



GÖTEBORGS  
UNIVERSITET

# Den artificiella intelligensens begynnelse i skolväsendet

ChatGPT och validitetsproblemet med  
summativ skrivbedömning i gymnasieskolan

David Nyqvist

Ämneslärarprogrammet med inriktning mot  
gymnasieskolan





GÖTEBORGS  
UNIVERSITET

“The reader must accept it as a fact that a digital computer can be constructed, and indeed have been constructed, according to the principles we have described, and that they can in fact mimic the actions of a human computer very closely.” – Alan Turing, 1950.

Uppsats/Examensarbete: 15 hp  
Kurs: LGSV2A  
Nivå: Avancerad nivå  
Termin/år: VT/2023  
Handledare: Lena Rogström  
Examinator: Louise Holmer

---

Nyckelord: ChatGPT, AUA, validitet, gymnasieskolan, svenska, fusk, bedömning, summativ bedömning, skrivbedömning, AI.

Keywords: ChatGPT, AUA, validity, upper secondary school, Swedish, cheating, assessment, summative assessment, writing assessment, AI.

Title in English: The beginning of artificial intelligence in the school system – ChatGPT and the validity problem of writing assessment in upper secondary school.

## Abstract

The present essay examines the apparent problem in the validity of writing assessment in the Swedish upper secondary school. The problem lies in the introduction of ChatGPT, which the students can use to write abstract texts of high quality. The main analyzing tool to examine this problem is Bachman's (2005) assessment use argument (AUA). The empirical material comes from seven interviews of secondary school teachers in Swedish. The result is comprised by the answers from the interviews applied to the assessment use argument. The results suggests that there are two main ways of solving the validity problems in the assessment argument. The first way is that teachers can control the assessment situations. This is done by having the students write in a locked writing program, and only having them write in the classroom. This ensures that the teachers know that ChatGPT was not used to write the texts. The second way is that teachers can change the writing construct that they wish to grade. The teachers mean that in the future this is a likely consequence of the introduction of ChatGPT. The assessment would then shift focus from the product, the text, to the students writing process. These results are then discussed, and two conclusions are formed to help develop a solution to the validity problem in writing assessment presented by ChatGPT.

These conclusions are as follows:

- The teachers can use a secure writing program to solve the validity problem.

- The teachers can also change the writing construct and the target domain to solve the validity problem, however that would require more information on how to do that, as well as changes in the policy document of the subject of Swedish.

# Innehållsförteckning

<b>1</b>	<b>Inledning .....</b>	<b>1</b>
<b>2</b>	<b>Syfte och frågeställningar .....</b>	<b>2</b>
<b>3</b>	<b>Forskningsbakgrund .....</b>	<b>4</b>
3.1	Artificiell intelligens och ChatGPT .....	4
3.2	Validitet, summativ bedömning och skrivbedömning .....	6
3.2.1	Validitet .....	6
3.2.2	Summativ bedömning .....	8
3.2.3	Skrivbedömning .....	9
3.2.4	Sammanfattning av forskningsbakgrunden .....	11
<b>4</b>	<b>Teori .....</b>	<b>12</b>
4.1	Inledning .....	12
4.2	Turingtestet .....	12
4.3	Bachmans assessment use argument .....	13
4.3.1	Validitetsargumentet .....	15
4.3.2	Användningsargumentet .....	15
4.3.3	Användandet av Bachmans assessment use argument i den här undersökningen .....	16
<b>5</b>	<b>Metod och material .....</b>	<b>18</b>
5.1	Intervjumetod .....	18
5.1.1	Intervjuguide .....	18
5.2	Urval .....	19
5.3	Genomförandet av intervjuerna .....	21
5.3.1	Turingtestet .....	21
5.3.2	Bachmans assessment use argument .....	21
5.3.3	Turingtestet och Bachmans assessment argument i kombination .....	21
5.3.4	Beskrivning av texterna .....	22
5.4	Analys av intervjuer .....	23
5.4.1	Tematisk analys .....	23
5.5	Etiskt förhållningssätt .....	24
5.6	Litteratursökning i forskningsbakgrunden .....	24
<b>6</b>	<b>Resultat .....</b>	<b>25</b>
6.1	Frågeställning I .....	25
6.2	Frågeställning II .....	26
6.2.1	Subteman .....	26

6.2.2	Teman .....	28
6.2.3	Informanternas svar applicerade på Bachmans assessment argument .....	29
6.3	Frågeställning III .....	31
6.3.1	Teman .....	31
6.3.2	Informanternas svar applicerade på Bachmans assessment argument .....	35
6.3.3	Låst skrivvyta.....	35
6.3.4	Nytt skrivkonstrukt .....	35
6.3.5	Didaktiska konsekvenser .....	36
<b>7</b>	<b>Diskussion .....</b>	<b>39</b>
7.1	Turingtestet.....	39
7.2	Validitetsproblemet.....	39
7.3	Lösningar på validitetsproblemet .....	39
7.4	Didaktiska konsekvenser .....	40
7.4.1	Låst skrivvyta.....	40
7.4.2	Separation mellan övning och examination .....	41
7.4.3	Bedömning av skrivprocessen .....	41
7.5	Måldomänen oförenlighet med ChatGPT .....	42
7.6	Slutsatser för att garantera att ChatGPT inte påverkar validiteten i skrivbedömningen.....	42
7.7	Diskussion om uppsatsens begränsningar .....	42
<b>8</b>	<b>Vidare forskning.....</b>	<b>44</b>
<b>9</b>	<b>Referenslista.....</b>	<b>46</b>
<b>10</b>	<b>Bilagor .....</b>	<b>49</b>
10.1	Bilaga 1: ChatGPT:s text. I undersökningen kallad för text 2.....	49
10.2	Bilaga 2: Elevtext 1 .....	51
10.3	Bilaga 3: Elevtext 2 .....	52
10.4	Bilaga 4: Intervjuguide .....	53
10.5	Bilaga 5: Informationsbrev .....	55
10.6	Bilaga 6: Samtyckesblankett.....	56

# 1 Inledning

Under hösten 2022 läste jag på ett forum om en elev som enligt egen utsago låtit en textgenerator som drivs av artificiell intelligens skriva flera texter åt både sig själv och andra klasskamrater. Dessa texter resulterade i godkänt betyg. Eleven skrev texterna i utbyte mot pengar, och det tog bara några minuter att skriva varje text. Om elever i gymnasieskolan har tillgång till ett verktyg som kan skriva texter kan problem uppstå med bedömningen av dessa texter och i förlängningen med betygsättningen. Skrivbedömningen i gymnasieskolan äventyras om svensklärare inte vet om en textgenerator eller eleven har skrivit en text som lämnats in i syfte att bedömas.

I en verklighet där elever kan producera inlämningsuppgifter av hög kvalitet med hjälp av ett program tillgängligt för allmänheten måste skolväsendet reagera. Gymnasieskolans betyg är ett urvalsverktyg för högre studier. Hur ska svensklärare och i förlängningen skolväsendet säkerställa att detta urvalsverktyg inte äventyras av textgeneratorer och artificiell intelligens?

Den här uppsatsen ämnar specifikt diagnostisera problemen rörande bedömning av texter i gymnasieskolan, i ljuset av den nya situationen med textgeneratorer och artificiell intelligens. Användandet av textgeneratorer kommer förmodligen att vara allmän kunskap hos landets gymnasieelever inom en snar framtid. Gymnasielärare tvingas därför att ta ställning till den nya tekniken. Frågeställningarna fokuseras på den problematik som tycks ha uppstått på många av landets gymnasieskolor under hösten och vintern 2022. Svensklärare vet inte längre om eleverna har skrivit texten, eller om en textgenerator har gjort det för dem. De måste därför garantera genom bedömningssituationen att det är eleverna som skriver. Under min VFU under våren 2023 har både jag och mina VFU-kamrater bevittnat de panikinsatta mötena om ChatGPT. Det är problematiken, inte möjligheterna, med textgeneratorer och AI som representerar den artificiella intelligensen begynnelse i skolväsendet.

Eftersom forskningen om AI i skolan fortfarande är i sin linda finns inte mycket tidigare forskning inom forskningsfältet. Den här undersökningen är därför högst teoretisk, och bygger på modeller och teorier om skrivbedömning.

## 2 Syfte och frågeställningar

Med bakgrund i inledningen som beskriver den nya problematiken i skrivbedömningen i gymnasieskolan kan ett syfte med uppsatsen formuleras. Syftet är att belysa hur ChatGPT skapar validitetsproblem i svensklärares summativa skrivbedömning och vad svensklärare har gjort för att lösa detta problem. Idag finns AI med hög textkompetens lättillgänglig på internet, och lärare kan därför stöta på texter skrivna av datorer i stället för elever.

Svensklärare måste därför antingen kunna se skillnad på en datorgenererad text och en elevtext, eller på annat säkerställa att datorgenererade texter inte bedöms summativt. Detta beror på att skolan har utöver en kunskapsförmedlande roll också en roll som urvalsinstrument. Svensklärare måste förhålla sig till den nya teknologin för att säkerställa värdet i de summativa bedömningarna som urvalsinstrument.

För att undersöka uppsatsens syfte används följande frågeställningar som hjälp:

I: Hur många lärare kan identifiera den datorgenererade texten bland de autentiska elevtexterna?

II: Hur kan Bachmans assessment use argument användas för att analysera validitetsproblemet som uppstått i samband med ChatGPT i den summativa skrivbedömningen?

III: Hur kan Bachmans assessment use argument användas för att analysera de lösningar på samma validitetsproblem som svensklärare har använt för att säkerställa validiteten i den summativa skrivbedömningen?

Dessa frågeställningar besvaras med hjälp av en intervjustudie där sju lärare intervjuas om både problemen med skrivbedömningen i förhållande till datorgenererad text, såväl som hur de har gjort för att lösa problemen. Intervjuerna börjar med ett test, där svensklärarna får läsa två elevtexter och en text skriven av ChatGPT. Testet fungerar som en språngbräda för intervjun, och visualiserar validitetsproblemet. Sedan frågas svensklärarna om hur ChatGPT är ett problem för valid skrivbedömning och hur de har förändrat sina bedömningsituationer i undervisningen för att garantera att eleverna skriver texterna. Testet där lärarna urskiljer elevtexten från den datorgenererade texten besvaras med en kort icke-generaliserbar

kvantitativ analys. Intervjuerna analyseras med hjälp av Bachmans assessment use argument för att hitta problemen men samtidigt också lösningarna i validitetsproblemet med summativ skrivbedömning.

## 3 Forskningsbakgrund

I forskningsbakgrunden presenteras verk från forskningsfälten artificiell intelligens och skrivbedömning. Först kommer artificiell intelligens och ChatGPT definieras och förklaras. Sedan beskrivs skrivbedömning som forskningsfält. För att förklara skrivbedömning beskrivs också validitet och summativ bedömning.

### 3.1 Artificiell intelligens och ChatGPT

För att beskriva och förklara vad ChatGPT är krävs en förklaring om vad artificiell intelligens, eller AI innebär. Artificiell intelligens är svårdefinierat från en filosofisk synvinkel. Ett provisoriskt svar om vad AI är kan ändå ges. AI är ett forskningsfält där program byggs upp för att kunna uppvisa beteenden. Dessa beteenden speglar det som vi människor anser innebära att inneha ett sinne. Artificiell intelligens är således ett konstgjort sinne, där det som vi människor kallar för intelligens återskapas. Frågan om vilka program som kan anses ha artificiell intelligens är filosofisk, eftersom intelligens och sinne kan definieras på olika sätt från olika filosofiska perspektiv (Arkoudas & Bringsjord 2014:34). Ett sätt att avgöra om huruvida ett program eller en maskin kan anses inneha artificiell intelligens beskrivs av Alan Turing (1950). Testet går ut på att en människa och en dator befinner sig i varsitt rum. I ett tredje rum finns en annan människa, som fungerar som en domare. Domaren kan inte se de två andra rummen och således inte i vilket rum människan respektive datorn befinner sig i. Domaren kan ställa frågor via textmeddelanden till både datorn och människan. Om domaren inte kan urskilja datorn från människan via svaren på textmeddelandena till en högre grad än 50%, anses datorn inneha artificiell intelligens (Turing 1950:433f; Arkoudas & Bringsjord 2014:34f).

Efter att artificiell intelligens har definierats kan ChatGPT beskrivas och förklaras. ChatGPT är ett program som kan interagera med människor på ett sätt som efterliknar ett samtal. När en människa interagerar med ChatGPT skapas en dialog, som tillåter ChatGPT att exempelvis svara på frågor och generera text på befallning. ChatGPT fungerar genom inmatningar som kallas för *prompts*. Dessa inmatningar resulterar i ett resultat, eller ett svar. Inmatningarna är således instruktioner, och ChatGPT:s syfte är att följa dessa instruktioner och generera ett resultat som svarar mot inmatningen. Grunden i ChatGPT är en språkmodell, som antingen

kan vara GPT-3,5 eller GPT-4 (OpenAI 2022). Eftersom GPT-4 är den nyaste modellen och använts i den föreliggande undersökningen beskrivs endast den här nedan.

GPT-4 är en multimodal storskalig språkmodell baserad på artificiell intelligens. Den kan hantera inmatningar i form av både bild och text. GPT-4 har testats på flera amerikanska akademiska prov, som exempelvis Uniform Bar Exam. UBE är ett prov som jurister i USA skriver för att bli advokater. På UBE presterade GPT-4 ett svar som bedömdes i nittionde percentilen, det vill säga ett svar som är bättre än 90% av alla andra svar på provet. GPT-4:s fick högsta betyg på de amerikanska gymnasiekurserna AP Environmental Science, AP Statistics, AP US History, och flera andra gymnasiekurser. På AP English Language and Composition fick GPT-4 dock bara 2/5 i betyg. Sammanfattningsvis uppvisar GPT-4 resultat som både är jämförelsebara med och överträffar mänskliga resultat (OpenAI 2023:1–6).

GPT-3, som är GPT-4:s föregångare, har tidigare testats i summativa bedömningssammanhang. Clark m.fl. (2021) skriver om hur väl människor identifierar skillnaden mellan mänsklig och datorgenererad text. Undersökningen genomfördes genom att ge fem textpassager till 780 respondenter. Därefter läste respondenterna texten, och gav sedan ett betyg på texten. Texterna handlade om olika ämnen, bland annat recept och nyhetsartiklar. Betygsskalan gick från 1–4, där 1 representerar att texten definitivt är skriven av en människa, 2 att texten förmodligen är skriven av en människa, 3 att texten förmodligen är skriven av en dator, och 4 att texten är skriven av en dator. Resultatet visade att när respondenterna skulle skilja mellan text skriven av GPT-3 och text skriven av en människa kunde de gissa rätt 49,9% av alla gånger (s.7282–7285).

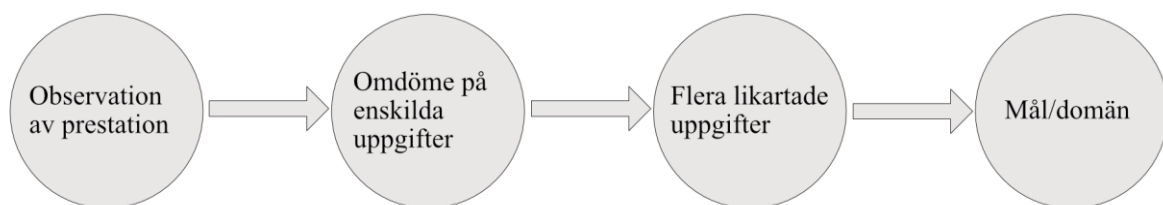
Det kan vara svårt att skilja mellan en universitetsstudents text och en datorskriven text. Därför är det komplicerat för universitetslärare att faktiskt bedöma hur väl en student kan svara på frågor rörande ett undervisningsområde (Cotton m.fl. 2023:3). GPT-3 tenderar att skriva på ett sätt som många universitetsstudenter har svårt för. GPT-3 hanterar genrer, refererar och tematiserar utan större problem. Däremot har GPT-3 svårt för att upprätthålla en argumentation i en längre text. GPT-3 har vidare svårt för grammatik och vanlig logik (Elkins & Chun 2020:3).

## 3.2 Validitet, summativ bedömning och skrivbedömning

Den föreliggande undersökningen ska behandla validiteten i gymnasieskolans summativa skrivbedömning i förhållande till ChatGPT. Därför krävs också en beskrivning om vad validitet, summativ bedömning och skrivbedömning innebär. Inledningsvis kommer begreppet validitet förklaras. Sedan beskrivs begreppet summativ bedömning. Till sist förklaras skrivbedömning.

### 3.2.1 Validitet

Validitet betyder giltighet, och i bedömningssammanhang innebär validitet således om bedömningen är giltig och träffsäker i relation till det som avses att mätas. En hög validitet innebär att bedömningen svarar mot skolans styrdokument. Bedömningen, och i förlängningen betyget, måste därför mäta det som styrdokumentet avser att mäta. (Klapp 2015:81). Det som styrdokumentet avser mäta kallas för *måldomän*, och brukar inom utbildning vara bred. När en elev skriver en text om exempelvis den litterära perioden romantiken bedöms inte bara elevens kunskaper om litteraturhistoria, men också dess kunskaper om grammatik och språkriktighet. Måldomänen är därmed det som styrdokumentet stipulerar att läraren ska bedöma, och den utgör ett sista steg i en bedömningskedja där flera olika prestationer vägs samman och resulterar i ett betyg. Denna bedömningskedja kan visualiseras med hjälp av en broanalogi (Kane m.fl. 1999:5–10).



Figur 1: Kanes m.fl. (1999) broanalogi. Översättning från engelska till svenska från Skar (2013).

I broanalogin finns tre pilar. Dessa pilar representerar *inferenser*, vilka fungerar som broar mellan de olika stegen i bedömningskedjan. En inferens är en slutledningsprocess som binder samman ett steg med ett annat. I bedömning kan en inferens bestå av steget från observation av en prestation till en tolkning av prestationen. Den första inferensen i broanalogin visar steget mellan en prestation och ett omdöme på en uppgift. Inferensens lämplighet baseras på två premisser. Den första premissen handlar om att prestationen har bedömts med lämpliga

kriterier och den andra premissen handlar om att prestationen har utförts under omständigheter som är förenliga med tolkningen av prestationen. Den andra inferensen kopplar samman omdömen på enskilda uppgifter med ett mer generaliserat omdöme. Detta baseras på flera uppgifter. Således uppnås en mer generaliserad tolkning av elevens prestation. Den tredje inferensen binder samman tolkningen av flera likartade uppgifter med måldomänen. En stark sådan inferens förutsätter att de likartade uppgifterna träffar och täcker hela måldomänen (Kane m.fl. 1999:9f). Inferenserna som Kane använder i broanalogin kommer sedan att ses i Bachmans (2005) modell AUA.

Inferenser och validitet kan kombineras i begreppet *construct validity*, som kan förstås som all data och bevis för en elevs förmåga som behövs för att skapa inferenser. Validiteten skapas av att det finns nog med relevant teoretiskt och empiriskt stöd för att bedöma ett konstrukt på ett träffsäkert vis. Ett konstrukt är ett på förhand definierat begrepp som validiteten ska testas på. (Huot, O'Neill och Moore 2010:506). Ett skrivkonstrukt är således en på förhand definierad uppfattning om vad skrivande innebär. Inom skrivbedömning krävs ett väldefinierat skrivkonstrukt för att kunna bedöma en elevs skrivkompetens med en hög validitet (Evensen 2010:14).

För att kunna avgöra vad som ingår i en måldomän måste bedömaren vara noga med att definiera vad måldomänen innebär. Om elevens skrivförmåga ska bedömas måste skrivkonstruktet således vara definierat. I skrivbedömningens fall finns följande formulering i Skolverkets (2023) ämnesplan för ämnet Svenska.

Eleven kan skriva argumenterande text och andra typer av texter, som är sammanhängande och begripliga samt **till viss del** anpassade till syfte, mottagare och kommunikationssituation. Eleven kan i huvudsak följa skriftspråkets normer för språkriktighet.

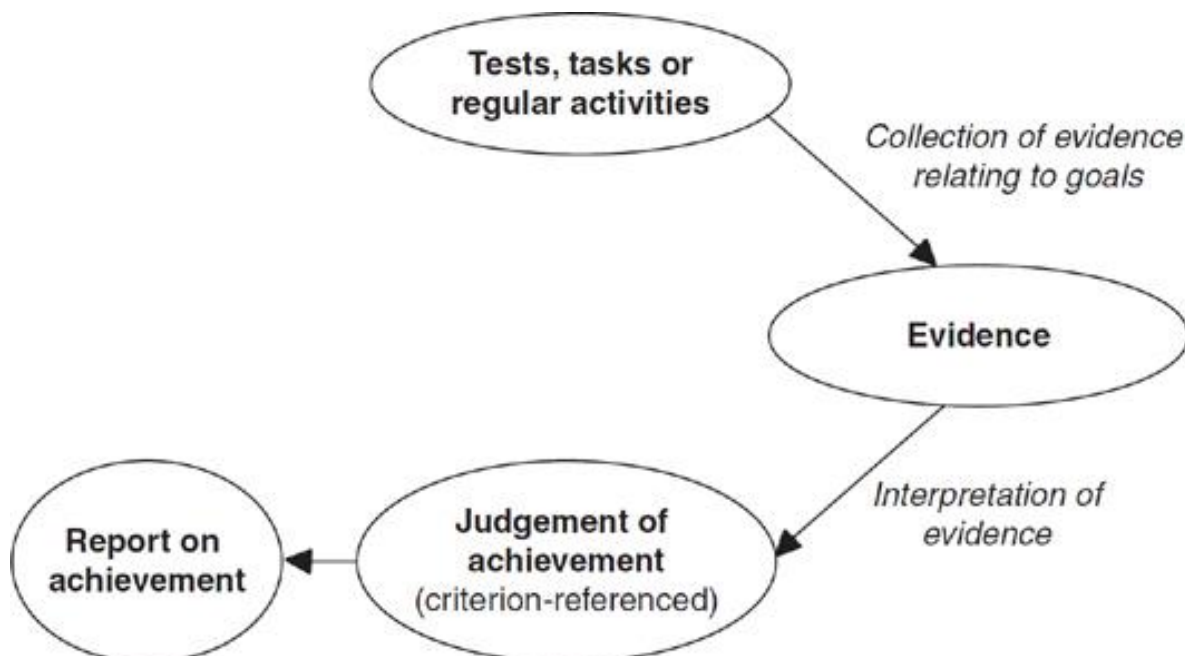
Eleven kan läsa, reflektera över och göra **enkla** sammanfattningar av texter samt skriva egna texter **som anknyter till det lästa**. I sitt arbete värderar och granskar eleven **med viss säkerhet** källor kritiskt samt tillämpar grundläggande regler för citat- och referatteknik.

Formuleringarna kommer från betygsriterierna för E-nivå i kursen Svenska 1. I måldomänen får det definierade skrivkonstruktet anses vara tämligen brett. Det står inte i detalj vad eleverna gör när de skriver, och hur skrivkonstruktet ska vara definierat för den bedömande läraren. Den bedömande läraren definierar skrivkonstruktet i bedömningen på egen hand.

### 3.2.2 Summativ bedömning

Bedömning kan vara antingen formativ eller summativ. Vid formativ bedömning är syftet att ge feedback till eleverna för att främja elevernas lärande. Den summativa bedömningens syfte är i stället att summera elevernas kunskaper, och kan exempelvis bestå av betyg. Den summativa bedömningen, och betygen som den genererar, är en central del av statens utvärderingsverktyg. De används både till förbättringar av skolsystemet och för att skapa en urvalsgrund för vidare studier (Klapp 2015:16).

Harlens (2012) modell illustrerar den summativa bedömningen på ett överskådligt sätt. Summativ bedömning består av fyra olika delmoment. Momenten illustreras nedan med hjälp av figur 2.



Figur 2: Harlens (2012) modell för summativ bedömning.

Det första steget i den summativa bedömningen är ett test eller ett prov. Detta leder till insamlande av bedömningsunderlag i relation till aktuella mål och styrdokument.

Bedömningsunderlaget tolkas sedan mot styrdokumentet, alltså mot den aktuella måldomänen, varpå ett omdöme eller ett betyg kan produceras (s.7f).

### 3.2.3 Skrivbedömning

En teoretisk grund till skrivbedömning kan hittas i begreppet *performance assessment*. Det är ett ord som i sig är svårt att tolka. Rimligtvis grundas de allra flesta bedömningar inom skolväsendet på en prestation. Det är dock inte det som ordet betyder. En prestationsbedömning är beroende av att själva bedömningen faktiskt mäter den prestation som den avser att mäta. Prestationsbedömningen måste mäta rätt kunskap, eller rätt konstrukt (Kane m.fl. 1999:5f). Observera att prestationsbedömning inte är unik för skrivbedömningen. Den summativa bedömningen som beskrivits tidigare är allmäntdidaktisk i det avseende att den går att applicera på alla ämnen i den svenska skolan. Nu till skrivbedömning. Följande kapitel syftar till att ge en överblick över skrivbedömningen som forskningsfält.

Skrivbedömning är en bedömning av elevers skrivförmåga, som svarar mot lärarens uppfattning om skrivkonstrukten. Skar (2013) menar att det finns två huvudsakliga aspekter i skrivbedömning. Den första aspekten handlar om skrivbedömning som ett redskap. Skrivbedömning är ett sätt att ta reda på vad elever kan, genom exempelvis prov och bedömningsuppgifter. Målet är att utvärdera elevens kunskaper och sätta ett betyg. Den andra aspekten handlar om betygen som fastslås i den första aspekten. Dessa betyg, eller mått på kunskap inom skrivande, används för att jämföra exempelvis elever med varandra och skolor med andra skolor. Dessa mått, som baseras på betygen, används som urvalsfaktorer när elever söker till högre utbildning. Skars avhandling undersöker primärt den första aspekten av skrivbedömning (Skar 2013:4f).

Skar (2013) skriver i sin avhandling om skrivbedömning och validitet i gymnasieskolan. Avhandlingen utgår från tre explorativa undersökningar av validitet i skrivbedömningen. Skar (2013) använder Bachmans modell AUA för att analysera svensklärares skrivbedömning. Tre lärare intervjuas och deras bedömningspraktik analyseras. Skar (2013) diskuterar tre huvudsakliga hot mot validiteten i den summativa skrivbedömningen. För det första finns hotet mot tolkningens validitet. Skar (2013) formulerar fem hypoteser inom detta hot. Hypotes A handlar om att lärare inte förklarar vad provpoäng innebär. Därmed är läraren inte

helt säker på vad som har bedömts. Hypotes B behandlar svårigheten med att elevens uppfattning om vad som bedömts inte stämmer överens med lärarens. Eleverna kan därför förlora förtroendet för lärarens bedömning. Hypotes C redogör för möjligheten om att lärarens uppfattning av elevens skrivförmåga inte stämmer överens med skrivkonstrukturen som avses att bedömas. Hypotes D innebär att det finns en diskrepans mellan skrivkonstrukturen som avses att bedömas och bedömningskriterierna. Hypotes E menar att lärarens skrivbedömning inte svarar mot tolkningen av elevsvaret. Det finns därmed en irrelevant varians i bedömningen (s.190f).

Skar (2013) menar att det andra hotet mot validiteten ligger i användningsargumentet i Bachmans assessment use argument. Även där formuleras hypoteser. Hypotes F står för hotet om dålig kommunikation mellan läraren och klassen om hur bedömningsunderlag ska användas. Hypotes G står för en diskrepans mellan bedömningsunderlag och användandet av den i relation till mätningen, och hypotes H att elever ibland inte är förberedda på att fatta beslut om deras utbildning (s.192f). Skar (2013) skriver fram det tredje hotet som ett hot mot likvärdigheten, som trots att det tas upp i Bachmans modell behandlas separat. Hypotes I handlar om att lärarens bedömning av elevernas texter inte är likvärdig (s.193).

En annan svensk avhandling om skrivbedömning är Blomqvist (2018). Den andra artikeln i avhandlingen handlar om svensklärares beslut om summativ bedömning, och de normer som värderas högst när elever bedöms i svenska 1. Resultatet som framkom visade att lärarna såg elevtexters kvalitet i relation till betygskriterierna. Textens kvalitet måste genomgående motsvara ett betygsstegs kriterier för att kunna bedömas med det betyget. Därmed skapas en tendens av att svensklärarna snarare tittar på texternas brister än dess styrkor. Vidare fanns flera kompensatoriska aspekter som kunde väga upp eventuella brister i texternas kvaliteter. De kompensatoriska aspekterna fick en stor inverkan på den summativa bedömningen (Blomqvist 2018:51).

Aldrin (2015) skriver om hur praktiserande svensklärare ser på textbedömning. Resultatet visar att majoriteten av lärarna menar att det är medelsvårt att bedöma texter. Elevtextbedömning är svårast för lärarna som arbetat kortast och även om värdet som indikerar detta inte är statistiskt signifikant kan det stödjas med internationella forskningsresultat. Bedömningens svårighet kopplas till kritiken mot svensklärarytbildningarnas bedömningsträning. Lärarna tillfrågas också om de ser några särskilda problem med bedömning av elevtexter. Den mest vanligt förekommande svårigheten

är att dra gränser mellan olika betyg. Alla betygssteg anses vara lika svåra att skilja mellan, vilket kan bero på att enkäten som undersökte lärarnas svar utkom när det nya betygssystemet som introducerades vid LGY11 inte hade använts så länge. En faktor som underlättar bedömning av elevtexter menade 50% av lärarna i undersökningen var sambedömning (Aldrin 2015:68–72). Aldrin (2015) menar vidare att lärarna fick en text tilldelad, och ombads bedöma den samt motivera varför. I analysen av de bedömningar som svensklärarna rapporterade upptäckes en relativt stor variation i bedömningen. Variationen innebar en skillnad på nästan ett helt betygssteg. Det finns vidare ingen majoritetsuppfattning i omdömet, utan svaren i enkäten visar nästan jämnt antal procent på E, D, och C. Undersökningen visar att lärarens sociala kontext, erfarenhet och miljö påverkar bedömningen (Aldrin 2015:8–81).

En äldre svensk avhandling är Östlund-Stjärnegårdh (2002) som skriver om bedömning av gymnasieelevers texter i svenska. Resultatet visar att det finns en tydlig varians i bedömningen av texter. Reliabiliteten mellan bedömare är relativt låg. Vidare visar resultatet att ju längre text, desto högre betyg (Östlund-Stjärnegårdh 2002:94f). Liknande resultat återfinns också i Hultman & Westman (1977).

### **3.2.4 Sammanfattning av forskningsbakgrunden**

Forskningsfälten som beskrivits är betydligt större än det som sammanfattas i forskningsbakgrunden. Målet med forskningsbakgrunden är snarare att sätta uppsatsen i perspektiv till det som redan har skrivits inom de flera olika forskningsfält som uppsatsen rör sig inom. AI och undervisning i kombination är något nytt, och det märks i de litteratursökningar som har utförts.

## 4 Teori

### 4.1 Inledning

I det här avsnittet presenteras den teori som använts för att analysera resultatet i undersökningen. Först beskrivs den teoretiska grunden för den första frågeställningen. Detta sker under rubrik 4.2. Sedan beskrivs den teoretiska grunden som används för den andra samt tredje frågeställningen, detta sker under rubrik 4.3.

En premiss som antogs som självklar innan undersökningen påbörjades var att elever på gymnasiet i de flesta fall skriver texter på en dator. Dessa texter får skrivas på lektionstid, men också på annan plats. Texterna bedöms summativt av svenskläraren.

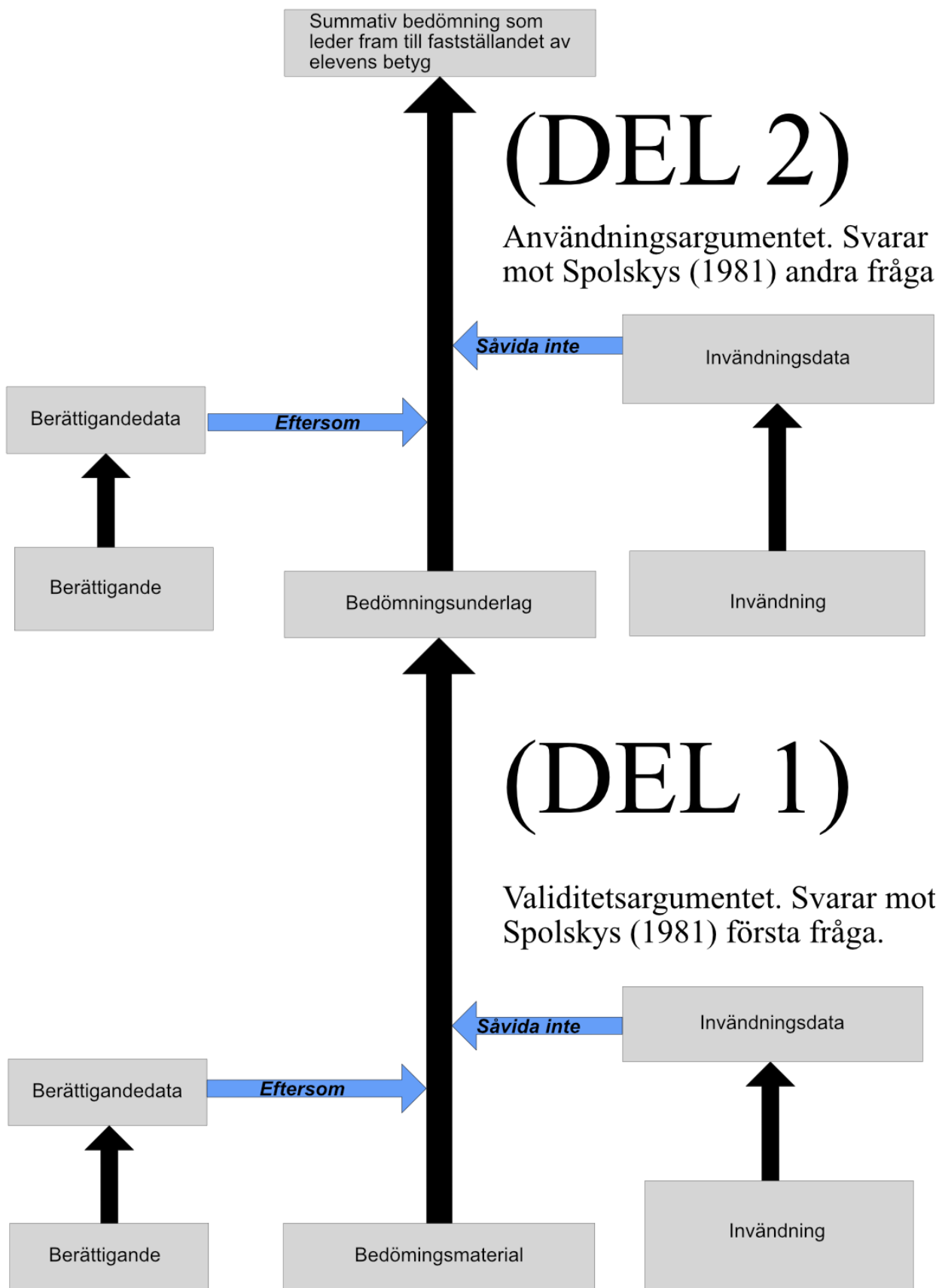
### 4.2 Turingtestet

*The Imitation Game* är ett test för att avgöra om en maskin eller en programvara kan anses inneha artificiell intelligens. Testet är utformat av Alan Turing år 1950. Testet går ut på att låta en människa gå in i ett rum med endast en teleprinter, alltså en sorts elektrisk skrivmaskin som kan skicka och ta emot meddelanden. I ett andra rum finns en dator som är utrustad med en teleprinter. I ett tredje rum finns ytterligare en människa som också har tillgång till en teleprinter. De tre parterna kan kommunicera med varandra anonymt. Det är den första människans uppdrag att avgöra vem av de två andra som är en dator och vem som är människa (Turing 1950:433f). Om den första människan inte kan skilja mellan datorn och människan till en högre grad än 50% anses datorn inneha artificiell intelligens (Arkoudas & Bringsjord 2014:35). Föreliggande undersökning ska utföra en variant av detta test, eftersom den ställer anonyma texter från både elever och en text från en textgenerator mot gymnasielärare i svenska. Läraren ska sedan försöka urskilja textgeneratorns text från elevtexterna, och därmed avgöra om ChatGPT kan anses inneha artificiell intelligens. Testets verkliga namn är *The Imitation Game*, men för enkelhetens skull kallas testet härnäst för Turingtestet. Testet används för att besvara den första frågeställningen.

### 4.3 Bachmans assessment use argument

Bachmans (2005) modell *assessment use argument*, förkortat AUA, är en central teori för undersökningen. Den används för att svara på både den andra och den tredje frågeställningen. Modellen är konstruerad för att kunna användas av yrkesgrupper som behöver en teori för att skapa validitet i bedömningen av ett konstrukt. I föreliggande undersökning är det specifikt skrivkonstruktet som lärarna definierar och validerar. Modellen ska kunna användas praktiskt, och är utvecklad specifikt för språkbedömning.

AUA:s teoretiska grund kommer från Spolskys (1981) två frågor om provanvändning. Båda de frågorna modifieras och kombineras i en modell som konstrueras av Bachman. Spolskys (1981) frågor omformuleras till följande frågor: ”Hur övertygande är anledningen att använda bedömningen på det här viset?”, samt ”Hur trovärdigt är beviset som stödjer det här argumentet?” (Bachman 2005, s.5). Bachman (2005) använder vidare Toulmins (2003) diagram om validitetsargumentation som grund för *assessment utilization argument*. Tillsammans bildar Toulmins validitetsargument (2003) och Bachmans (2005) *assessment utilization argument* den sammanfogade modellen *assessment use argument*. Toulmins (2003) diagram utgör alltså validitetsargumentet i strukturen, och Bachmans (2005) *assessment utilization argument* utgör användningsargumentet. Tillsammans utgör de en komplett kedja av inferenser från bedömningen av ett individuellt test, i det här fallet en elevtext, till den slutgiltiga summativa bedömningen (Bachman 2005:14). Spolskys första fråga svarar således mot den första delen i Bachmans *assessment argument*, alltså validitetsargumentet. Se del 1 i figur 3 nedan. Spolskys andra fråga svarar mot den andra delen i Bachmans AUA, alltså användningsargumentet. Se del 2 i figur 3 nedan. I figur 3 har Bachmans AUA anpassats till den här undersökningen. En mer detaljerad beskrivning av hur *assessment use argument* fungerar följer här nedan.



Figur 3: Bachmans (2005) assessment use argument (AUA). Översättningen av modellen från engelska till svenska är hämtad från Skar (2013)

### 4.3.1 Validitetsargumentet

Validitetsargumentet utgör den första halvan av assessment use argument och svarar mot Spolskys (1981) första fråga. Det syns som del 1 i figur 3 ovan. Modellen börjar i bedömningsmaterial, som Bachman (2005) kallar för assessment performance. I validitetsargumentet ska bedömningsmaterialets validitet testas för att se om det kan användas som bedömningsunderlag (Bachman 2005, s.9f). Bedömningsmaterialet i undersökningens sammanhang är elevtexter, och dessa testas om de kan vara valida nog att användas som bedömningsunderlag.

I validitetsargumentet finns en eller flera *warrants*, eller berättigandedata, vilka är argument för att använda bedömningsmaterialet som bedömningsunderlag. Se vänster sida i figur 3 i del 1. Det är berättigandedatan som möjliggör inferensen, som representeras av den stora pilen i mitten av modellen. Inferensen i det här sammanhanget utgör steget från att se elevtexter som bedömningsmaterial till att se dem som bedömningsunderlag. Berättigandedata grundar sig i *backing*, eller berättigandet. (Bachman 2005, s.9f). I undersökningens sammanhang kan berättigandet exempelvis vara att elever oftast skriver sina texter utan hjälp från andra personer. Berättigandedatan blir då att den enskilda eleven också förmodligen skrivit sin egen text, utan hjälp från andra personer. Berättigandedatan skapar ett argument för validiteten.

I diagrammet finns också *rebuttals*, eller invändningar, som består av alternativa förklaringar till varför hot mot validiteten skulle kunna uppstå. *Rebuttal data*, eller invändningsdata, består av bevis som stärker anledningen till att validiteten kan hotas (Bachman 2005, s.9f). I undersökningens sammanhang skulle en invändning kunna vara att eleven har haft tillgång till ChatGPT, och invändningsdata kan då bli att läraren inte kan veta om eleven har skrivit texten utan assistans från ChatGPT.

### 4.3.2 Användningsargumentet

Nästa steg i Bachmans (2005) AUA är *assessment utilization argument*. Denna modell undersöker användandet av den data som framkommit i den första delen av modellen. Assessment utilization argument svarar mot Spolskys (1981) andra fråga. Den andra delen av modellen börjar i slutet av den första. Det är med andra ord bedömningsunderlaget som är

data i den andra modellen. (Bachman 2005:18ff). Denna data testas om den kan användas för en summativ bedömning av elevens skrivförmåga

I användningsargumentet finns precis som i validitetsargumentet berättigandedata. Dessa grundar sig i berättiganden (Bachman 2005:18ff). I den här undersökningens sammanhang kan ett berättigande vara att läraren har konstruerat en skrivuppgift som svarar mot måldomänen, och berättigandedata att elevens text också då bör svara mot måldomänen.

Vidare finns precis som i validitetsargumentet också rebuttals, eller invändningar i *assessment utilization argument*. Det finns två typer av invändningar. Den första typen behandlar en anledning av att inte använda tolkningen för att göra anspråket eller ta beslutet. Den andra typen av invändning representerar de eventuella oanade konsekvenser som kan uppstå när tolkningen används för att göra anspråk på något (Bachman 2019:8ff). En invändning i den föreliggande undersökningen skulle kunna vara att läraren inte har kontrollerat om eleverna har använt ChatGPT, och invändningsdatan blir då att ChatGPT har påverkat validiteten.

När de här två modellerna sätts samman menar Bachman (2005) att de bildar ett *assessment use argument*. Se figur 3 ovan. Det är ett argument som modellerar en grund för både provdesign och tolkningen av provets svar. Bachman (2005) menar vidare att det är viktigt att poängtera att även fast AUA innehåller både ett användningsargument och ett validitetsargument, är argumenten inte likvärdiga. Tolkningen av endast en elevtext leder mycket sällan enbart till en slutgiltig summativ bedömning (s.24). Därmed måste validitetsargumentet ses som ett av många, och att det inte finns ett jämnt förhållande mellan validitetsargument och användningsargument.

### **4.3.3 Användandet av Bachmans *assessment use argument* i den här undersökningen**

Följande scenario kan användas för att förstå Bachmans modell bättre. En svensklärare låter eleverna i en klass skriva utredande texter. Efter några lektioner lämnas 30 elevtexter in. Detta representerar inledningen i validitetsargumentet. Svenskläraren måste då avgöra om dessa texter är valida att användas som bedömningsunderlag. Det finns argument för och mot att använda texterna som bedömningsunderlag. Generellt sett brukar eleverna skriva sina egna

texter. Läraren har dessutom sett eleverna skriva texterna på lektionen, även om det är svårt att hålla koll på 30 elever samtidigt när de skriver. Utöver det får eleverna skriva hemma. Detta representerar berättigandet i validitetsargumentet. Motargumentet skulle kunna vara att några av eleverna har fått hjälp hemifrån att skriva texterna. Någon av eleverna har kanske betalat någon att skriva den utredande texten för eleven. Även om svenskläraren inte rimligtvis kan anses vara tvungen att kontrollera det här problemet, representerar det ändå invändningen i validitetsargumentet. Om berättigandedata väger tyngre än invändningsdata anses bedömningsunderlaget inneha en god validitet. Svenskläraren kan i så fall anse att texterna är valida och kan användas som bedömningsunderlag.

Nästa del av Bachmans modell, alltså användningsargumentet kan därmed påbörjas. Validiteten i huruvida bedömningsunderlaget kan användas i en summativ bedömning ska prövas. Med andra ord undersöks bedömningsunderlagets validitet i relation till den måldomän som avses att undersökas. Grunden för användningsargumentet är bedömningsunderlaget. Precis som i validitetsargumentet finns argument för och mot. Svenskläraren kan hävda att uppgiften och texterna svarar väl mot styrdokumentet. Det är i så fall ett berättigande. På samma sätt som i validitetsargumentet vägs berättiganden mot invändningar. Ett motargument, och därmed en invändning kan vara att svenskläraren inte kan garantera att några elever inte har fuskat. Därmed kan inte bedömningsunderlaget anses vara korrekt och kan inte användas för summativ bedömning. Svenskläraren måste väga invändningar mot berättiganden och ta ett beslut om bedömningsunderlaget ska kunna anses ha en god validitet och därmed kunna användas i summativ bedömning.

## 5 Metod och material

I det här kapitlet beskrivs metoden som har använts för att samla in materialet i undersökningen. Först kommer intervjumetoden att beskrivas. Sedan beskrivs urval och de val som gjorts i samband med det. Efter att urvalet förklarats beskrivs hur intervjuerna har genomförts. I samband med det beskrivs också texterna som använts i undersökningen. Kapitlet avslutas med en beskrivning av det etiska förhållningssättet som har applicerats på uppsatsen samt en genomgång om hur litteratursökningen i forskningsöversikten har gått till.

### 5.1 Intervjumetod

I den föreliggande undersökningen har en intervjuundersökning använts för att samla in materialet. Samtliga frågeställningar besvaras med hjälp av intervjuundersökningen. Materialet i undersökningen består av semistrukturerade intervjuer. Christofferson & Johannesen (2015) menar att den semistrukturerade intervjun har fördelen att resultera i fylligare svar från informanterna än standardiserade intervjuformer. Vidare får intervjun en ökad flexibilitet. Nackdelarna med semistrukturerade intervjuer är att kodningen blir svårare och att flexibiliteten inte är lika hög som vid ostrukturerade intervjuer (Christoffersen & Johannesen 2015:5). Anledningen till att en semistrukturerad intervjuguide valdes som metod för resultatinsamling var att ChatGPT är ett nytt fenomen i den svenska skolan. Således krävdes en relativt öppen ingång för att skapa möjlighet till en bredd i svaren. Lärarna kan ha haft vitt skilda upplevelser av ChatGPT, och intervjun anpassades efter det. Samtidigt var undersökningens analysmetod redan bestämd innan intervjuerna påbörjades. Det ledde till att intervjuformen och intervjuguiden inriktades mot en specifik analysmetod. Därav valdes semistrukturerad intervjuform, eftersom formen både ger möjlighet för informanterna att uttrycka sig på ett fritt vis lämpligt för ett nytt fenomen, samtidigt som förberedda frågor kan ställas. Dessa frågor var konstruerade för att svara mot analysmetoden.

#### 5.1.1 Intervjuguide

Intervjuguiden återfinns som bilaga 4 och skrevs enligt Christofferson & Johannesens (2015) anvisningar. De menar att forskaren bör ställa inledande frågor som är lätta att svara på i början av intervjun. Informanten får inte provoceras (Christofferson & Johannesen 2015:86f).

I intervjuguiden användes frågor om informanternas utbildning och vilka ämnen de undervisar i som inledande frågor. Nyckelfrågorna bör utgöra den största delen av en kvalitativ intervju. Halva tiden läggs därför åt dessa frågor. Om forskaren ställer känsliga frågor bör forskaren vara försiktig, och beakta informantens känslor. Till sist ställs de avslutande frågorna, där informanten får möjlighet att uttrycka sig på ett mer fritt sätt. Forskaren bör förvarna om att intervjun närmar sig sitt slut genom att säga till i slutet av intervjun, exempelvis genom att indikera att det bara är två frågor kvar i intervjun (Christoffersen & Johannesen 2015:86).

## 5.2 Urval

I den föreliggande undersökningen har ett explorativt urval använts eftersom undersökningen utforskar ett ämne som hittills är relativt outforskat. I ett explorativt urval är det inte nödvändigt att få ett tvärsnitt av befolkningen. Därför har inte informanterna i undersökningen sorterats efter ålder, kön eller erfarenhet i yrket. Ett bekvämlighetsurval har använts för att välja ut informanter. Det finns flera nackdelar med bekvämlighetsurval, som risken att forskaren väljer informanter enbart utifrån bekvämlighet (Denscombe 2015:71f). Så var emellertid inte fallet i den föreliggande undersökningen. På grund av att undersökningen utfördes på våren var det svårt att få tag på informanter. Detta beror på lärares arbetsbörda i skolan under den här tiden, eftersom de nationella proven genomfördes. Av 75 informanter svarade 7 att de kunde medverka på en intervju. Många potentiella informanter kontaktades, men bara ett fåtal kunde medverka. Därav fanns inget utrymme för att sålla bort informanter.

Empirin består av sju intervjuer. Svensklärarna kontaktades via mejl. Intervjuerna spelades in och transkriberades sedan manuellt i Microsoft Word utan dikteringshjälp. Informanterna har arbetat mellan 5 och 30 år i gymnasieskolan. I samband med att svensklärarna bjöds in till att medverka i studien bifogades ett informationsbrev som informerade om att informanterna har möjlighet till att dra sig ur studien när som helst. Se bilaga 5.

De sju intervjuerna är mellan 24 och 58 minuter långa. Samtliga informanter svarade på alla frågor.

Tabell 1: Information om intervjuerna

Intervju nr	Längd i minuter	Svarat på alla frågor
1	24	Ja
2	58	Ja
3	45	Ja
4	35	Ja
5	37	Ja
6	45	Ja
7	28	Ja

De sju informanterna hade mellan 5 och 30 års erfarenhet i läraryrket. De hade alla svenska i sin ämneskombination. Samtliga lärare hade språk- eller humanistiska ämnen utöver svenska i sin ämneskombination.

Tabell 2: Information om informanterna

Informant nr.	År i läraryrket	Ämnen	Skolform
Informant 1	5 år	Svenska, spanska	Kommunal skola
Informant 2	25 år	Svenska, religion	Kommunal skola
Informant 3	26 år	Svenska, engelska	Kommunal skola
Informant 4	30 år	Svenska, tyska	Kommunal skola
Informant 5	28 år	Svenska, engelska	Kommunal skola
Informant 6	8 år	Svenska, historia	Kommunal skola
Informant 7	11 år	Svenska, historia	Kommunal skola

I resultatet kommer tre begrepp att användas för att indikera hur många informanter som anser något. Begreppet *en informant* motsvarar just en informant. *Flertalet informanter* motsvarar två till sex informanter. *Samtliga informanter* innebär att alla informanter tycker samma sak.

## 5.3 Genomförandet av intervjuerna

### 5.3.1 Turingtestet

Efter att de inledande frågorna ställts till informanterna ställdes de inför Turingtestet. Informanterna fick tre texter och en uppgiftsinstruktion. Texterna återfinns som bilagor 1–3 i uppsatsen. Informanterna ombads läsa dessa. Tiden som informanterna använde för att läsa texterna sträckte sig mellan två och sex minuter. Efter att informanterna läst texterna fick de svara på om de tyckte att någon text var annorlunda än de andra. Därefter ombads de identifiera texten som ChatGPT skrivit.

Ett problem med att utföra testet på det här viset är att autenciteten i bedömningsituationen minskar med det faktum att informanterna inte känner eleverna som de bedömer. Vanligtvis har gymnasielärare observerat sina elevers prestationer under en längre tid. Det innebär att de är medvetna om sina elevers sätt att skriva och formulera sig. Därför kan en eventuell skarp gräns urskiljas mellan en elevs tidigare prestation utan ChatGPT och en elevs prestation med ChatGPT.

### 5.3.2 Bachmans assessment use argument

Enligt Bachmans (2005) grundmodell applicerades senare svensklärarnas tankar om ChatGPT:s effekt på validiteten i undervisningen. Modellen konstruerades från validitetsargument till användningsargument, som Bachman (2005) menar är ett sätt att undersöka validitet i tolkningen och användningen av en bedömning (Bachman 2005:25).

### 5.3.3 Turingtestet och Bachmans assessment argument i kombination

Turingtestet ger värdefull information om huruvida svensklärare kan identifiera en datorskriven text bland elevtexter, men erbjuder framför allt en språngbräda för vidare diskussion om validitet i skrivbedömningen. Eftersom ChatGPT fortfarande är en ny företeelse i skolan är inte alla lärare medvetna om dess kapacitet. Därför krävs det att ChatGPT:s förmågor demonstreras innan en diskussion om validitet i skrivbedömningen kan påbörjas. Diskussionen analyseras med hjälp av Bachmans AUA. De två teoretiska

förhållningssätten är därmed sammankopplade – Turingtestet skapar en grund för assessment use argument.

### 5.3.4 Beskrivning av texterna

Text 1 och 3 är autentiska elevtexter som skrevs 2015 som en del av nationella provet i Svenska 1. De bedömdes med betyget A. Anledningen till att de två elevtexterna valdes ut var för att de innehöll relativt få stavfel och grammatiska fel. ChatGPT gör vanligtvis inga stavfel eller grammatiska fel såvida den inte instrueras att göra det. Texterna kommer från Uppsala universitets korpus över nationella provet i ämnet Svenska. Skälet till att 2015 års nationella prov användes var för att det är det senaste provet vars uppgiftsinstruktion inte längre är sekretessbelagd. Anledningen till att ett nationellt prov användes var för att få en standardiserad och för testets skull neutral uppgiftsinstruktion. Text 1 är 476 ord lång och text 3 är 592 ord lång. Texterna skrevs ut på A4-papper med typsnittet Times New Roman, typstorlek 12 och radavstånd 1,5. Detta resulterade i att båda texterna var lite mer än en sida långa vardera. Elevtexterna kan inte bifogas i sin helhet eftersom de är belagda med sekretess. Texterna fick visas för informanterna, men får inte publiceras. Däremot får utdrag ur texterna publiceras. Vid opponeringen visas texterna för examinator. Inför opponeringen får opponenter tillgång till texterna. För utdrag ur elevtexterna, se bilaga 2 och 3.

Text 2 är en text skriven av ChatGPT. Texten genererades med ChatGPT-4, vilket vid tidpunkten för genereringen var den nyaste versionen av ChatGPT. För att ha tillgång till ChatGPT-4 krävs en prenumeration som kostar 25 amerikanska dollar i månaden. Faktumet att en betaltjänst användes för att generera texten skulle kunna innebära att trovärdigheten för testet minskar. Det är förmodligen mindre troligt att en elev skulle betala för en så relativt dyr tjänst jämfört med att använda alternativet som är gratis. Anledningen till att ChatGPT-4 användes var för att utvecklingen snabbt går framåt, och det är därför möjligt att ChatGPT-4 är gratis eller mer lättillgängligt inom kort. Vidare framstod det som relevant att testa informanterna med den senaste tekniken för att illustrera ChatGPT:s fulla potential. Texten är 377 ord lång (se bilaga 1). Uppgiftsinstruktionen var utförlig, vilket är en fördel när ChatGPT instrueras. Detta tillät möjligheten att bara kopiera in uppgiftsinstruktionen i ChatGPT, utan att ge den allt för många egna instruktioner. På så sätt influeras inte instruktionerna av forskarens kunskaper om text, utan en elev hade kunnat ge samma instruktion. Den enda

ytterligare instruktionen som ChatGPT gavs var den om att skriva som en gymnasieelev, och att göra några fel. På så sätt genererades en text som var mer lik den som en elev kan tänkas skriva i gymnasieskolan.

ChatGPT gavs följande *prompt*, eller instruktion för att generera text 2. Först angavs uppgiftsinstruktionen som också givits till eleverna när de skrev provet 2015. ChatGPT genererade då en text, och denna text formaterades i Microsoft Word. Precis som elevtexterna formaterades text 2 till Times New Roman, typstorlek 12 och radavstånd 1,5.

I skolan och arbetslivet är det mycket vanligt med muntliga presentationer, men många människor känner obehag inför sådana situationer. Detta kan bli ett problem i ett modernt informationssamhälle, skriver Peter Letmark i artikeln ”Därför blir vi skräckslagna” (Dagens Nyheter 4.5.2012). Varför tycker många att det är jobbigt att tala inför andra? På en webbplats om kommunikation diskuteras frågan om talängslan. Skriv ett inlägg som kan publiceras på webbplatsen. Syftet med inlägget är att väcka intresse i frågan och ge läsarna perspektiv på ämnet. Resonera kring varför muntliga presentationer kan upplevas som obehagliga. Diskutera hur talängslan kan arbetas bort. Presentera något från Letmarks artikel som är relevant för ditt resonemang. Tänk på att dina läsare inte har läst artikeln. Rubrik: Inför allas blickar. Skriv som en gymnasieelev och gör några fel.

## 5.4 Analys av intervjuer

### 5.4.1 Tematisk analys

Analysen av intervjuerna utfördes med metoden *thematic analysis*. Hädanefter kommer metoden kallas för tematisk analys. Det finns sex steg i tematisk analys. Det första steget handlar om att bekanta sig med data som insamlats. Intervjuerna transkriberas ortografiskt och anteckningar förs över initiala tankar om intervjumaterialet. I det andra steget skapas koder, alltså gemensamma fragment i intervjuerna. Informanterna måste vara rörande överens om något för att det ska räknas som en kod. All data arbetas genom systematiskt och samtliga koder markeras i materialet. I det tredje steget påbörjas sökandet efter teman, där koderna

används för att stödja övergripande teman. Koderna kombineras och formar tillsammans teman, eller subteman. Subteman är teman inom teman. Subteman kan sedan ingå i större teman. Tankekartor används med fördel för att visualisera temana. I det fjärde steget granskas temana. Vissa teman är kanske inte hållbara, andra är ologiska. Temana ska vara logiska och ska passa ihop. I steg fem ska samtliga teman namnges och tydligt definieras. Grunden eller essensen i temat ska kopplas till en aspekt i data. Varje tema ska också analyseras, temana ska kopplas till den övergripande frågeställningen. I steg sex ska den tematiska analysens rapport skrivas. Läsaren ska övertygas om validiteten i analysen. Data ska inte bara beskrivas – utan argumenteras för (Braun & Clarke 2006:87–93). I analysen av intervjuerna har de sex stegen utförts.

## 5.5 Etiskt förhållningssätt

Det finns fyra etiska riktlinjer som bör betäckas innan en intervjuundersökning påbörjas. Dessa är *informerat samtycke*, *konfidentialitet*, *konsekvenser* och *forskarens roll* (Kvale & Brinkmann 2014:107–111). Samtliga riktlinjer har använts i den här undersökningen för att skapa en etiskt hållbar studie. Riktlinjen *informerat samtycke* har mötts genom att en samtyckesblankett delats ut till informanterna innan intervjun påbörjades. Vidare blev informanterna informerade om studiens syfte via e-post när de initialt kontaktades. Riktlinjen *konfidentialitet* har mötts genom att informanternas identitet har hemlighållits. Riktlinjen *konsekvenser* har beaktats genom att potentiella konsekvenser av informanternas deltagande i studien har uppmärksamats och bemötts i undersökningens utformning. Riktlinjen *forskarens roll* har bemötts genom att eftersträva vetenskaplig kvalitet och korrekta resultat.

## 5.6 Litteratursökning i forskningsbakgrunden

Källorna som används i kapitlet Forskningsbakgrund har hittats via en kombination av två sökmetoder. Först användes sökmotorerna Supersök och Google Scholar med sökorden ”validitet” + ”summativ”+ ”skrivundervisning”. Sedan kedjesöktes dessa källor tills de aktuella forskningsfälten kunde beskrivas på ett adekvat sätt.

## 6 Resultat

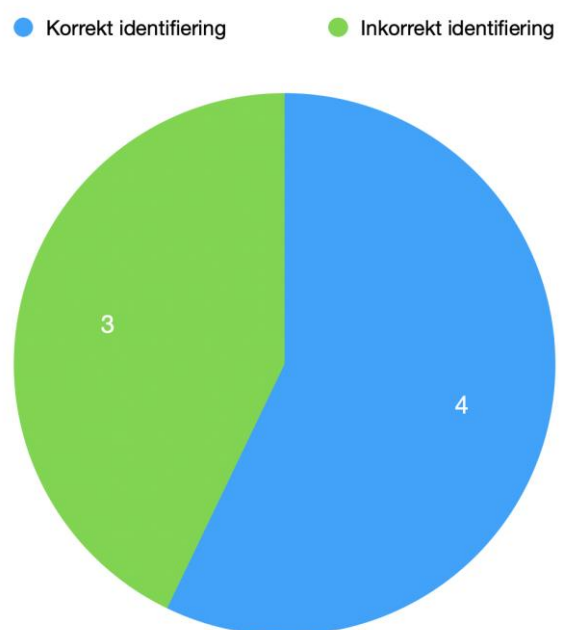
I det här kapitlet redovisas de resultat som framkommit i undersökningen. Resultaten kommer att behandlas i samma ordning som frågeställningarna. Först presenteras resultatet på frågeställning I, sedan följer resultatet för frågeställning II och sist presenteras resultatet på frågeställning III.

### 6.1 Frågeställning I

Frågeställning I behandlar frågan om hur många svensklärare är kapabla till att identifiera en text skriven av ChatGPT bland elevtexter.

Intervjuerna visade att fyra av sju informanter kunde identifiera texten skriven av ChatGPT bland de tre texterna. Samtliga informanter menade dock att det var svårt att hitta texten skriven av ChatGPT. Svensklärarna har en missidentifieringsgrad på 42%, vilket innebär att de inte alltid kan identifiera en text skriven av ChatGPT bland totalt tre texter. Det här testet kan däremot inte anses fastställa att svensklärare inte kan identifiera texter skrivna av ChatGPT i en autentisk

bedömningssituation. Barmark & Djurfeldt (2020) menar att en förutsättning för att kunna generalisera resultat är att det finns ett representativt stickprov, som uppnås genom att urvalet av respondenter är slumpmässigt (Barmark & Djurfeldt 2020:65f). Det har inte den här undersökningen beaktat. Därmed kan inte resultatet generaliseras. Ett representativt stickprovs storlek måste vidare svara mot frågeställningen (Barmark & Djurfeldt 2020:69). Frågeställningen i den här undersökningen handlar bara om informanterna och huruvida de kan identifiera en text skriven av ChatGPT. Undersökningens stickprov är därmed inte heller tillräckligt stort för att kunna dra några generaliserbara slutsatser om svensklärares generella förmåga att identifiera en datorskriven text bland elevtexter. Det får däremot anses bekräfta Clarks m.fl. (2021) resultat om att lärare kan se skillnad på elevtexter och texter skrivna av GPT-3 ungefär 50% av gångerna (s.7282). Turingtestet stipulerar att om människan inte kan



Figur 4: Svensklärarnas missidentifieringsgrad

skilja mellan dator och människa till en högre grad än 50% innehar datorn artificiell intelligens (Arkoudas & Bringsjord 2014:35). Därmed blir ChatGPT underkänd i Turingtestet i den här undersökningen.

## 6.2 Frågeställning II

Frågeställning II handlar om det validitetsproblem som uppstått i den summativa bedömningspraktiken i samband med ChatGPT:s lansering. Med bakgrund i det faktum att svensklärarna uppvisade stora svårigheter i att identifiera texten skriven av ChatGPT, diskuterades validitet i bedömningen av skrivuppgifter och hur ChatGPT har påverkat den. Efter att den tematiska analysen hade utförts på intervju svaren framkom följande teman och subteman. Temana samt subtemana presenteras här nedan som underrubriker. Först står de koder som använts för att konstruera subtemat eller temat. Därefter följer en beskrivande text om hur informanterna resonerar om ämnet som subtemat eller temat handlar om. Till sist följer 1–3 citat som illustrerar exempel på hur informanterna har svarat rörande ämnet som temat handlar om. Efter att samtliga subteman och teman presenterats appliceras informanternas svar på Bachmans assessment argument.

### 6.2.1 Subteman

Här nedan presenteras två subteman som framkom i den tematiska analysen om frågeställning 1. De kommer sedan att tillsammans utgöra ett gemensamt tema.

#### **Subtema 1: ChatGPT producerar abstrakt text av hög kvalitet som är svår att urskilja från elevtexter.**

##### **Koder:**

1. ChatGPT skriver abstrakt text av hög kvalitet.
2. ChatGPT följer uppgiftsinstruktionen på ett mycket träffsäkert vis.
3. Informanterna uppvisar stora svårigheter med att identifiera ChatGPT:s text bland de andra elevtexterna
4. ChatGPT kan instrueras att efterlikna elevtexter.
5. Texten som ChatGPT har producerat är i stort sett likvärdiga med elevtexterna sett till textkvalitet.

Samtliga informanter menade att texten som ChatGPT genererar tenderar att vara abstrakt och inte specifik. Texten beskrivs av flertalet informanter som allmän och beige. Detta motsvarar informanternas tidigare erfarenheter av ChatGPT i undervisningen. Vidare menar samtliga informanter att ChatGPT:s text följer uppgiftsinstruktionen på ett mycket träffsäkert vis. Ett av sätten som informanterna identifierade texten skriven av ChatGPT var att uppgiftsinstruktionen åtföljdes på ett noggrant sätt som elever sällan är kapabla till. Flertalet informanter uppvisade också stora svårigheter med att identifiera texten skriven av ChatGPT. Det var bara en informant som verkade säker omedelbart efter läsningen av texterna. Flertalet informanter menade också att ChatGPT kan instrueras att efterlikna elevtexter. Därmed uppstår ytterligare ett hinder. Till sist menade flertalet informanter att texten som ChatGPT producerade var i stort sett likvärdig med de två andra elevtexterna. Informanternas svar kan sammanfattas med att ChatGPT producerar abstrakt text av hög kvalitet som är svår att urskilja från elevtexter.

### **Exempel på citat från informanterna**

Den här (text 2, skribentens anmärkning) är också väldigt bra, naturligtvis, lite småsaker som kan förbättras förstås. Jag hade bedömt den till minst godkänt. (Intervju med informant 1)

Dom är nog ganska likvärdiga, ingen som jag tänker att den här skulle jag sätta F på direkt. Utan det skulle bli godkänt. (Intervju med informant 3).

Jamen den (text 2, skribentens anmärkning) är lite plattare. Det finns mer driv i språket i de andra, det glimrar till lite ibland, det gör det inte här. Den är liksom mer beige. Sen är strukturen jättegod, språkligt bra (...). (Intervju med informant 5).

Citaten visar att informanterna tycker att texten skriven av ChatGPT är sticker ut något när det gäller textkvalitet, men att den i övrigt är relativt likvärdig med de andra texterna.

## **Subtema 2: Svensklärarna kan inte garantera att eleven har skrivit texten om eleven har haft tillgång till ChatGPT**

### **Koder:**

1. ChatGPT:s texter är unika. De kan inte spåras av Urkund eller andra programvaror.

2. Lärarna vet inte vem som har skrivit texten om eleverna har tillgång till ChatGPT när de skriver.
3. Eleverna behöver bara skriva in uppgiftsinstruktionen i ChatGPT. De behöver inte visa några kunskaper i skrivande eftersom ChatGPT skriver texten åt dem.
4. Lärarna vet inte hur utspritt användandet av ChatGPT är hos eleverna.

Samtliga informanter menade att ChatGPT:s texter är unika, och kan därför inte spåras av Urkund eller andra programvaror. Svensklärarna kan därför inte heller veta om eleverna skrivit texten eller om ChatGPT har skrivit texten, förutsatt att eleverna har haft tillgång till ChatGPT när de har skrivit texten. Vidare menade samtliga informanter att eleverna inte måste uppvisa några kunskaper i skrivande när de använder ChatGPT, eftersom ChatGPT är kapabel till att skriva texten genom att enbart uppgiftsinstruktionen skrivs in i programvaran.

#### **Exempel på citat från informanterna**

Jag fick in en väldigt bra text som jag läste nu på påsklovet, väldigt bra diskuterat igen, så jag skickade in till Urkund, det gav inget. Jag tror att det var äkta, jag kan aldrig riktigt vara nöjd med något. (Intervju med informant 2).

Enstaka individer är det än så länge. Jag tror att de flesta är lite rädda för det. Så. Så det är inte... sen så kanske är jag grundlurad och det är 90% av mina elever som... nej. (Intervju med informant 5).

Dessa citat visar att lärarna inte är säkra på vilka elever som använder ChatGPT, samtidigt som Urkund, vilket är ett program som fungerar som en plagiatkontroll, inte kan kontrollera texter skrivna av ChatGPT

### **6.2.2 Teman**

Här nedan presenteras de teman som framkom i den tematiska analysen om frågeställning II.

#### **ChatGPT utgör en risk för validiteten i bedömningen av skrivuppgifter.**

De två tidigare subtemana utgör punkt 1 samt 2 i det här temat. Den tredje punkten är en kod.

#### **Subteman och koder:**

1. ChatGPT producerar abstrakt text av hög kvalitet som är svår att urskilja från elevtexter
2. Svensklärarna kan inte garantera att eleven har skrivit texten om eleven har haft tillgång till ChatGPT.
3. Svensklärarna menar att ChatGPT utgör en risk för validiteten i skrivbedömningen.

Eftersom ChatGPT producerar abstrakt text av hög kvalitet som är svår att urskilja från elevtexter och svensklärarna inte har något sätt att fastställa huruvida en elev eller ChatGPT har skrivit en text, utgör ChatGPT en risk för validiteten i bedömningen av skrivuppgifter.

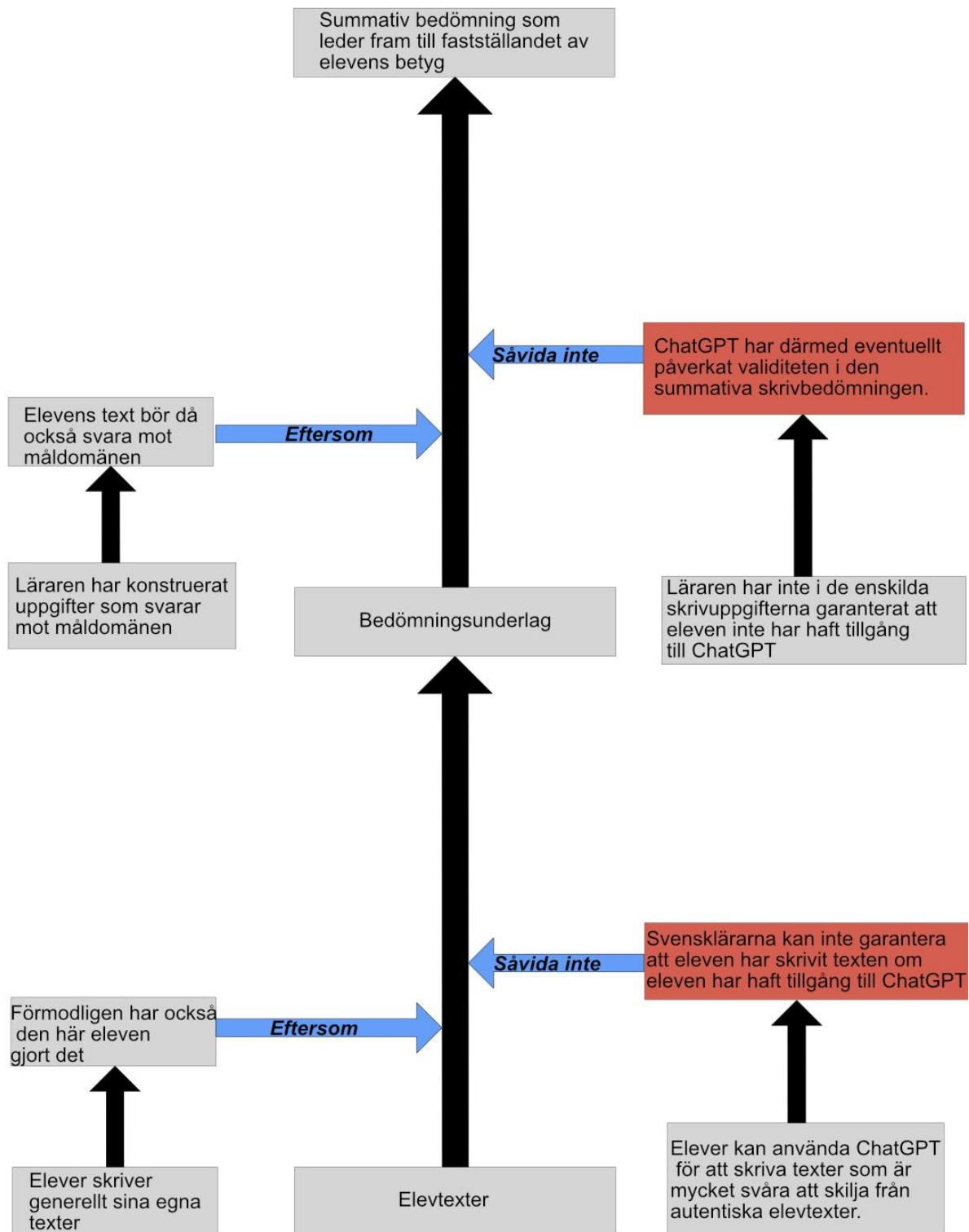
### **Exempel på citat från informanterna**

Jag vet inte vem som har skrivit texten. Så enkelt är det. Och det är elevens kunskaper som ska behandlas, inte hur väl de kan använda ett tekniskt hjälpmedel. (Intervju med informant 4).

Citatet visar att informanten inte kan veta vem som har skrivit texten om eleverna har haft tillgång till ChatGPT. Det är ett problem eftersom det är elevens kunskaper i skrivande som ska bedömas, inte hur väl de har använt ChatGPT.

### **6.2.3 Informanternas svar applicerade på Bachmans assessment argument**

Frågeställning II handlar om validitetsproblemet som ChatGPT innebär för den summativa skrivundervisningen. De teman som framkommit rörande den här frågeställningen har applicerats på Bachmans assessment argument nedan i figur 5. I validitetsargumentet syns en röd ruta på invändningsdatapositionen. Se höger sida i figur 5. Den indikerar validitetsproblemet, vilket svensklärarna menar är att de inte kan garantera att eleven har skrivit texten om eleven har haft tillgång till ChatGPT. Samma validitetsproblem indikeras igen på invändningsdatapositionen i användningsargumentet. Se höger sida i figur 5. Användningen av bedömningsunderlaget blir osäker när invändningsdatapositionen i validitetsargumentet pekar på ett stort problem i validiteten i bedömningsunderlaget. Därmed blir också användningsargumentet svagt i AUA. Det blir svårt för bedömaren, i det här fallet svenskläraren att motivera att bedömningen i fråga är valid.



Figur 5: Validitetsproblemet enligt Bachmans assessment argument

I modellen syns validitetsproblemet på två platser, på invändningsdatapositionen i validitetsargumentet såväl som på invändningsdatapositionen i användningsargumentet.

## 6.3 Frågeställning III

Frågeställning II handlar om validitetsproblemet i den summativa skrivbedömningen som uppkommit i samband med ChatGPT. Detta validitetsproblem har redan illustrerats i Bachmans assessment argument. Frågeställning III handlar om de lösningar som svensklärarna har vidtagit för att lösa validitetsproblemet. Dessa lösningar presenteras likt i frågeställning II i underrubriker som indikerar teman. Först presenteras de koder som temat bygger på. Sedan kommenteras koderna. Till sist presenteras 1–2 citat från informanterna som koderna har byggt på. Efter att temana har presenterats sätts de in i Bachmans assessment argument, och på så sätt kan både lösningarna illustreras i samma modell som problemen. I undersökningen framkom också att svensklärarna begrundat didaktiska konsekvenser i samband med en av lösningarna. Dessa bifynd har trots att de inte specifikt ingår i frågeställningar eller syfte inkluderats i frågeställning III, eftersom de är av värde för den framtida svenskundervisningen.

### 6.3.1 Teman

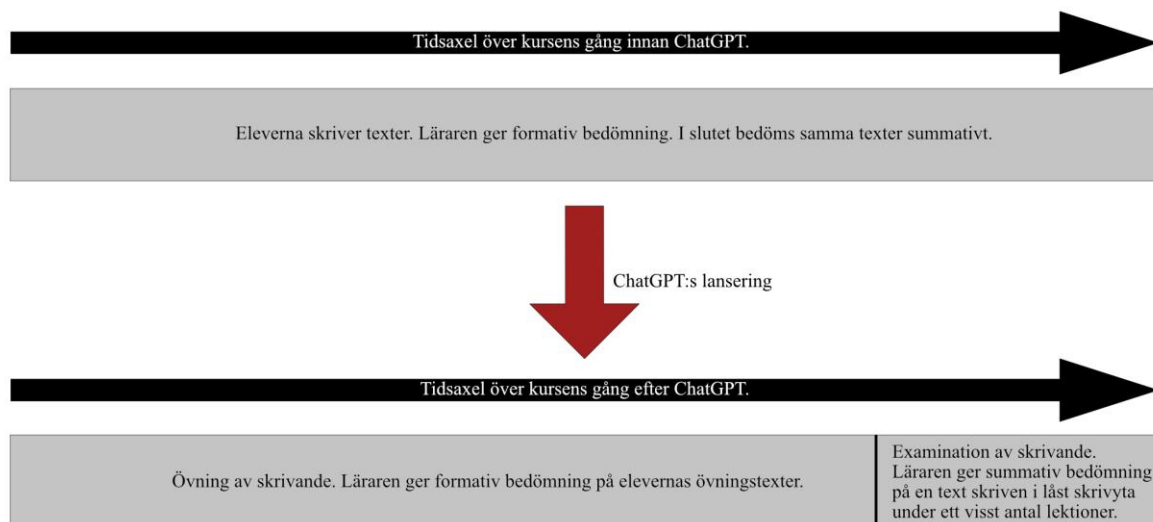
#### **Svensklärarna kontrollerar bedömningssituationen**

##### **Koder:**

1. Svensklärarna har delat upp övning och examination i skrivundervisningen.
2. Examinationsformerna har anpassats. Skrivandet sker i låst skrivyta.
3. Undervisningen är mer lärarstyrd och strukturerad.
4. Flera av informanterna har använt låst skrivyta redan innan ChatGPT lanserades.

Flertalet informanter har som en konsekvens av ChatGPT ändrat på sina rutiner rörande examination. Sedan ChatGPT lanserades har flertalet informanter valt att utföra sin examination i en s.k. låst skrivyta. Elevernas texter skrivs då på lektionstid i ett låst program där eleven inte kan använda någon annan av datorns funktioner utöver den kontrollerade skrivytan. Flertalet informanter menar att de i högre grad än innan har delat upp övning och examination. Innan ChatGPT lanserades skrevs elevernas texter progressivt i klassrummet. Läraren gav formativ bedömning i klassrummet. Till slut lämnades elevernas texter in och de bedömdes. Efter att ChatGPT lanserades använder lärarna sig av skrivuppgifter där eleverna får skriva hemma utanför den låsta skrivyta. Dessa texter bedöms då bara formativt. De texter

som bedöms summativt skrivs bara i låst skrivvyta på lektionstid. Skillnaden mellan de två sätten att examinera och undervisa skrivande skulle kunna illustreras med följande modell.



Figur 6: Modell över skillnaden i relationen mellan undervisning och examination före och efter ChatGPT.

I modellen syns två tidsaxlar. Den övre tidsaxeln och tillhörande rutan representerar situationen innan ChatGPT. Då bedömdes texterna som eleverna skrev hemma, och läraren gav formativ feedback under processens gång. Den undre tidsaxeln representerar situationen efter ChatGPT. Övning av skrivande och den formativa bedömningen har separerats från examination av skrivande och den summativa bedömningen. Examinationen sker i en låst skrivvyta.

Flertalet informanter menade att de redan innan ChatGPT lanserades hade använt sig av en liknande uppdelning av undervisning och examination i skrivande. Därmed kan inte ChatGPT anses vara den enda faktorn som påkallar en separation mellan övning och examination i skrivandet. Däremot menade samtliga informanterna att denna separation i framtiden är nödvändig för att säkerställa validiteten i skrivundervisningen.

### Exempel på citat från informanterna

Det har ju lett till att skrivningarna numera blir på en låst skrivvyta på lektionstid. Det är dom tider man har på sig att skriva. (Intervju med informant 1)

Ja men om man tidigare gav elever mer tid att bearbeta text och sätta sig in i materialet, så har jag blivit mer aktiv. Jag delar upp texterna, jag bearbetar dom på ett annat sätt tillsammans med eleverna, alltså jag är mer aktiv, jag har en mer aktiv

läsundervisning, eleverna tvingas slå upp ord, diskutera med varandra om slutsatser, källor, sådana här små saker hela tiden. Och när det gäller skrivandet så är det mer uppstyrt utifrån man gör det mer låst, mer strukturerat. (Intervju med informant 7)

De två citaten visar på svensklärarnas första lösning på validitetsproblemet. Svensklärarna har tvingats förändra sin undervisning såväl som sin bedömning för att lösa validitetsproblemet.

### **Den låsta skrivytan har didaktiska konsekvenser.**

#### **Koder:**

1. Eleverna presterar inte lika väl när de enbart skriver i klassrummet, vilket beror på att vissa elever inte kan koncentrera sig lika lätt i klassrummet. Eleverna kan också bli nervösa och stressade.
2. Examinationsformerna har anpassats. Eleverna skriver texter som används för summativ bedömning enbart på lektionstid och i låst skrivyta.
3. Vissa elever gynnas av att skriva på låst skrivyta, andra missgynnas.
4. Språk behöver bearbetas under en längre tid.
5. För att låta elevernas språk bearbetas under en längre tid måste lärarna dela upp skrivandet på flera lektioner.

Samtliga informanter menar att elever inte skriver lika bra texter när de skriver i en låst skrivyta. Det beror delvis på att vissa elever har lätt för att bli nervösa och stressade när de tvingas skriva i en klassrumsmiljö. Det beror också på att elever kan ha svårt för att koncentrera sig i klassrummet. De hade hellre skrivit hemma. Vidare kan elevernas skrivlust minskas som en konsekvens av att de måste producera text i stunden. En annan didaktisk konsekvens som skrivandet i den låsta skrivytan frambringar menar informanterna är att språk måste bearbetas under en längre tid. Därmed måste lärarna dela upp skrivandet på flera lektioner. Flertalet informanter har låtit eleverna ta med sig anteckningar från övningslektionerna till de examinerande lektionerna, i syfte att minska de didaktiska konsekvenserna.

#### **Exempel på citat från informanterna**

Jag vet att många elever känner sig stressade, därför att dom bara ska hinna, bara två timmar på sig att skriva en och en halv sida. (Intervju med informant 1)

Och då tänkte jag på det här fyrkantiga sättet. Det blir... skrivande på beställning i stunden. (...) Det dödar lusten, tror jag. Och vi är så olika också. En del tycker om att skriva kort och snabbt. (Intervju med informant 2)

Citaten visar att svensklärarna har reflekterat över de didaktiska konsekvenser som den låsta skrivytan för med sig. De är medvetna om att lösningen på validitetsproblemet skapar andra problem.

## **I framtiden kan svensklärarna examinera skrivprocessen i högre grad än förut**

### **Koder:**

1. Svensklärarna menade att de i framtiden kan fokusera på elevernas skrivprocess snarare än texten i den summativa bedömningen.
2. Svensklärarna menar att användandet av ChatGPT inte kan anses vara en del av skrivandet i styrdokumentens nuvarande form.

Flertalet informanter menade att fokusera och examinera elevernas skrivprocess snarare än själva texten kan vara ett sätt att i framtiden kringgå ChatGPT:s textkompetens. Det är däremot inte något som sker just nu. Detta beror delvis på att ett sådant arbetssätt hade varit tidintensivt. Vidare ansåg samtliga informanter att användandet av ChatGPT inte kan anses vara en del av skrivprocessen. Skrivande är ett hantverk, och ChatGPT utför hantverket åt eleven. Därmed är ChatGPT inkompatibel med informanternas syn på elevers skrivande.

### **Exempel på citat från informanterna**

Det kanske är så att vi kommer lägga mer fokus på processen och inte den slutgiltiga produkten. Alltså att processen i sig blir viktigare, inte slutprodukten. (Intervju med informant 5).

Citatet visar på en reflektion som flertalet informanter gjorde. Processen kan bli en viktigare del i skrivbedömningen i framtiden eftersom det är en lösning på validitetsproblemet som ChatGPT innebär. Läraren bedömer alltså på elevens skrivprocess under en längre tid snarare än att bedöma en färdig produkt.

### **6.3.2 Informanternas svar applicerade på Bachmans assessment argument**

Sett från Bachmans (2005) assessment argument finns två lösningar på validitetsproblemet som ChatGPT innebär. Lösningarna illustreras i Bachmans assessment argument, se figur 7 nedan.

### **6.3.3 Låst skrivyta**

Lärarna menar att en möjlig lösning på validitetsproblemet med ChatGPT är att låta eleverna skriva i låsta skrivytor på lektionstid. I Bachmans modell AUA passar lösningen, alltså det som återigen berättigar bedömningen, bäst in på berättigandedata i validitetsargumentet. Se figur 7 nedan. Lösningen markeras med grön färg. Bedömningen av texterna garanteras validitet genom att eleverna skriver på lektionstid i en skrivmiljö som läraren kontrollerar. Det faktum att eleverna får skriva i låst skrivyta har enligt informanterna didaktiska konsekvenser. Detta markeras med gul färg.

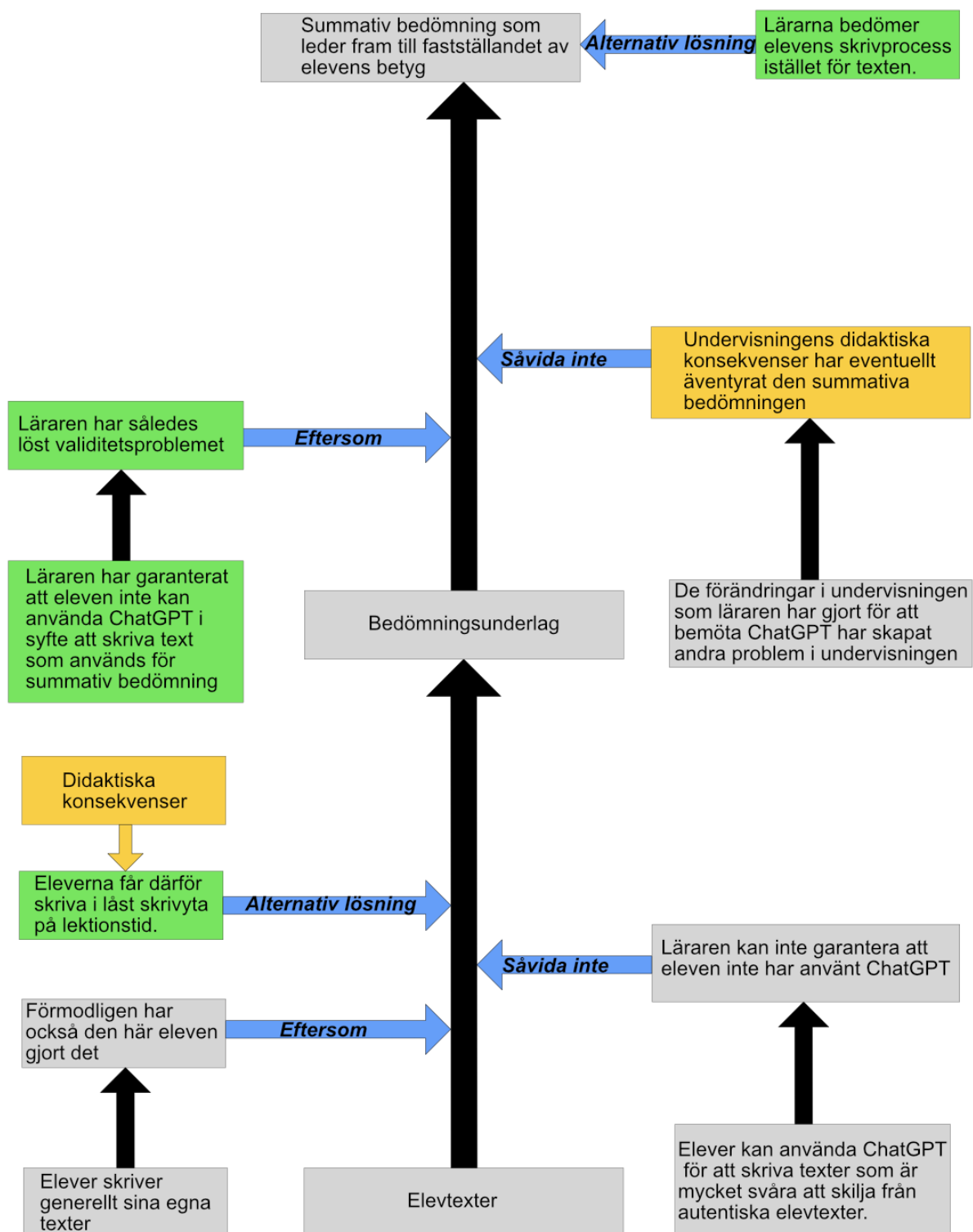
### **6.3.4 Nytt skrivkonstrukt**

En annan möjlig lösning på validitetsproblemet med ChatGPT menade flertalet informanter var att förändra anspråket, det vill säga det som läraren faktiskt ska bedöma. Detta markeras med röd färg, högst upp i användningsargumentet. Informanterna vill med andra ord förändra konstruktet som avses att bedömas. Flera av informanterna resonerade om att det i framtiden är det kanske inte längre texten, själva produkten, som bedöms. Det är kanske snarare elevens skrivprocess. Eftersom ChatGPT inte kan uppvisa en skrivprocess på samma sätt som elever kan, är detta ett potentiellt sätt att bemöta validitetsfrågan i skrivbedömningen. Informanterna är emellertid överens om att bedömandet av elevens skrivprocess inte sker just nu.

### 6.3.5 Didaktiska konsekvenser

När lärarna förändrar sin bedömning och undervisning som en konsekvens av ChatGPT får detta didaktiska konsekvenser. De didaktiska konsekvenserna syns på två platser i modellen. Först indikeras de didaktiska konsekvenser som kan inträffa när eleverna måste skriva i låst skrivvyta. De syns i validitetsargumentet, i den undre delen av modellen. Dessa didaktiska konsekvenser markeras med gul färg, och är en reaktion på det berättigande som den låsta skrivvyta innebär. Lärarna menar bland annat att elever med koncentrationssvårigheter kan få svårt att prestera under en kort period i ett klassrum. Således finns det didaktiska konsekvenser i att använda en låst skrivvyta.

Vidare syns samma konsekvenser igen, men den här gången som invändning i användningsargumentet. Invändningen hade benämnts som typ 1 i Bachmans (2005) modell, alltså att det finns en klar anledning för att antingen inte använda tolkningen eller att göra en annan tolkning baserad på data som finns (Bachman 2005:21). De didaktiska konsekvenserna blir en invändning eftersom de representerar ett argument mot att använda bedömningsunderlaget i den summativa skrivbedömningen.



Figur 7: Lösningar på validitetsproblemet applicerat på Bachmans assessment argument

I figur 7 ovan syns två lösningar på validitetsproblemet som ChatGPT innebär för den summativa skrivbedömningen. Läraren kan låta eleverna skriva sina texter i låst skrivyta, och därmed garantera att ChatGPT inte används. Detta syns till vänster i validitetsargumentet. Ovanför finns de didaktiska konsekvenser som tidigare nämnts. Dessa återkommer i användningsargumentet som en invändning på höger sida.

Lösningen med låst skrivyta garanterar också validitet i användningsargumentet, vilket syns i berättigande samt berättigandedata i användningsargumentet. Läraren kan också bedöma elevernas skrivprocess snarare än texten i sig. Detta syns högst upp, i grön färg med en pil som pekar på måldomänen som bedömningen sker mot. Med hjälp av de lösningar som har föreslagits i figur 7 ovan kan bedömaren göra en mer valid bedömning. Berättigandedata väger tyngre än invändningsdata.

## 7 Diskussion

I följande avsnitt diskuteras resultatet. Inledningsvis diskuteras det validitetsproblem som ChatGPT innebär för den summativa skrivbedömningen samt informanternas lösningar på detta validitetsproblem. Vidare diskuteras de didaktiska risker som lösningarna kan innebära. Sedan följer en reflektion om ChatGPT och den nuvarande måldomänen. Efter reflektionen formuleras två slutsatser rörande att garantera validitet i skrivbedömningen. Till sist diskuteras uppsatsens begränsningar och vidare forskning föreslås.

### 7.1 Turingtestet

Resultatet i testet blev att 42% eller tre av sju svensklärare inte klarade av att identifiera texten skriven av ChatGPT. Svensklärarna tyckte att testet var svårt. Förmodligen blir testet mycket svårare om en mer realistisk situation skapas. Svensklärare kan få in 30 texter som ska bedömas i stället för 3. Då krävs det mer av svensklärarnas textanalysförmåga för att kunna urskilja en datorskriven text bland elevtexter. ChatGPT blev också underkänd på Turingtestet, men det borde delvis bero på att urvalet litet. Om bara en till informant hade svarat fel på Turingtestet hade resultatet blivit 50%, och således hade ChatGPT klarat Turingtestet.

### 7.2 Validitetsproblemet

Undersökningen har pekat på ett validitetsproblem i den summativa skrivbedömningen. Problemet grundar sig i att svensklärare inte längre kan urskilja datorskriven text från elevtext, vilket både syns i Turingtestet såväl som i informanternas svar i intervjuerna. Informanterna menar att det här validitetsproblemet inte har skapats när ChatGPT lanserats, utan att det har funnits tidigare. Problemet tycks däremot ha blivit markant större och tvingat fram lösningar.

### 7.3 Lösningar på validitetsproblemet

ChatGPT utgör en risk för validiteten i skrivbedömningen. Om ChatGPT skulle tillåtas vara en risk för validiteten i skribedömningen äventyras den urvalsroll som gymnasieskolan har för

högre utbildning. Därför krävs lösningar på problemet. Informanterna finner två huvudsakliga lösningar på validitetsproblemet. Den första lösningen är relativt okomplicerad. Lösningen består av att eleverna får skriva i en låst skrivvyta. Det är en lösning som redan använts innan ChatGPT lanserades, varför det också fanns en beredskap hos skolorna när validitetsproblemet med ChatGPT uppstod. Det fanns redan digitala skrivprogram hos skolorna som kunde låsas. Informanterna menar att anledningen till att dessa program redan fanns är att validitetsproblemet i skrivbedömningen inte är nytt. Eleven har sedan länge kunnat ta hjälp av en familjemedlem eller på annat sätt få utomstående hjälp. Anledningen till att det här validitetsproblemet åtgärdades innan ChatGPT lanserades är inte klarlagt för informanterna. Däremot verkar problemet ha blivit värre efter ChatGPT, och då tvingat svensklärare att använda låsta skrivvytor.

En annan lösning på validitetsbetyget var enligt informanterna att förändra konstruktet som ska bedömas. I stället för att bedöma texten som eleverna skrivit ska svensklärarna bedöma skrivprocessen. Informanterna var inte överens om hur en sådan bedömning skulle gå till. Det är möjligt att det skrivkonstrukt som finns i styrdokumentet för ämnet svenska inte fungerar med det skrivkonstrukt eller måldomän som svensklärarna menar när de talar om att bedöma skrivprocessen. En möjlig anledning skulle kunna vara att ämnesplanen i svenska stipulerar att eleverna ska *skriva* texter. Därför är det möjligt att informanterna inte har tolkningsutrymme i ämnesplanen för att motivera en skrivundervisning som främst bedömer skrivprocessen snarare än elevernas texter. En sådan bedömning hade missat måldomänen.

## 7.4 Didaktiska konsekvenser

I det här avsnittet diskuteras de möjliga didaktiska konsekvenserna som kan frambringas av svensklärarnas lösningar på validitetsproblemet.

### 7.4.1 Låst skrivvyta

Den låsta skrivvytan kan ha flertalet didaktiska konsekvenser för eleverna. Informanterna menade att den låsta skrivvyta gynnar en viss typ av elever, och missgynnar de som har svårt för att koncentrera sig i klassrummet. Det finns således en didaktisk risk med den låsta skrivvytan. Vidare menar flertalet av informanterna att skrivprocessen måste få ta tid. Därav väljer flera av informanterna att låta eleverna skriva under flera lektioner. Många av lärarna

lät eleverna ha med sig anteckningar på lektionerna de examinerande. Problemet med att skriva under flera lektioner är att elever kan ta med sig anteckningar från ChatGPT med färdiga formuleringar in på lektionen. Risken är att om en elev kopierar in en formulering från ChatGPT syns inte detta i Urkund. Således finns även där en didaktisk risk när det kommer till att kontrollera ifall lärande verkligen har skett.

### **7.4.2 Separation mellan övning och examination**

En annan didaktisk risk skulle kunna ligga i den nya separationen mellan övning och examination i skrivundervisningen. Lärarna är överens om att det finns didaktiska problem med att enbart låta elever skriva texter som ska bedömas summativt på lektionstid. Eleverna kan bli stressade och inte prestera på samma nivå som om de fick skriva hemma. Separationen mellan övning och examination innebär ett störningsmoment för eleverna, men en garant för validiteten i skrivbedömningen.

### **7.4.3 Bedömning av skrivprocessen**

I resultatet framkom det inte om bedömning av skrivprocess snarare än produkt skulle ha någon didaktisk risk. Det bör däremot finnas didaktiska risker med att i grunden förändra vad som bedöms i elevers skrivande. Den största didaktiska risken med bedömning av skrivprocessen borde vara komplexiteten i att bedöma en skrivprocess. När läraren summativt bedömer en färdig produkt, som en text, är texten statisk. Den förändras inte, och det ger läraren en stabil grund att basera sin summativa bedömning på. En elevs skrivprocess är däremot flödande och dynamisk. Den borde därför vara svår att fånga i en summativ bedömning. Den här didaktiska risken konsekvensen syns inte figur 7, men hade kunnat markeras vid den gröna lösningen längst upp i modellen. Det kommer att krävas en ny måldomän i ämnesplanen för att bedöma skrivprocessen i stället för texter. Ett nytt skrivkonstrukt kommer att behöva formuleras.

## 7.5 Måldomänen oförenlighet med ChatGPT

Lärarna menar att de inte kan se en giltig skrivprocess där ChatGPT har assisterat med skrivandet. Den passar inte ihop med måldomänen, alltså både lärarnas uppfattning och måldomän om skrivande såväl som Skolverkets ämnesplans måldomän för skrivande. Flera av informanterna menar att eleverna i framtiden kommer att använda program som ChatGPT som en hjälp i skrivandet. Debatten kan jämföras med den om miniräknaren. En lösning hade kunnat vara att uppdatera styrdokumentet, och då förändra skrivkonstrukten och måldomänen för skrivande. Kanske ChatGPT kommer att vara förenlig med skrivande i gymnasieskolan i framtiden?

## 7.6 Slutsatser för att garantera att ChatGPT inte påverkar validiteten i skrivbedömningen

Slutsats 1: Lärare kan använda en låst skrivyta för att garantera att eleverna inte har tillgång till ChatGPT och därmed garantera lösa validitetsproblemet med ChatGPT.

Slutsats 2: Lärare kan också ändra skrivkonstrukten som avses att bedömas för att svara mot problemet som ChatGPT utgör i skrivbedömningen. Det är däremot inte möjligt i nuläget, eftersom styrdokumentet inte tillåter den tolkningen av skrivkonstrukten. Det finns inte tolkningsutrymme i måldomänen för ett sådant skrivkonstrukt.

## 7.7 Diskussion om uppsatsens begränsningar

En begränsning i uppsatsens användning av Turings test ligger i att informanterna inte sedan tidigare är bekanta med den elevgrupp som de ska försöka avgöra om huruvida de har använt ChatGPT eller inte. Om lärarna sedan innan känner till eleverna språk hade de möjligtvis lättare kunnat identifiera texterna, eftersom det hade funnits en tidigare uppfattning om just den elevens förmågor. Om en elev lämnar in en avsevärt mycket bättre text än vad den brukar skriva finns det en större anledning att misstänka fusk.

En annan begränsning i uppsatsen ligger i bristen på material i forskningsbakgrunden. På grund av att ChatGPT fortfarande är en ny företeelse i skolan, finns inte mycket tidigare forskning. Därmed krävdes det en bred infallsvinkel i delen tidigare forskning, där textbedömning, validitet och ChatGPT:s textförmåga kombineras.

## 8 Vidare forskning

I detta avslutande kapitel av uppsatsen föreslås vidare forskning som kan bygga på den undersökning som har genomförts. Ett inledande förslag är att göra en betydligt större kvantitativ studie där svensklärare testas i sin förmåga att kunna se skillnad på datorskriven text och elevtext. Resultatet i den här uppsatsen visar på att svensklärare har stora svårigheter med att identifiera datorskriven text. En större studie hade skapat ett generaliserbart resultat, som också hade kunnat få skolpolitiska konsekvenser.

Ett centralt resultat som framkom i undersökningen var att svensklärare numera delar upp skrivandet på flera lektioner. Detta har didaktiska konsekvenser. Men frågan är vilka didaktiska konsekvenser, och om de är positiva eller negativa för elevernas lärande. Detta är något som hade kunnat undersökas vidare. En initial tanke om detta är att undersöka problematiken med att läraren bedömer samma text summativt som läraren sedan tidigare bedömt formativt. Hur mycket av lärarens egna idéer, tankar och kunskaper har satts in i elevens text när den bedöms formativt i klassrummet? Separationen mellan formativ och summativ bedömning i undervisningen eliminerar risken att läraren bedömer en text som den delvis skrivit själv.

Ett annat resultat som framkom var att informanterna resonerade om skrivkonstruktet i samband med ChatGPT. Om skolväsendet står inför en teknisk revolution i form av textgeneratorer finns det anledning att begrunda vad svensklärare faktiskt ska lära våra elever. Förr bedömdes förmodligen andra förmågor i svenskämnet än vad som bedöms idag, vilket kan bero på den tekniska revolution som samhället deltagit i. Elever lär sig inte att skriva PM för hand, utan de lär sig att skriva dem en dator. Om samhället i övrigt använder sig av textgeneratorer på ett naturligt sätt finns det kanske anledning att förändra skrivkonstruktet, och för att göra det krävs en förändring av måldomänen. Frågan bör utredas, och tillsättas vidare forskning.

En annan aspekt som hade kunnat utforskats mer är den socioekonomiska dimensionen som ChatGPT för med sig. På sätt och vis kan ChatGPT anses spela en utjämnande roll i undervisningen eftersom programmet är gratis och tillgängligt för allmänheten. Som svensklärarna i studien ger uttryck för är validitetsproblemet i skrivbedömningen inte nytt. Externa faktorer som exempelvis föräldrar med högre utbildning har sedan länge kunnat

hjälpa eleven att skriva. Dessa föräldrar bidrar till en naturlig fördel för eleven eftersom de är mer vana vid att skriva texter i genrer som bedöms i skolan och kan då hjälpa eleven hemma. När alla elever har tillgång till ett så kraftfullt verktyg som ChatGPT borde elevens familjs utbildningsnivå spela mindre roll i sammanhanget summativ skrivbedömning. Om det verkligen är så är något som bör utforskas vidare.

Svensklärarna i undersökningen var överens om att ChatGPT kan bli en viktig resurs i undersökningen. De hade däremot ännu inte lärt sig hur ChatGPT kan användas i undervisningen, och kanske till och med användas i bedömningen av undervisningen. Detta kan undersökas vidare, exempelvis genom att använda Bachmans (2005) assessment argument åt andra hållet, det vill säga med intentionen att designa prov och tester. Då hade ChatGPT blivit ett verktyg, snarare än den problematik som har präglat den artificiella intelligensens begynnelse i skolväsendet.

## 9 Referenslista

- Arkoudas, A. Bringsjord, S. (2014). Philosophical foundations. I Frankish, K & Ramsey, W. (Red.) *The Cambridge Handbook of Artificial Intelligence* (s.15–33). Cambridge University Press. <https://doi.org/10.1017/CBO9781139046855>
- Aldrin, E. (2015). Textbedömning i svenskämnet: Attityder, erfarenheter och variation. *Forskning om undervisning och lärande*, (14), 62–83.
- Bachman, L.F. (2005). Building and Supporting a Case for Test Use. *Language Assessment Quarterly: An International Journal*, 2 (1), 1–34. [https://doi-org.ezproxy.ub.gu.se/10.1207/s15434311laq0201\\_1](https://doi-org.ezproxy.ub.gu.se/10.1207/s15434311laq0201_1)
- Barmark, M., & Djurfeldt, G. (2020). *Statistisk verktygslåda 0: Att förstå och förändra världen med siffror* (Andra upplagan ed.). Lund: Studentlitteratur.
- Blomqvist, P. (2018). *Samtal om skrivbedömning: Lärares normer, beslut och samstämmighet*. Doktorsavhandlingar i språkdidaktik, 11. [Doktorsavhandling]. Stockholms universitet. <http://su.diva-portal.org/smash/record.jsf?pid=diva2%3A1197840&dswid=6322>
- Braun, V. Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <http://dx.doi.org/10.1191/1478088706qp063oa>
- Cotton, R. E. Cotton, P. A. Shipway, J. R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. I *Innovations in Education and Teaching International*. 1–12. <https://doi.org/10.1080/14703297.2023.2190148>
- Christoffersen, L., Johannessen, A., & Andersson, S. (2015). *Forskningsmetoder för lärarstudenter* (1. uppl. ed.). Lund: Studentlitteratur.

- Clark, E. August, T. Serrano, S. Hauong, N. Gururanga, S. Smith, N, A. (2021). All That's 'Human' Is Not Gold: Evaluating Human Evaluation of Generated Text. (7282–7296). <https://doi.org/10.48550/arXiv.2107.00061>
- Denscombe, M. (2018). *Forskningshandboken: För småskaliga forskningsprojekt inom samhällsvetenskaperna* (Fjärde upplagan ed.). Lund: Studentlitteratur.
- Elkins & Chun (2020). Can GPT-3 pass a Writers Turing Test? I *Journal of Cultural Analytics*. (s.1–16). 10.22148/001c.17212
- Evensen, L. S. (2010) En gyldig vurdering av elevers skrivekompetanse? I J. Smith (red.), *Rammer for skrivning: om skriveutvekkling i skole og yrkesliv* (s.13–31). Trondheim: Tapir akademisk forlag.
- Harlen, W. (2012) On the Relationship between Assessment for Formative and Summative Purposes. I Gardner, J. (Red.), *Assessment and Learning* (s.2--9). Sage Publications Ltd.
- Huot, B., O'Neill, P., & Moore, C. (2010). A Usable Past for Writing Assessment. *College English*, 75(5), 495–517. Hämtad från <http://www.ncte.org/library/NCTEFiles/Resources/Journals/CE/0725-may2010/CE0725Usable.pdf>
- Hultman, T., & Westman, M. (1977). *Gymnasistsvenska*. Lund: Liber Läromedel.
- Kane, M. T., Crooks, T. J., & Cohen, A. S. (1999). Validating Measures of Performance. *Educational Measurement: Issues and Practice*, 18(2), 5–17. doi:10.1111/j.1745-3992.1999.tb00010.x
- Klapp, A. (2015). *Bedömning, betyg och lärande* (1. uppl. ed.). Lund: Studentlitteratur.

Kvale, S., Brinkmann, S., & Torhell, S. (2014). Den kvalitativa forskningsintervjun (Tredje [reviderade] upplagan ed.).

OpenAI (2023). *GPT-4 Technical Report*. 1–100. <https://doi.org/10.48550/arXiv.2303.08774>

OpenAI (2022) *Introducing ChatGPT*. OpenAI.com. Hämtad 2023-06-12 från <https://openai.com/blog/chatgpt>

Skar, G. (2013). *Skrivbedömning och validitet. Fallstudier av skrivbedömning i svenskundervisning på gymnasiet*. [Doktorsavhandling, Stockholms universitet]. DiVA. <https://www.diva-portal.org/smash/get/diva2:623658/fulltext01.pdf>

Turing, A. (1950). COMPUTING MACHINERY AND INTELLIGENCE. I *Mind. A Quarterly Review* (236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>

*Ämne – Svenska*. [Ämnesplan] (2022). Skolverket. <https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/gymnasieprogrammen/amne?url=-996270488%2Fsyllabuscw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DSVE%26version%3D8%26tos%3Dgy&sv.url=12.5dfee44715d35a5cdfa92a3>

Östlund-Stjärnegårdh, E. *Godkänd i svenska?: Bedömning och analys av gymnasieelevers texter*. Skrifter utgivna av Institutionen för nordiska språk vid Uppsala universitet, 57. [Doktorsavhandling]. Uppsala universitet. DiVA. [Godkänd i svenska?: Bedömning och analys av gymnasieelevers texter \(diva-portal.org\)](https://www.diva-portal.org/urn:nbn:se:uu:diva-17111)

# 10 Bilagor

## 10.1 Bilaga 1: ChatGPT:s text. I undersökningen kallad för text 2.

### Inför allas blickar

Hej allihopa!

Jag ville bara ta upp en intressant fråga som jag nyligen läste om i en artikel av Peter Letmark i Dagens Nyheter (4.5.2012) med titeln "Därför blir vi skräckslagna". Artikeln handlar om varför många människor känner obehag inför muntliga presentationer i skolan och arbetslivet. Detta är en viktig fråga att diskutera eftersom vi lever i ett modernt informationssamhälle där kommunikation är nyckeln till framgång.

En av huvudorsakerna till varför människor tycker det är jobbigt att tala inför andra är rädslan för att göra bort sig eller att bli dömd av åhörarna. Många av oss är rädda för att säga något fel, stamma eller glömma vad vi skulle säga mitt i vår presentation. Det kan också vara så att vi känner oss obekväma med att bli iaktagna av andra, vilket gör oss mer självmedvetna och ängsliga.

I sin artikel nämner Letmark att detta obehag kan bli ett problem i vårt samhälle eftersom vi förväntas vara kapabla att uttrycka oss väl och på ett effektivt sätt. Men det finns sätt att arbeta bort talängslan. För det första kan vi öva på att tala inför andra genom att delta i aktiviteter som kräver muntlig kommunikation, såsom debattklubbar, teater eller att hålla tal vid olika evenemang. Ju mer vi övar, desto mer bekväma kommer vi att känna oss i dessa situationer.

För det andra kan vi också använda oss av avslappningstekniker och andningsövningar för att hantera vår nervositet. Genom att fokusera på vår andning och vara närvarande i stunden kan vi bli mindre självmedvetna och mer fokuserade på vårt budskap.

Slutligen kan vi också förbereda oss väl inför våra presentationer genom att öva innehållet, strukturera det väl och ta hjälp av visuella hjälpmedel för att stödja vår presentation. När vi

känner oss väl förberedda blir det lättare att känna sig trygg och säker när vi står inför en publik.

Sammanfattningsvis är talängslan något som många av oss brottas med, men genom att öva, använda avslappningstekniker och förbereda oss väl kan vi övervinna denna rädsla och bli bättre på att kommunicera våra idéer och åsikter. Låt oss ta itu med detta problem tillsammans och bli bättre på att uttrycka oss inför andra!

Tack för att ni läste mitt inlägg, och jag hoppas att det gav er några perspektiv på ämnet.

## 10.2 Bilaga 2: Elevtext 1

Rädslan att tala inför publik är något många drabbas utav. Adrenalinet pumpar, rösten darrar och rädslan för att framstå inkompetent drar självsäkerheten i botten. I dagens samhälle krävs det av oss att kunna framträda och presentera inför andra. Den sociala rädslan har dock ökat de senaste femtio åren vilket medför begränsningar. Undersökningar visar på att var fjärde svensk anser att det är ”rejält obehagligt” att tala inför andra. Många olika faktorer påverkar och förklarar vårt beteende under talsituationer men lösningar på problemet finns. (...)

## 10.3 Bilaga 3: Elevtext 2

Tänk dig att du står på en scen. Framför dig sitter massor med människor. Du ska precis hålla ett tal och alla fokuserar på dig. Känns det inte lite läskigt? Majoriteten av alla människor är mer eller mindre rädda för att prata inför folk. Ändå är det något man tvingas göra många gånger i livet. Redan som barn förväntas man hålla presentationer och tal i skolan och genom hela livet ska man kunna presentera och argumentera för sina idéer. Men hur kommer det sig att så många tycker att det är obehagligt? Finns det någon lösning på problemet? (...)

## 10.4 Bilaga 4: Intervjuguide

# Intervjuguide

### Inledande frågor

Dessa frågor är tänkta att värma upp intervjun.

1. Hur länge har du arbetat som lärare?
2. Vilken utbildning har du?
3. Vilka ämnen förutom svenska undervisar du i?

### Identifiering – *TEXTERNA DELAS UT, LÄRARNA LÄSER TEXTERNA*

Dessa frågor behandlar identifieringen av texterna.

4. Är du medveten om att det finns en programvara som heter ChatGPT som är tillgänglig för allmänheten?
5. En av texterna är skriven av ChatGPT. Kan du se vilken det är?

*LÄRAREN GISSAR → RÄTT TEXT PEKAS UT.*

6. Vad tycker du om textkvalitén som ChatGPT genererar?
7. Om du hade varit tvungen att sätta ett betyg på texterna, vilka betyg hade du satt på respektive text?
8. Hur motiverar du betyget? Kan du ange några språkliga drag som motiverar betyget?

### Validitetsfrågor – *DISKUSSIONEN OM VALIDITET OCH SKRIVBEDÖMNING*

9. Innebär ChatGPTs textkompetens ett problem för validiteten i bedömningen av texter?
10. Innebär användandet av ChatGPT plagiat?
11. Hur ser du på elevers användande av ChatGPT?
12. Hur ser du på bedömningen av elevtexter i samband med ChatGPT? Värderar du dem på samma sätt som förut i din summativa bedömning.
13. I ämnesplanen under Svenska 1 står det att elever ska skriva egna argumenterande och andra typer av texter. Hur kan vi se på begreppet "skriva" i samband med ChatGPT?
14. Tycker du att vi kan rättfärdiga ChatGPT som ett verktyg för eleverna?
15. Om vi kan rättfärdiga ChatGPT som ett verktyg för eleverna, likt miniräknaren för matematiken, kan vi då säga att texter skrivna med hjälp av ChatGPT är valida och bör användas som bedömningsunderlag?
16. Vilken funktion har de nationella proven i relation till den nya situationen med AI med hög textkompetens som ChatGPT?
17. Vilken funktion har traditionella inlämningsuppgifter i och med ChatGPT?
18. Finns det några möjligheter för elever med skrivsvårigheter i ett eventuellt användande av ChatGPT? Hjälper eller stjälper ChatGPT?

## **Explorativa frågor om skrivundervisningen**

19. Har du förändrat din undervisning efter att du blev medveten om ChatGPT? På vilket sätt? Hur har du hanterat validitetsfrågan?
20. Hur utspritt bland elever är användandet av ChatGPT tror du?
21. Vilka konsekvenser får dina förändringar på din undervisning?
22. Vad innebär ChatGPT för den framtida skrivbedömningen?
23. Har ni diskuterat på skolan hur ni gemensamt bör se på ChatGPT? Finns det några konkreta åtgärder som gjorts?
24. Hur bör svensklärare se på ChatGPT? Som en begränsning eller ett verktyg?

## 10.5 Bilaga 5: Informationsbrev

Hej!

Jag heter David Nyqvist och jag läser min sista termin på ämneslärarprogrammet vid Göteborgs universitet. Jag ska nu skriva mitt examensarbete och söker informanter som skulle kunna tänka sig ställa upp på intervju.

Syftet med examensarbetet är att undersöka hur svensklärare ser på bedömningspraktiken i samband med ChatGPT.

Medverkan innebär ett intervjutillfälle om cirka 30 minuter på en tid som passar er. Intervjuerna kommer att vara semistrukturerade.

- Medverkan i undersökningen är frivillig och kan när som helst avbrytas. Du som medverkat har också rätt att återta information du lämnat.
- Enligt konfidentialitetskravet kommer deltagarnas uppgifter att behandlas med skydd då ingen obehörig har rätt att ta del av materialet.
- Det inspelade materialet transkriberas och kommer, enligt nyttjandekravet, endast användas för denna uppsats syfte och kommer efter examinationen raderas.
- Samtliga informanter ges ett fiktivt namn.

**Vid eventuella frågor vänligen kontakta mig!**

David Nyqvist

Mejl: [gusnyqvda@student.gu.se](mailto:gusnyqvda@student.gu.se) Mobil: 0736446038

Med vänliga hälsningar, David Nyqvist.

## 10.6 Bilaga 6: Samtyckesblankett

### Samtyckesblankett för informanter

Mitt namn är David Nyqvist och jag läser min sista termin på ämneslärarprogrammet vid Göteborgs universitet. Jag ska nu skriva examensarbete. Syftet med examensarbetet är att undersöka validitet i bedömningen av texter i relation till ChatGPT.

Intervjutillfället sker på plats och beräknas ta cirka 30 minuter. Medverkan i undersökningen är frivillig och kan när som helst avbrytas. Du som medverkat har också rätt att återta information du lämnat. Dina uppgifter kommer att behandlas konfidentiellt och du kommer ges ett fiktivt namn. Intervjuerna kommer att spelas in och därefter transkriberas. När examensarbetet godkänts vid examination kommer det inspelade materialet raderas.

- Jag samtycker till att medverka i denna intervjustudie om validitet i bedömningen av texter i relation till ChatGPT.
- Jag har skriftligen informerats om studiens syfte och hur informationen samlas in och hanteras.
- Jag har informerats om att mitt deltagande är frivilligt och kan när som helst avbrytas.
- Jag samtycker till att medverka i denna studie.

Namnsteckning \_\_\_\_\_

Namnförtydligande \_\_\_\_\_

Ort och datum \_\_\_\_\_

Vid eventuella frågor kontakta mig!

Mail: [gusnyqvda@student.gu.se](mailto:gusnyqvda@student.gu.se)

Mobil: 0736446038