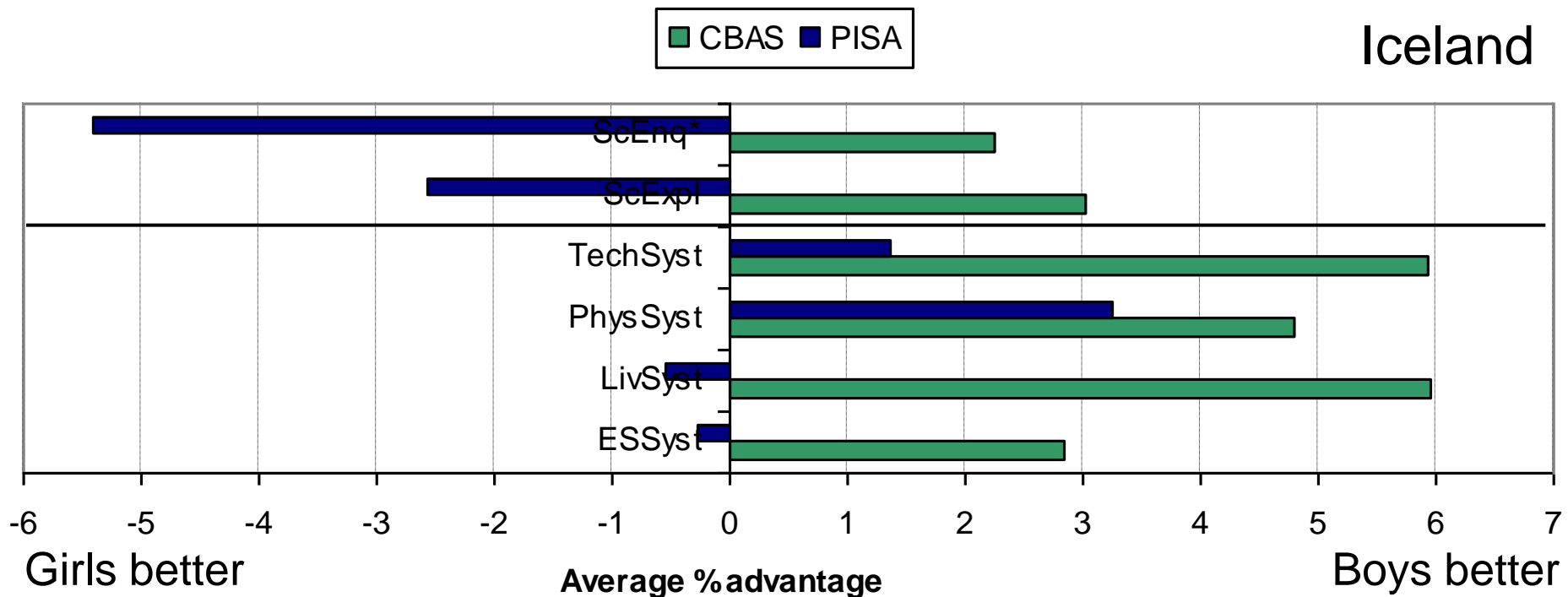
# Gender difference in PISA P&P and CBAS

- Gender difference in advantage across tests and countries
- Scale: Mean=5, St.dev=2



# Sub-domains of science in PISA

- Enquiry or Explanation.
- Systems: Physical, Living, Earth and Space, Technology.



# Sub-domains of science

- Domain coverage across test modalities is similar but boys outperform girls in all domains on the computer-based items.

# Results from the PISA ICT student questionnaire

- **Boys** are more motivated on the CBAS test and they enjoy it more than girls,
- **Boys** have more experience with computer-based games, internet, games-type software that would be similar to the flash animations, video footage and interactive functions of CBAS items,
- **Boys** are more confident on ICT tasks and may therefore approach the test with more confidence,
- **Boys** use computers outside the home more than girls which may contribute to greater confidence in skills transference and greater familiarity with different keyboards, screens and software.

# ICT familiarity

- Effects of ICT familiarity on performance for boys and girls across countries.
- Point difference:

	<i>Advantage for ICT familiar girls</i>	<i>SE</i>	<i>Advantage for ICT familiar boys</i>	<i>SE</i>
Denmark	0.21	0.21	<b>0.46</b>	0.19
Iceland	0.07	0.10	<b>0.38</b>	0.16
Korea	<b>0.29</b>	0.13	<b>0.43</b>	0.16

Scale:

Mean=5

Stdev=2

\*Significant differences are displayed in bold





# ICT familiarity

- Correlations between ICT familiarity science scores on the CBAS and P&P tests are small to non-existent.

---

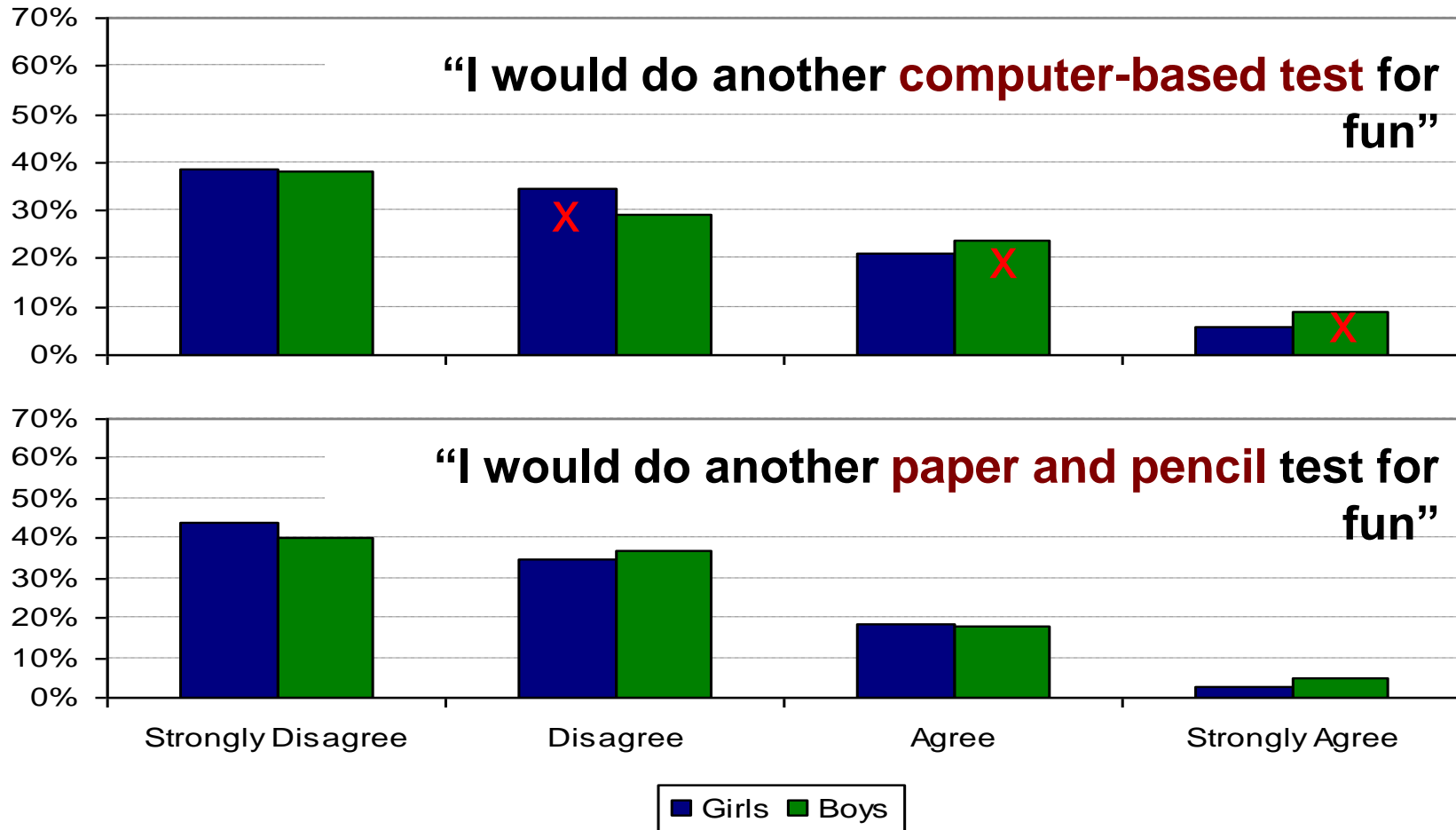
	<i>P&amp;P</i>		<i>CBAS</i>	
	<i>Girls</i>	<i>Boys</i>	<i>Girls</i>	<i>Boys</i>
Denmark	0.03	0.09	0.06	0.11
Iceland	-0.04	0.07	-0.02	0.10
Korea	0.05	0.07	0.07	0.10

---



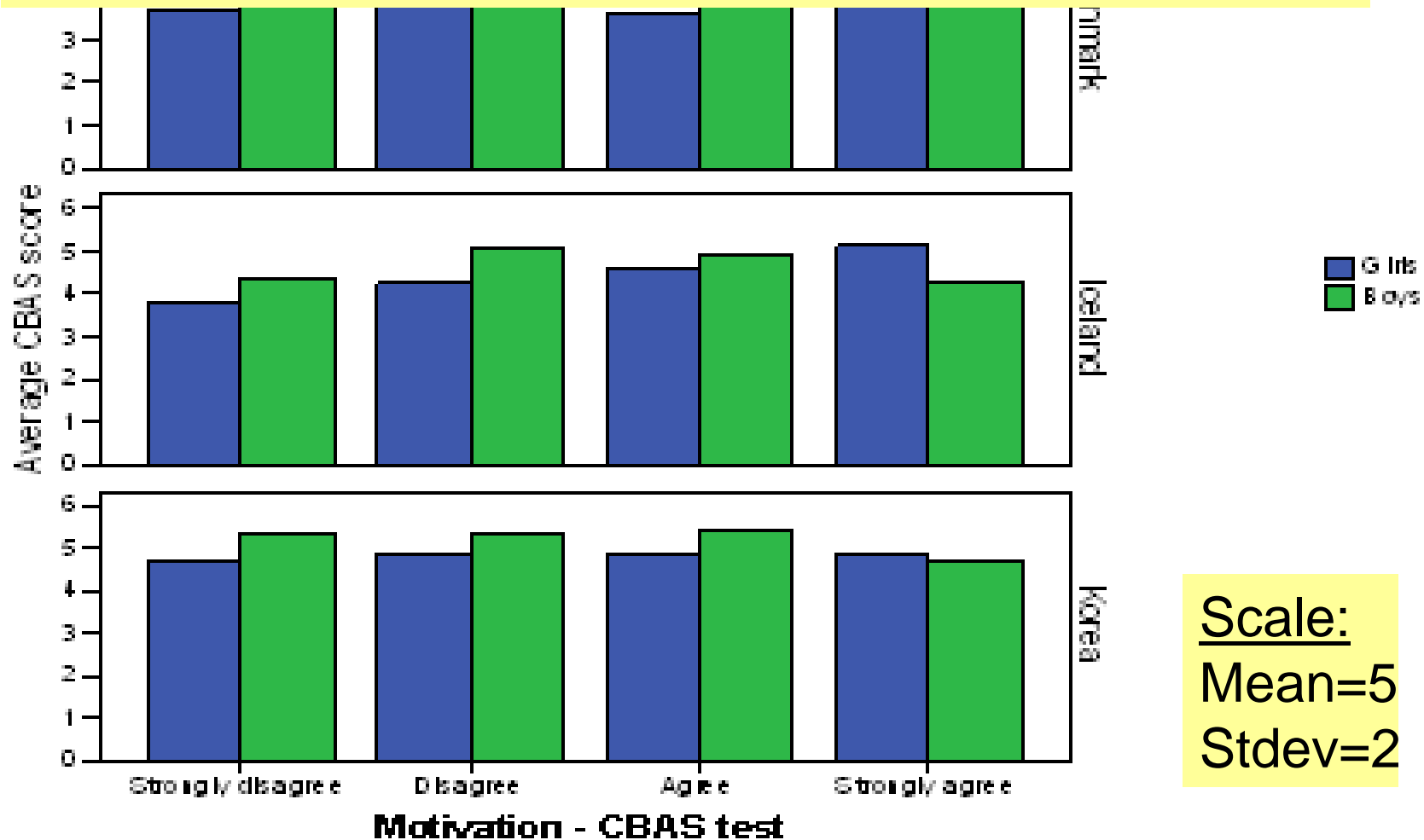
# Motivation

- Iceland: Boys are more motivated than girls on the CBAS test.



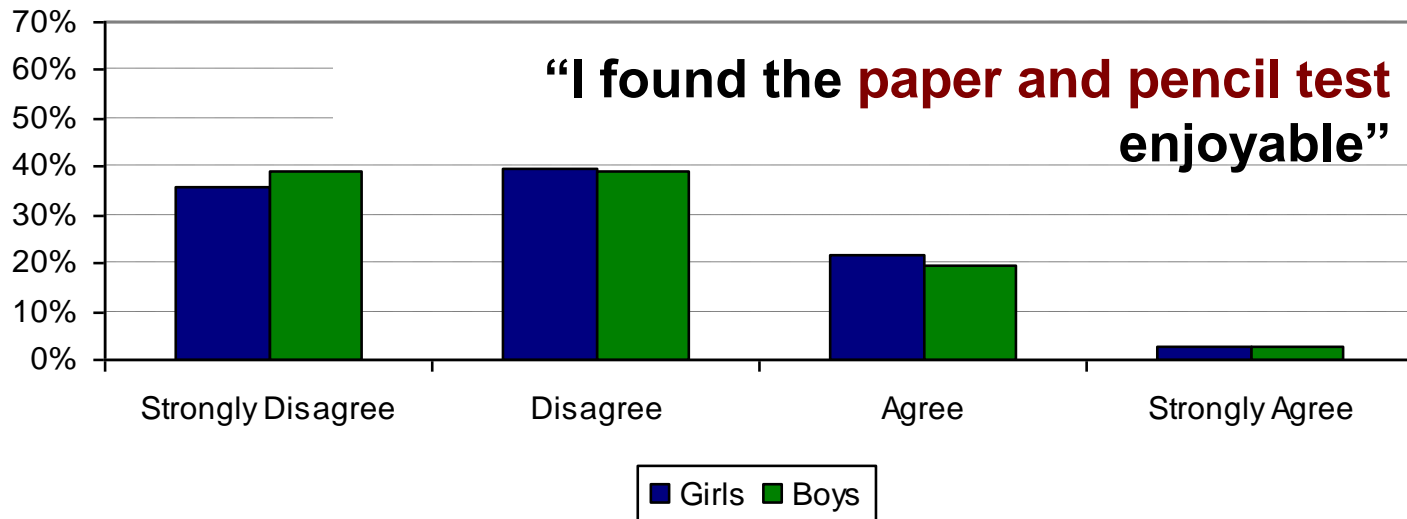
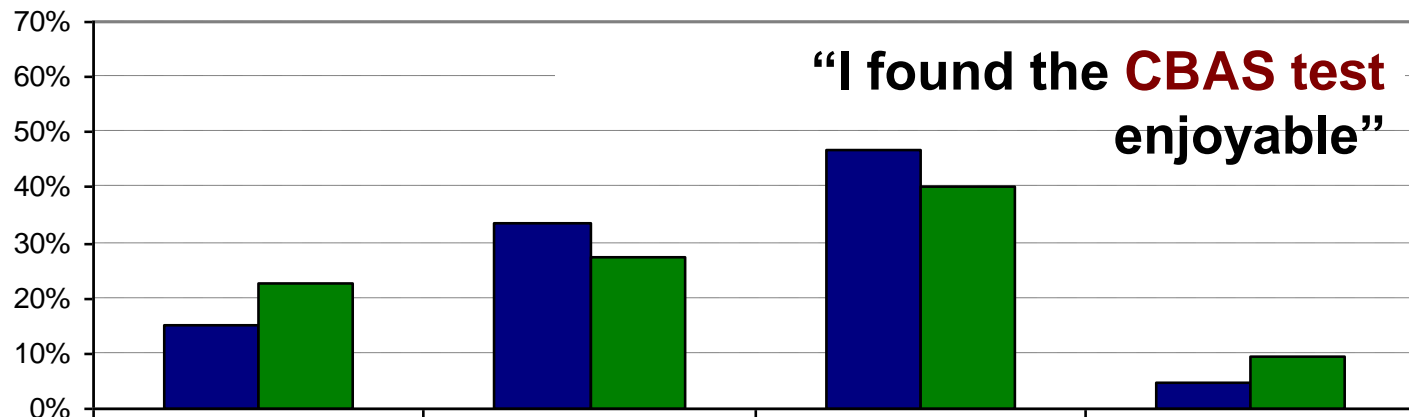
# Motivation and CBAS achievement

Overall, there is no association between achievement and test motivation.



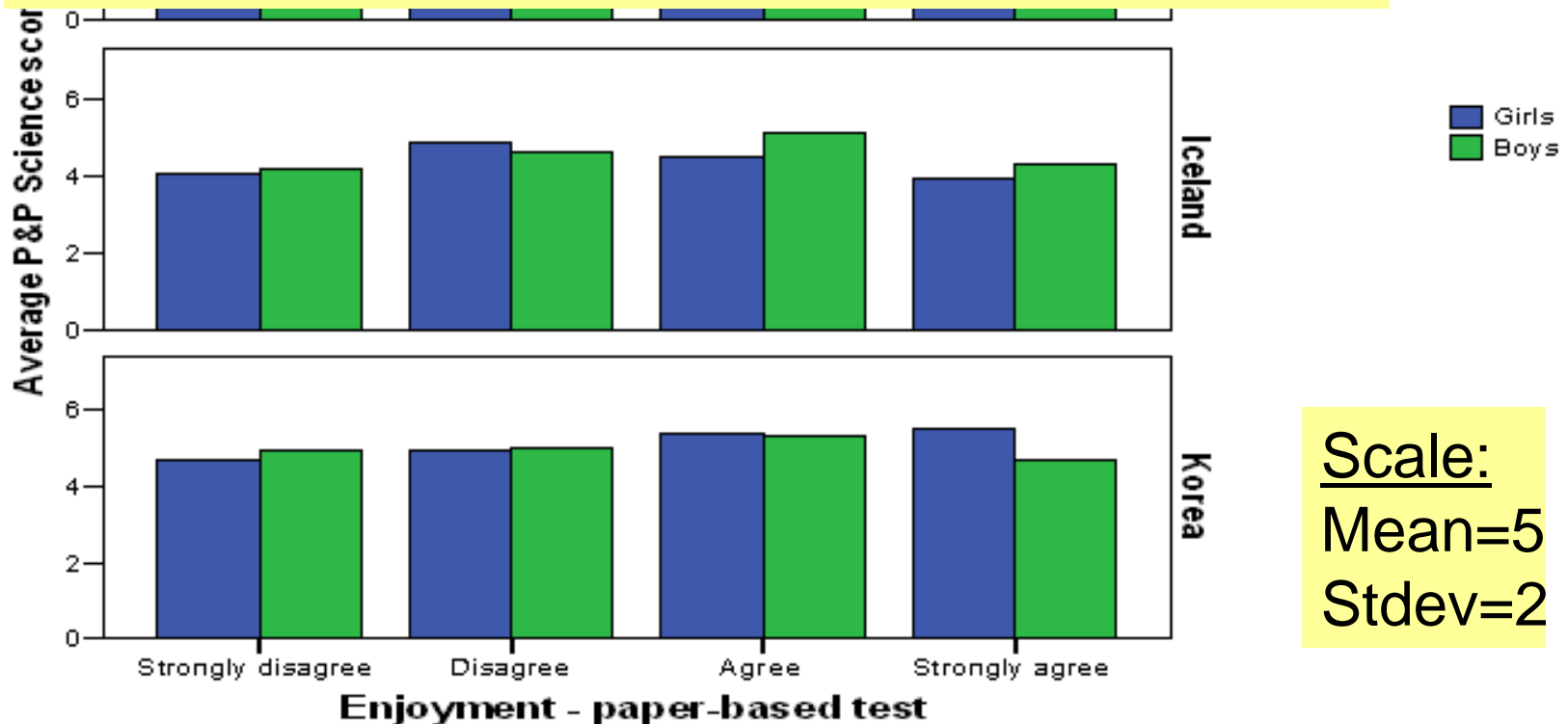
# Enjoyment

- Iceland: Girls enjoyed the P&P test slightly more than boys.



# Enjoyment and CBAS achievement

Overall, it is clearly not enjoyment that is influencing the better performance for boys on CBAS than PISA p&p.



Scale:  
Mean=5  
Stdev=2



# Results from the PISA ICT questionnaire and CBAS achievement

- Correlations do not reveal any significant associations between these questionnaire factors and CBAS achievement.
- Consequently, we must consider other factors that may influence performance.

# The PISA effort thermometer

## Identical in PISA P&P and CBAS

*How much effort did you invest?*

Please try to imagine an actual situation (at school or in some other context) that is highly important to you personally, so that you would try your very best and put in as much effort as you could to do well.

In this situation you would mark the highest value on the "effort thermometer", as shown below:

Compared to the situation you have just imagined, how much effort did you put into doing this test?

How much effort would you have invested if your marks from the test were going to be counted in your school marks?



- 10
- 9
- 8
- 7
- 6
- 5
- 4
- 3
- 2
- 1

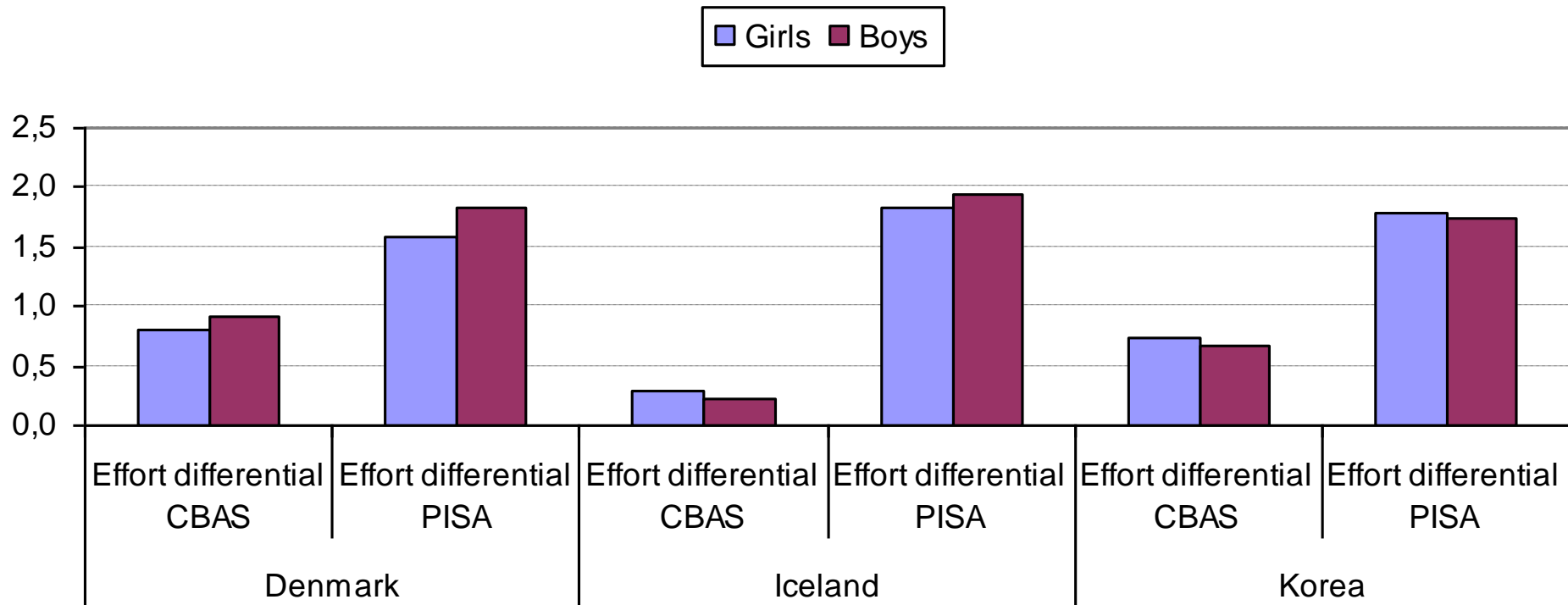
- 10
- 9
- 8
- 7
- 6
- 5
- 4
- 3
- 2
- 1

- 10
- 9
- 8
- 7
- 6
- 5
- 4
- 3
- 2
- 1



# Effort

- Gender difference in effort in P&P and CBAS



Effort differential can range from 0 to 9 points





# Effort

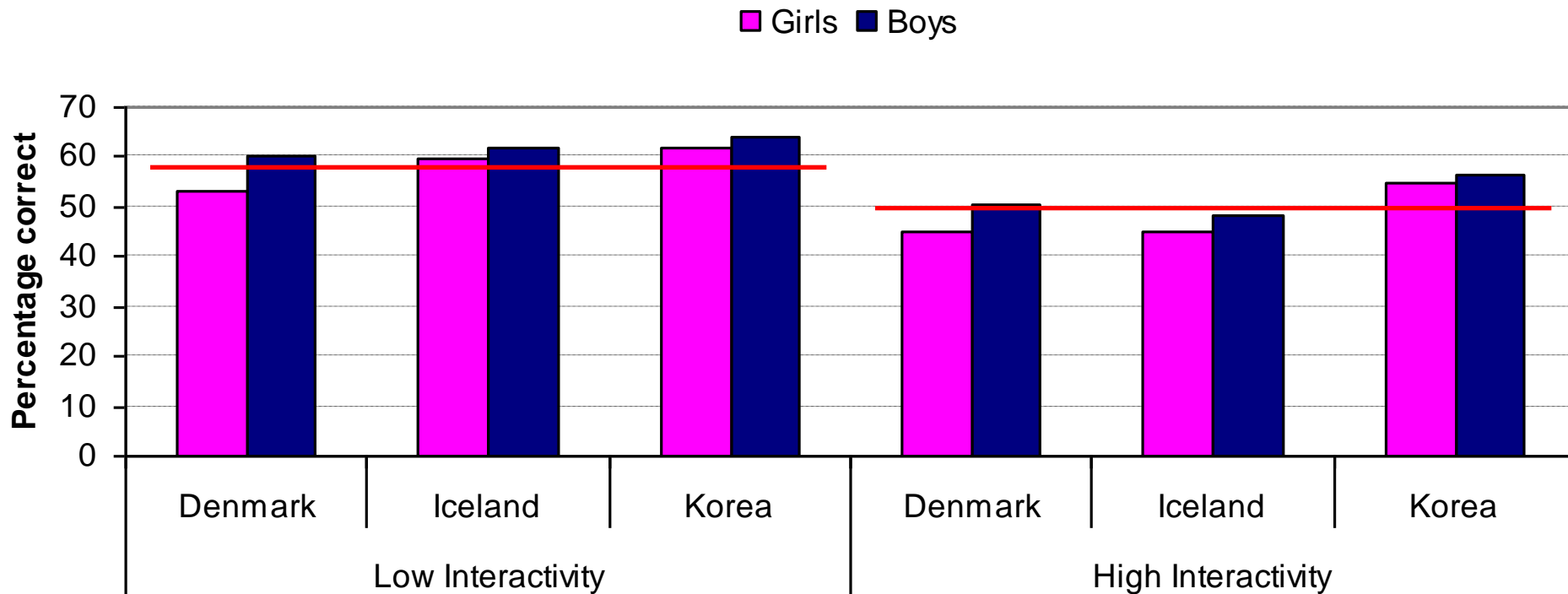
- ‘Which test did you put more effort into?’
  - A higher proportion of **boys** than girls in all three countries report that they put more effort into the CBAS test
  - A higher proportion of **girls** than boys in Denmark and Iceland report that they put more effort into the P&P test

# CBAS interactivity

- Level of interactivity
  - Low interactivity
    - Example: “Assembly Line”, a short video of an automated car assembly line is shown and the question is about the role of robots in society.
  - High interactivity
    - Example: “Plant Growth”, the student is required to move buttons up and down a scale, performing a limited number of experimental trials on optimal temperature and soil acidity levels for growing wheat and asked to report which levels give the best outcome.

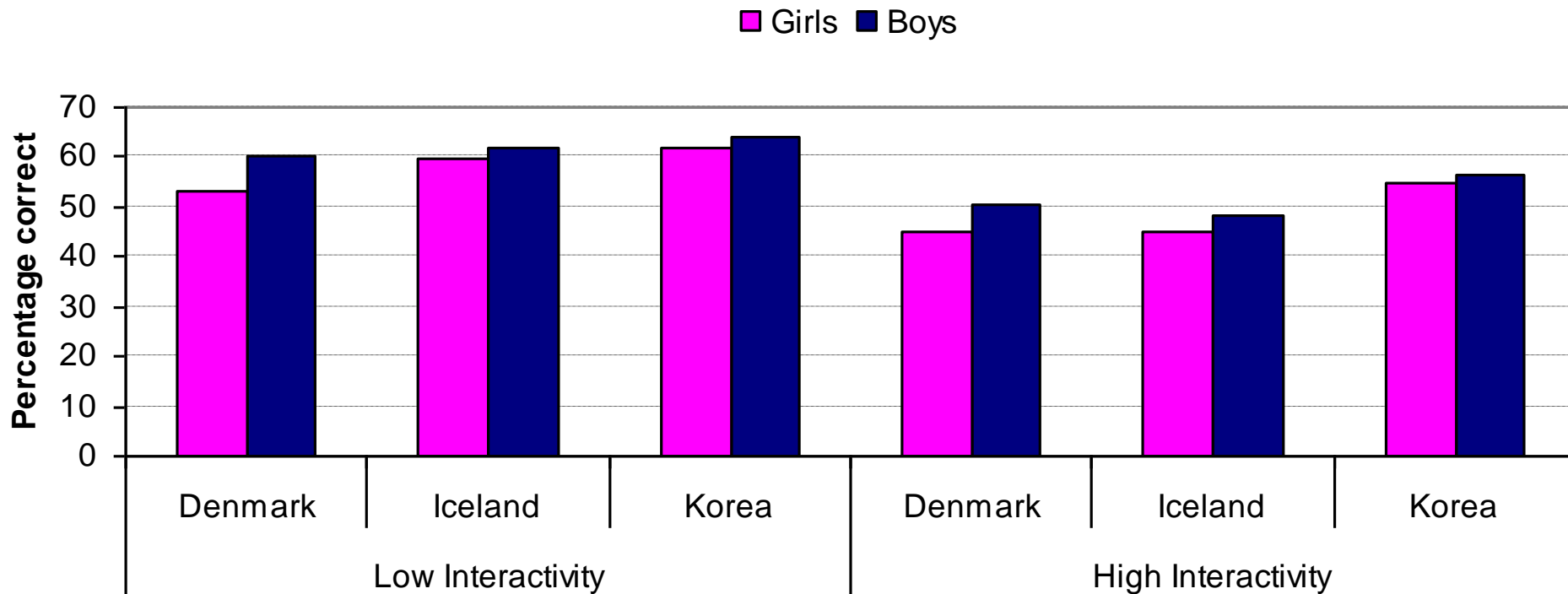
# CBAS interactivity

- For both genders and in all three countries, high interactivity items were more difficult than low interactivity items.



# CBAS interactivity

- Not a consistent gender gap.
- The types of ICT skills necessary to answer the CBAS questions are low level skills and well within the grasp of most 15 year old boys and girls.

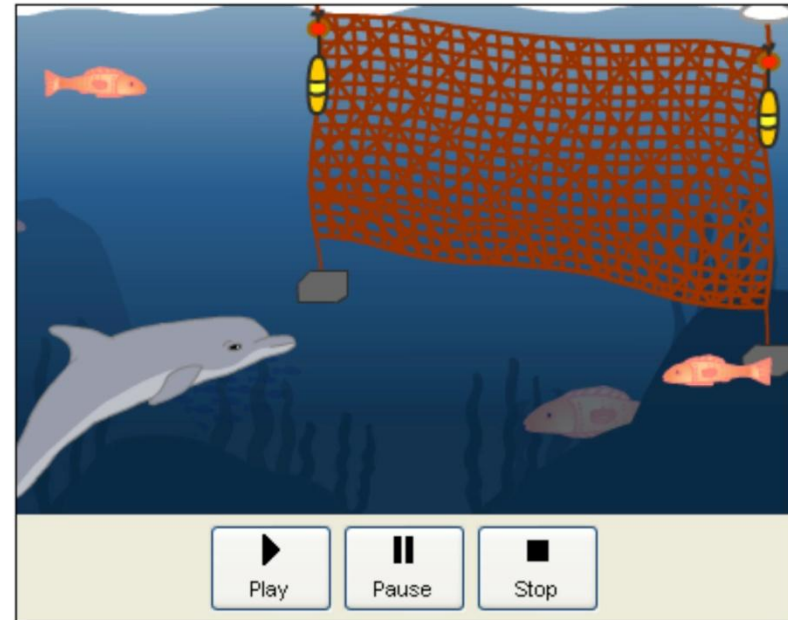
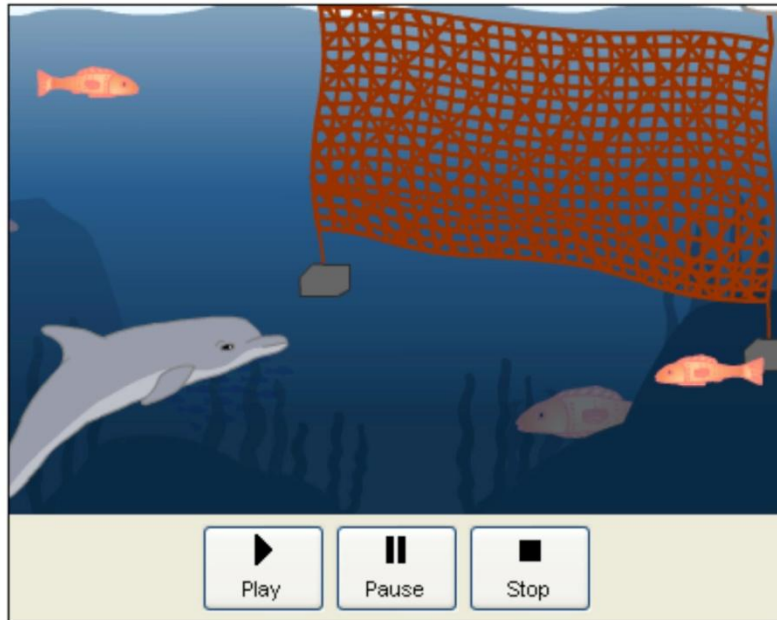


# Reading load

- Word per CBAS item were counted
  - The number of words in the stimulus, embedded in the image, in the question stem and in the multiple choice response options.
  - Items grouped:
    - Low reading load
    - Medium reading load
    - High reading load

## Question 9: Echolocation

A major hazard for dolphins is getting trapped in fishing nets. Scientists are testing warning devices attached to nets to deter dolphins. The devices send out a sound signal every few seconds. The number of dolphins touching a net was counted over three weeks. In the first week the warning devices were not attached to the net (left). In weeks 2 and 3 the devices were attached (right).

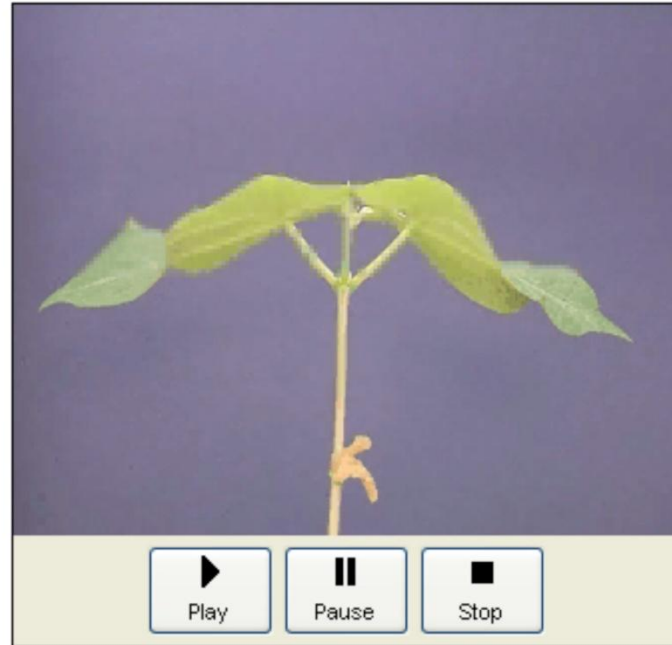


What was the purpose of week 1 of this experiment?

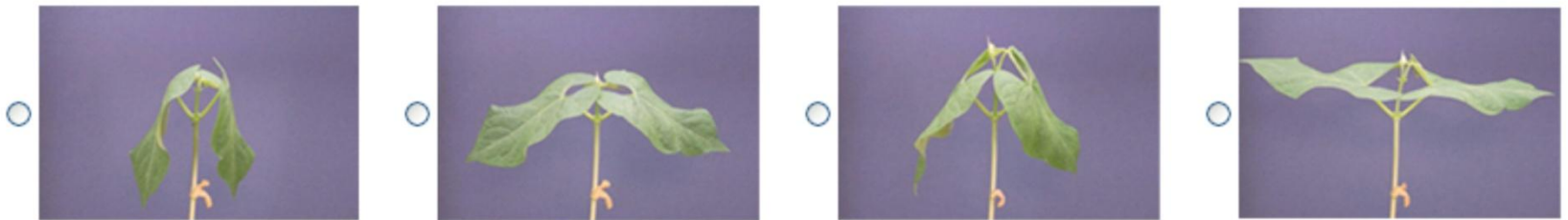
- To test whether sounds attract dolphins.
- To observe the behaviour of dolphins near warning devices.
- To collect information for comparison purposes.
- To have data on the number of fish eaten.

## Question 2: Bean Leaves

The time lapse movie shows the movements of the leaves of a bean plant at intervals of one hour over a 48 hour period. The plant is provided with adequate water throughout this period.

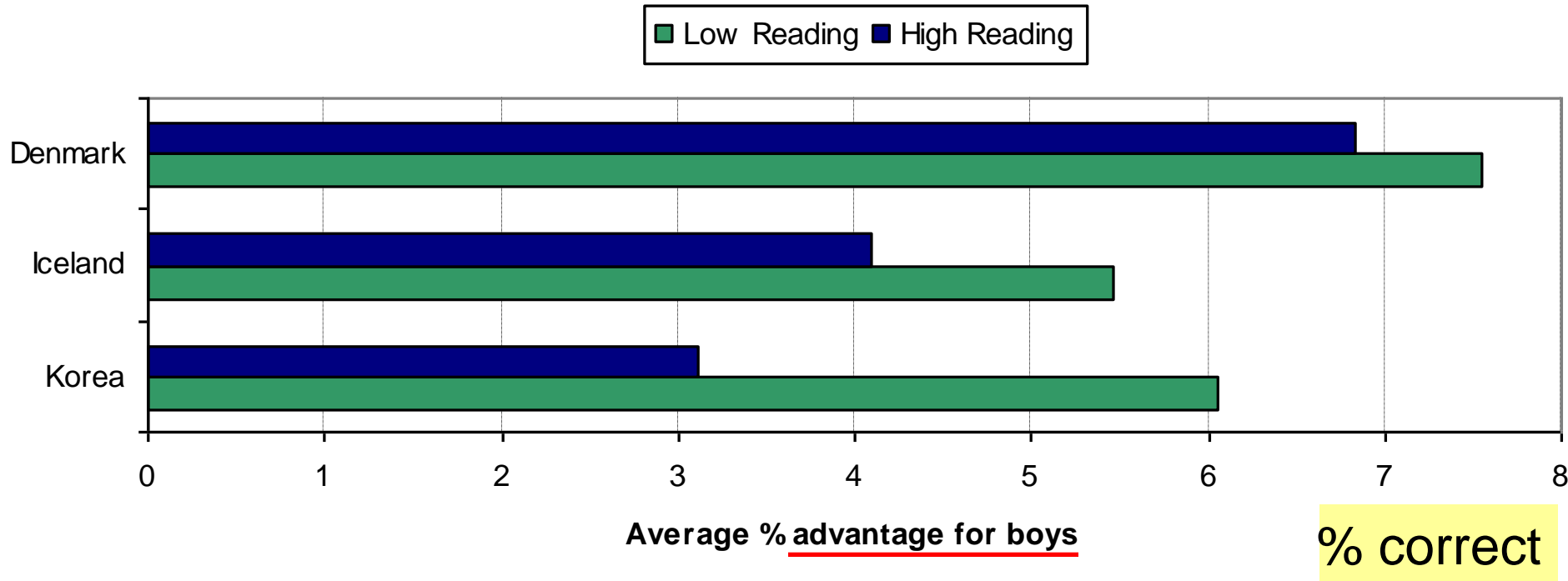


Which picture shows the bean plant when the light intensity is greatest?



# Reading load

- Boys outperform girls on CBAS items of both low and high reading load.
- The effect is consistently stronger for LOW reading load items.





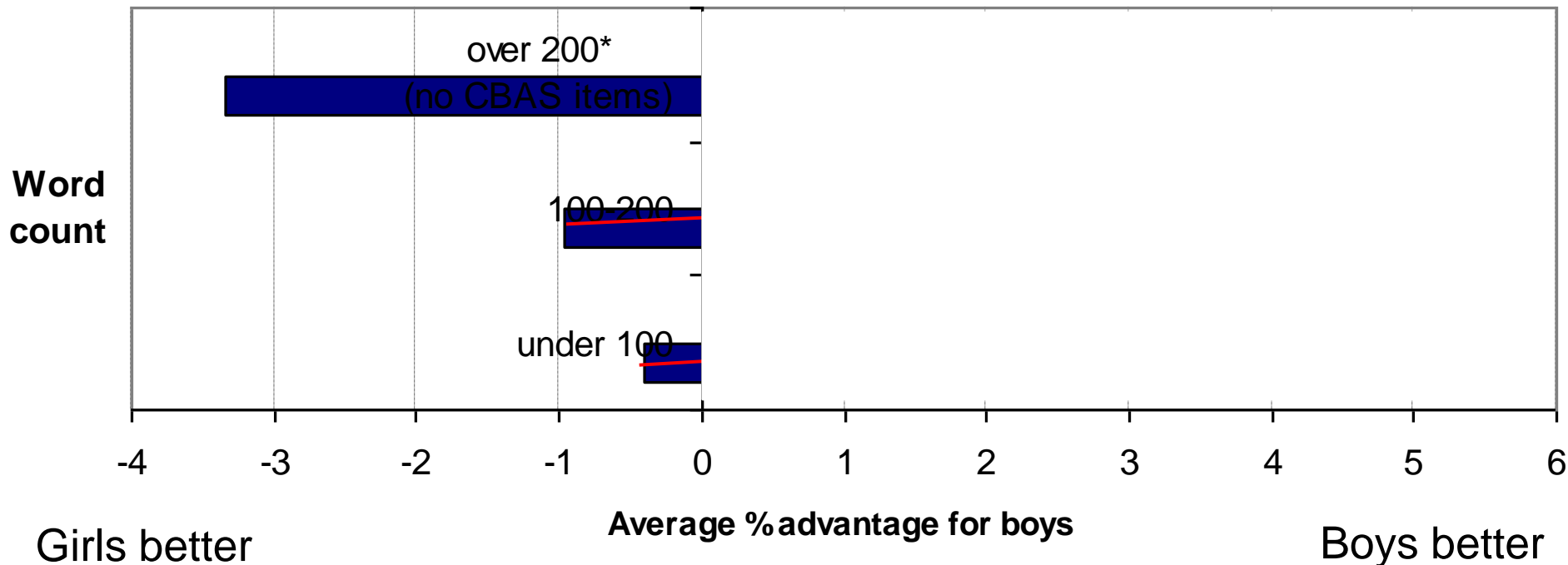
# Reading load

## Comparing P&P and CBAS

- The higher reading load on P&P science items appears to disadvantage boys.
- Gender difference in performance by item reading load in P&P and CBAS, in Iceland:

■ PISA Science items ■ CBAS items

% correct



# Reading load

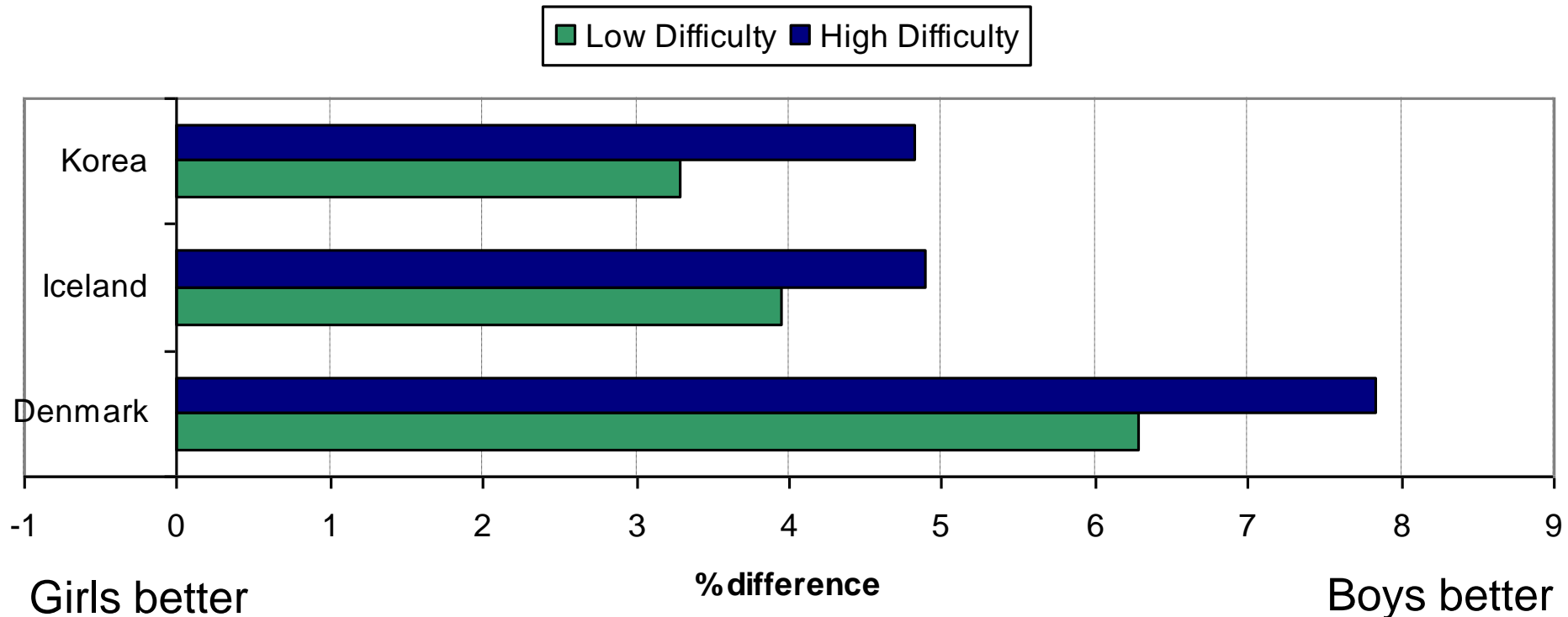
- Gender difference in P&P favoring girls in Iceland is removed in items of low reading load.
- Boys may be disadvantaged by the length of some of the longer P&P items.
- However, reading load cannot explain fully the advantage for boys on CBAS items.

# Item difficulty

- Overall the CBAS items were easier than the paper-based items.
  - % correct is much higher for CBAS than for P&P in the three countries.
- There may be an interaction between item difficulty and gender.
  - Are the easier CBAS items comparatively easier for boys than girls?
  - Do P&P science items **fatigue** the boys and encourages them to 'give up' more than girls?

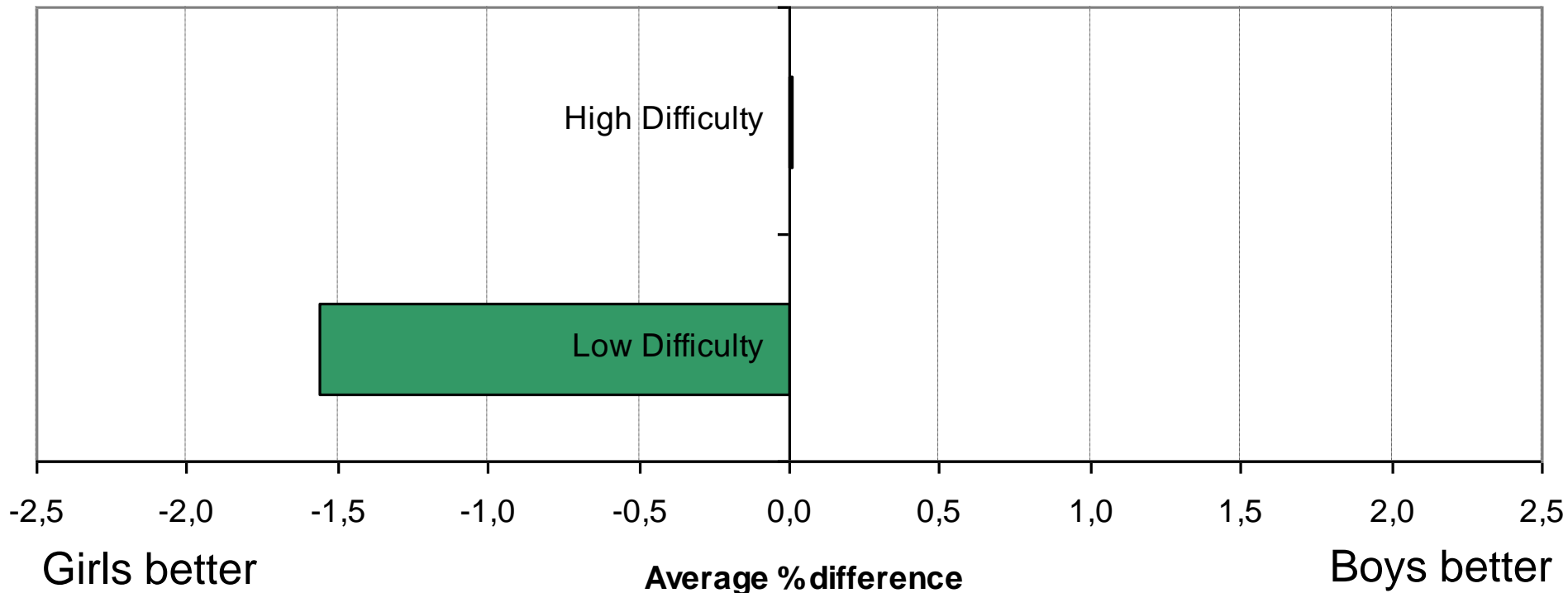
# Item difficulty

- Boys' performance advantage over girls is greater for the high difficulty CBAS items than on the low difficulty CBAS items.
- Across all three countries.



# Item difficulty

- Compare CBAS with the PISA P&P test.
  - Icelandic boys do comparatively better on the more difficult P&P science items.
  - However, girls only outperform boys on low difficulty items.



# Item difficulty

- Boys do comparatively better on more difficult P&P and CBAS items.
- A plausible explanation of why boys outperform girls in CBAS and not in PISA P&P may be that overall the CBAS items were easier than the P&P items.
- **It is possible that the sheer number of P&P science items of low difficulty may fatigue boys and encourages them to ‘give up’ more so than girls.**
  1. In CBAS boys assert more effort than girls but in P&P girls assert more effort
  2. In CBAS, boys’ advantage is higher for high difficulty than low difficulty items
  3. In P&P, girls’ advantage is only present for low difficulty items
  4. The CBAS test took 1 hour but the PISA P&P took 2 hours

# Was CBAS a test for boys?

- There are 9 videos over 5 units showing boys performing certain activities where they are specifically named in the text and sighted in the video footage.
- There are a further two items in one unit where a boy is named and illustrated as the principal actor in the animated scene.
- On the girls' side, there are no items showing girls performing activities by video and only one item that refers to a girl by name.
- The lack of girls in the items may lead to a lower level of engagement with the test for the girls and a consequently lower level of performance.

- Identifying with the gender of the actor may be important for engagement but is it important for performance?



# Summary

- Boys in all three countries outperformed girls when a science test was presented via computer.
- The gender differences in performance cannot easily be explained by motivation, computer item interactivity, enjoyment or familiarity with computers.
- Neither can item-by-item analysis interactivity.
- The increase in boys' performance in CBAS may, at least to some extent, be explained by lower reading load and by boys' greater test fatigue on numerous low difficulty items.
- Can the rest be explained by a gender bias in the test items themselves?
  - The FT item pool for CBAS was not gender biased.
  - READING LITERACY IN DIGITAL MEDIA
    - We're looking forward to analysing the PISA 2009 ERA results, with a wider difficulty range and balanced gender representation in items.

