



Ett svenskt akademiskt vokabulärtest – revidering, validering och förslag på vidareutveckling

Downloaded from: <https://research.chalmers.se>, 2023-03-25 23:10 UTC

Citation for the original published paper (version of record):

Warnby, M., Malmström, H., Bardel, C. (2022). Ett svenskt akademiskt vokabulärtest – revidering, validering och förslag på vidareutveckling. *Nordic Journal of Literacy Research*, 8(1)

N.B. When citing this work, cite the original published paper.

Ett svenskt akademiskt vokabulärtest – revidering, validering och förslag på vidareutveckling

Marcus Warnby,^{1*} Hans Malmström² & Camilla Bardel¹

¹Stockholms universitet; ²Chalmers tekniska högskola, Sverige

Sammandrag

Ett akademiskt ordförråd betraktas generellt som oundgängligt för akademisk litteracitet, t.ex. för att läsa och förstå akademiska texter. Bedömning av akademisk ordkunskap kan vara relevant i en rad utbildningssammanhang. Syftet med den här studien var att validera ett flervalstest av receptivt svenskt akademiskt ordförråd. Ur ett befintligt frekvensbaserat ordnivåtest extraherades ett svenskt akademiskt vokabulärtest (SAVT). SAVT mäter akademisk ordkunskap utifrån En svensk akademisk ordlista, ESAO. Förstudier följdes av revideringar av SAVT. Därefter administrerades testet tillsammans med en enkät med frågor om social och språklig bakgrund samt självskattad språklig kompetens till 551 gymnasieelever i högskoleförberedande program. Statistiska och lexikala analyser av resultaten genomfördes. Reliabilitet och variation var tillräckligt goda för att kunna diskriminera testdeltagares olika kunskapsnivåer på gruppnivå. En positiv samvariation mellan testresultat och självskattad läsförmåga samt testresultatskillnader mellan grupper av olika kön och förstaspråk stärkte validitetsargumentet. De lexikala analyserna visade att ordens frekvens och polysemi, liksom nyansskillnader i betydelse mellan vissa alternativord, kunde innebära särskilda svårigheter för testdeltagarna, vilket kan ses som typiskt för en testgrupp som utvecklar sin akademiska litteracitet. Förslag till vidareutveckling av testet (och av den underliggande listan av akademiska ord), bland annat med utbyggnad av antalet item, diskuteras.

Nyckelord: Akademiskt ordförråd; svenska; första- & andraspråk; test; validitet

Abstract

A Swedish Academic Vocabulary Test – revision, validation, and suggestions for further development

Academic vocabulary is generally regarded as indispensable for academic literacy, e.g., for reading and understanding academic texts. Assessment of academic vocabulary knowledge can be relevant in a range of educational contexts. The aim of this study was to validate a multiple choice test of receptive Swedish academic vocabulary. A Swedish academic vocabulary test (SAVT) was extracted from an existing frequency-based word levels test. SAVT measures knowledge of words from a Swedish academic word list. Pilot studies were followed by minor revisions to the extracted test. Subsequently, the test, a sociolinguistic survey and a linguistic self-estimation instrument were administered to 551 students in university preparatory upper secondary education. Statistical and lexical analyses were conducted based on the results. Reliability and variance were sufficient and discriminated different proficiency levels among test-takers at group level. Positive covariance

*Korrespondans: Marcus Warnby, e-post: marcus.warnby@su.se

© 2022 Marcus Warnby, Hans Malmström & Camilla Bardel. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

Citation: M. Warnby, H. Malmström & C. Bardel. "Ett svenskt akademiskt vokabulärtest – revidering, validering och förslag på vidareutveckling" *Nordic Journal of Literacy Research*, Vol. 8(1), 2022, pp. 106–127. <http://dx.doi.org/10.23865/njlr.v8.3282>

between test scores and self-reported reading ability and test score differences between groups of different gender and first languages added to the validity argument. The lexical analyses showed that word frequency, polysemy and meaning nuances between certain words could pose difficulties for some students, which is typical for a test group in the process of developing their academic literacy. Suggestions for further development of the test (and the underlying list of academic words), including the addition of more items, are discussed.

Keywords: *academic vocabulary; Swedish; first and second language, test; validity*

Ansvarlig redaktör: Arild Michel Bakken

Mottatt: Juni, 2021; Antatt: Februar, 2022; Publicert: Mai, 2022

Inledning

Nybörjarstudenter inom högre utbildning ställs inför en väsentlig utmaning i form av den stora mängd akademisk text som de behöver tillgodogöra sig (Porter, 2018). Läsning av akademisk text förutsätter ett rikt ordförråd (Qian, 2002; Schmitt et al., 2001) som omfattar 1) ett stort basordförråd, 2) ett tekniskt och domänspecifikt ordförråd och 3) ett akademiskt ordförråd, dvs. ord som är särskilt vanligt förekommande inom de flesta ämnesområden (Nation, 2001). Tidigare forskning i den nordiska kontexten visar att det finns skäl att anta att gymnasieskolan eventuellt inte förbereder eleverna tillräckligt för universitetsstudier (Hellekjær, 2005; Malmström, 2017) och att test i akademiskt ordförråd kan hjälpa lärare att upptäcka elevers språkliga kunskapsluckor (t.ex. Skjelde & Coxhead, 2020).

Studier från den svenska ungdomsskolan visar att lågfrekventa, oftast tekniska, ord får uppmärksamhet i ämnesklassrummet medan det akademiska ordförrådet sällan ägnas samma uppmärksamhet (t.ex. Reierstam, 2020), vilket innebär att elever förväntas förvärva centrala delar av det akademiska språket på egen hand (t.ex. Uddling, 2019). Flera forskare (t.ex. Lim Falk & Holmberg, 2016; Snow & Uccelli, 2009) hävdar dock att det akademiska språket är så specifikt att det kräver medveten undervisning oavsett de studerandes förstaspråk (L1) eller andraspråk (L2).

Då det akademiska ordförrådet spelar en central roll för förståelsen av akademisk text är det viktigt att fastslå vilka ord som är akademiska och att utveckla instrument som gör det möjligt att mäta akademiskt ordförråd, inte minst i syfte att avgöra när pedagogiska insatser för att utveckla ordförrådet kan vara aktuella. För svenskt vidkommande representerar *En svensk akademisk ordlista* (ESAO) (Ribeck et al., 2014) en konstruktion av svenskt akademiskt ordförråd, som bl.a. har använts för att analysera gymnasieelevers textproduktion utifrån akademiskt språkbruk (Lim Falk & Holmberg, 2016; Ohlsson, 2021). ESAO riktar sig i första hand till mer avancerade inlärare med svenska som andraspråk (talare hädanefter kallade L2S), men också till dem med svenska som förstaspråk (hädanefter L1S) som är ovana vid det akademiska språket (Ribeck et al., 2014).

För att mäta effekten av språk- och ämnesintegrerad undervisning på gymnasieelevers svenska ordkunskap konstruerade Lindberg och Johansson (2019) ett

ordnivåtest – *Swedish Levels Test* (hädanefter SLT) – som enligt upphovspersonerna är en jämförbar (”comparable”) svensk version (Lindberg & Johansson, 2019, s. 242) av det välkända engelskspråkiga *Vocabulary Levels Test* (VLT) (Nation, 1990; Schmitt et al., 2001), i vilket det finns en sektion som prövar akademiska ord. Utöver att mäta kunskap om de 10 000 mest frekventa orden i en korpus prövar SLT även ord från ESAO. Lindberg och Johansson (2019) använde SLT för att mäta generell ordförståelse, inte för att uttala sig om individers akademiska ordkunskap, trots att testet torde erbjuda en sådan möjlighet. Till skillnad från den engelskspråkiga förlagan saknar SLT en separat akademisk sektion och ESAO-orden är utspridda över testet. Om ändamålet är att enbart mäta akademiskt ordförråd behövs ett test eller en sektion som gör just det.

Syftet med den föreliggande studien var att validera ett flervalstest av receptivt svenskt akademiskt ordförråd. Detta innebar att en akademisk sektion i SLT, jämförbar med den som finns i VLT (Schmitt et al., 2001), först behövde identifieras. Studien bidrar också till den empiri av valideringsstudier som vokabulärtestningen behöver, då brister inom fältet avseende en systematisk validitetskritik nyligen har lyfts fram (Schmitt et al., 2020).

Studien motiveras av ett reellt behov av ett test som är enkelt att administrera och som ger en indikation om testdeltagares akademiska ordkunskap. Ett sådant test kan fungera både som mätinstrument och pedagogiskt verktyg i olika utbildningskontexter och i forskning om lärande och utveckling av svenskt ordförråd. I den här artikeln beskrivs hur ett svenskt akademiskt vokabulärtest (SAVT) extraherades ur SLT. SAVT:s giltighet, framtida utveckling och användning diskuteras utifrån statistiska och lexikala analyser.

Bakgrund

Validitet och validering

Eftersom den här studien syftar till att validera ett test behöver begreppen validitet och validering, liksom reliabilitet, ges visst utrymme innan testet och dess konstrukt beskrivs och analyseras.

Generellt kan validitet definieras som giltigheten i de tolkningar och inferenser som kan göras av mätresultat, och för att utvärdera denna giltighet krävs en valideringsprocess. Vi ansluter i denna studie till Messick (1989), vars utgångspunkt är att reliabilitet är en del av validitetsbegreppet och som sådan en nödvändig, men inte tillräcklig, förutsättning för validitet. Validering kan inte enbart bygga på matematiska beräkningar utan kräver dessutom teoretiska argument och subjektiva beslut, varför vi här tillämpar en argumentbaserad approach som beaktar flera dimensioner av validitet (Kane, 2006).

Validitetsbegreppet har definierats på olika sätt (Lissitz, 2009). *Samtidig validitet* rör olika instruments resultatöverensstämmelse, där exempelvis ett akademiskt ordkunskapstest som samvarierar positivt med läsförmåga kan ses som ett validitetsargument.

Innehållsvaliditet utreder testinnehållets giltiga estimering av måldomänen, där den operationaliserade domänen (här svenskt akademiskt vokabulärtest) bör överensstämma med den teoretiska (här ESAO). *Konstruktvaliditet* ses som ett övergripande helhetsbegrepp för validitet (t.ex. Kane, 2006) och olika former av evidens behövs för de inferenser som görs utifrån testresultat. Två centrala hot mot konstruktvaliditeten som behöver kontrolleras är *underrepresentation av konstruktet*, dvs. att testet samlar för lite information om hela konstruktet, och *konstrukt-irrelevant varians*, där testresultatet påverkas av fler dimensioner än det avsedda konstruktet.

Validering är den process med vilken testresultat ges betydelse. En validering utgående från argumentbaserade överväganden kring konstruktvaliditet bygger på två slags argument enligt Kane (2006):

- 1) *tolkningsargument* som uttrycker de påståenden som kan göras baserat på tolkning av testresultat.
- 2) *validitetsargument* som fungerar som stöd för att hävda att tolkningen är giltig, valid.

Med avstamp i en argumentbaserad validering analyseras ett svenskt akademiskt vokabulärtest i denna studie.

Akademiskt ordförråd

Att behärska det akademiska språket är centralt för en individs gymnasiala och postgymnasiala utbildningsframgång. I en anglosaxisk kontext avser *akademiskt språk* ofta det språk som en individ möter i samtliga utbildningsformer (Nagy & Townsend, 2012). I svensk kontext talas bl.a. om *skolspråk* (Hyltenstam et al., 2012), *ämnesspråk* (Ribeck, 2015) eller *akademiskt språk* (Ask, 2007).

Vokabulärforskning (t.ex. Nation, 2001) har visat att läsare behöver förstå 95–98 % av orden i en text för att förstå texten som helhet (vilket innebär att ungefär vart tjugonde ord är okänt). Det är känt att ordförståelse samvarierar med läsförmåga (se t.ex. Skolverket, 2001, där flickors större ordkunskap korrelerade med en bättre läsförmåga i relation till pojkar) samt påverkar läsförståelsen i fackämnen (t.ex. Skolverket, 2004). Således kan ett validitetsargument för ett receptivt akademiskt ordkunskapstest vara att testet samvarierar positivt med mått på läsförmåga. För acceptabel textförståelse inom högre utbildning krävs enligt Nation (2001) kunskap om *basord*, *tekniska* ord och *akademiska* ord. Basord förekommer ofta och i alla typer av texter (Jakobsen et al., 2018). Tekniska ord, t.ex. *elektrolyt*, *fotosyntes*, *reformationen* (Lindberg, 2006), förekommer mer sällan och huvudsakligen inom ett specifikt ämnesområde. Akademiska ord, av vissa benämnda som allmänakademiska eller ämnesneutrala (Enström, 2004; Lindberg & Johansson Kokkinakis, 2007; Olsson, 2016), är ”ord som, till skillnad från tekniska termer, utgör en betydande del av löpande akademisk text, oavsett ämnesområde”, t.ex. *perspektiv*, *innehåll*, *respektive* (Ribeck, 2015, s. 24).

Forskning om akademisk vokabulär har blivit framträdande under de senaste decennierna, i synnerhet avseende engelska (Coxhead, 2016). Det akademiska ordförrådet definieras och operationaliseras på aningen olika sätt, vilket resulterat i ett antal akademiska ordlistor (främst på engelska). Utgångspunkten för skapandet av sådana ordlistor är inte sällan pedagogisk och kan sägas vila på forskning som hävdar att ordförrådet är en central förklarande faktor för skolframgång (t.ex. Lindberg, 2006; Saville-Troike, 1984). De akademiska orden i sådana ordlistor utgör cirka 10–14 % av orden i engelsk akademisk text (Jansson et al., 2013; Therova, 2020). Mot bakgrund av detta omfång av akademiska ord i akademisk text, samt det faktum att man behöver ha kännedom om 95–98 % av orden i en text för att förstå den, blir vikten av akademisk ordförståelse tydlig: utan god kunskap om akademiska ord är det svårt att läsa och förstå akademisk text.

Begreppet akademisk vokabulär har kritiserats i litteraturen. Ett problem rör ordens polysemi; akademiska ord kan ha skilda betydelser i olika ämnesdiscipliner (Hyland & Tse, 2007), eller ändra betydelse i en vardaglig kontext (Lindberg, 2006). Vissa, bl.a. Hyland och Tse (2007), anser att ämnesområdet, snarare än en ”allmän” akademisk vokabulär, har större betydelse för vad som är ett relevant ordförråd. Icke desto mindre visar forskning att de (listade) akademiska orden omfattar en ansevärd mängd av orden i akademiska texter. Således kan akademiska ordlistor potentiellt fungera som pedagogiska verktyg, förutom att användas för utveckling av testinstrument kopplade till akademiskt ordförråd.

Forskning gällande akademiska ordlistor på andra språk är inte lika omfattande. I anslutning till akademiskt ordförråd nämns här projektet *OrdiL* (Lindberg & Johansson, 2007) som bl.a. skapade en korpus över *ord i läromedel*. Denna korpus använde Holmegaard för att jämföra ordkunskapsskillnader mellan L1S och L2S tillsammans med andra instrument. Hon fann att L1S hade bättre resultat än L2S och hävdade bl.a. att ”(m)ånga elever som inte har svenska som modersmål har stora problem med att förstå ord och begrepp i skolans texter” (2007, s. 154). *OrdiL* utgör dock inte ett renodlat akademiskt ordförråd, men Holmegaards resultat är likvärdigt av relevans för vår valideringsprocess.

Nämnvärt är också det skandinaviska projektet *Language Use in Nordic Academic Settings* som resulterade i att det idag finns en *Akademisk Ordliste – Bokmål* (Johannessen et al., 2016), en *dansk akademisk ordliste* (Jakobsen, 2018) och *En svensk akademisk ordlista* (ESAO) (Ribeck et al., 2014).

Pionjärarbetet *ESAO*, som finns tillgängligt via Språkbanken¹, innehåller 652 lemmor som kännetecknar svensk akademisk textgenre. Av de 652 orden utgörs 42 % av substantiv, 24 % av verb, 14 % av adjektiv och 8 % är adverb. *ESAO* täcker knappt 12 % av orden i akademiska texter (Jansson et al., 2013).

ESAO är extraherad ur korpusen svensk akademisk korpus (SveAk, tillgänglig via korpusverktyget *KORP*²), innehållande 25,4 miljoner textord från akademiska texter

¹ <https://spraakbanken.gu.se/ao/index.html>

² www.spraakbanken.gu.se/korp

publicerade inom humaniora och samhällsvetenskap. Då användningen av svenska är begränsad när det gäller akademiska publiceringsändamål inom naturvetenskaperna inkluderades inte texter från naturvetenskaperna i korpusen, vilket kan anses vara tveksamt utifrån konstruktdefinitionen.

För att ett ord skulle inkluderas i ESAO, dvs. klassas som akademiskt, krävdes att det var frekvent förekommande i SveAk, jämnt spritt över och oberoende av ämnesområden, inte alltför frekvent i skönlitteratur och dessutom inte ett av de 1000 mest frekventa lättlästa svenska orden. Utifrån dessa parametrar genererades ESAO (se Ribeck et al., 2014, för en mer detaljerad beskrivning).

ESAO:s konstruktörer (Ribeck et al., 2014) nämner flera framtida utvecklingsbehov. Bland annat diskuteras dilemman med polysemi då vissa betydelser av ett ord kan vara vanligare än andra i akademiska texter: "[E]tt exempel är verbet *uppfatta* som enligt ordboken kan betyda 'förstå, tolka' och 'lyckas höra'. Enligt vår bedömning är det främst den första betydelsen som är aktuell i SveAk" (Ribeck et al., 2014, s. 380). Tyvärr möjliggjorde ESAO-projektets ramar inte en ingående granskning av polysemiproblemet. En översiktlig genomgång av ESAO, utförd av försteförfattaren (Warnby, 2019a), visar att flera av orden i listan kan ha en mer akademisk betydelse (ofta abstrakt), en mer vardaglig (ofta konkret) och ibland även en mer domänspecifik betydelse, t.ex. *bära*, *bild*, *föreställning* och *knyta*. En annan genomgång av ESAO-orden i konkordansverktyget visar även att vissa utländska ord taggats som svenska, exempelvis *social* och *art* (Warnby, 2019b). För Ohlsson (2021) och Johansson et al. (2017) framstår vissa ESAO-ord som vardagliga eller allmänspråkliga.³

Observationer av det här slaget väcker frågor om validiteten hos ESAO, även om ordlistans många fördelar överväger. Innan en mer förfinad version utvecklats bör ESAO kunna "tjäna som underlag för utveckling av framtida ordtester" (Ribeck et al., 2014, s. 371), t.ex. för indikativa mått på receptiv akademisk ordkunskap.

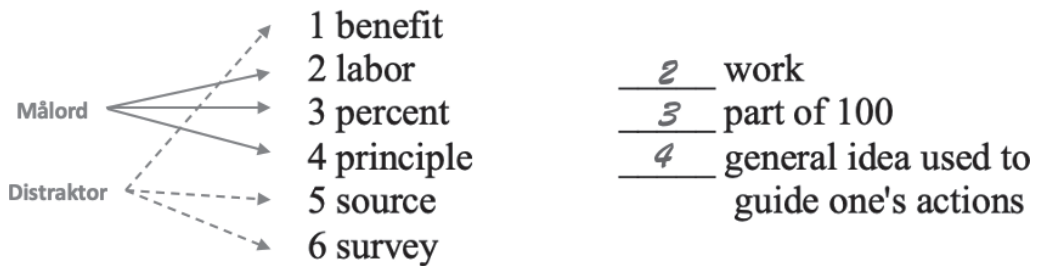
Ordförrådstest

Ett välkänt exempel på ett test av receptiv ordkunskap är VLT, konstruerat av Nation (1990) och sedermera reviderat och validerat av Schmitt et al. (2001). VLT är frekvensbaserat och syftar till att uppskatta storleken på en individs allmänna och akademiska ordförråd. Fyra sektioner är benämnda utifrån de frekvensband om tusen ord som orden i sektionen prövar (2K, 3K, 5K, och 10K). En femte sektion prövar ord från en akademisk ordlista.

Receptiva ordkunskapstest har ofta ett igenkänningsformat där en testdeltagare utgår från ett ord som ges (s.k. *bjudord*) och därefter väljer, från en flervallista, det ord som främst speglar bjudordets betydelse. VLT är ett igenkänningstest, men har ett matchningsformat där orden är grupperade i kluster med en vänsterställd lista

³ Det bör dock noteras att Ohlsson (2018) först eliminerade nästan hälften av de 652 ESAO-orden (extraherade ur en humaniora och samhällsvetenskap) eftersom hon ansåg dem vara allmänspråkliga, och därefter använde sin egen reducerade lista som akademisk ordförrådsnorm i analysen av bl a naturvetenskaplig textproduktion.

med sex ord varav tre är korrekta (målord) och tre är fel (distraktorer), och där testdeltagaren ska matcha målorden med passande ordförklaring (synonymer/förklaringar) i en högerställd lista (se figur 1). Eftersom det rör sig om ett enspråkigt testformat där målordkunskapen mäts med hjälp av andra ord i språket kan hypotesen göras att testdeltagare med målspråket som L2 kommer att uppvisa lägre resultat än de med målspråket som L1 p.g.a. ett mindre utvecklat ordförråd (jfr Holmegaard, 2007). Vi argumenterar nedan för att sådana skillnader i resultat skulle utgöra evidens för validitet.



Figur 1. Exempel på ett korrekt ifyllt kluster i den akademiska sektionen i VLT (Schmitt et al., 2001) med gråmarkerade angivelser av målord och distraktorer.

Nations ursprungliga VLT med sex kluster och 18 målord per sektion har en maxpoäng på 90 poäng, medan det reviderade VLT (Schmitt et al., 2001) har 150 poäng fördelade på 10 kluster och 30 målord per sektion. Ett målord är alltid mindre frekvent än ordförklaringarna. VLT utgår från att testdeltagare ska kunna visa partiell kunskap om ett ord. Därför är orden inom samma kluster tydligt separerade semantiskt.

I den svenska kontexten har ordförrådstest som på olika sätt fokuserar ett utbildningsrelevant språkligt register funnits länge; t.ex. innehåller högskoleprovet sedan slutet 1970-talet en ORD-del⁴. Såvitt vi vet finns det dock i dagsläget bara en dokumenterad studie där svenskt akademiskt ordförråd, såsom det definieras i ESAO, prövas receptivt, nämligen Lindberg och Johansson (2019).

Ett svenskt ordförrådstest

Som en svensk version av VLT konstruerade Lindberg och Johansson (2019) ett frekvensbaserat ordförrådstest (SLT). Förutom ord från frekvensband 1K–10K inkluderades ett urval ord från ESAO. Testkonstruktionen är inte tydligt beskriven och författarna diskuterar inte eventuell validering av testet. SLT är inte publicerat i sin helhet någonstans. Emellertid har Lindberg och Johansson via personlig kommunikation delat med sig av SLT.

Testformatet i SLT liknar VLT i så måtto att det är ett matchningstest med ord i en vänsterkolumn som ska matchas med ordförklaringar. I SLT markerar dock

⁴ <https://www.uhr.se/studier-och-antagning/Hogskoleprovet/Prov-i-atta-delar/>

testdeltagaren den cirkel som passar för vart och ett av alla sex orden (se figur 2). Av de sex orden ska tre (målord) matchas med en ordförklaring, medan övriga tre (distraktorer) aktivt ska markeras som 'inget ord passar'. Enligt Lindberg och Johansson (2019) är det en fördel att alla sex ord i klustret testas även om endast tre ord har uttryckta betydelser.

	1. aktuell	2. betydelsefull	3. påfallande	4. inget ord passar
283. väsentlig	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
284. rådande	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
285. oberoende	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
286. förebyggande	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
287. uppenbar	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
288. verksam	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

Figur 2. Exempel på SLT-kluster från Lindberg & Johansson. Korrekta svar markerade.

SLT innehåller 48 kluster totalt, vilket innebär en totalpoäng om 288 poäng eftersom alla sex ord i varje kluster testas. I motsats till VLT är SLT inte indelat i sektioner för de olika frekvensbanden eller för de akademiska orden; i stället är de akademiska orden spridda över klustren i testet.

Skapandet av ett separat akademiskt test utifrån SLT

Eftersom SLT prövar ord från ESAO var det möjligt att extrahera ett antal ord som endast mäter akademisk ordkunskap. Ambitionen var att dra nytta av det arbete som Lindberg och Johansson (2019) bidragit med genom att i största möjliga mån använda intakta kluster från SLT, t.ex. utan att göra förändringar med avseende på målorden. Däremot eftersträvades att testet skulle likna förebilden VLT– ett väletablerat internationellt riktmärke.

Identifiering av akademiska ord i SLT

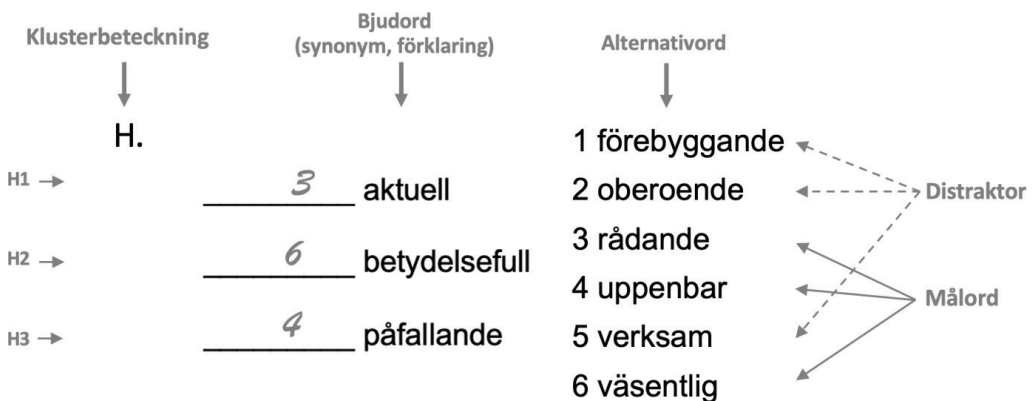
I ett första steg identifierades samtliga akademiska alternativord i SLT:s 48 kluster. Bland testets totalt 288 alternativord ingår 96 ESAO-ord spridda över 33 kluster, där inget, ett, två eller alla tre målorden är akademiska. Därför identifierades

alla kluster där samtliga målord är akademiska; åtta kluster uppfyllde dessa villkor.

De åtta identifierade akademiska klustren i SLT återfinns på olika frekvensnivåer. Det första klustret är hämtat en tredjedel in i SLT, dvs. någonstans i 3K. Det åttonde klustret är det sista i SLT, dvs. från 10K. De övriga sex klustren placerar sig däremellan. Det är känt att frekvensnivå i relation till akademiska ord inte är oproblematiskt (Schmitt et al., 2001). I det här sammanhanget indikerar klustrens frekvensnivåer snarare en sannolikhet att de tidigare orden i testet är mer kända än de senare.

Förstudier och nödvändiga förändringar

Före användning i en fullskalig datainsamling användes de åtta klustren i fyra förstudier som också innebar vissa förändringar. Bland annat föll det sig inte naturligt för testdeltagarna att utgå från de sex orden i vänsterkolumnen för att därefter söka efter en betydelse (eller avgöra om 'inget ord passar'). De utgick i stället från ordförklaringarna för att söka ett passande målord bland de sex orden i vänsterkolumnen. Det blev även uppenbart att man, om man är säker på och svarar rätt på de tre målorden, automatiskt får tre poäng även för 'inget ord passar' utan att ha visat kunskap om distraktorernas betydelse. I en reviderad version av testet placerades därför ordförklaringarna som bjudord i en vänsterkolumn, medan de sex alternativorden placerades i en högerkolumn; svar fick skrivas med en siffra, dessutom ströks 'inget ord passar'. Därmed liknade formatet VLT (se figur 3). Sålunda bestod testet av åtta kluster benämnda A – H och omfattade totalt 24 item (A1, A2, A3, ... H1, H2, H3) med följande ordklassfördelning: 9 substantiv, 9 verb och 6 adjektiv. Två kollegor till författarna (båda akademiker) besvarade avslutningsvis testet med 100 % korrekthet. Förstudierna visade vidare att det tog knappt 10 minuter för de flesta elever (n = 83) att besvara de åtta akademiska klustren; ingen behövde avbryta av tidsskäl. Flera elever undrade om gissning var tillåtet, vilket föranledde ett förtydligande i instruktionen (*Gissa bara om du tror att du kanske vet*).



Figur 3. Exempel på korrekt ifyllda item H1, H2 och H3 i kluster H i SAVT med gråmarkerade angivelser av klustrets beteckning, ordens benämning och funktion.

Ur SLT extraherades alltså ett fristående test i svensk akademisk vokabulär (SAVT). De förändringar som genomfördes under förstudierna kan sammanfattas enligt följande:

1. Alternativet 'inget ord passar' från originalversionen eliminerades och poäng gavs bara för visad kunskap om målorden.
2. Alternativorden placerades i en högerkolumn; bjudorden i en vänsterkolumn.
3. Alla alternativord ordnades alfabetiskt.
4. Enskilda bjudord ordnades alfabetiskt.
5. Flerordsförklaringar ordnades efter antal ord och föregicks av de enskilda bjudorden.

Klustrens egentliga innehåll, dvs. orden, är detsamma som i SLT medan de genomförda formatförändringarna gör att SAVT i större utsträckning liknar VLT:s format. Emellertid bör några kvarvarande skillnader mellan SAVT och den akademiska sektionen i VLT noteras. För det första utgår VLT från ordfamiljsbegreppet medan SAVT bygger på ESAO som utgår från lemma. För det andra är medellängden för en definition i SAVT 1,7 ord, vilket är avsevärt kortare än i VLT där medellängden är 3,8.⁵ För det tredje består SAVT av åtta kluster, vilket är färre än de tio klustren i det reviderade VLT (Schmitt et al., 2001), men samtidigt fler än de sex i det ursprungliga VLT (Nation, 1990). I en empirisk studie undersöktes hur många item som kan behövas för att reliabelt uppskatta en testdeltagares ordkunskap av ett 1000-ord frekvensband (Gyllstad et al., 2020). Deras slutsatser visar att det är bättre att mäta färre frekvensband till förmån för fler item per frekvensband, och att 30 eller fler item är att föredra per 1000 ord. Omräknat i procent utgör ett sample på 30 item av 1000 ord 3 %. SAVT använder 24 item från ESAO:s 652 ord, vilket är 3,7 %.

På det hela taget föreföll SAVT fungera så väl att det kunde utgöra grunden för en större validering. Hela SAVT återfinns i bilagan där samtliga kluster anges i alfabetisk ordning.

Metod

För att validera SAVT gjordes först en större datainsamling där deltagare med begränsad erfarenhet av akademiskt språk besvarade SAVT och en tillhörande enkät med frågor om social och språklig bakgrund samt självskattad läsförmåga.

⁵ Det reviderade VLT (Schmitt et al., 2001) existerar i två likvärdiga versioner. Medellängden på definitionerna för version 1 och 2 är 4 resp. 3,6 ord, vilket ger den sammanslagna rapporterade medellängden 3,8 ord.

Deltagardata analyserades statistiskt. Lexikala analyser föranledda av de statistiska analyserna och av deltagarreaktioner skapade ytterligare underlag för en kvalitativ diskussion om SAVT:s möjligheter att mäta testdeltagares akademiska ordkunskap.

I datainsamlingen deltog 551 gymnasieelever (15–18 år, $M = 16.06$ ($SD = .27$) år) från samtliga sex högskoleförberedande program; eleverna var spridda över 22 klasser och sju skolor i Mellansverige. Genomförandet skedde höstterminen 2019 med papper och penna vid ett lektionstillfälle per klass i början av årskurs 1. Information om studien hade skickats i förväg till skolorna och gavs även muntligt vid datainsamlingen i varje klass av försteförfattaren, som på plats även kunde hantera frågor och reaktioner. Eleverna (samtliga >15 år) fattade själva beslut om samtycke. Studien har genomgått godkänd etikprövning.

I enkäten uppgav eleverna bl.a. kön och L1. Utifrån en femgradig skala fick de även skatta sin allmänna läsförmåga i svenska samt sin läshastighet och läsförståelse av svensk skoltext. Ett högre värde på skalan innebar en ökning av den efterfrågade förmågan/färdigheten på så sätt att 1 = *Svag/Väldigt långsamt/Mycket svårt att förstå* och 5 = *Utmärkt/Väldigt snabbt/Allt är lätt att förstå*.

I internationella mätningar är flickors läsförståelse signifikant högre än pojkars (Skolverket, 2018) och elever med utländsk bakgrund har ”en lägre genomsnittlig nivå jämfört med elever med inhemsk bakgrund” (s. 30). Sådana skillnader i SAVT-resultat mellan flickor och pojkar, samt mellan L1S och L2S, kan utgöra validitetsargument. På liknande vis kan SAVT-resultat som korrelerar positivt med den självskattade läsförmågan vara evidens för validitet. Självskattad förmåga är nära förknippat med *self-concept* och forskare hävdar att *academic self-concept* kan vara en förklarande faktor för elevers utbildningsresultat (t.ex. Bong & Clark, 1999). Elevers självskattningar kan vidare användas som summativ bedömning av förmåga (Taras, 2009), i synnerhet om elever har insikt i målen och förmågorna. I svensk utbildningskontext är elevers självskattande förmåga redan från tidig ålder ett explicit uttryckt mål i läroplanen där det exempelvis står att varje elev ”utvecklar förmågan att själv bedöma sina resultat och ställa egen och andras bedömning i relation till de egna arbetsprestationerna och förutsättningarna” (Skolverket, 2019, kap. 2.7).

För de statistiska analyserna matades alla data in i IBM SPSS Statistics, Version 27.0, där en itemanalys utfördes för att utforska testets inre konsistens. Itemanalysen tillsammans med beskrivande statistik kan inte garantera god validitet utan bidrar till att identifiera mättekniskt problematiska item. För att validera testet utifrån argumenten om kön, L1 och självskattad läsförmåga utfördes T-test och korrelationsanalys.

För lexikala analyser föranledda av den statistiska analysen och av deltagarreaktioner gjordes frekvensjämförelser mellan målord och bjudord samt analyser av lexikal närhet mellan alternativordens betydelser. Via KORP användes en korpus från

det nätbaserade diskussionsforumet Flashback⁶ (3,2 miljoner textord) som referens för ett vardagligt skriftspråksbruk för att göra jämförelser mellan bjudordens och målordens möjliga frekvens i ungdomars skriftspråkliga vardagsliv. För att undvika subjektiva tolkningar om ords betydelser användes fyra standardiserade lexikografiska verktyg som norm: Svensk ordbok (SO), Svenska Akademiens ordlista (SAOL), Lexins svenska lexikon (Lexin; användes också av ESAO:s konstruktörer) samt synonymtjänsten Synonymer.se.⁷

Analys efter administrering av SAVT

Itemanalys

En itemanalys, inkluderande beskrivande statistik för hela testet och för enskilda item samt mått på reliabilitet, genomfördes utifrån testdeltagarnas svar på alla 24 item.⁸ I tabell 1 presenteras översiktlig beskrivande statistik av testresultat från hela urvalsgruppen (N = 551) med mått på medelpoäng, lösningsfrekvens (LF, dvs. medelvärde uttryckt i procent), poängspridning, nedre och övre kvartiler, fördelningsmått för skevhet och toppighet samt andelen nollsvar, dvs. uteblivna svar.

Tabell 1. Beskrivande statistik för SAVT-resultat: Medelpoäng (M), standardavvikelse (SD), lösningsfrekvens (LF), fördelningsvärden för hela urvalsgruppen och andelen nollsvar

N	M	SD	LF	Min	Max	Nedre kvartil	Övre kvartil	Skevhet	Toppighet	Nollsvar i %
551	16.47	3.69	.69	1	24	14	19	-.695 (.104)	.665 (.208)	10

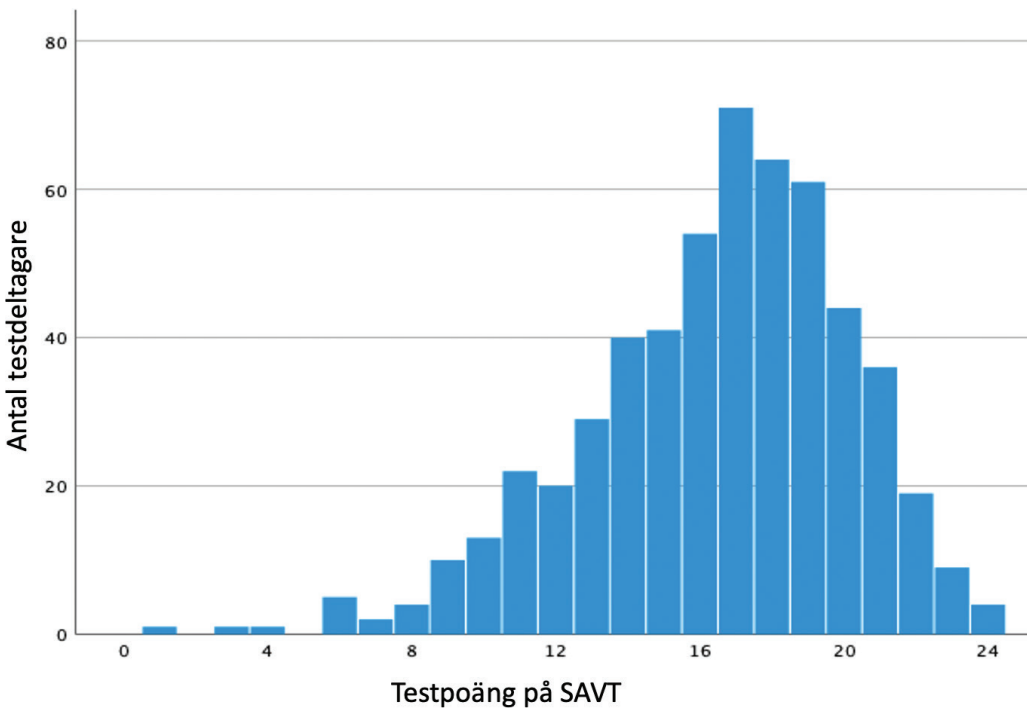
Not: Maxpoäng på SAVT = 24 poäng.

Även om det finns en förskjutning av de relativt toppiga resultaten åt det övre intervallet visar fördelningen av testresultat i stapeldiagrammet nedan (figur 4) att poängen är måttligt normalfördelade. Ingen takeffekt kan noteras.

⁶ KORP innehåller även andra korpusar kategoriserade under sociala medier, t.ex. Familjeliv och Twittermix. Flashback-korpusen ansågs passande för föreliggande studies syfte.

⁷ De fyra tjänsterna nås via internet: SO och SAOL: <https://svenska.se/>; Lexin: <http://lexin.nada.kth.se/lexin/>; Synonymer.se: <https://www.synonymer.se/> och konsulterades 2020-11-16 för en slutkontroll av sökningarna.

⁸ Även om man kan anta att ett målord inom ett kluster inte är oberoende av de övriga två orden i samma kluster visar studier med samma testformat att målorden kan analyseras som enskilda item (Schmitt et al., 2001).



Figur 4. Stapeldiagram över den måttliga normalfördelningen av SAVT-resultaten.

Antalet korrekta svar per item samvarierar negativt med antalet nollsvaret per item från början till slutet. Trenden är en fallande lösningsfrekvens genom testet med högsta lösningsfrekvensen på .99 (item C2) och lägsta på .19 (item H3). Alla item med låg lösningsfrekvens (<.35), totalt fyra stycken, återfinns från mitten och mot slutet av testet, medan fyra item med mycket hög lösningsfrekvens (>.95), återfinns i testets två första kluster. När det gäller antalet nollsvaret per item, som uppgår till i genomsnitt 10 % över hela testet, är trenden ett ökande antal nollsvaret genom testet. Noterbart är att item G1 och H3 som är de två item med lägst lösningsfrekvens har fler nollsvaret än korrekta svar.

Cronbachs alpha är tillfredsställande, $\alpha = .762$, i synnerhet för användningar på grupp-nivå. Reliabilitetsmåttets påverkan av eliminering av enskilda item och diskrimineringsindex utifrån ett korrigerat punkt-biseriellt samband undersöktes. I analysen identifierades item D1 (LF = .51) som eventuellt problematiskt med ett lågt diskrimineringsindex (.093) och en marginellt positiv påverkan på α om D1 eliminerades ($\alpha = .771$). Övriga item ligger inom ett intervall med godtagbar till god diskriminering (.216 – .463). Ett samband mellan ett lägre diskrimineringsindex och en hög lösningsfrekvens kan skönjas, förutom för tre item med DI < .3 (A2, D1 och G2).

En utgångspunkt för diskrimineringsindex för ett item är att den grupp av testdeltagare som matchat bjudordet med målordet har en högre medelpoäng på hela testet än grupper som valt ett felaktigt alternativord. En avvikelse från denna utgångspunkt

kan innebära ett hot mot testets validitet. För item D1, som uppvisat lågt diskrimineringsindex, presenteras i tabell 2 svarsfördelningen: antalet testdeltagare som valt olika alternativ eller lämnat blankt, samt dessa gruppers medelpoäng på hela testet.

Tabell 2. Fördelning av svarsbeteende på item D1: antal testdeltagare (n), medelpoäng (M) och standardavvikelse (SD)

Svarsalternativ	n	M	SD
1	206	16,17	3,67
2	26	14,23	3,67
3	9	12,11	4,65
4 (målord)	280	17,3	3,22
5	1	11,0	–
6	5	13,2	2,17
Nollsvar	24	14,46	5,2

För item D1 med bjudordet *förklaring* har distraktor 1 (*analys*) valts av 206 (37 %) testdeltagare som också har en medelpoäng (16,17) på testet som är lägre än hela urvalgruppens medelpoäng, medan det korrekta alternativet 4 (*tolkning*) attraherat flest testdeltagare (n = 280) vilka också har en medelpoäng (17,3) som överstiger medelpoängen. Svarsalternativ 2 och nollsvar samlar vardera cirka 5 % av testdeltagarna som dock har medelpoäng i nivå med nedre kvartilen. Analysen visar därmed att item D1 diskriminerar reliabelt, även om den stora gruppen i svarsalternativ 1 har en tydlig inverkan på testets inre konsistens.

Korrelationsanalyser

En korrelationsanalys utfördes för att undersöka hur testet samvarierar med ett närliggande konstrukt, nämligen självskattad läsförmåga. Analysen visade svaga men positiva signifikanta samband mellan testresultat och tre självskattningar: 1) läsförmåga i svenska ($r(523) = .259, p < .001$), 2) förståelse av svensk skoltext i skolan ($r(523) = .241, p < .001$) och 3) läshastighet ($r(523) = .107, p < .005$). Ett sammansatt index av de två första variablerna användes i en kompletterande analys som också gav ett signifikant positivt samband, $r(523) = .303, p < .001$.

För att undersöka eventuell evidens för att SAVT kan kopplas till gruppskillnader utfördes separata oberoende T-test mellan testresultat och de två självrapporterade dikotoma variablerna kön (n = 550) och L1, där endast de som uppgett sig ha ett enda L1 inkluderades (n = 482), här separerade som L1S och L2S. Av gruppen L2S (n = 75) rapporterades 26 olika förstaspråk, där de vanligast förekommande var kurdiska (n = 13), somaliska (n = 11) och arabiska (n = 11).

Flickors resultat (n = 300, M = 17.04, SD = 3.25) är högre än pojkars (n = 250, M = 15.81, SD = 4.05). Även om Cohens d (.34) indikerar en måttlig effekt, vilket betyder att det finns en stor överlappning, är skillnaden statistiskt signifikant ($p < .001$).

En statistiskt signifikant skillnad ($p = <.001$) framkommer även mellan de två språkgrupperna, där L1S har ett högre testresultat ($n = 407, M = 17.22, SD = 3.24$) än L2S ($n = 75, M = 13.32, SD = 3.89$). Cohens d (1.16) indikerar en mycket stor effekt, vilket betyder att en betydande andel av L2S-testdeltagare (givet liknande kontext) sannolikt skulle få ett testresultat som ligger i nivå med nedre kvartilens testmedelpoäng, vilket i sin tur stödjer validitetsargumentet.

Lexikala analyser

Förmedlade reaktioner på enskilda item under genomförandet var mycket ovanliga. De enda tre handuppräkringar som gjordes är värda att uppmärksammas, framför allt eftersom de kan kopplas till item med låga lösningsfrekvenser. Två elever (L1S) menade att de förstod samtliga alternativord i kluster A, men inte bjudordet *jakande* i item A2, vilket visade sig ha en hög andel nollsvär. En annan elev (L2S) menade att det saknades matchande svarsalternativ till bjudordet *rymma* som ska matchas med målordet *omfatta* i kluster G. Samma item, G1, visade sig vara ett av de item med lägst lösningsfrekvens. Den sistnämnda reaktionen aktualiserar polysemiproblematiken där ordens betydelse i en akademisk kontext kan skilja sig från den i en annan kontext, vilket i sig inte är ett validitetshot mot testet, men pekar på att andraspråkstalares, men sannolikt även förstaspråkstalares, tillgodogörande av kunskapsinnehåll skulle kunna hindras av brister i djupkunskap av vissa ord.

En annan fråga som reaktionerna aktualiserar är ordens frekvens. Om ett målord är mer frekvent i ungdomars språkliga värld än ett bjudord kan det bli svårt att hävda att det är målordet som testas. Av det skälet analyserades de uppmärksammade ordens frekvens i Flashback-korpusen. Bjudordet *rymma* är i sin akademiska betydelse långt mer frekvent förekommande i korpusen än målordet *omfatta*. Det motsatta förhållandet gäller dock för *jakande* som ingår i bjudorden ”svara jakande”, vilket ska matchas med målordet *bekräfta*. I korpusen förekommer *bekräfta* ungefär 140 gånger per enskild förekomst av *jakande*. Huruvida övriga bjudord är mer frekventa än målorden undersöks inte här, men de två exemplen belyser problematiken.

Ett stort antal elever valde att matcha bjudordet *förklaring* i item D1 med distraktorn *analys*, medan majoriteten valde målordet *tolkning* (se tabell 2). Validitetsfrågan handlar här om de två alternativorden: ligger de så nära varandra semantiskt att de bör ses som synonymer? För bjudordet *förklaring* finns emellertid inget i ordböckerna (SO, SAOL eller Lexin) eller på synonymer.se som kan kopplas till eller påminner om distraktorn *analys*. Visserligen kan det sägas föreligga ett logiskt samband mellan orden på så sätt att en *analys* föregår en *förklaring* men att de har samma betydelse kan inte hävdas och därmed kan det inte heller ses som partiell kunskap om man väljer *analys*. Följaktligen verkar item D1 vara ett item som tydligt innebär kunskap om nyansskillnaden hos ord, vilket kan vara viktigt vid akademiska studier. Den statistiska analysen visade även att elever med hög medelpoäng svarade korrekt, medan de med lägre medelpoäng valde en distraktor, vilket sammantaget stärker den innehållsliga validiteten, trots negativ inverkan på den inre konsistensen.

En analys som den ovan gjordes på samtliga ord inom varje kluster där alternativord och målord kontrollerades för semantisk närhet. Av SAVT:s samtliga alternativord visar sig kluster F ha två alternativord, *innebära* och *utgöra*, vilka dyker upp som möjliga synonymer till varandra på synonymer.se. Det är dock enbart målordet *innebära* – och inte distraktorn *utgöra* – som kan kopplas till bjudordet *medföra* i item F2.

Diskussion och slutsatser

Utifrån testet – SLT – framtaget av Lindberg och Johansson (2019) identifierade denna studie en akademisk sektion jämförbar med den som finns i VLT (Nation, 1990; Schmitt et al., 2001) med syftet att validera denna sektion som ett fristående indikativt test av svensk akademisk vokabulär (SAVT). Här prövas receptiv kunskap om 24 svenska akademiska ord från ESAO.

Testresultat och enkätsvar användes för en kritisk analys utifrån argumentbaserade överväganden om testets möjligheter att göra valida mätningar av testdeltagares akademiska ordkunskap.

Studien vilar på konstruktet svensk akademisk vokabulär såsom det definieras i ordlistan ESAO, som täcker de flesta ämnesdiscipliner men inte omfattar de naturvetenskapliga disciplinerna. Även om ESAO kan kritiseras anser vi att inkluderingskriterierna (Ribeck et al., 2014) är explicita och replikerbara. Framför allt utgör ESAO-orden oomtvistligt en betydande del av akademisk text.

Inför datainsamlingen förändrades testformatet jämfört med hur det såg ut i SLT så att det mer liknade VLT. Bland annat ströks SLT-alternativet 'inget ord passar' eftersom vi menar att det inte visar kunskap om distraktorordens betydelse och därmed kan leda till konstrukt-irrelevant varians.

Resultaten visar att det finns variation i mätdata och att SAVT diskriminerar mellan kunskapsnivåer såväl inom urvalsgruppen som mellan de undersökta undergrupperna. Den interna konsistensen anses tillfredsställande åtminstone för denna studies gruppanalyser. Lösningfrekvenserna per item är fallande från början till slut, vilket ses som ett valitt och förväntat resultat eftersom SLT, ur vilket SAVT är extraherat, är frekvensindelad. Testresultaten samvarierade positivt med självskattningar avseende läsförmåga vilket indikerar valida dimensioner av receptiv färdighet. Analyserna visade att testets konstruktion kan hantera kunskap om ords betydelser men blottade också vissa utmaningar i vissa kluster avseende frekvens, polysemi och nyansskillnader mellan ord.

Även om effektstorleken var liten hade flickorna i urvalsgruppen ett signifikant högre resultat i jämförelse med pojkarna, vilket stärker validitetsargumentet utifrån erkända läsförståelseskillnader mellan könen.

En markant stor effektstorlek framkom när vi undersökte skillnader i testresultat mellan undergrupperna med L1S och L2S. Även här hävdar vi att resultaten är valida, utifrån tolkningsargumentet att SAVT mäter kunskap om en kombination av två ord (akademiskt målord + bjudord) och eftersom svenska är L1 för den högre

presterande undergruppen och det, som vi nämnt, finns skäl att anta att L2S har ett mindre utvecklat ordförråd (Lindberg & Johansson, 2007).

Ett kritiskt resonemang kan här föras kring testbias och risk för ojämlika förhållanden, i synnerhet för L2S-inlärare om fallet är som i Reierstam (2020, s. 246) där en ”relativt osynlig språkpedagogik där fokus framför allt låg på ämnesbegrepp och inte så mycket på undervisning av ämnesövergripande språk” användes i undervisning med L2S-inlärare. Ett test där bjudorden ges på L2S-testdeltagares L1 skulle eventuellt ge en säkrare bild av deras kunnande om ESAO-orden, men samtidigt inte bli jämförbart med SAVT. Samtidigt bör det noteras att begränsad kontakt med det akademiska språkliga registret kan förekomma såväl hos L2S som L1S. Vi har också tidigare i studien konstaterat att det akademiska ordförrådet sällan ägnas den explicita undervisning som skulle behövas (Lim Falk & Holmberg, 2016; Reierstam, 2020; Snow & Uccelli, 2009; Uddling, 2019).

Begränsningar

SAVT har vissa begränsningar som innebär att resultat från testet bör tolkas med försiktighet. Till att börja med testar det i sin nuvarande version endast 24 ord från ESAO, dvs. 3,7 % av totalt 652 ord, vilket kan vara ett fall av underrepresentation av konstruktet. Visserligen ligger det över den föreslagna tröskelnivån av 3 % av samplade item från ett frekvensband (Gyllstad et al., 2020). En sådan tröskelnivå får betraktas som ett slags minimum eftersom det är allmänt känt att en ökning av antalet item också i större grad ökar reliabiliteten. Därutöver bör det noteras att Gyllstad et al. (2020) utgick från ett 1000-ords frekvensband. De akademiska orden i ESAO utgör inte ett frekvensband i strikt bemärkelse.

Vidare har VLT-formatet, på vilket SAVT bygger, tidigare fått utstå kritik bland annat för att flervalformatet möjliggör gissning (bl.a. Kamimoto, 2008). Trots att instruktionen var att bara gissa när man trodde sig veta frågade testdeltagare om de skulle gissa (blint). Detta kan tolkas som att eleverna brukar gissa blint i flervalstest, och vi kan inte ta för givet att alla deltagare alltid svarat efter bästa förmåga. Det spridda antalet nollsvaret i hela testet tolkas dock som att testdeltagarna som grupp genomfört testet så gott de kunnat och följt instruktionen att inte gissa blint.

Ett problem med SAVT kan vara att vissa bjudord är mindre frekventa än sina målord. Om testdeltagare förstår målord men inte bjudord kan de tappa tilltro till testets giltighet. Ett känt fenomen i receptiva ordförrådstest som VLT är att icke-vised förståelse inte kan tolkas som icke-förståelse endast av målordet utan även av bjudordet (Read, 2000; Enström, 2004). VLT-formatet kan därmed inte sägas pröva enbart målord utan snarare den ordkombination som målordet utgör tillsammans med bjudord. Oavsett denna problematik när igenkänningstest används bör en konstruktionsprincip som utgår från att bjudorden är mer frekventa än de avsedda målorden generera ett robustare test.

Även om den positiva och signifikanta korrelationen mellan testresultat och läsförmåga är evidens för validitet kan den svaga korrelationen bero på att självskattningen

utgick från testdeltagarnas allmänna läsförmåga och deras läsförståelse av skolrelaterade texter, dvs. inte explicit av akademisk text.

Framtida utveckling

Validering bör ses som en ständigt pågående process (Schmitt et al., 2020). I vår avsikt att validera SAVT har vi identifierat några brister, och vi föreslår vidareutveckling av konstruktet akademiskt ordförråd och av SAVT framför allt avseende följande:

- Undersök möjligheter att inkludera naturvetenskapliga texter i ESAO.
- Bygg ut SAVT med fler kluster som på ett balanserat sätt representerar ESAO:s ordklassfördelning och olika frekvensnivåer för ökad representation av konstruktet.
- Analysera hur en revidering av item A2, D1 och G2 samt tillägg av item kan bidra till testets inre konsistens.
- Använd bjudord som är mer frekventa än målord.

Att mäta ordförråd på den mikronivå som görs i ett test som SAVT förutsätter vissa antaganden. De separata orden testas, ges poäng och aggregeras för att säga något generellt om individers akademiska ordkunskap. SAVT prövar enbart 24 av ESAO:s 652 ord. För att uppskatta vad ett testresultat hypotetiskt kan betyda i termer av kunskap om hela ESAO får testpoängen multipliceras med 27,16. Urvalsgruppens medelpoäng (16,48) indikerar en kunskap om cirka 448 ESAO-ord, dvs. att drygt två tredjedelar av orden är kända. Eftersom ESAO täcker in cirka 12 % av orden i akademisk text kan vi spekulera vad akademisk ordkunskap i nivå med den som uppmättes här får för betydelse för läsning av akademisk text. Om två tredjedelar (dvs. 8 %) av de akademiska orden är kända skulle det innebära att ≈ 96 % av alla orden i en akademisk text är kända ($100 \% - 12 \% + 8 \%$); det här tankeexperimentet bygger emellertid på att läsaren kan *alla andra ord* i texten, högfrekventa såväl som lågfrekventa (ofta tekniska ord) vilket långt ifrån alltid kommer att vara fallet. I praktiken skulle detta kunna innebära att en stor grupp elever (i synnerhet de elever som har en akademisk ordkunskap som är lägre än medelvärdet i urvalsgruppen) kan komma att uppleva läsning av akademisk text som utmanande eftersom det krävs kunskap om 95–98 % av orden i en text för att förstå den. För att göra sådana uttolkningar, s.k. prediktiv validitet, bör framtida studier kombinera test för att säkra deras betydelse för akademisk läsförståelse. Vi menar att SAVT i dagsläget ensamt inte kan utgöra underlag för sådana kunskapsanspråk.

Den här versionen av SAVT utgör ett första viktigt steg i utvecklingen av ett instrument för att estimerar ett högskolerelevant ordförråd, såväl för högskoleförberedande gymnasieelever som högskolestudenter som är i början av sin akademiska språkutveckling. Testet kan rimligen redan användas för att tentativt uttala sig om huruvida grupper av individer saknar en minsta tillräcklig ordförrådsstorlek för att läsa akademisk text. Det betyder inte att kunskap om ESAO eller full poäng på SAVT är

tillräckligt för att läsa och tillgodogöra sig akademisk text, men utan den kunskapen kommer läsningen med största sannolikhet att påverkas negativt.

Det akademiska språket har så många egna särdrag och konventioner att man inte kan anta att det tillägnas och automatiseras utan fokuserad undervisning. Fördjupade studier vad gäller testresultat och deras samband mellan t.ex. olika testdeltagargrupper och olika utbildningsnivåer/-inriktningar samt test av parallella litteracitetsförmågor behövs därför för att vidare undersöka utbildningsvetenskapliga dimensioner. Vidare bör det noteras att SAVT prövar det svenska akademiska skriftspråkets vokabulär, men forskning om tillägnande av talat akademiskt språk saknas, vilket är en nog så viktig dimension av det akademiska språk studenter möter i svensk högskolekontext.

Syftet med framtida vidareutveckling av SAVT förblir oförändrat: SAVT ska kunna användas för att uttala sig om storleken på akademisk vokabulär hos såväl grupper som individer.

Författarbiografi

Marcus Warnby är doktorand i språkdidaktik inom forskarskolan SEMLA för lärarutbildare vid Stockholms universitet. Marcus avhandlingsarbete behandlar akademisk ordkunskap hos gymnasieelever i relation till språkdidaktiska och utbildningsvetenskapliga dimensioner.

Hans Malmström är biträdande professor i vetenskapens kommunikation och lärande vid Chalmers tekniska högskola. Hans forskning handlar främst om engelska som undervisningsspråk i högre utbildning, och om hur integrering av ämnesinnehåll och kommunikation påverkar undervisning och lärande och utvecklingen av akademisk och disciplinär litteracitet.

Camilla Bardel är professor i moderna språk med didaktisk inriktning och vetenskaplig ledare för forskarskolan SEMLA vid Stockholms universitet. Camillas forskningsintressen ligger inom språkdidaktik och flerspråkighet, med fokus främst på tredjespråkets ordförråd och grammatik.

Referenser

- Ask, S. (2007). *Vägar till ett akademiskt skriftspråk*. Växjö University Press.
- Bong, M. & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational psychologist*, 34(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1
- Coxhead, A. (2016). Acquiring academic and disciplinary vocabulary. I K. Hyland & P. Shaw (Red.), *The Routledge handbook of English for academic purposes* (s. 177–190). Routledge. <https://doi.org/10.4324/9781315657455>
- Enström, I. (2004). Ordförråd och ordinläring – med särskilt fokus på avancerade inlärare. I K. Hyltenstam & I. Lindberg (Red.), *Svenska som andraspråk – i forskning, undervisning och samhälle* (s. 171–195). Studentlitteratur.
- Gyllstad, H., McLean, S. & Stewart, J. (2020). Using confidence intervals to determine adequate item sample sizes for vocabulary tests: An essential but overlooked practice. *Language Testing*, 38(4), 558–579. <https://doi.org/10.1177/0265532220979562>

- Hellekjær, G. O. (2005). The acid test: Does upper secondary EFL instruction effectively prepare Norwegian students for the reading of English textbooks at colleges and universities? [Doktorsavhandling, Universitetet i Oslo]. <http://urn.nb.no/URN:NBN:no-11346>
- Holmegaard, M. (2007). Långa ord – en svårighet för flerspråkiga studerande? I I. Lindberg & S. Johansson-Kokkinakis (Red.), *OrdiL – en korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år* (s. 135 – 170). Göteborgs universitet.
- Hyland, K. & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL quarterly*, 41(2), 235–253. <https://doi.org/10.1002/j.1545-7249.2007.tb00058.x>
- Hyltenstam, K., Axelsson, M. & Lindberg, I. (2012). *Flerspråkighet: en forskningsöversikt*. Vetenskapsrådet.
- Jakobsen, A. S. (2018). *Danish academic vocabulary: Four studies on the words of academic written Danish*. [Doktorsavhandling]. Köpenhamns universitet.
- Jakobsen, A. S., Coxhead, A., & Henriksen, B. (2018). General and academic high frequency vocabulary in Danish. *Nordand*, 2(01), 64–89. <https://doi.org/10.18261/issn.2535-3381-2018-01-04>
- Jansson, H., Ribbeck J., & Sköldberg, E. (2013). *Från aspekt till övergripande – en nätbaserad lista över svenskt akademiskt ordförråd*. [Konferensbidrag] NFL – Nordiska Föreningen för Lexikografi. https://spraakbanken.gu.se/ao/pdf/Jansson_Ribbeck_Sk%C3%B6ldberg.pdf
- Johannessen, J. M., Saidi, A. & Hagen, K. (2016). Constructing a Norwegian Academic Wordlist. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (s. 1457–1462). European Language Resources Association. <https://aclanthology.org/L16-1232.pdf>
- Johansson, S., Hagen, K. & Johannessen, J. B. (2017). A bilingual academic word list: the merging of a Norwegian and a Swedish list. *Oslo Studies in Language*, 9(3), 1457–1462. <https://doi.org/10.5617/osla.5854>
- Kamimoto, T. (2008). Nation’s vocabulary levels test and its successors: a re-appraisal. ProQuest Dissertations Publishing.
- Kane, M. (2006). Validation. I R. L. Brennan (Red.), *Educational measurement. Fourth Edition* (s. 17–64). Praeger – Greenwood publishing group.
- Lim Falk, M. & Holmberg, P. (2016). Paths to academic writing in a globalized world: A longitudinal study of content and language integrated learning in upper secondary school in Sweden. I S. Plane, C. Bazerman, F. Rondelli, C. Donahue, A. N. Applebee, C. Boré, P. Carlino, M. Marquilló Larruy, P. Rogers & D. R. Russell (Red.), *Research on writing: multiple perspectives* (s. 161–182). The WAC Clearinghouse. <https://doi.org/10.37514/int-b.2017.0919.2.09>
- Lindberg, I. (2006). Bedömning av skolrelaterat ordförråd. I M. Olofsson (Red.), *Symposium 2006: Bedömning, flerspråkighet och lärande* (s. 83–107). HLS förlag.
- Lindberg, I. & Johansson Kokkinakis, S. (2007). *OrdiL: en korpusbaserad kartläggning av ordförrådet i läromedel för grundskolans senare år*. Göteborgs universitet.
- Lindberg, I. & Johansson, S. (2019). The Development of Swedish receptive vocabulary in CLIL: A multilingual perspective. I L. K. Sylvén (Red.), *Investigating content and language integrated learning: insights from Swedish high schools* (s. 236–258). Multilingual Matters Bristol.
- Lissitz, R. W. (Red.) (2009). *The concept of validity: Revisions, new directions and applications*. IAP.
- Malmström, M. (2017). *Syner på skrivande: Föreställningar om skrivande i mediedebatter och gymnasieskolans läroplaner*. [Doktorsavhandling]. Lunds universitet.
- Messick, S. (1989). Validity. I R. L. Linn (Red.), *Educational measurement. 3rd edition* (s. 13–103). Macmillan Publishing.
- Nagy, W. & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91–108. <https://doi.org/10.1002/rrq.011>
- Nation, P. (1990). *Teaching and learning vocabulary*. Heinle & Heinle Publishers.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Ohlsson, E. (2021). *Den synliggjorda vokabulären och praktiken: Gymnasieelevers akademiska skrivande på svenska*. [Doktorsavhandling, Göteborgs universitet]. <http://hdl.handle.net/2077/68077>
- Olsson, E. (2016). *On the impact of extramural English and CLIL on productive vocabulary*. [Doktorsavhandling]. Göteborgs universitet.
- Porter, H. D. (2018). Constructing an understanding of undergraduate disciplinary reading: An analysis of contemporary scholarship. *Journal of college reading and learning*, 48(1), 25–46. <https://doi.org/10.1080/10790195.2017.1362970>
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language learning*, 52(3), 513–536. <https://doi.org/10.1111/1467-9922.00193>
- Read, J. A. S. (2000). *Assessing vocabulary*. Cambridge University Press.

- Reierstam, H. (2020). *Assessment in Multilingual Schools: A comparative mixed method study of teachers' assessment beliefs and practices among language learners-CLIL and migrant students*. [Doktorsavhandling]. Stockholms universitet.
- Ribeck, J. (2015). *Steg för steg: naturvetenskapligt ämnesspråk som räknas*. [Doktorsavhandling]. Göteborgs universitet.
- Ribeck, J., Jansson, H. & Sköldberg, E. (2014). Från aspekt till övergripande – en ordlista över svensk akademisk vokabulär. *Nordiske Studier i Leksikografi* (12), 370–384. <https://tidsskrift.dk/nsil/article/view/21029>
- Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, 18(2), 199–219. <https://doi.org/10.2307/3586690>
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88. <https://10.1177/026553220101800103>
- Schmitt, N., Nation, P. & Kremmel, B. (2020). Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation. *Language Teaching*, 53(1), 109–120. doi.org/10.1017/S0261444819000326
- Skjelde, K. M. & Coxhead, A. (2020). Mind the gap: Academic vocabulary knowledge as a predictor of English grades. *Acta Didactica Norden* 14(3): 20 sidor. <https://doi.org/10.5617/adno.7975>
- Skolverket. (2001). *PISA 2000: svenska femtonåringars läsförmåga och kunskaper i matematik och naturvetenskap i ett internationellt perspektiv*. Skolverket. <https://www.skolverket.se/getFile?file=904>
- Skolverket. (2004). *PISA 2003: svenska femtonåringars kunskaper och attityder i ett internationellt perspektiv*. Skolverket. <https://www.skolverket.se/publikationer?id=1374>
- Skolverket. (2018). *PISA 2018: 15-åringars kunskaper i läsförståelse, matematik och naturvetenskap*. Skolverket. <https://www.skolverket.se/publikationer?id=5347>
- Skolverket. (2019). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011: reviderad 2019*. Skolverket. <https://www.skolverket.se/getFile?file=4206>
- Snow, C. E. & Uccelli, P. (2009). The challenge of academic language. I N. Torrance & D. R. Olsson (Red.), *The Cambridge handbook of literacy* (s. 112–133). Cambridge university. <https://doi.org/10.1017/cbo9780511609664.008>
- Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of further and higher education*, 33(1), 57–69. <https://doi.org/10.1080/03098770802638671>
- Therova, D. (2020). Review of Academic Word Lists. *TESL-EJ*, 24(1), 1. <http://www.tesl-ej.org/wordpress/issues/volume24/ej93/ej93a5/>
- Uddling, J. (2019). *Textsamtalens möjligheter och begränsningar i språkligt heterogena fysikklassrum*. [Doktorsavhandling] Stockholms universitet.
- Warnby, M. (2019a). *Piloting, investigating and analyzing an academic Swedish word test*. [Konferensbidrag] QRM 2019 – Quantitative Research Methods in Education Conference 2019, Göteborgs universitet, Sverige.
- Warnby, M. (2019b). *Exploring the Swedish and English passive academic vocabulary of upper secondary students*. [Konferensbidrag] EuroSLA29, Lund, Sverige.

BILAGA. SAVT

INSTRUKTIONER

- Koppla ihop förklaringen till vänster med det ord i listan till höger som du tycker passar bäst. Skriv siffran (1–6) på linjen framför varje förklaring.
- Arbeta så snabbt du kan!
- Flera ord kan vara svåra, men se till att göra så gott du kan!
- Om du inte vet vad ett ord betyder, så gissa inte blint! Gissa bara om du tror att du kanske vet vad ordet betyder!

Här är ett exempel:

<u> 6 </u> något som används för att åka på vatten	1 skola
<u> 3 </u> ett djur med fyra ben	2 klocka
<u> 5 </u> hål	3 hund
	4 bil
	5 grop
	6 båt

ORDFÖRRÅDSTEST I SVENSKA

Gör så gott du kan!

A. <u> </u> börja	1 bekräfta	E. <u> </u> följd	1 avseende
<u> </u> svara jakande	2 bära	<u> </u> omfattning	2 belopp
<u> </u> ha på sig	3 inleda	<u> </u> standard	3 gods
	4 slippa		4 konsekvens
	5 stämma		5 norm
	6 utreda		6 utsträckning
B. <u> </u> en positiv sak	1 beteende	F. <u> </u> influera	1 analysera
<u> </u> när man är borta	2 drama	<u> </u> medföra	2 innebära
<u> </u> sätt att uppföra sig	3 egendom	<u> </u> upphöra	3 påverka
	4 frånvaro		4 relatera
	5 fördel		5 utgå
	6 samtal		6 utgöra
C. <u> </u> avsiktlig	1 ekologisk	G. <u> </u> rymma	1 avse
<u> </u> äkta	2 enorm	<u> </u> vara	2 beröra
<u> </u> ensam i sitt slag	3 kommersiell	<u> </u> visa sig	3 bestå
	4 medveten		4 beteckna
	5 naturlig		5 framträda
	6 unik		6 omfatta
D. <u> </u> förklaring	1 analys	H. <u> </u> aktuell	1 förebyggande
<u> </u> synvinkel	2 begrepp	<u> </u> betydelsefull	2 oberoende
<u> </u> som börjar ifrån	3 perspektiv	<u> </u> påfallande	3 rådande
	4 tolkning		4 uppenbar
	5 utbrott		5 verksam
	6 utgångspunkt		6 väsentlig