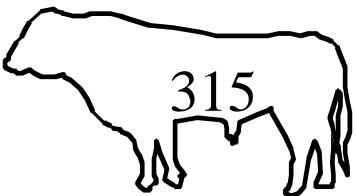# Estimating Breeding Values

- ## Principle
  - how is it estimated?

- ## Properties
  - Accuracy
  - Variance
  - Prediction Error

- ## Selection Response
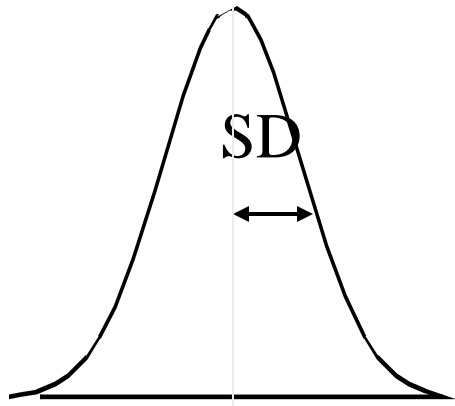 - select on EBV

GENE422/522
Lecture 2

Observed	Phen. Dev.	Genetic Value	Env. Effects
Underlying Model	P	G	E
315	+15	+ 3	+12
307	+ 7	+ 7	0
303	+ 3	- 2	+ 5
294	- 6	+ 8	-14
287	-13	- 6	- 7
Standard Deviation	12	6	10

	Phen. Dev.	Genetic Value	Env. Effects
	P	G	E
	+15	+3	+12
	+7	+7	0
	+3	-2	+5
	-6	+8	-14
	-13	-6	-7
SD	12	6	10

Genetic variance

Environmental variance

SD
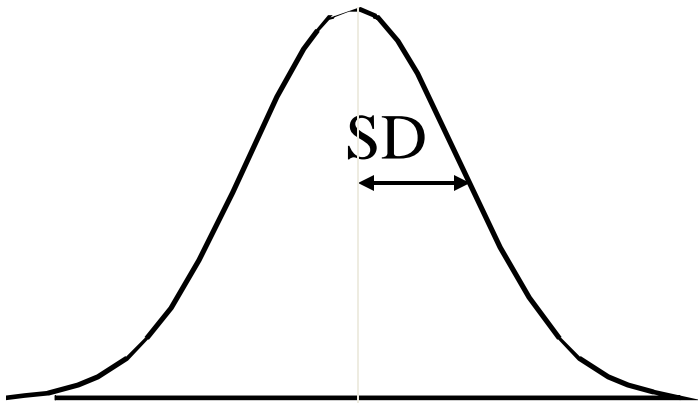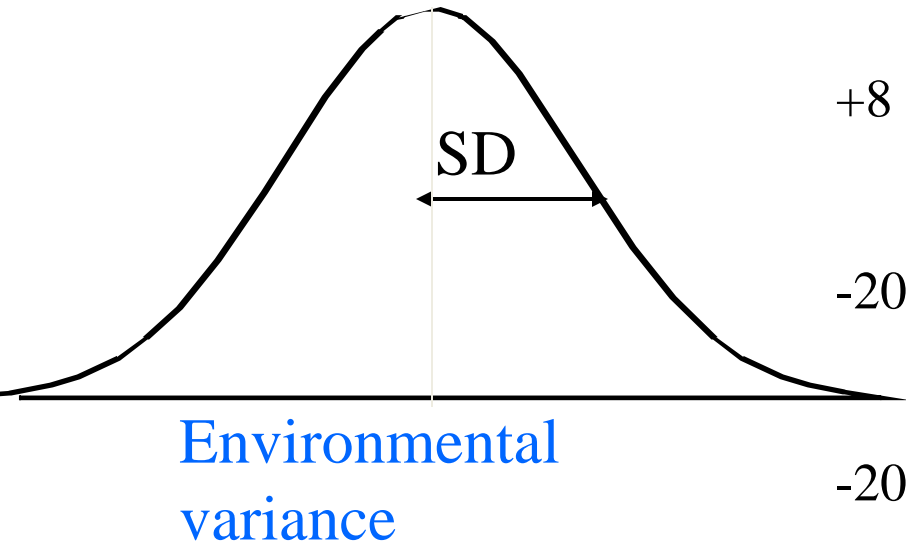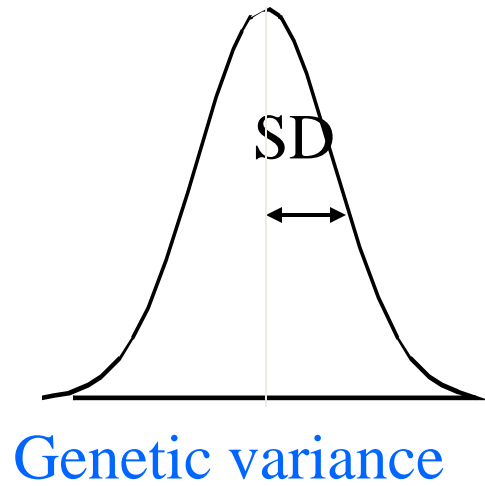
	Phen. Dev.	Genetic Value	Env. Effects
	P	G	E
	+27	+3	+24
	+7	+7	0
	+8	-2	+10
	-20	+8	-28
	-20	-6	-14
SD	21	6	20

Genetic variance

Environmental variance

SD

Components of variation (phenotypic)
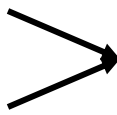
Variation: → **Genetic**

in phenotype → **Environmental**

$$\begin{matrix} V_A \\ V_E \end{matrix} \Big> V_P$$
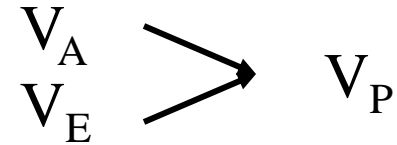
Heritability is proportion of variation that is (additive) genetic

$$h^2 = V_A / V_P$$

When do we use heritability?

- **Predicting Breeding Values from phenotypic differences**
  - **Weighting the value of phenotypic information on various relatives**
  - **Breeding values are used to rank and select animals in order to achieve genetic improvement**
- **Predict the result of selection**

IN REAL LIFE

We can only see P   We want to estimate A.

need a model

Phenotype   =   Genetics   + Environment
(Phenotypic difference   is due to   Additive Genetic Effect + Residual Effect)

P   =   A   + E

Estimated Breeding Value = EBV = $h^2.P$

(if based on own performance)

…is that part of the phenotypic difference

that you believe is due to breeding value!

The larger heritability, the more of phenotypic differences are due to (additive) genetic value
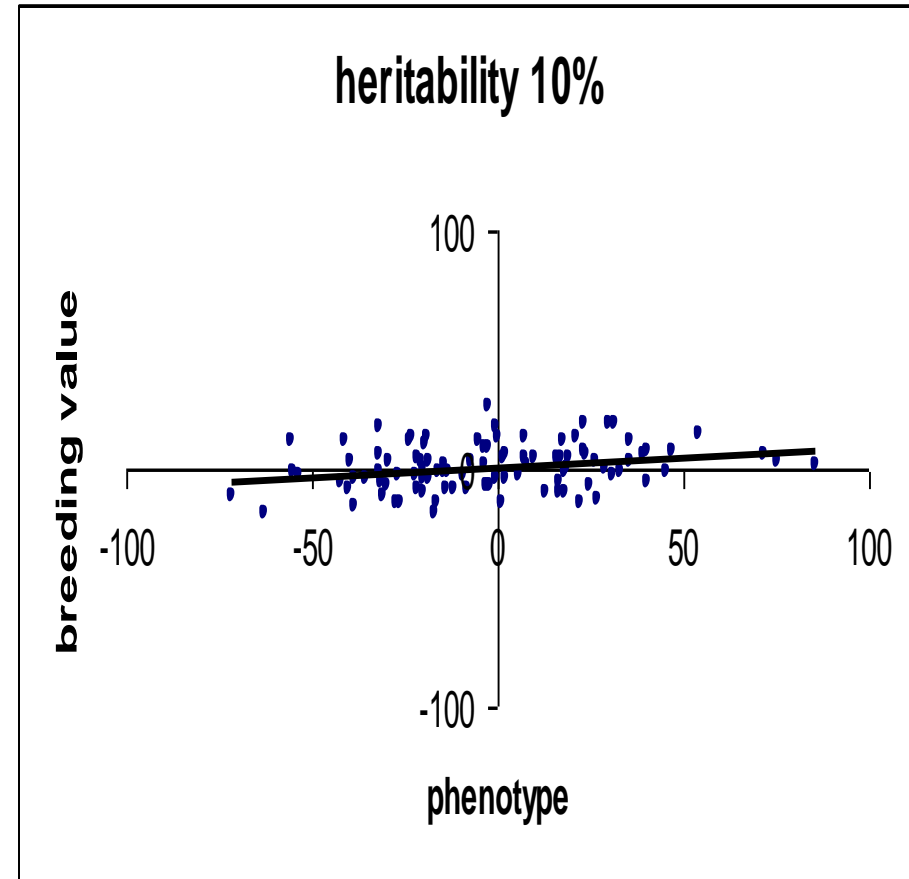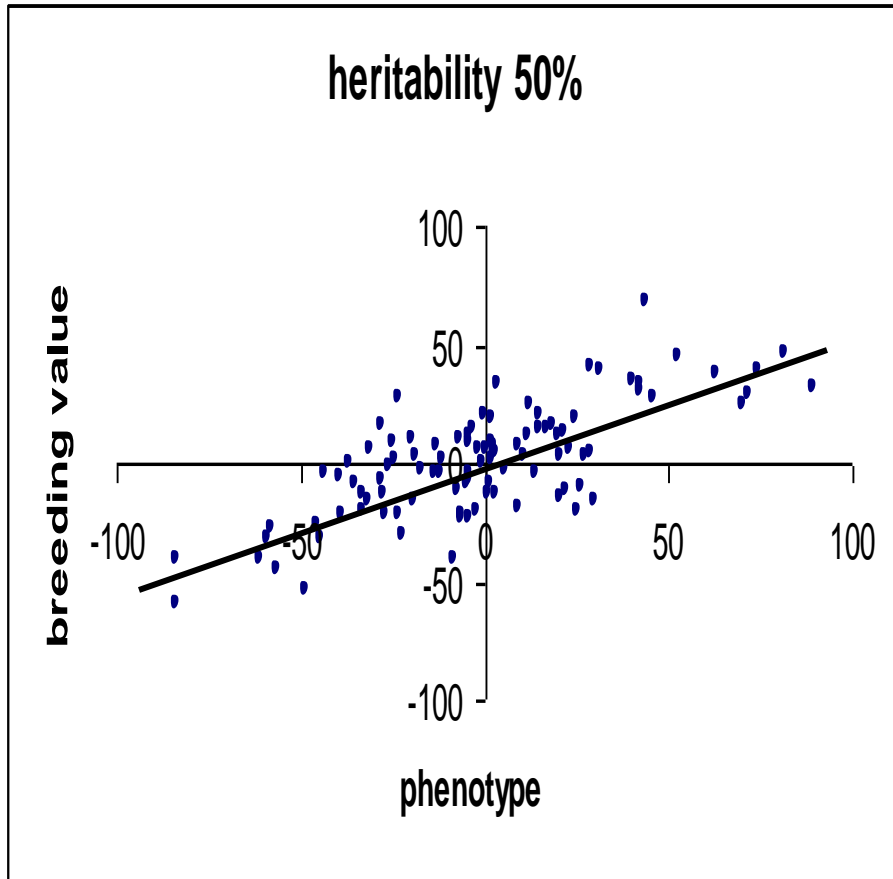
Slope is equal to heritability

Principle of estimating breeding value

- Based on regression

- Predicting difference in breeding value from phenotypic differences

Correction for fixed effects

- We use a phenotypic deviation from a contemporary mean

  - population mean
  - herd or flock mean
  - mean of all animals born in 2009 in herd A
  - Management group
  - All males who are born in November in the flock

- is a way to correct for non-genetic effects

Example of contemporary groups

Bull	YW	Herd Ave	P	EBV h²=40%
Bert	330	300	+30	+12
Folssy	300	260	+40	+16

•Note that the ranking in P differs from ranking in EBV
•Need to assume that herd differences are non-genetic

One fixed effect:
A model for fixed effect correction

$$Y = \mu + BT + A + E$$

Birth Type	Mean Weaning Weight Kg.
Single	25
Twin	23
Triple	21

Strategy: Express phenotypes as deviations from their group means.

A 25 Kg. twin: 25 – 23 = +2Kg.
A 25 Kg triple: 25 – 21 = +4Kg.

Several fixed effects:
A model for fixed effect correction

$$Y = \mu + BT + Herd + A + E$$

Additive model

Birth Type	Mean Weaning Weight Kg.		
	Herd A	Herd B	Overall
Single	23	27	25
Twin	21	25	23
Triple	19	23	21
Overall	21	25	23

Strategy:
Correct phenotype for each class effect separately.

25 Kg. Twin in herd A: 25 - 23 -   0 – (-2) = +4Kg.
25 Kg Triple in herd B: 25 – 23 –(-2) – (+2) = +2Kg.

Correction for birth type   Correction for herd

A model for fixed effect correction

$$Y = \mu + BT * Herd + A + E$$

Birth Type	Mean Weaning Weight Kg.		
	Herd A	Herd B	Overall
Single	23	27	25
Twin	21	25	23
Triple	19	23	21
Overall	21	25	23

Strategy: Express phenotypes as deviations from their group means.

25 Kg. Twin in herd A: 25 - 21 = +4Kg.
25 Kg Triple in herd B: 25 – 23 = +2Kg.

A model for fixed (continuous)effect correction

$$Y = \mu + b.age + A + E$$

	Age (mo)	Mean Weaning Weight Kg.
A	11	280
B	13	295

Strategy: Correct phenotypes to a 'constant' age

Need: the growth per month (= 12kg/mo) → =b

Corrected weights

A: 280 + 12 = 292 kg

B: 295 – 12 = 283 kg

Conclusion about fixed effects

- Need to correct for Fixed Effect to avoid bias in EBV, taking out non-genetic effects

- In real life there may be several FE

- They may not all be balanced

- Need a (statistical) model to correct for FE (see Lecture 4, 5)

Further EBV properties

1. Accuracy of EBV's
2. Variance of EBV
3. Prediction Error
4. Predicting Response

EBV properties:

1.   Accuracy of EBV's

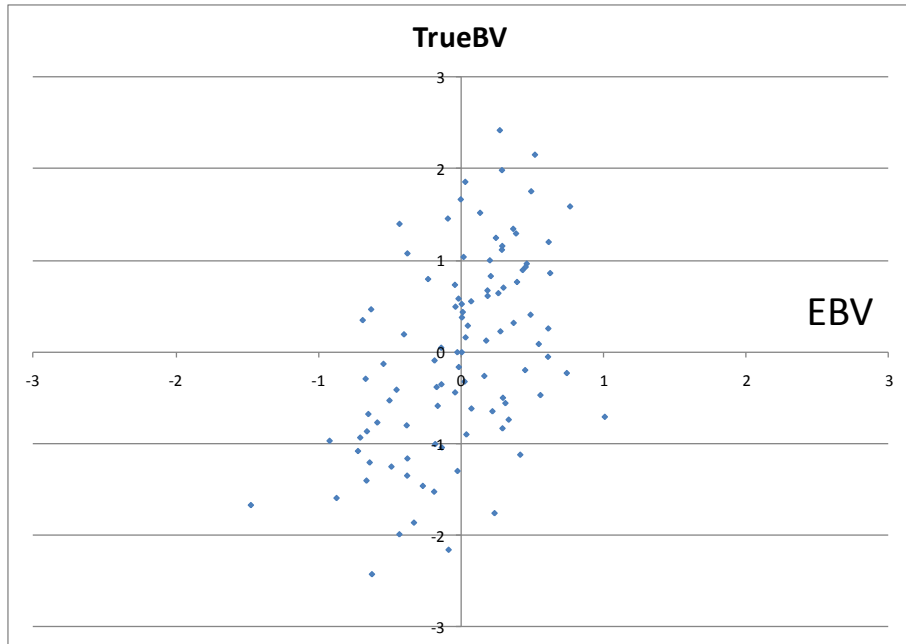Symbol: r_{IA} =   correlation

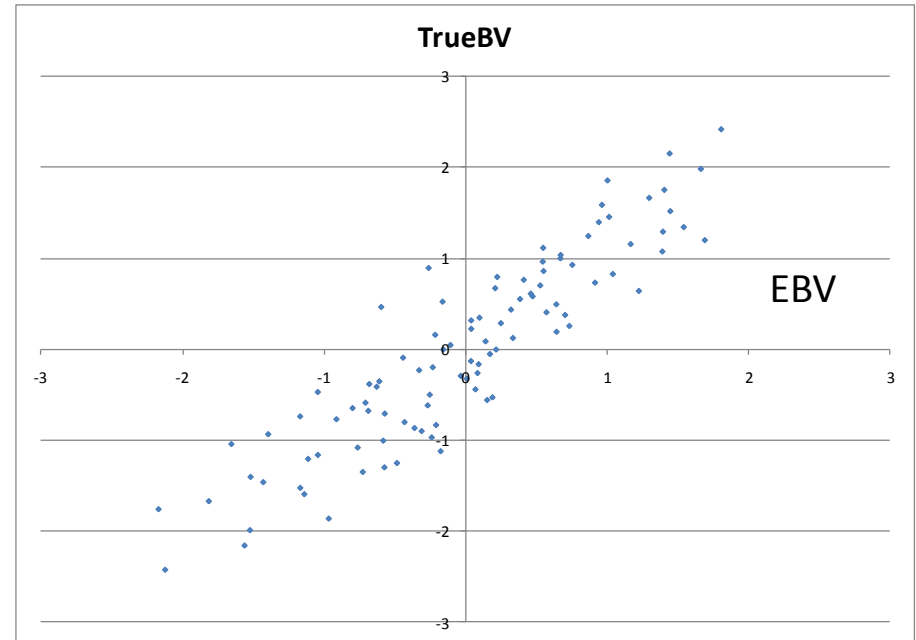between estimated and true breeding value

EBV = (Index = "I")

True BV = ("A")

(sometimes we use r_{IH} instead of r_{IA} where H is for a

multi trait breeding value $\rightarrow$ aggregate genotype)

accuracy of EBV = correlation with True BV

Accuracy = 45%

Accuracy = 90%

Accuracy of EBV's – some examples

| Information Used	Index (=EBV)	Accuracy
Own Performance	$EBV = h^2 P$	h
Progeny	$EBV_{sire} = 2n/(n+\alpha).$ ProgMean	$\sqrt{n/(n+\alpha)}.$
Generic	$EBV = b_1 P_1 + b_2 P_2 + ...$	Use Selection Index Theory

Look at the idea, not at the maths!

Examples of accuracies

	h²=0.1	h²=0.3	
own information	0.32	0.55	*equal to sqrt h²*
mean of 10 half sib	0.23	0.33	
mean of 1000 half-sibs	0.49	0.50	*max is sqrt 0.25 = 0.5*
mean of 1000 full-sibs	0.70	0.71	*max is sqrt 0.5 = 0.71*
mean of 100 progeny	0.85	0.94	*max approaches 1.0*

Accuracies of animal increase as they get older → (more info)

Accuracy of predicting a breeding value
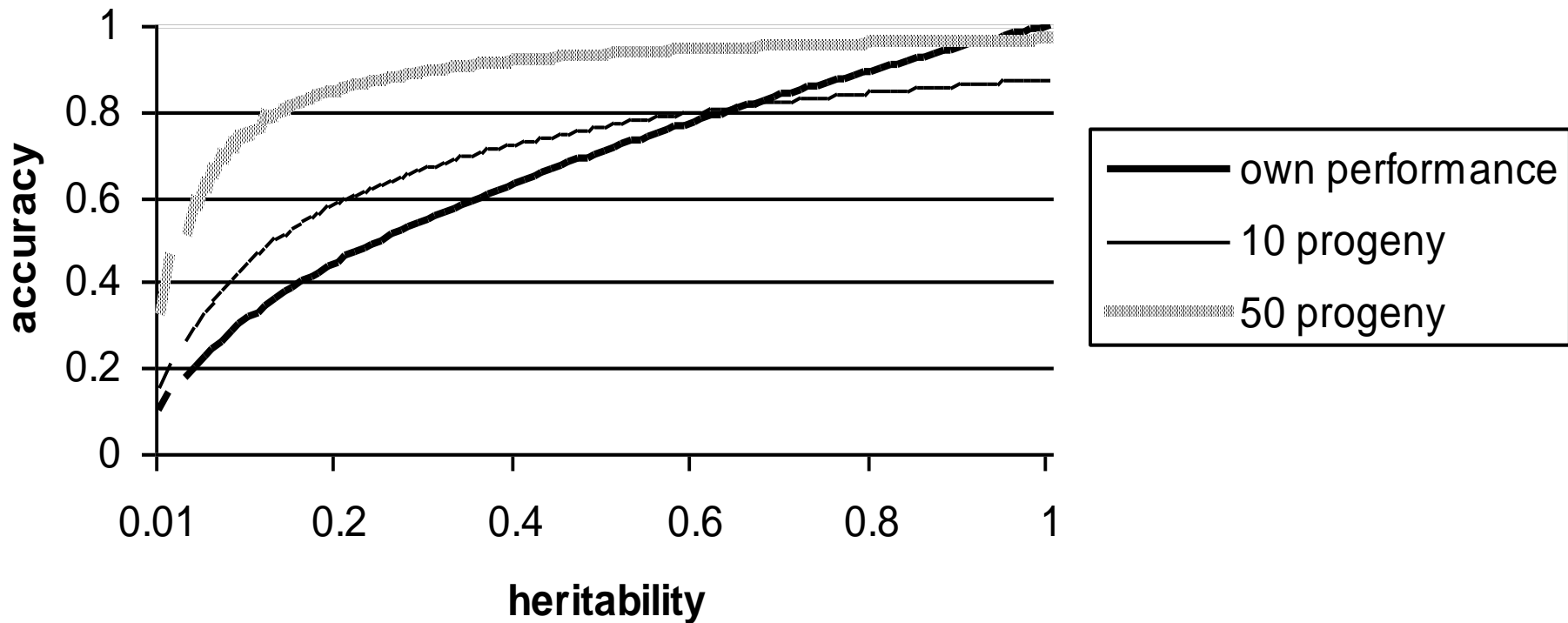## - increases as an animal gets older
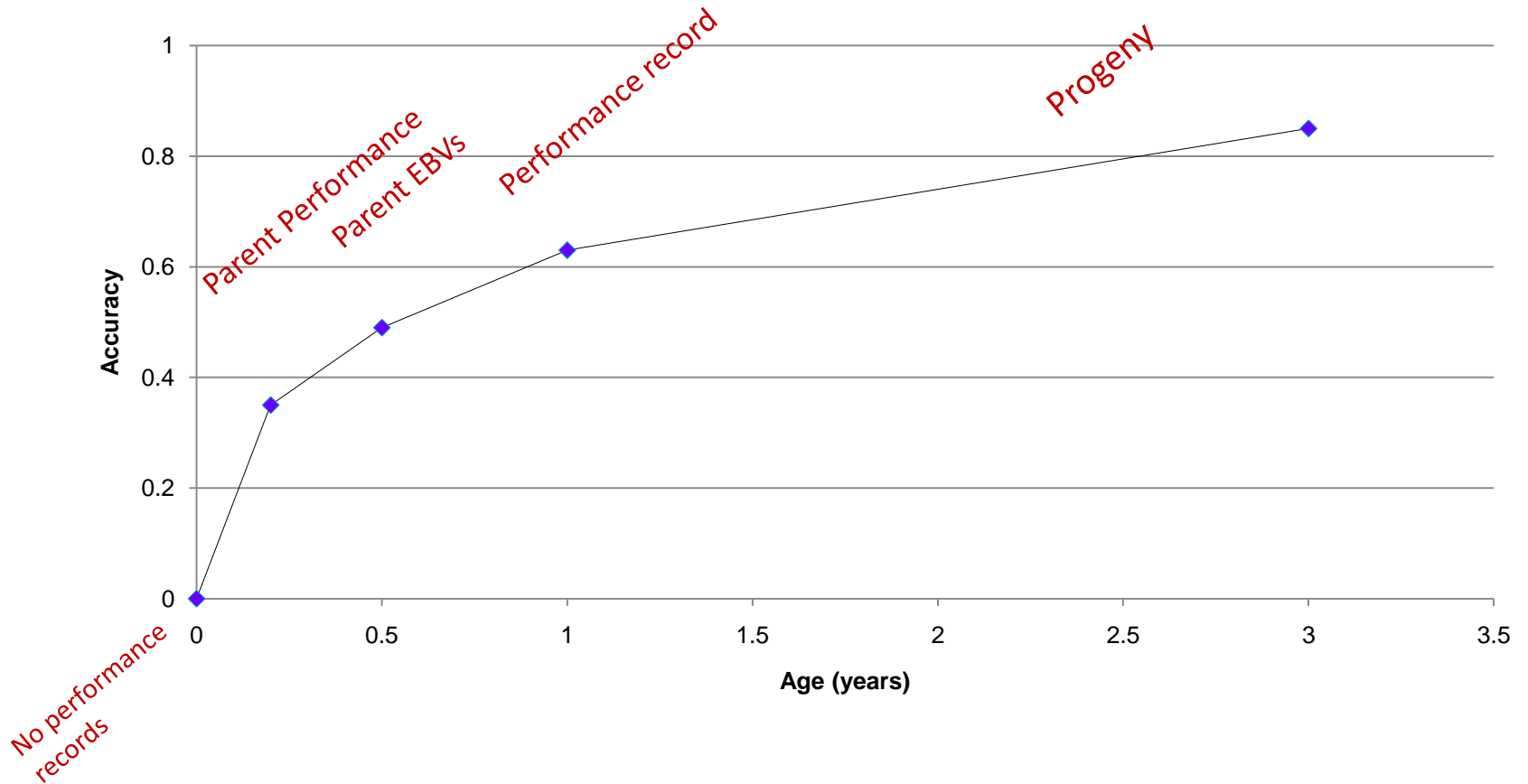
Assumed heritability = 25%

Example: Accuracy of progeny test

	Nr of progeny		
	5	50	100
h^2			
0.1	0.34	0.75	0.85
0.5	0.91	0.99	0.995

Progeny test is the only way to get a very high (near 1) accuracy

But more progeny are needed when heritability is low

Accuracy of own performance vs progeny test

Progeny test gives usually more accurate EBV then own performance

EBV properties:
2.Variance of EBV's
- how much they differ

- $Var(EBV) = r^2_{IA} V_A$
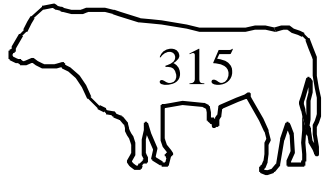
- $SD(EBV) = r_{IA} \sigma_A = $ Standard deviation of EBVs

Simply a function of accuracy

Note the extremes of Var(EBV) if $r_{IA} = 0$ vs $r_{IA} = 1$

Observed	Phen. Dev.	Genetic Value	Env. Effects	EBV (h^2=0.3)
	P	G	E	$r_{IA = 0.55}$
315	+15	+ 3	+12	4.5
307	+ 7	+ 7	0	2.1
303	+ 3	- 2	+ 5	0.9
294	- 6	+ 8	-14	-1.8
287	-13	- 6	- 7	-3.9
SD	11	6	10	3.3
		V_A		V_{EBV}

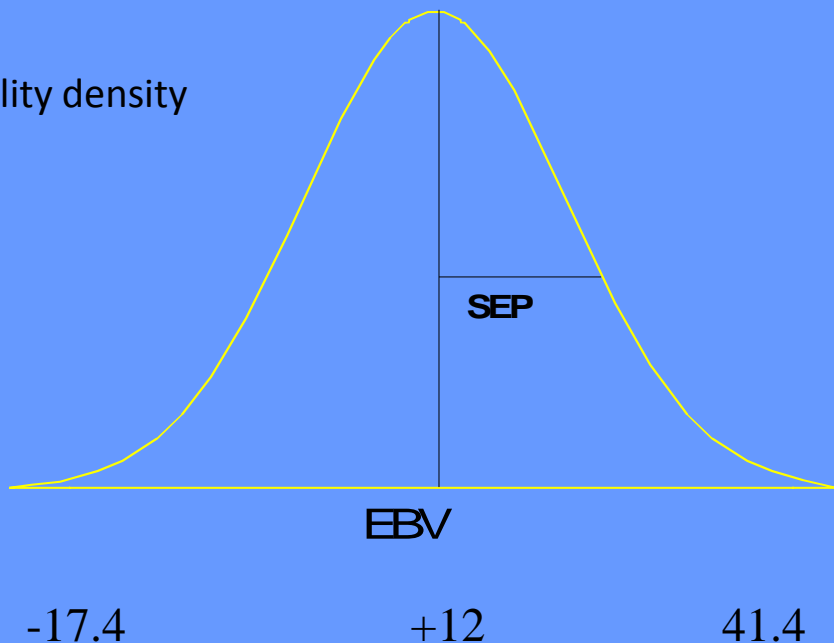| Phen. Dev.	Genetic Value	Env. Effects	EBV ($h^2=0.09$)
P	G	E	$r_{IA = 0.32}$
+27	+3	+24	2.4
+7	+7	0	0.6
+8	-2	+10	0.7
-20	+8	-28	-1.8
-20	-6	-14	-1.8
SD 21	6	20	1.8
	V_A		V_{EBV}

EBV properties:
Prediction Error Variance
- how much they still may change

- $PEV = var(EBV-TBV) = (1-r^2_{IA})V_A$ — Prediction Error Variance

- $SEP = sqrt(PEV) = \sqrt{(1-r^2_{IA})}\sigma_A$ — Standard Error of Prediction

Probability density
of TBV)

SEP

EBV

-17.4   +12   41.4

Let σ_a = 19 Kg, EBV = +12 Kg,

Own record: Accuracy^2 = h^2 = 0.4
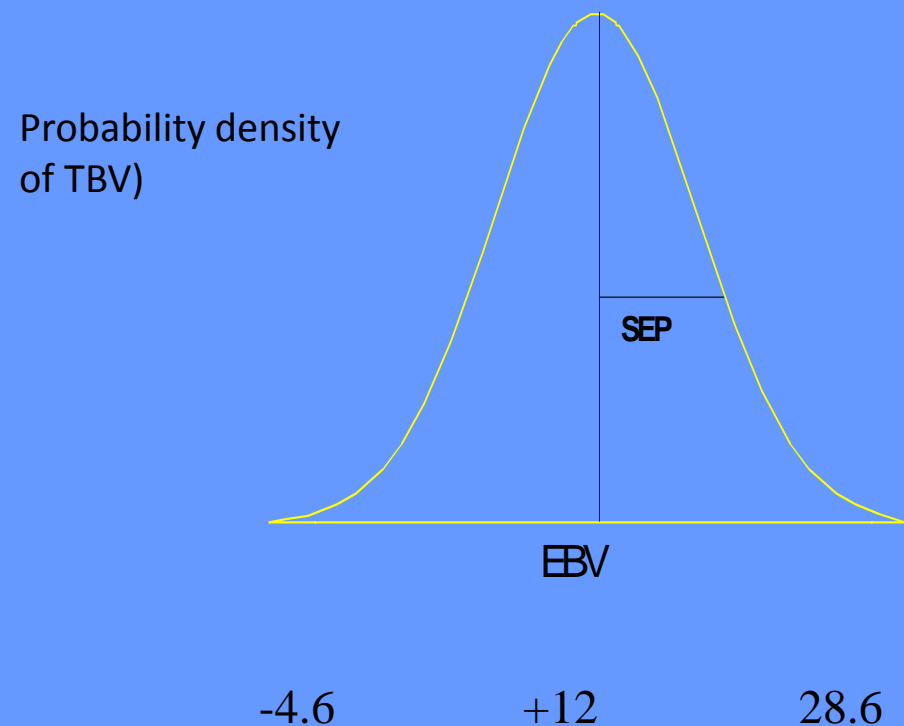
Then SEP = $\sqrt{(1-0.4)}.19 = 14.7$

Conf. Interval: EBV$\pm$ 2.SEP

Prediction Error Variance
- how much they still may change

- $PEV = var(EBV-TBV) = (1-r^2_{IA})V_A$

Probability density of TBV)

SEP

EBV

-4.6   +12   28.6

Let $\sigma_a = 19$, EBV = +12,

Progeny Test: Accuracy^2 = 0.81

Then SEP = $\sqrt{(1-0.81)}.19 = 8.3$

Conf. Interval: EBV $\pm$ 2.SEP

Prediction Error Variance
- how much they still may change

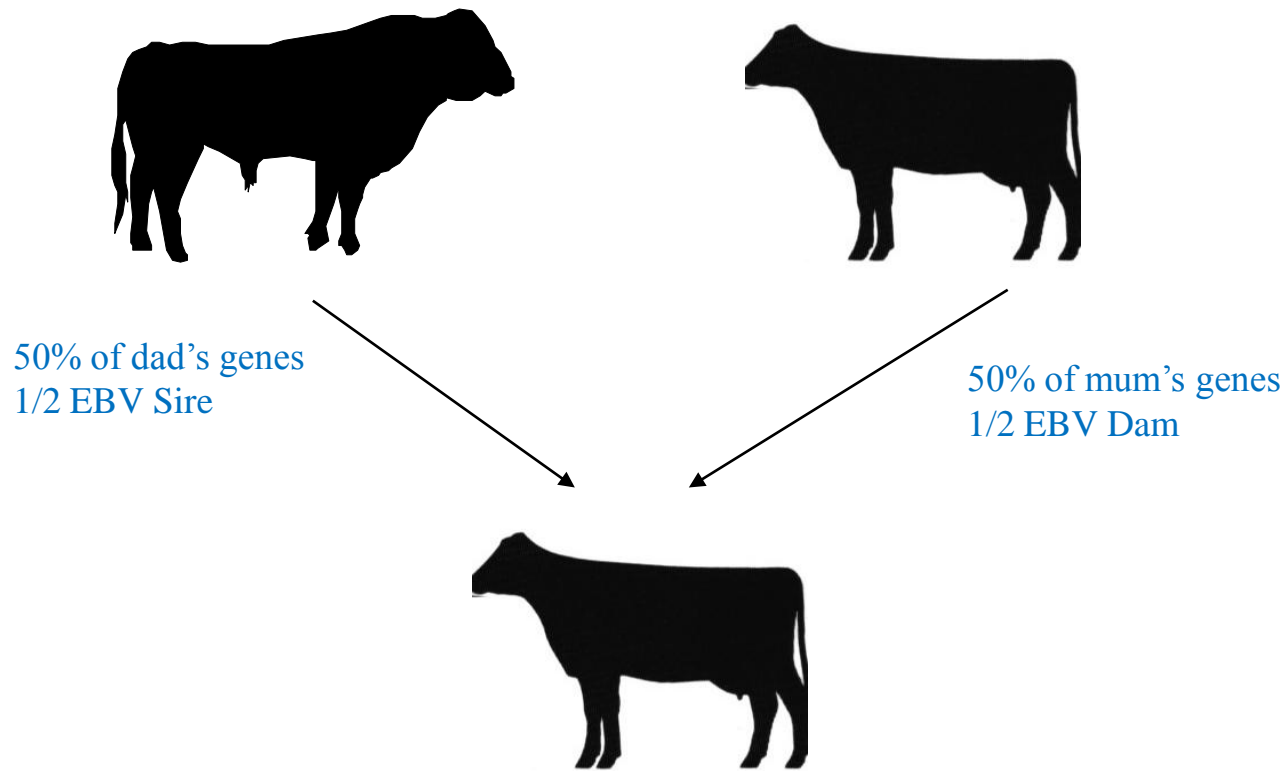**Consider Yearling weight EBV ($h^2 = 0.4$; $\sigma_a = 19$)**

EBV = +12

Information for EBV	Accuracy (r_{IH})	SEP	95%CI_ EBV
None	**0**	**19**	
Own performance	**0.63**	**14.7**	**-17.4 - 41.2**
Progeny Test n=40	**0.90**	**8.3**	**-4.6 - 28.6**
Progeny Test n=200	**0.98**	**3.8**	**4.4 - 19.6**

So Conf.Interval still quite large, even for very accurate EBVs
So in individual cases, variation around EBV can be quite large, but less so in selected groups.

Selection on EBV

50% of dad's genes
1/2 EBV Sire

50% of mum's genes
1/2 EBV Dam

Expected Value of progeny = 1/2 EBV$_{sire}$ + 1/2 EBV$_{dam}$

50% of dad's genes
1/2 BV Sire

50% of mum's genes
1/2 BV Dam

Expected Breeding Value of progeny = 1/2 EBV_sire + 1/2 EBV_dam

Prediction Error

Mendelian Sampling

Mendelian Sampling

Breeding Value Offspring

*Identifiable Environment Offspring
(herd, paddock, season, year)*

*Random Environment Offspring
('luck', disease, measurement error)*
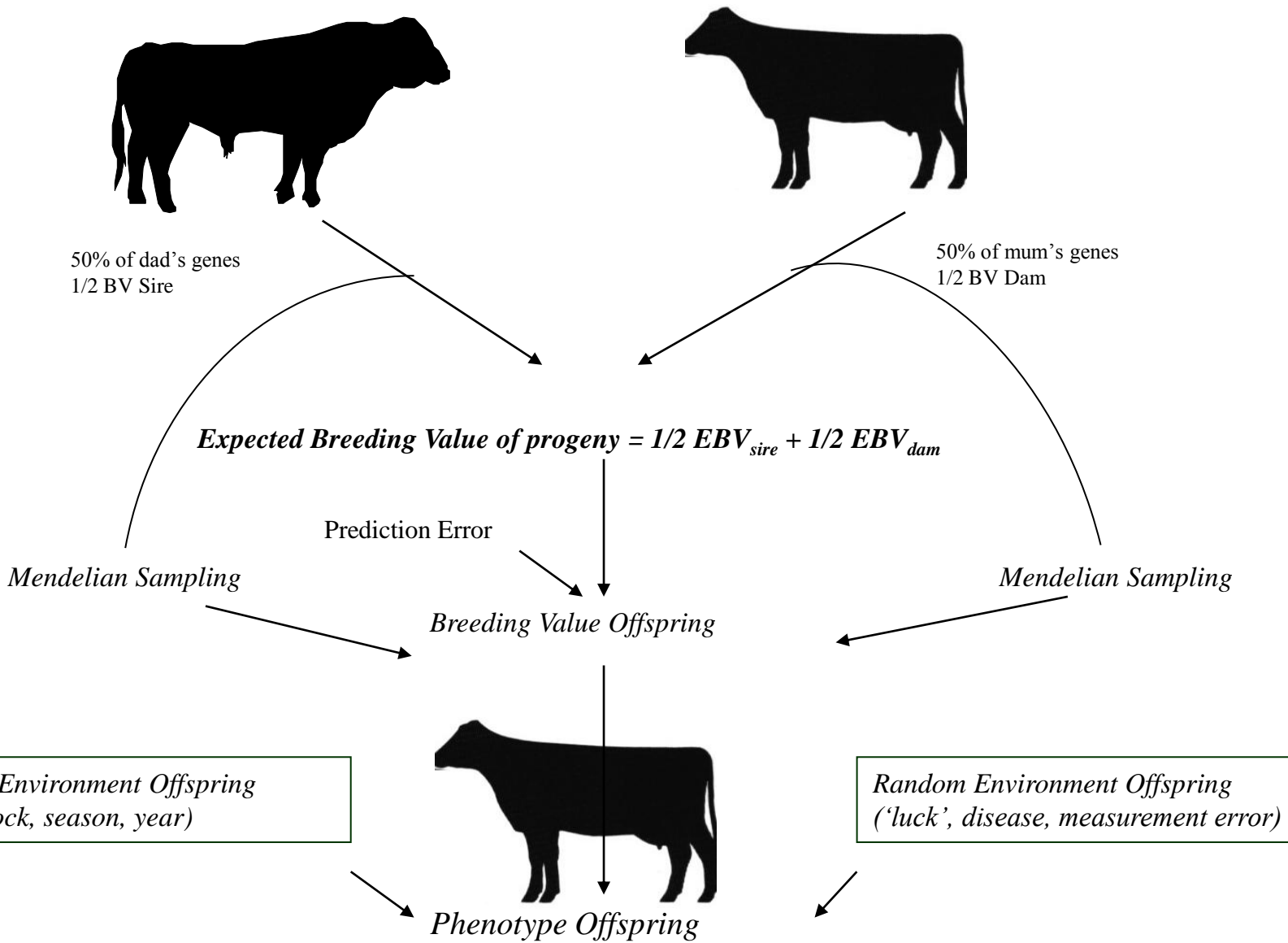
Phenotype Offspring

Figure 4.2. Genetic and environmental factors influencing a progeny's genotype

Some things to note

- EBV's on parents are additive

- Predicted performance of offspring does <u>not</u> depend on <u>accuracy</u> of the parents' EBVs
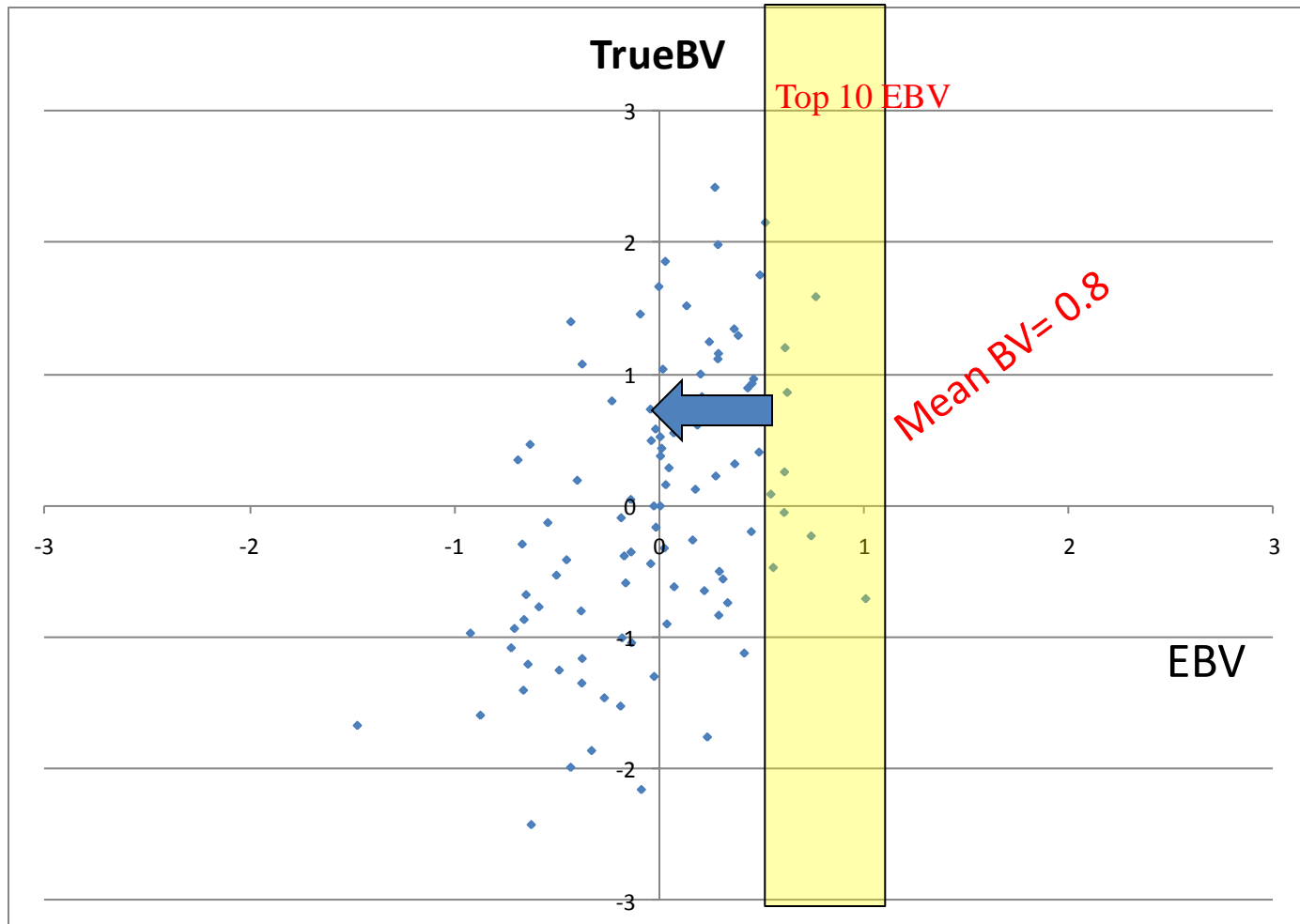
- Suppose   EBV_A   +56   r = 0.50

                 EBV_B   +56   r = 0.95

  select A or B?
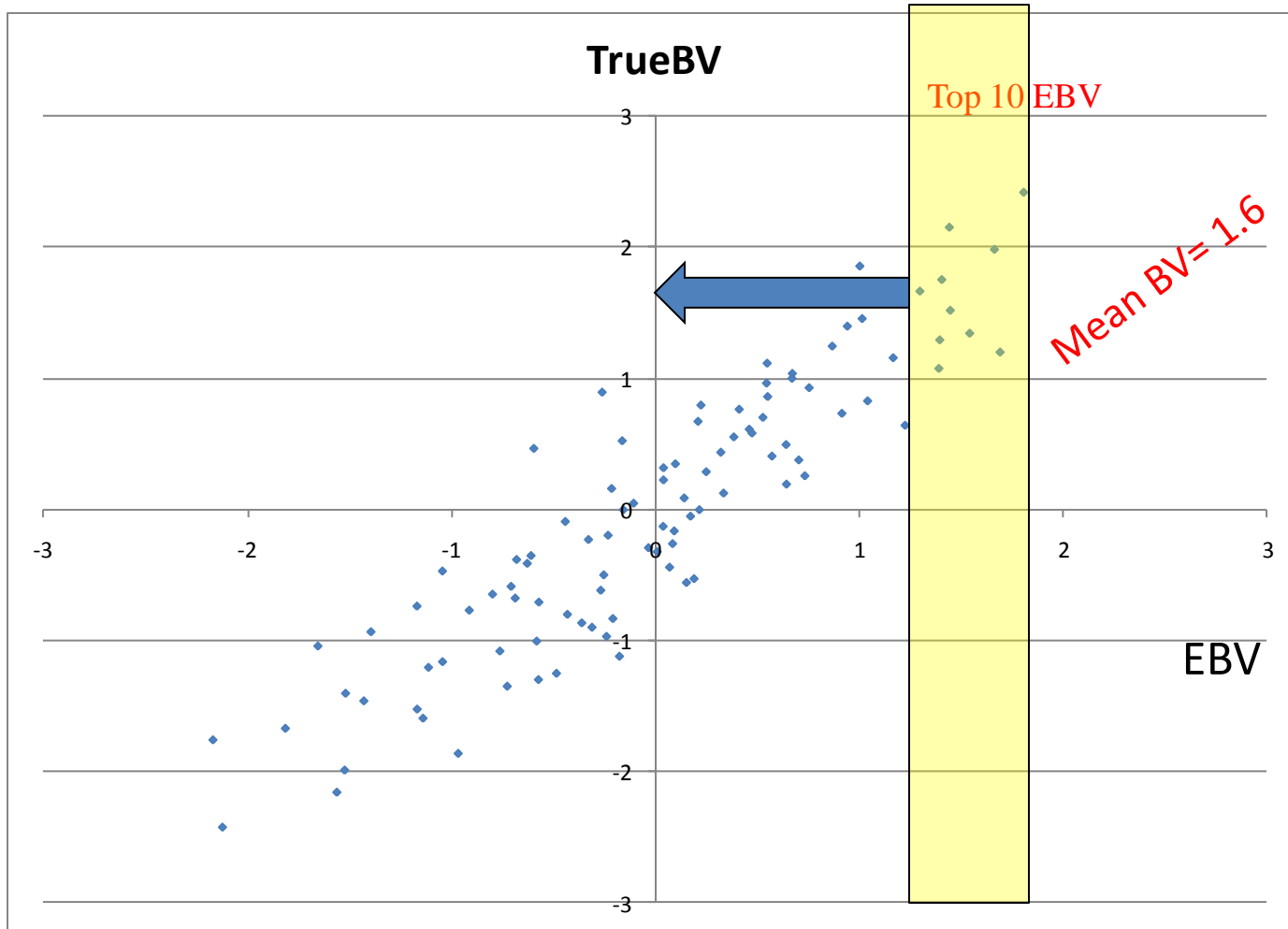
Answer: should not matter (if one is risk neutral)
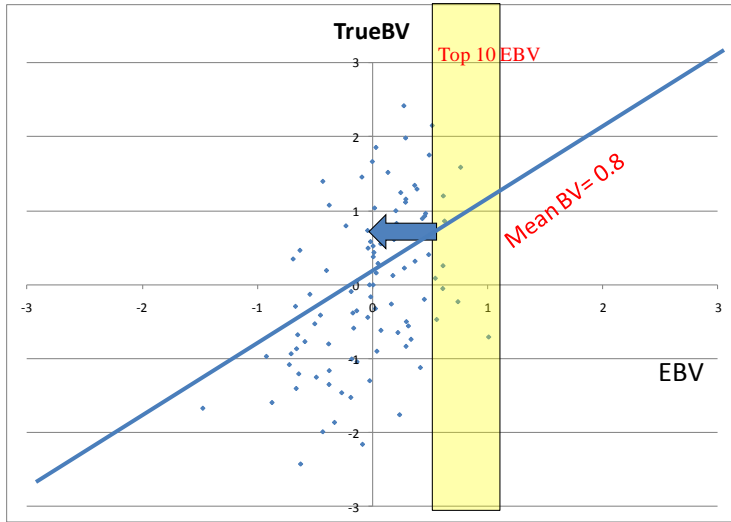
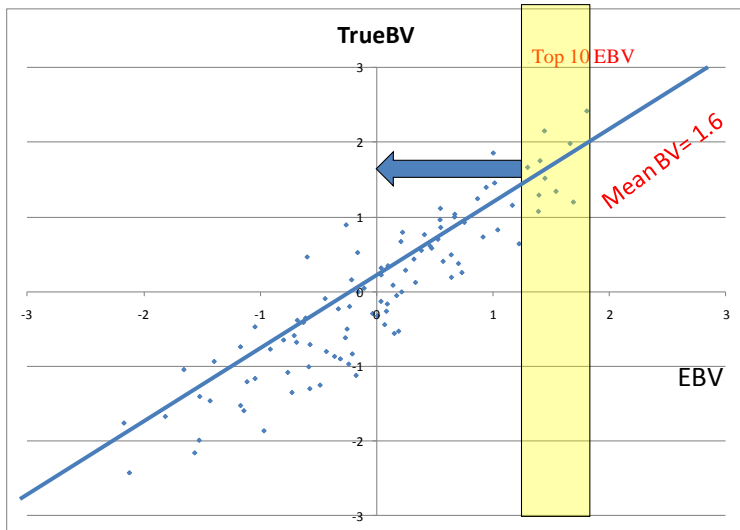# Select on EBV: accuracy related to response

Accuracy = 45%

Double accuracy gives double selection response!

TrueBV

Top 10 EBV

Mean BV= 1.6

EBV

Accuracy = 90%

Predicted Response

i = selection intensity (standard normal)

Regression of A on EBV = 1
i.e. slope is the same for accurate and inaccurate EBVs, see left

select on EBV's:

Response = i * SD(EBV)

$R = i * r * \sigma_a$

General to predict response per year

$$R_{yr} = \frac{i_{sires} \, r_{IA\,sires} + i_{dams} \, r_{IAdams}}{L_{sires} + L_{dams}} \sigma_A$$

Should optimize i and L

Should <u>maximize</u> r_{IA} ⟹ **Information from correlated traits**

⟹ **Information from relatives**