



UMEÅ UNIVERSITET

VÄLKOMMEN TILL KUNSKAPSVECKAN

WELCOME



UMEÅ UNIVERSITET

Bedömningsskalor och kravgränser

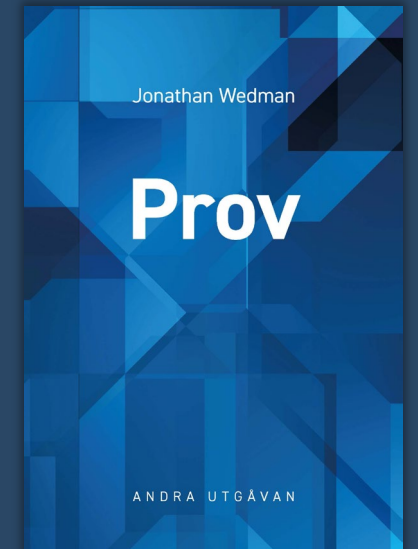
Vad finns det att välja på?

Jonathan Wedman

Fil. dr i Beteendevetenskapliga mätningar

Excellent lärare

TUV / UPL



"The Office"

Michael: "Vem har flest poäng?"

Pam: "Jag tror att de ligger lika. Du gav Jim 10 poäng, du gav Dwight en guldstjärna och du gav Stanley en tumme upp. Och jag vet inte riktigt hur jag ska jämföra de skalorna."

/Amerikanska "The Office", S3E23, ca 12:20

"The Office"

- **10 poäng:** Bra eller dåligt?

– 1-10 / 1-100 / 10-20 / -10 – +10 / 10-1 (inverterad)

- **En guldstjärna:** Bra eller dåligt?

– *En* guldstjärna



– *En guldstjärna*



- **En tumme upp:** Bra eller dåligt?

– *En* tumme upp



– *En* tumme *upp*



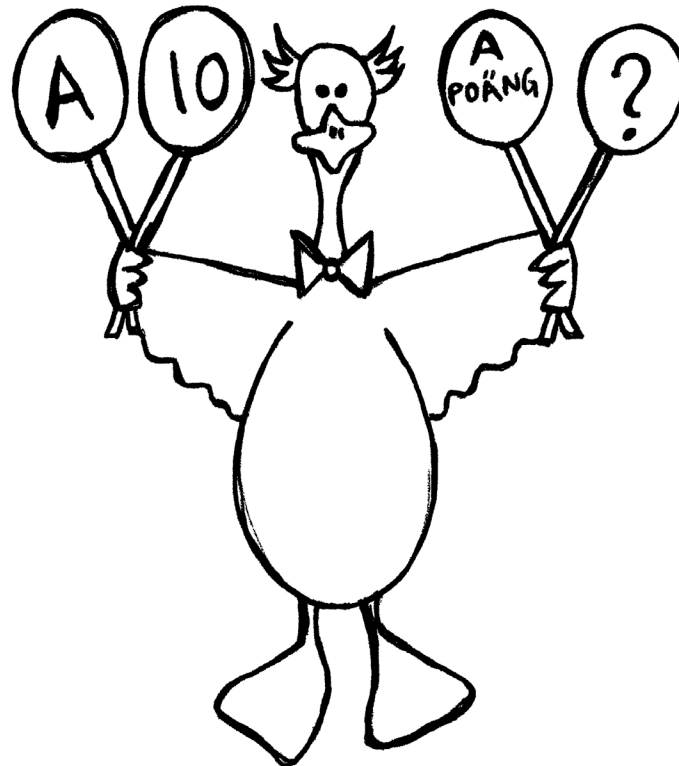
Skala för *betyg* och *omdömen* i skolan

- **Omdömen** (åk 1–5)
 - Mer än godtagbara kunskaper
 - Godtagbara kunskaper
 - **Otillräckliga kunskaper**
- **Termins**betyg**** (åk 6 – HT åk 9) och **slutbetyg** (VT åk 9)
 - A/B/C/D/E/**F**

OBS! Dessa är skalor för *betyg* och *omdömen*.
Det finns även fler **bedömnings**skalor (t.ex poäng).

Bedömningskala

Valfri skala som läraren använder för att värdera en elevprestation på en enskild uppgift eller ett prov.



Vanliga bedömningsskalor i skolan

Raka poäng

Nivåpoäng

Betygsbeteckning

(Procent)

Raka Poäng

- Numeriska samt enbart kvantitativa
 - Exempelvis 0–30 poäng
- Kompensatoriska
 - Alla poäng är likvärdiga
- Lätta att vikta vid den professionella bedömning lärare behöver göra vid betygssättning.
- **Avdramatisera provet**: Skriv ut poängen på varje fråga(?)
- Enkelt att konvertera till *procent*.
- Kan kombineras med en eller flera kravgränser

Raka Poäng

Skolverket  raka poäng

- "...kommer att leda till ett **mer tillförlitligt** resultat."
//Skolverket. (2022b, augusti). Digitalisering av de nationella proven – övergripande information
- "... borgar för den **mest tillförlitliga bedömningen** av proven."
// Skolverket. (2022c). Redovisning av uppdrag att digitalisera de nationella proven m.m. (Dnr 2021:63)

Raka poäng - exempel

Förklara, ur ett validitetsperspektiv, vad konsekvensen blir av att använda trickfrågor i ett prov. (4p)

Raka poäng - exempel

Totalpoäng på provet

• 1 rätt	1 poäng	}	= 1–3 poäng ger Underkänt
• 2 rätt	2 poäng		
• 3 rätt	3 poäng		
• 4 rätt	4 poäng	}	= 4–5 poäng ger Godkänt
• 5 rätt	5 poäng		

Men måste det vara så här?

Raka poäng - möjligheter

Totalpoäng på provet

• 1 rätt	13 poäng	}	= 1, 4 eller 13 poäng ger Underkänt
• 2 rätt	1 poäng		
• 3 rätt	4 poäng		
• 4 rätt	2 poäng	}	= 2 eller 7 poäng ger Godkänt
• 5 rätt	7 poäng		

Men kanske inte så pedagogisk?

Raka poäng vs värdeord

- Funkar poäng när progressionsuttrycken i BK är kvalitativa?
- Ja! 😊
- **E/C/A = Översiktligt / Utförligt / Utförligt och nyanserat**
- Ett prov med exempelvis 30 frågor.
- Olika frågor behandlar ämnesområdet ur olika perspektiv (nyanser).
- Fler korrekta svar visar att eleverna kan beskriva mer utförligt.
- Många korrekta svar visar att eleverna dessutom kan det nyanserat.

Raka poäng vs värdeord

Engelska åk 9, ur betygskriterierna

- "Eleven lyssnar samt förstår och tolkar...
 - **E)** huvudsakligt innehåll och tydliga detaljer
 - **C)** på ett välgrundat sätt huvudsakligt innehåll och väsentliga detaljer
 - **A)** på ett välgrundat och nyanserat sätt såväl helhet som detaljer
- ...i talat språk i måttligt tempo och i olika sammanhang."

Raka poäng vs värdeord

- Flervalsfrågor, 30st, med tre svarsalternativ vardera.
- Olika frågor har olika fokus
 - Huvudsakligt innehåll
 - Tydliga detaljer
 - Väsentliga detaljer
 - Mer och mindre välgrundade tolkningar
 - Nyanser (tolkningar ur olika perspektiv)
 - Helheten
- Kravgränser för olika *nivåer*: Låg=22 / Mellan=25 / Hög=27

Nivåpoäng

- Numeriska samt kombination av kvantitativa och kvalitativa
 - Exempelvis: 30 poäng varav 14 E-poäng, 8 C-poäng och 8 A-poäng (14/8/8)
- Icke-kompensatoriska
 - Olika poäng har olika värde
- Lätta att vikta vid den professionella bedömning lärare behöver göra vid betygssättning.
- **Avdramatisera provet:** Skriv ut nivåpoängen på varje fråga.
- Kan konverteras till *procent*.
- Kan kombineras med en eller flera kravgränser

Nivåpoäng – exempel

Förklara, ur ett validitetsperspektiv, vad konsekvensen blir av att använda trickfrågor i ett prov. (1/1/2)

Nivåpoäng vs värdeord

Engelska åk 9, ur betygskriterierna (påminnelse)

- "Eleven lyssnar samt förstår och tolkar...
 - **E)** huvudsakligt innehåll och tydliga detaljer
 - **C)** på ett välgrundat sätt huvudsakligt innehåll och väsentliga detaljer
 - **A)** på ett välgrundat och nyanserat sätt såväl helhet som detaljer
- ...i talat språk i måttligt tempo och i olika sammanhang."

Nivåpoäng vs värdeord

Engelska åk 9

- Flervalsfrågor, 30st, med tre svarsalternativ vardera.
 - 14st **E**-frågor fokuserar på **huvudsakligt innehåll** och **tydliga detaljer**
 - 8st **C**-frågor fokuserar på **huvudsakligt innehåll** och **väsentliga detaljer**
 - 8st **A**-frågor fokuserar på **helheten** och **tolkningar ur olika perspektiv**
- Kravgränser för olika *nivåer*:
 - Låg = 18p
 - Mellan = 22p, varav minst 6p från C-frågor
 - Hög = 27p, varav minst 6p från A-frågor

Betygsbeteckning

- Alfabetiska samt enbart kvalitativa
 - Exempelvis: E–A på frågan eller provet som helhet
- Antingen kompensatoriska eller icke-kompensatoriska
 - Väljs vid kravgränssättning
- Svårt att vikta vid den professionella bedömning lärare behöver göra vid betygssättning. Måste kvantifieras för att kunna viktas.
- **Avdramatisera provet**: Skriv ut betygsbeteckningen om den används.

Betygsbeteckning – exempel

Förklara, ur ett validitetsperspektiv, vad konsekvensen blir av att använda trickfrågor i ett prov. (E–C)

Betygsbeteckning

Betygskriterierna och betygsbeteckningar är övergripande och **inte** uppgiftsanpassade, vilket gör att de **inte** lämpar sig för väl avvägd och framåtsyftande återkoppling. Elevernas uppmärksamhet riskerar här att styras mot betygen i stället för vad de ska lära sig i ämnet eller kursen. (Skolverket, 2022, s. 20)

Det är viktigt att lärare klargör exempelvis att ... betygsbeteckningar som används som omdömen på prov och uppgifter **inte** med en enkel formel kan vägas samman till ett betyg i slutet av terminen eller kursen. Det är viktigt att eleverna förstår att betyget **inte** är ett medelvärde av resultaten på de prov och uppgifter som eleven genomfört fram till och med betygssättningen. (2022a, s. 31)

Betygssättning

Två steg

1. Se till delarna (Bedömningsskalor)

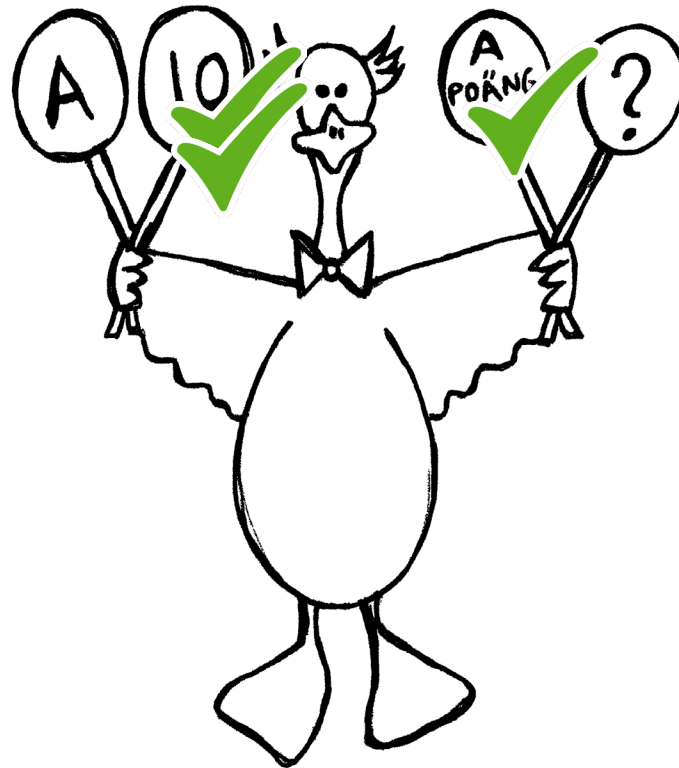
1. Enskilda bedömningsunderlag vägs och värderas mot **delar av** BK.
2. Mer tillförlitliga/relevanta = Mer vikt
3. Nyare underlag = Mer vikt
4. Större underlag = Mer vikt
5. "Låg/Låg/Medel/Hög" **kan bli** A-nivå

2. Se till helheten (Omdömes-/Betygsskala)

1. Måluppfyllelse av **hela** betygskriterier utvärderas.
2. Sedan 2022-07-01:
 1. **E-nivå**: Icke-kompensatoriskt = FEEE **blir** betyget F.
 2. **D-A-nivå**: Delvis kompensatoriskt = CAAA **kan bli** betyget A.

Vilken bedömningskala ska man välja?

Välj **helst** en av skalorna som använder poäng.



Kravgränser

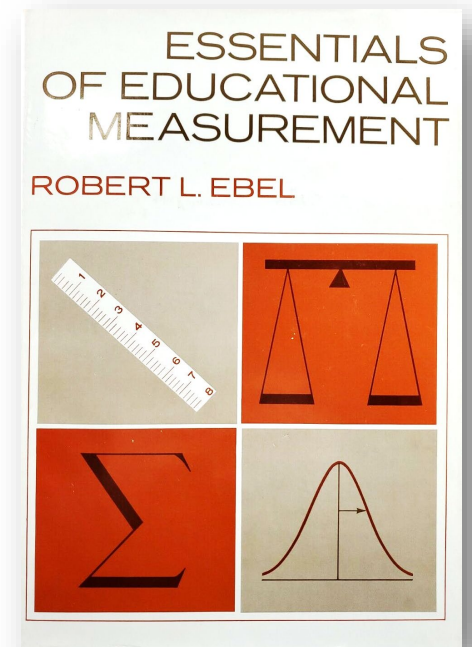


UMEÅ UNIVERSITET

Ebel (1972)

”Att bestämma en lägstanivå för en godkänd prestation
involverar alltid några tämligen godtyckliga
och inte helt tillfredsställande beslut.”

(p. 492; fritt översatt)

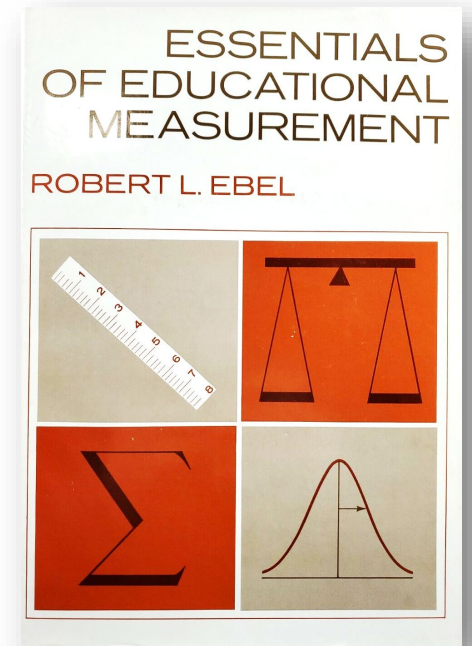


Ebel (1972)

”Att bestämma en lägstanivå för en godkänd prestation
involverar alltid några tämligen godtyckliga
och inte helt tillfredsställande beslut.”

(p. 492; fritt översatt)

”Den som förväntar sig att upptäcka den ”riktiga” gränsen för
godkänt ... är dömd att bli besviken

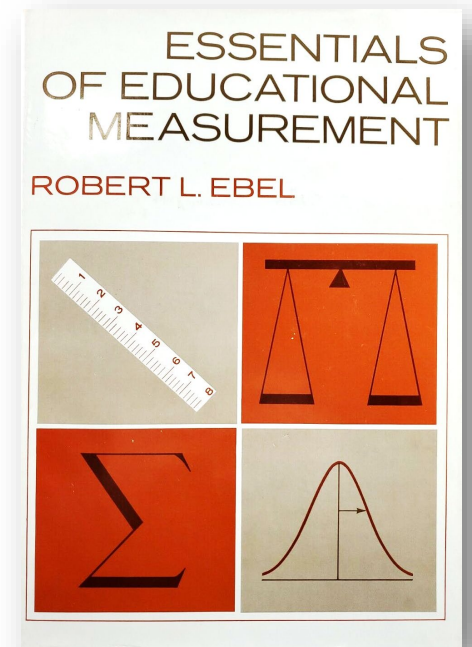


Ebel (1972)

"Att bestämma en lägstanivå för en godkänd prestation involverar alltid några tämligen godtyckliga och inte helt tillfredsställande beslut."

(p. 492; fritt översatt)

"Den som förväntar sig att upptäcka den "riktiga" gränsen för godkänt ... är dömd att bli besviken eftersom en "riktig" kravgräns inte existerar och därför inte går att upptäcka.

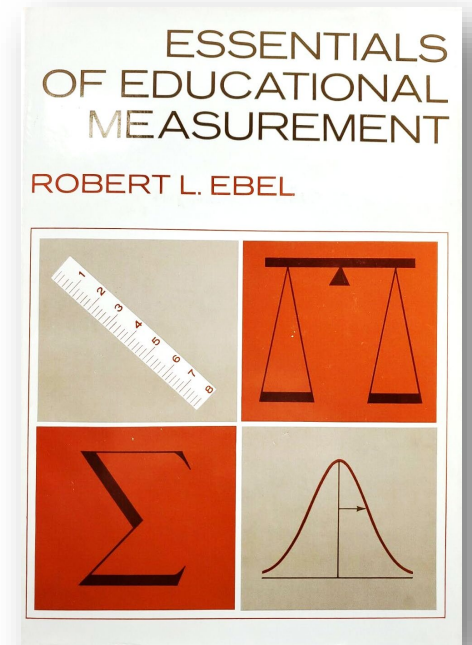


Ebel (1972)

"Att bestämma en lägstanivå för en godkänd prestation involverar alltid några tämligen godtyckliga och inte helt tillfredsställande beslut."

(p. 492; fritt översatt)

"Den som förväntar sig att upptäcka den "riktiga" gränsen för godkänt ... är dömd att bli besviken eftersom en "riktig" kravgräns inte existerar och därför inte går att upptäcka. Allt en [lärare], som måste fastställa en kravgräns, kan hoppas på

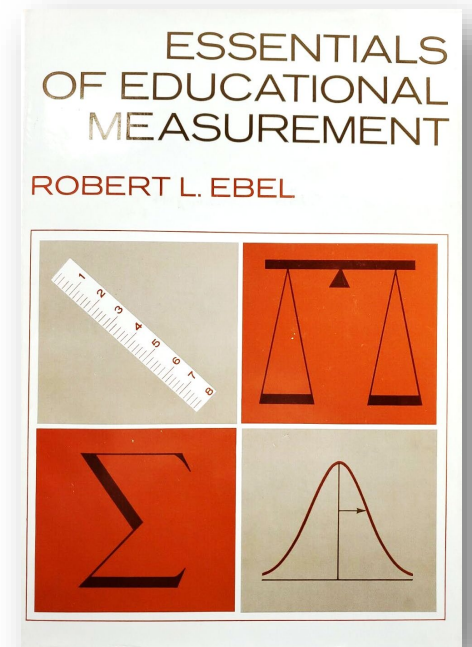


Ebel (1972)

”Att bestämma en lägstanivå för en godkänd prestation involverar alltid några tämligen godtyckliga och inte helt tillfredsställande beslut.”

(p. 492; fritt översatt)

”Den som förväntar sig att upptäcka den ”riktiga” gränsen för godkänt ... är dömd att bli besviken eftersom en ”riktig” kravgräns inte existerar och därför inte går att upptäcka. Allt en [lärare], som måste fastställa en kravgräns, kan hoppas på och allt dennes [elever] kan begära



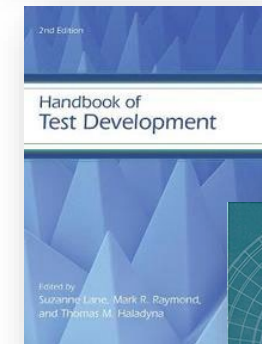
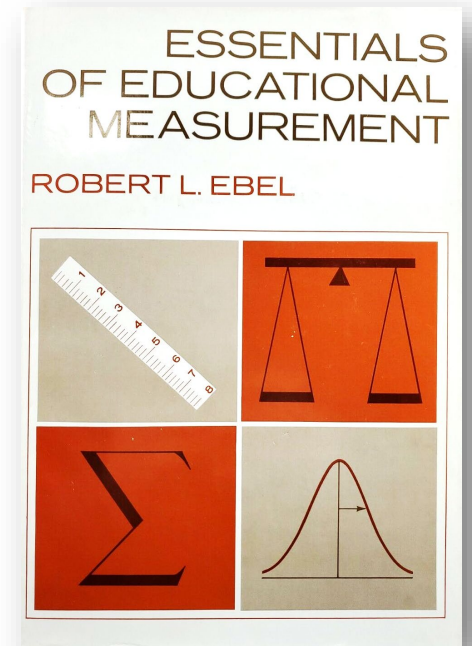
Ebel (1972)

"Att bestämma en lägstanivå för en godkänd prestation involverar alltid några tämligen godtyckliga och inte helt tillfredsställande beslut."

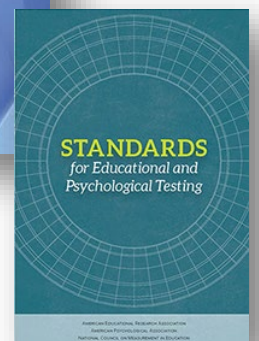
(p. 492; fritt översatt)

"Den som förväntar sig att upptäcka den "riktiga" gränsen för godkänt ... är dömd att bli besviken eftersom en "riktig" kravgräns inte existerar och därför inte går att upptäcka. Allt en [lärare], som måste fastställa en kravgräns, kan hoppas på och allt dennes [elever] kan begära är att grunden för kravgränsen är tydligt utskriven och så **rimlig** som möjligt."

(p. 496; fritt översatt)



Cizek, 2015

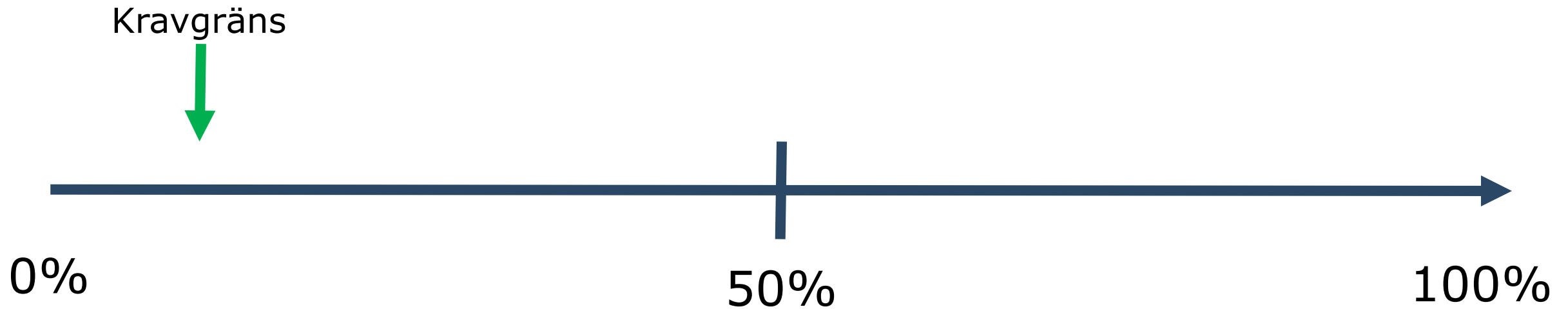


AERA, APA
& NCME, 2014

Ingen objektiv gräns

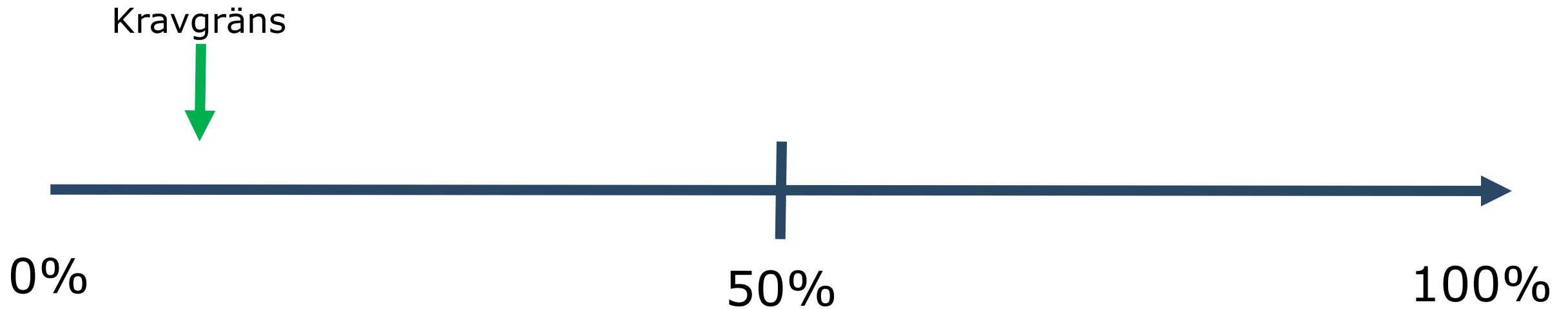
- Kravgränser är beroende av provets svårighetsgrad
 - **Svårt** prov > Låg gräns
 - **Medelsvårt** prov > Medelhög gräns
 - **Lätt** prov > Hög gräns

Ett för svårt prov



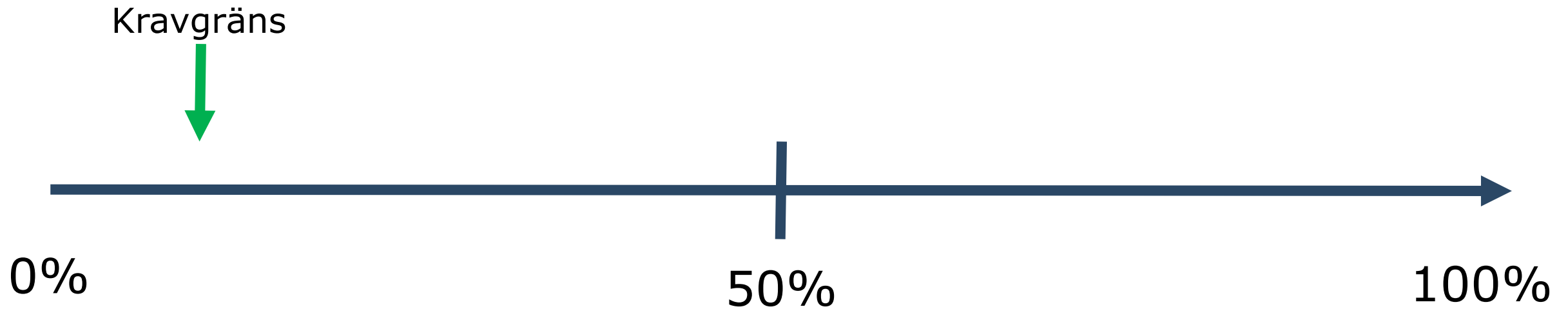
De flesta eleverna har presterat **lågt**.
De har spenderat merparten av sin tid och energi på att svara **fel**.

Ett för svårt prov



De flesta eleverna har presterat **lågt**.
De har spenderat merparten av sin tid och energi på att svara **fel**.
Det insamlade underlaget visar framförallt vad de **inte** kan.

Ett för svårt prov



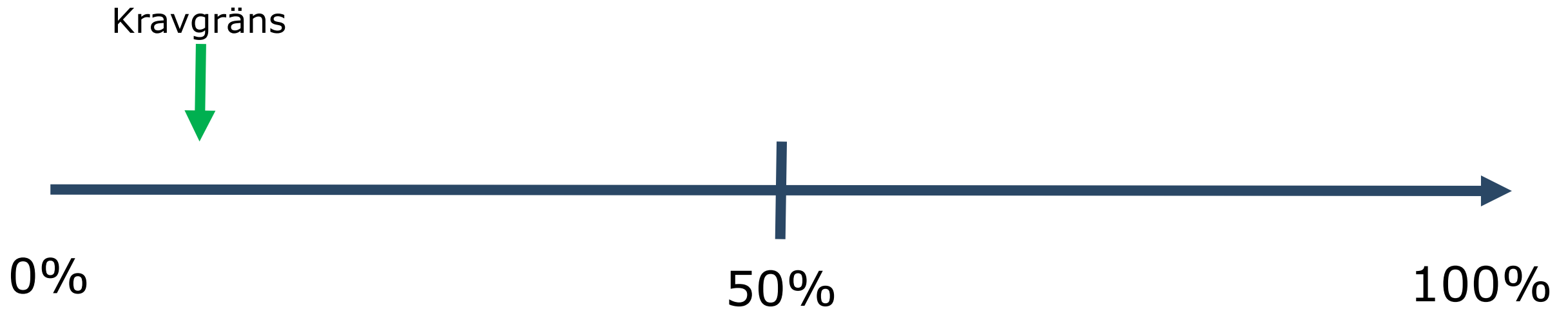
De flesta eleverna har presterat **lågt**.

De har spenderat merparten av sin tid och energi på att svara **fel**.

Det insamlade underlaget visar framförallt vad de **inte** kan.

Tillförlitligheten blir bristfällig pga **mindre** information om elevens kunnande.

Ett för svårt prov



De flesta eleverna har presterat **lågt**.

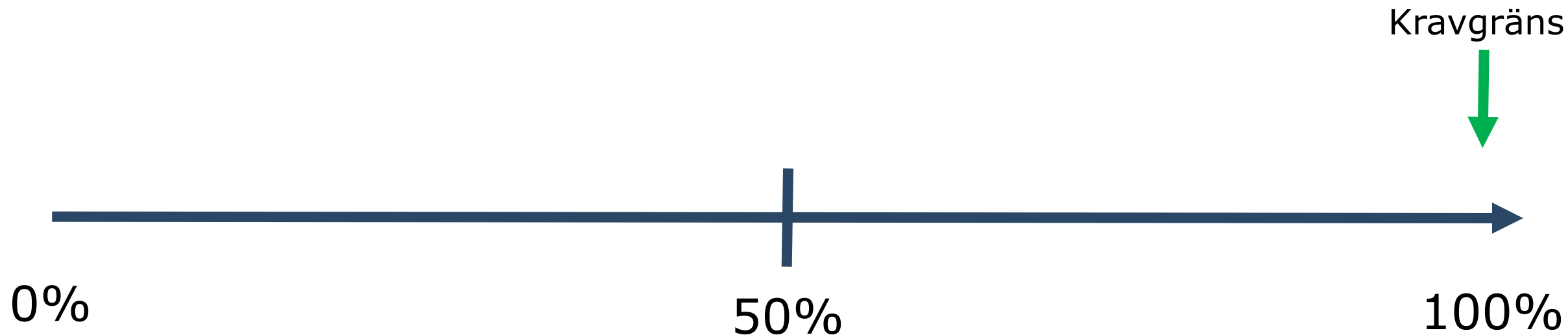
De har spenderat merparten av sin tid och energi på att svara **fel**.

Det insamlade underlaget visar framförallt vad de **inte** kan.

Tillförlitligheten blir bristfällig pga **mindre** information om elevens kunskande.

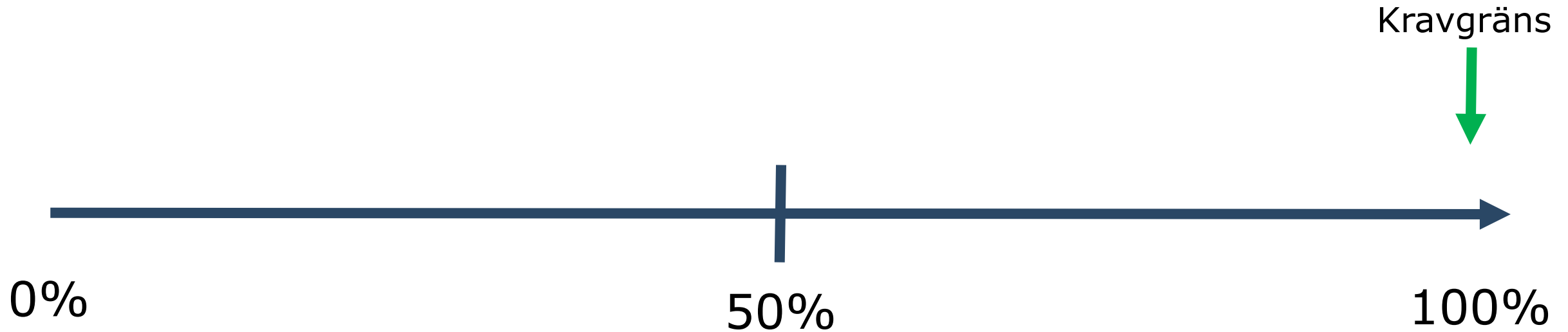
Den framåtsyftande återkopplingen riskerar att bli **ohanterlig**.

Ett för lätt prov



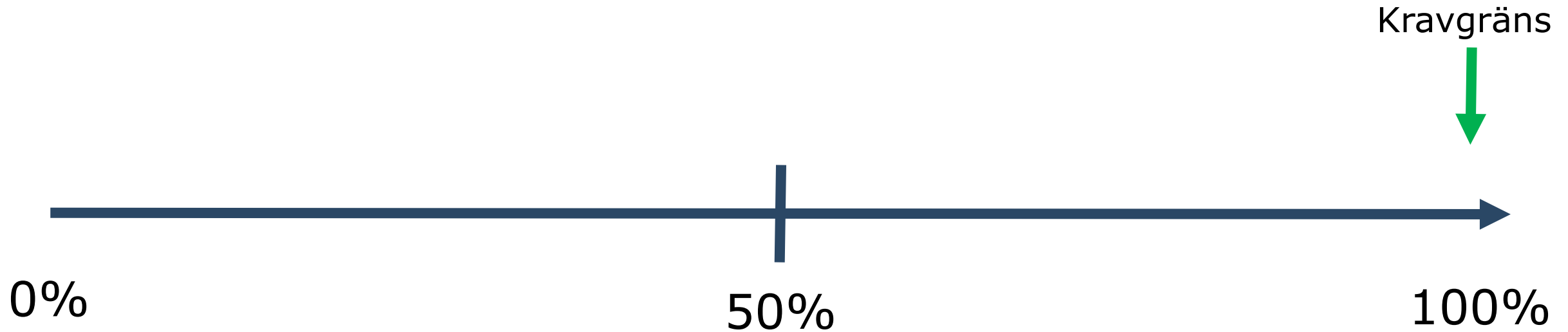
De flesta eleverna har presterat **högt**.
De har spenderat merparten av sin tid och energi på att svara **rätt**.

Ett för lätt prov



De flesta eleverna har presterat **högt**.
De har spenderat merparten av sin tid och energi på att svara **rätt**.
Det insamlade underlaget visar framförallt vad de **kan**.
Tillförlitligheten blir högre pga **mer** information om elevens kunskande.

Ett för lätt prov



De flesta eleverna har presterat **högt**.

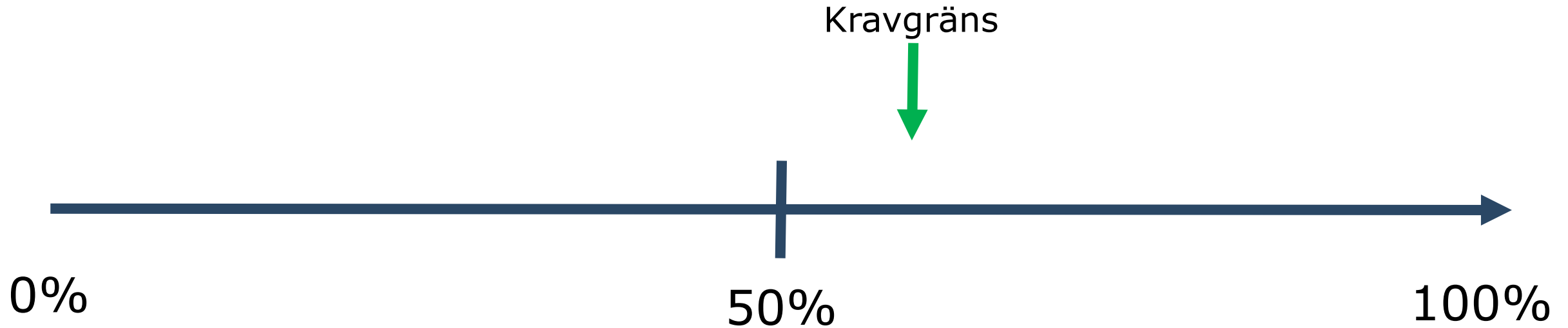
De har spenderat merparten av sin tid och energi på att svara **rätt**.

Det insamlade underlaget visar framförallt vad de **kan**.

Tillförlitligheten blir högre pga **mer** information om elevens kunnande.

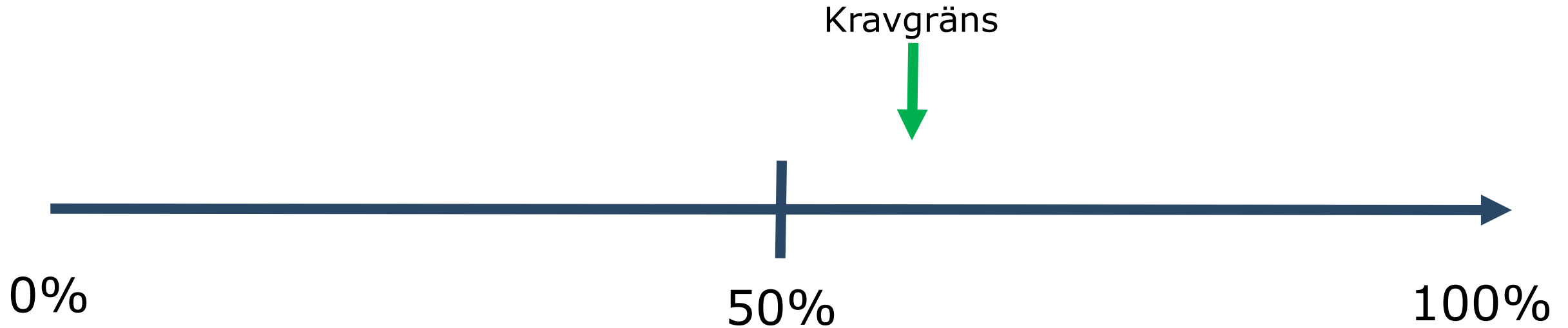
Men! Takeeffekten **begränsar** relevansen och försvårar återkoppling.

Ett lagom svårt prov



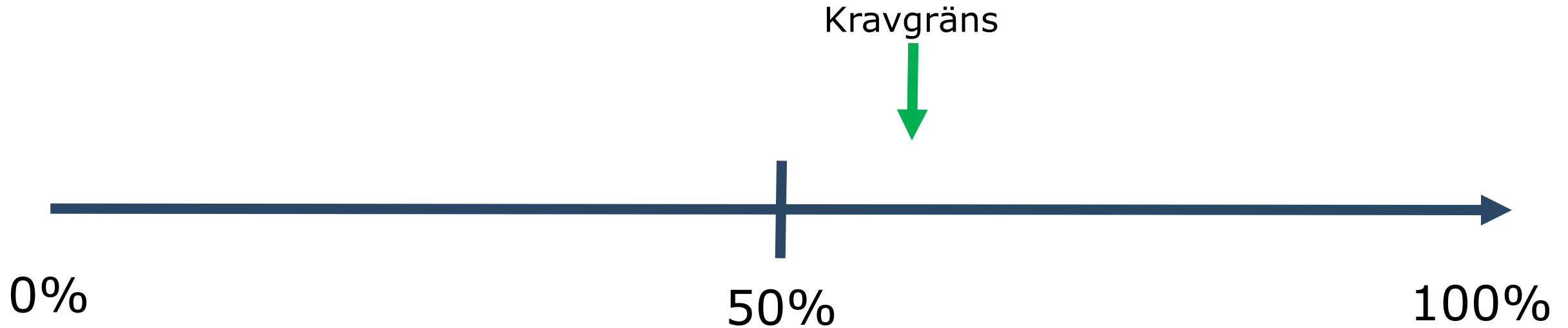
De flesta eleverna har presterat **över medel**.
De har spenderat merparten av sin tid och energi på att svara **rätt**.

Ett lagom svårt prov



De flesta eleverna har presterat **över medel**.
De har spenderat merparten av sin tid och energi på att svara **rätt**.
Det insamlade underlaget visar framförallt vad de **kan**.
Tillförlitligheten blir högre pga **mer** information om elevens kunskande.

Ett lagom svårt prov



De flesta eleverna har presterat **över medel**.

De har spenderat merparten av sin tid och energi på att svara **rätt**.

Det insamlade underlaget visar framförallt vad de **kan**.

Tillförlitligheten blir högre pga **mer** information om elevens kunskande.

Och! Förutsättningar för en **god** relevans och **konstruktiv** återkoppling.

Att konceptuellt tänka i procent

- **Heltäckande** prov

E \approx 60%

C \approx 75%

A \approx 85%

- **E**-prov

E \approx 75–80%

Canvas standard
D = 61
C = 70
B = 80
A = 90

Återigen: Dessa siffror är **inte** objektiva sanningar eller ens starka rekommendationer utan ödmjuka förslag/utgångspunkter.

Hur ska man kravgränssätta?

- Kravgränser är **subjektiva**
 - Utveckla gärna tillsammans i ett arbetslag
- Kravgränser bör **variera**
 - Utifrån provets svårighetsgrad
 - Utifrån var eleverna är i utbildningen
- Kravgränser för ett prov ger aldrig "*Betyget E/C/A*", utan handlar om hög/låg **nivå** eller E-/C-/A-**nivå**.
 - Läraren samlar underlag som ger stöd åt att eleven kan komma att nå måluppfyllelse för ett visst betyg vid tiden för betygssättning.



UMEÅ UNIVERSITY

Tack!

Jonathan Wedman

**Fil. dr i Beteendevetenskapliga mätningar
Excellent lärare**