



Pearson SEE mbo

Richtlijnenbundel

voor het gebruik van *Pearson SEE mbo*
als instellingsexamen





Beste lezer,

U hebt een flink document in handen.

Dat komt doordat we veel informatie met u willen delen over Pearson SEE mbo, het instellingsexamen Engels voor mbo-studenten, waar we zeer trots op zijn! Waarom?

Daar zijn vele redenen voor, maar de belangrijkste is dat Pearson SEE mbo precies bepaalt wat het eindniveau per vaardigheid is van uw studenten binnen de A2- of B1-schaal, zonder dat u dat veel tijd en moeite kost. Het examen duurt namelijk maar één uur en binnen drie uur hebt u de resultaten van alle studenten op uw computer.

Te mooi om waar te zijn? Nee. Pearson is een gecertificeerde examenleverancier. Dat betekent dat Pearson SEE een valide exameninstrument is en voldoet aan norm voor valide exameninstrumenten. De norm bestaat uit producteisen (congruent aan het Onderzoekskader 2017 van de inspectie) en uit proces- en organisatie-eisen.

In dit document kunt u onder andere lezen hoe u het examen het beste kunt gebruiken, aan welke voorwaarden u moet voldoen wat betreft ICT en techniek en leggen wij verantwoording af op het gebied van betrouwbaarheid en validiteit.

Wanneer u nog vragen hebt, kunt u uiteraard altijd contact met ons opnemen.

Pearson ELT



Inhoudsopgave

- 1 Indicatoren en exameneisen
- 2 Pearson SEE mbo Handleiding voor docenten
- 3 Pearson SEE mbo Handleiding voor ICT
- Appendix A: Pearson GSE-syllabus
- Appendix B: Progress Technical Report
- Appendix C: Progress Test Information Booklet



1. Indicatoren en exameneisen

De Inspectie van het Onderwijs heeft exameneisen opgesteld waaraan instellingsexamens moeten voldoen. Het beoordelen van het examen gebeurt aan de hand van examenstandaarden en indicatoren, zoals die zijn vastgelegd in Onderzoekskader 2017, Regeling standaarden examenkwaliteit mbo 2017.

Onder het kopje 'Aanwijzingen voor de adviseur exameninstrumenten' kunt u zien hoe deze eisen zijn toegepast op Pearson SEE mbo.

1.1 Indicator 2.1.2 Dekking van het kwalificatiedossier



Ministerie van Onderwijs, Cultuur en
Wetenschap

Portret (Talen)

Bij de inrichting van kwalificerende examens is sprake van volledigheid en passende complexiteit, om de beroepsuitoefening op het vereiste beheersingsniveau zoals omschreven in het kwalificatiedossier te examineren. Dit betekent dat:

- *de exameninstrumenten de generieke kwalificatievereisten voor talen zodanig afdekken dat recht wordt gedaan aan het ERK-niveau dan wel het referentieniveau.*
- *de beroepsgerichte eisen voor moderne vreemde talen, zoals geformuleerd in deel c van het kwalificatiedossier, in voldoende mate terug moeten komen in het examen. Als taaleisen terugkomen bij meer kerntaken en/of werkprocessen, dan mogen die desgewenst ook geïntegreerd geëxamineerd worden. Daarbij moet dan wel de link met de kerntaak/werkproces gelegd worden. Als de moderne vreemde taaleisen zeer specifiek zijn geformuleerd bij een bepaald werkproces, dan moeten deze terugkomen in de examinering van dat werkproces.*

Toelichting

- *ERK-niveaus. Voor Nederlands en Moderne vreemde talen is bepaald dat alle vijf taalvaardigheden (luisteren, lezen, gesprekken voeren, spreken en schrijven) per taal geëxamineerd dienen te zijn met examens die qua ERK-niveau en qua toetsvorm passen bij de taalvaardigheden waarover een beoordeling wordt gegeven.*
- *Referentieniveaus. Op basis van de brief van OCW van 29-1-2010 met het onderwerp: "Implementatie referentieniveaus taal en rekenen mbo. Kenmerk BVE/Stelsel/180716" gaat het toezicht op Nederlands voor opleidingen die na 1 augustus 2010 zijn gestart, vanaf 2010/2011, uit van het volgende: De instellingsexamens voor de vaardigheden spreken, gesprekken voeren, schrijven en taalverzorging dienen geëxamineerd te zijn met examens die qua referentieniveau en qua toetsvorm passen bij de taalvaardigheden waarover een beoordeling wordt gegeven. Het combineren van verschillende taalvaardigheden in één examen is mogelijk (bijvoorbeeld spreken en gesprekken voeren). Het examen moet de kandidaat wel de gelegenheid bieden om voor elke taalvaardigheid aan te tonen dat hij deze taalvaardigheid op het vereiste taalniveau beheerst. Voor de vaardigheden lezen en luisteren wordt alleen een verantwoording opgevraagd.*

Aanwijzingen voor de adviseur exameninstrumenten:

a. Wordt het juiste ERK-niveau dan wel referentieniveau gehanteerd, zoals opgenomen in het kwalificatiedossier?

De scores van Pearson SEE mbo worden aangegeven op de Global Scale of English (GSE), die gerelateerd is aan de ERK-niveaus. Op een schaal van 10 tot 90 geeft de GSE aan wat het niveau is van de taalvaardigheid op het gebied van Engels. Pearson SEE mbo geeft ook de scores weer die overeenkomen met de ERK-niveaus (bijvoorbeeld A1, A2, etc.) De toets is empirisch onderzocht en geeft met precieze gradaties het niveau en de vooruitgang van een kandidaat aan op het gebied van de ERK-normen.

Er is geen centraal accreditatieschema waar testontwikkelaars zich op kunnen beroepen om hun claims te ondersteunen dat de test overeenkomt met ERK. De Raad van Europa heeft echter wel een uitgebreide handleiding gepubliceerd (Raad van Europa, 2009) waarin procedures zijn beschreven die testontwikkelaars kunnen gebruiken om hun claims mee te ondersteunen. Deze procedures zijn geïmplementeerd op de noodzakelijke punten tijdens het ontwikkelproces van Progress. Meer informatie over dit proces vindt u in het Progress Technical Report. Pearson SEE mbo is een gecustomiseerde versie van Progress. Alle informatie in het Technical Report en Test Information Booklet is van toepassing op Pearson SEE mbo.

b. Is per taalvaardigheid de toetsvorm passend?

Om elke vaardigheid te testen, zijn zorgvuldig geconstrueerde toetsvragen nodig die werkelijk de vereiste taalvaardigheid van studenten testen en geen andere vaardigheden of technieken, zoals het geheugen, het afnemen van een test, het anticiperen, het raden van antwoorden, etc. De toetsonderdelen in Pearson SEE mbo zijn ontwikkeld door experts en zijn uitgebreid getest, waardoor de toetsresultaten ook het werkelijke taalniveau aangeven. Voor meer informatie over schrijven, selectie en de verschillende toetsonderdelen van Pearson SEE mbo, zie sectie 3 van het Progress Technical Report.

Pearson SEE mbo is een geheel geautomatiseerde en digitale toets. De toets is online toegankelijk via een stabiel internetplatform. De toets bevat zowel audioteksten, audiotekstfragmenten als geschreven teksten en tekstfragmenten. In de toets kan dus worden geluisterd (via de luidsprekers van de computer) en gelezen (via het beeldscherm van de computer). Ook bevat de toets onderdelen waarbij kan worden geschreven (met het toetsenbord van de computer). Verder bevat de toets onderdelen waarbij de kandidaat kan spreken (in de microfoon van de computer). Alle vaardigheden, zowel de receptieve als de productieve, komen dus aan bod.

Een lijst met testonderdelen van Pearson SEE mbo vindt u in het Progress Test Information Booklet. Met testonderdelen worden vaak meerdere vaardigheden tegelijkertijd getoetst. Bijvoorbeeld bij het toetsonderdeel 'Listen and then write' wordt zowel een score gegeven voor het luisteren als voor het schrijven. Deze geïntegreerde benadering om vaardigheden te toetsen, komt overeen met natuurlijk taalgebruik. Daarom geeft Pearson SEE mbo een zeer

accuraat beeld van de werkelijke taalvaardigheid. Pearson SEE mbo is een zeer passende toetsvorm voor het meten van taalvaardigheid.

c. Sluiten de opdrachten en teksten aan bij het vereiste ERK-niveau dan wel referentieniveau?

Opdrachten in Pearson SEE mbo zijn gebaseerd op de Pearson GSE-syllabus. Deze syllabus geeft een overzicht van de verschillende soorten taken die nodig zijn voor de diverse ERK-niveaus. Dit betekent dat kandidaten die op A2 en/of B1 ERK-niveau moeten presteren, alleen taken moeten uitvoeren die passen bij dat niveau. Onderwerpen, woordenschat en grammatica, gebruikt in de Pearson SEE mbo-toets, zijn gebaseerd op de ERK Functions and Notions documents (Van Ek & Trim 1990a, 1990b). Deze bevatten thema's, algemene begrippen en taalfuncties voor verschillende ERK-niveaus.

1.2 Indicator 2.1.3 Cesuur



Ministerie van Onderwijs, Cultuur en
Wetenschap

Portret (Talen)

Er is een verantwoorde procedure gevolgd voor het bepalen van de cesuur, zowel voor de examenonderdelen als voor het examen als geheel. De cesuur in de exameninstrumenten ligt op een niveau, waarop de student gemiddeld voldoet aan het vereiste ERK-niveau dan wel referentieniveau. De cesuur is zodanig opgesteld dat de beoordeling in totaal niet lager uitpakt dan voor de generieke eisen voor taal en rekenen noodzakelijk is. Er is sprake van een evenwichtige, niet oneigenlijke punttoekenning die recht doet aan het belang van de taalvaardigheden.

Er is sprake van een weging van de examenopdrachten/-toetsen onderling, die geen afbreuk doet aan de genoemde cesuur.

Toelichting

- *Een passende cesuur is afgestemd op de eisen waaraan voldaan moet worden en de moeilijkheidsgraad van de toets. Het combineren van verschillende taalvaardigheden in één examen is mogelijk (bijvoorbeeld spreken en gesprekken voeren). Het examen moet de kandidaat wel de gelegenheid bieden om voor elke taalvaardigheid in het examen aan te tonen dat hij deze taalvaardigheid op het vereiste taalniveau beheerst.*
- *ERK-niveau. Voor deze opleidingen geldt dat voor Nederlands en Moderne vreemde talen is bepaald dat alle vijf taalvaardigheden (luisteren, lezen, gesprekken voeren, spreken en schrijven) per taal geëxamineerd dienen te zijn.*
- *Referentieniveau: Daarnaast heeft het toezicht op Nederlands vanaf cohort 2010/2011 betrekking op het referentieniveau: hiervoor geldt dat alle vier domeinen - mondelinge taalvaardigheid (per subdomein), leesvaardigheid, schrijfvaardigheid en taalverzorging - geëxamineerd dienen te zijn. Een passende cesuur is afgestemd op de eisen waaraan voldaan moet worden en de moeilijkheidsgraad van de toets.*
- *Er komt een cijfer voor de decentrale examens en een cijfer voor de centrale examens.*

Aanwijzingen voor de adviseur exameninstrumenten:

Beoordeel of een student die slaagt voor een taalvaardigheid het vereiste ERK- of referentieniveau beheerst. Het ontbreken van een expliciete cesuur betekent niet dat automatisch een onvoldoende beoordeling volgt. Een cesuur kan eventueel herleid worden uit de aangeboden informatie.

Let op de volgende punten:

a. Is duidelijk wanneer een student slaagt voor iedere taalvaardigheid?

- Instellingsexamen op A2:

Een student moet een GSE-score van 30 behalen om te voldoen aan de A2-eisen voor spreken en schrijven. Pearson SEE mbo geeft deze scores duidelijk aan per vaardigheid. Pearson heeft een schaal van 1–10 ontwikkeld, die bij elk GSE-score past. Deze schaal van 1–10 gaat uit van een minimale score van 5,5 om aan de eisen van het A2 ERK-niveau te voldoen. De reikwijdte van 5- 8 op de 1–10 schaal, past bij het A2 ERK-niveau, en de scores van 9–10 op de 1–10 schaal, stijgen boven het vereiste ERK-niveau uit.

- Instellingsexamen op B1:

Een student moet een GSE-score van 43 behalen om te voldoen aan de B1-eisen voor spreken en schrijven. Pearson SEE mbo geeft deze scores duidelijk aan per vaardigheid. Pearson heeft een schaal van 1–10 ontwikkeld, die bij elk GSE-score past. Deze schaal van 1–10 gaat uit van een minimale score van 5,5 om aan de eisen van het B1 ERK-niveau te voldoen. De reikwijdte van 5- 8 op de 1–10 schaal, past bij het B1 ERK-niveau, en de scores van 9–10 op de 1–10 schaal, stijgen boven het vereiste ERK-niveau uit.

Een uitgebreide uitleg vindt u in hoofdstuk 2, Handleiding voor docenten. Hier kunt u ook de bijpassende ERK-niveaus per GSE-score vinden.

b. Lig de cesuur zodanig dat de student aan het niveau voldoet?

Elk ERK-niveau komt overeen met een scala van scores van de GSE. Bijvoorbeeld: het ERK-niveau A2 bevat GSE-scores 30-42. Als een kandidaat op het gebied van spreekvaardigheid A2-niveau moet halen, dan betekent dat dat hij op dat gebied een minimale score van 30 op de GSE-schaal moet halen. Op de ERK-schaal zou de kandidaat op A2-niveau blijven totdat hij een B1 minimale score zou behalen van 43. De brede schaal van de GSE laat dus heel duidelijk zien dat de kandidaat aan het vereiste niveau voldoet.

c. Is van alle toetsen de cesuur gegeven, dus ook van bijvoorbeeld een portfolio?

Pearson SEE mbo is een geïntegreerde toets van alle taalvaardigheden samen. Het heeft per niveau en vaardigheden een eenduidige cesuur.

- d. Is bij het gebruik van meer dan één toets duidelijk aangegeven hoe de toetsen onderling gewogen worden?*

Pearson SEE mbo is een geïntegreerde toets. Deze bevat afzonderlijke onderdelen voor luisteren, spreken, lezen en schrijven en geeft per vaardigheid een aparte score. De toets geeft ook een algemene score voor de competentie in het Engels waarin verschillende taalvaardigheden worden gemeten. Hoe de onderlinge verdeling is opgesteld, kunt u lezen in het Progress Technical Report, sectie 2, Test Description.

- e. Zijn de te behalen punten/scores/beoordelingen in overeenstemming met het te behalen ERK-niveau of referentieniveau?*

Pearson heeft ook een schaal van 1-10 ontwikkeld, zodat deze bij elk ERK-niveau past. Deze schaal van 1-10 gaat uit van een minimale score van 5,5 om aan de eisen van A2 of B1 ERK-niveau te voldoen. De reikwijdte van cijfer 5-8 past bij het A2 of B1 ERK-niveau, en de scores van 9-10 stijgen boven de vereiste ERK-niveaus uit. Een uitgebreide uitleg vindt u in hoofdstuk 2, Handleiding voor docenten.

1.3 Indicator 2.4 Beoordelingswijze



Ministerie van Onderwijs, Cultuur en
Wetenschap

Portret (Talen)

Het exameninstrumentarium is voorzien van een beoordelingsvoorschrift dat een zo objectief mogelijke beoordeling waarborgt. Dit betekent dat:

- het beoordelingsvoorschrift is voorzien van richtlijnen die leiden tot een onafhankelijke beoordeling;*
- de kenmerken van de taaluitvoering (of vergelijkbare beoordelingscriteria) herleidbaar in het beoordelingsvoorschrift zijn opgenomen en een correct antwoordmodel of beoordelingsschema beschikbaar is;*
- de beoordelaar eruit kan opmaken welke (deel)waardering aan een prestatie van een student moet worden toegekend;*
- de beoordelaar eruit kan opmaken welke eindwaardering moet worden gegeven;*
- de beoordelaar de gegeven waarderingen conform het beoordelingsvoorschrift herleidbaar kan onderbouwen. Uit de beoordelingsrichtlijnen blijkt dat de beoordeling plaatsvindt op het niveau van de taalvaardigheid. De beoordeling wordt gebaseerd op beoordelingscriteria/kenmerken van taakuitvoering uit het Raamwerk of referentiekader.*

Toelichting

- Antwoordmodel: Een opsomming van goede, soms van minder goede en foute antwoorden bij open of gesloten vragen.*
- Beoordelingsschema: Een schema dat dient als richtlijn bij de beoordeling van antwoorden op opgaven waarbij geen eenduidig antwoordmodel op te stellen is. In het beoordelingsschema worden criteria vermeld aan de hand waarvan het antwoord beoordeeld moet worden. Deze criteria kunnen zowel op de inhoud als op de structuur van het antwoord betrekking hebben.*
- Scoringsvoorschrift: In het voorschrift wordt vermeld: de maximaal haalbare toetsscore en de scorepunten per vraag/opdracht, en eventueel nog de aftrekpunten van goede of gedeeltelijk goede antwoorden.*

- *Beoordelaarinstructie: De beoordelingsinstructie heeft betrekking op zaken als: de beoordeling, het voortijdig beëindigen van een examen, het komen tot een eenduidige beoordeling in geval van 2 of meer beoordelaars.*

Aanwijzingen voor de adviseur exameninstrumenten:

- a. Algemeen: wordt uit de informatie duidelijk dat er een objectieve betrouwbare beoordeling plaatsvindt?*

Bij Pearson SEE mbo worden zowel de schriftelijke als de mondelinge antwoorden automatisch gescoord. De techniek van automatisch scoren maakt gebruik van kunstmatige intelligentie, waarbij juiste en onjuiste antwoorden herkend worden. Het automatisch scoren van Pearson SEE mbo is zeer uitgebreid getest. Uit de testen kwam naar voren dat de relatie tussen automatisch scoren door de machine en menselijke scores zeer sterk was met een algehele correlatie van .97. Een overzicht van de correlatie per vaardigheid vindt u in het Progress Test Information Booklet.

- b. Is er een antwoordmodel of een beoordelingsschema (met beoordelingscriteria)?*

De antwoordmodellen die gebruikt zijn om scores te koppelen aan de antwoorden in Pearson SEE mbo zijn gebaseerd op duizenden antwoorden, die verzameld zijn tijdens de testontwikkeelfase van Progress. De antwoorden zijn gekoppeld aan ERK-niveaus, vastgesteld door deskundige beoordelaars. De antwoorden zijn in een database opgenomen, die gebruikt wordt als beoordelingscriterium voor automatisch scoren. Voor meer informatie hierover, zie Progress Technical Report, Section 5, Field Testing.

- c. Is er een scoringsvoorschrift dat beschrijft welke antwoorden goed, gedeeltelijk goed en welke als fout moeten worden beschouwd, inclusief de daarbij behorende waardering (punten)?*

Het automatische scoresysteem kan op twee manieren punten toekennen aan de prestaties van studenten, afhankelijk van het type taak dat zij moeten verrichten. Vragen waarbij slechts één antwoord mogelijk is (bijvoorbeeld bij multiple choice of vul het juiste antwoord in) worden door het systeem als juist of onjuist aangemerkt. Wanneer meerdere antwoorden mogelijk zijn, worden deze vergeleken met honderden antwoorden die een score gekregen hebben op een schaal van 1–3. Deze score van 1–3 is gebaseerd op de Progress score-rubrieken die een beschrijving geven van mogelijke antwoorden met verschillende

eigenschappen (bijvoorbeeld inhoud, nauwkeurigheid, spreekvaardigheid, etc). De antwoorden krijgen een score van 1–3 afhankelijk van de mate van overeenkomst met de voorbeeldantwoorden. Dus antwoorden met meerdere mogelijkheden worden niet als juist of onjuist aangemerkt, maar geven een score van 1–3. De Pearson SEE mbo-scorerubrieken kunnen op aanvraag geleverd worden, maar omdat de antwoorden van Pearson SEE mbo automatisch in een score worden omgezet, zijn deze rubrieken en andere beoordelingsmodellen niet nodig voor docenten.

d. Is beschreven hoe eventuele deelwaarderingen leiden tot een eind-waardering?

Pearson SEE mbo geeft zowel een totaalscore aan als een individuele score per vaardigheid. De totaalscore wordt berekend op basis van alle toetsvragen. De scores voor de vaardigheden en de scores voor grammatica en woordenschat worden berekend op basis van relevante vaardigheden of kennis. De scores worden apart berekend; het zijn geen gemiddelde scores. Het toekennen van scores aan toetsonderdelen is gebaseerd op twee criteria: de prestatie van de student bij het uitvoeren van een taak (berekend op een schaal van 1–3) en de moeilijkheid van de taak. De scores (1, 2 en 3) voor de toetsonderdelen worden gewogen, afhankelijk van de moeilijkheidsgraad van de toetsonderdelen. De moeilijkheidsgraad van de onderdelen is gekoppeld aan de ERK Can Do Statements. Toetsonderdelen die overeenkomen met hogere ERK/GSE-niveaus worden zwaarder berekend.

e. Is er een heldere beoordelaarinstructie met algemene richtlijnen over de wijze waarop de prestaties van studenten beoordeeld moeten worden?

Doordat er automatische scoring plaatsvindt, is er geen training of beoordelingsinstrument nodig.

1.4 Indicator 2.1.5 Transparantie



Ministerie van Onderwijs, Cultuur en
Wetenschap

Portret (Talen)

Alle betrokkenen hebben een helder beeld van hoe het examen eruit ziet en hoe het wordt uitgevoerd en beoordeeld. Studenten en beoordelaars zijn geïnformeerd over de beoordelingscriteria en de wijze waarop het examen zal plaatsvinden en over de beoordeling. Het exameninstrumentarium is voorzien van afnamecondities en aanwijzingen voor de beoordelaars en de studenten voor de uitvoering van het examen.

Toelichting

- *De beschrijving van de opdracht en/of gewenst resultaat, de beoordelingscriteria, de wijze van beoordeling en de cesuurbepaling zijn bekend voor de student.*
- *De student moet vooraf over die informatie beschikken die nodig is om het examen af te kunnen leggen; de student mag eigenlijk geen vragen meer hebben over de voorbereiding, uitvoering en beoordeling van het examen.*
- *De omschrijving van de context/setting van de examenopdracht, de te treffen voorbereidingen, de toegestane hulpmiddelen en de inrichting van de examenlocatie zijn bekend voor de student en beoordelaar.*
- *Voor de student en beoordelaar is duidelijk wat zij moeten doen voorafgaand aan het toetsmoment/aanvang van de beoordeling. Daarnaast is het voor de student en beoordelaar duidelijk welke hulpmiddelen een student tijdens de toets mag gebruiken of juist niet mag gebruiken.*

- *Wijze waarop het examen zal plaatsvinden: de omschrijving van het geheel van de omstandigheden waaronder de toets wordt afgenomen is bekend bij de student en beoordelaar. Dit zorgt ervoor dat de student goed aan de opdracht kan voldoen en dat de beoordelaar aan de hand van de gestelde beoordelingscriteria kan beoordelen.*
- *De beoordelaarinstructie heeft betrekking op zaken als: algemene richtlijnen over de wijze waarop de prestaties van studenten beoordeeld moeten worden, het geven van aanwijzingen aan de student tijdens de afname of beoordeling, het voortijdig beëindigen van een examen, het komen tot een eenduidige beoordeling in geval van twee of meer beoordelaars.*

Aanwijzingen voor de adviseur exameninstrumenten:

- a. Zijn de beschrijving van de opdracht en/of het gewenste resultaat, de beoordelingscriteria, de wijze van beoordeling en de cesuurbepaling bekend bij de student?***

Opdrachtbeschrijvingen zijn opgesteld in eenvoudig, toegankelijk Engels (bijvoorbeeld: *Listen to the recording and type what you hear*). De opdrachten en taken van Pearson SEE mbo komen overeen met de opdrachten en taken die studenten in het klaslokaal hebben uitgevoerd, vooral als het curriculum is gebaseerd op functionele en communicatieve begrippen beschreven in het ERK. De wijze van beoordeling (automatische toewijzing van een cijfer) wordt door docenten uitgelegd in de informatiesessie over Pearson SEE mbo. Deze sessie moet plaatsvinden voor de afname van het examen. Zie bijlage 1, Tijdlijn invoering Pearson SEE mbo als instellingsexamen van hoofdstuk 2, Pearson SEE mbo Handleiding voor docenten.

De beoordelingscriteria (inhoud, uitspraak, beklemtoning, spreekvaardigheid, gepast taalgebruik, enz.) kunnen studenten terugvinden in de can-do-statements voor A2 en B1. Deze zijn opgenomen in de Pearson GSE-syllabus. De cesuurbepaling voor A2 en B1 is duidelijk gegeven door de GSE. Studenten moeten een GSE score van 30 of hoger halen, wat overeenkomt met het A2-niveau of een GSE score van 43 of hoger halen, wat overeenkomt met het B1-niveau. Dit is duidelijk te zien in Het Pearson SEE mbo-platform als studenten inloggen op hun account.

- b. Heeft de student vooraf informatie die nodig is om het examen af te kunnen leggen?***

Elke student moet minimaal een oefentoets en proefexamen maken voordat hij het instellingsexamen doet. De oefentoets is een deel van een kopie van de gehele toets. De

bedoeling hiervan is om er zeker van te zijn dat de apparatuur goed werkt en de student door de test heen kan navigeren. Het proefexamen is een exacte kopie van het uiteindelijke instellingsexamen. Studenten krijgen te maken met elk type vraag en beleven het verloop van de toets. Ze krijgen scores en feedback gebaseerd op ERK-richtlijnen, net als bij het instellingsexamen. De feedback op het proefexamen geeft studenten vooral informatie over welke onderdelen ze nog moeten oefenen om uiteindelijk het gewenste niveau van de Pearson SEE mbo-toets te behalen. Na de afname van het proefexamen krijgen studenten de mogelijkheid om vragen te stellen in een vraag-en-antwoordsessie die door hun docent wordt gegeven.

- c. *Hebben de studenten voldoende informatie gekregen over de omschrijving van de context/setting van de examenopdracht, de te treffen voorbereidingen, de toegestane hulpmiddelen en de inrichting van de examenlocatie?*

Studenten moeten een oefentoets en proefexamen doen voordat zij het instellingsexamen doen. Deze toetsen moeten nauwkeurig de context nabootsen van de examenopdracht en de inrichting van de examenlocatie. Deze toetsen moeten ook duidelijk maken welke voorbereiding noodzakelijk is en wat van de studenten verwacht wordt, zowel wat betreft de inhoud als de vorm. Pearson SEE mbo is een test die ontworpen is om de competentie van een taalverwerver te meten om communicatieve functies te kunnen uitvoeren, die zijn beschreven in het ERK. Het is de verantwoordelijkheid van het betreffende instituut om de studenten op dit soort taken voor te bereiden. Er zijn geen hulpmiddelen toegestaan bij het maken van de Pearson SEE mbo-toets.

- d. *Is de wijze waarop het examen zal plaatsvinden bekend bij de student en de beoordelaar? Dit betreft een omschrijving van het geheel van de omstandigheden waaronder de toets wordt afgenomen. Dit zorgt ervoor dat de student aan de opdracht kan voldoen en dat de beoordelaar aan de hand van de gestelde beoordelingscriteria de student kan beoordelen.*

De Pearson SEE mbo-toets moet in een setting plaatsvinden die goedgekeurd is door een IT-medewerker van de school. Zowel de oefentoets, het proefexamen en het instellingsexamen moet in deze omgeving worden afgenomen. Hierdoor raken studenten gewend aan het feit dat ze een koptelefoon op zullen hebben en dat ze in een microfoon moeten praten in plaats van met een medestudent of docent. Dit zorgt ervoor dat studenten hun opdrachten goed kunnen uitvoeren.

- e. *Heeft de beoordelaarinstructie betrekking op zaken als: algemene richtlijnen, de wijze waarop de prestaties van studenten beoordeeld worden, het geven van aanwijzingen aan studenten tijdens de afname, het voortijdig beëindigen van een examen, het komen tot een eenduidige beoordeling bij twee of meer beoordelaars?*

In hoofdstuk 2, Handleiding voor docenten, is alle noodzakelijke informatie voor de beoordelaar opgenomen.

2. Handleiding voor docenten en medewerkers van het Examenbureau

In dit hoofdstuk vindt u informatie over de inhoud van het examen en de berekening van de verschillende examenonderdelen. Daarnaast geven we adviezen hoe u uw studenten het beste kunt voorbereiden op het examen en wat wel en niet toegestaan is tijdens het examen. Voordat u daadwerkelijk met Pearson SEE mbo gaat werken, krijgt u een training, zodat u vertrouwd raakt met het gebruik van Pearson SEE mbo.

Deze handleiding is samengesteld op basis van de ervaringen van mbo-instellingen die Pearson SEE mbo gebruiken. Wij verzoeken u de beschreven procedures voor ieder toetsmoment in uw Pearson SEE mbo-pakket zo nauwkeurig mogelijk te volgen, zowel voor het proefexamen als voor het instellingsexamen.

2.1 Wat toetst u met Pearson SEE mbo?

Pearson See mbo kan als instellingsexamen worden ingezet voor de onderdelen:

- Lezen
- Luisteren
- Spreken/Gesprekken voeren
- Schrijven.

De onderdelen 'Lezen' en 'Luisteren' kunt u dus toetsen met Pearson See mbo, maar het zijn de resultaten van deze onderdelen van het Centraal Examen (CE) die voor de diplomering meetellen. Studenten moeten deze vaardigheden ten minste op ERK-niveau B1 afleggen.

De onderdelen 'Spreken/Gesprekken voeren' en 'Schrijven' toetst u eveneens met Pearson See mbo, en de resultaten hiervan gebruikt u voor het officiële instellingsexamen. Studenten moeten deze vaardigheden ten minste op ERK-niveau A2 afleggen (zowel het centraal examen als het instellingsexamen mag op een hoger niveau worden afgelegd).

Pearson SEE mbo is een geïntegreerde taaltoets, waarbij alle vaardigheden aan bod komen. Vaak worden meerdere vaardigheden tegelijkertijd getoetst. Zo wordt bij het examenonderdeel *Listen and then write* voor het onderdeel Luisteren én voor het onderdeel Schrijven een score gegeven. Deze benadering komt overeen met natuurlijk taalgebruik, waarbij ook meerdere vaardigheden gelijktijdig worden gebruikt.

Het eerste deel van het examen is adaptief. Niet elke student krijgt dan ook hetzelfde aantal vragen. Afhankelijk van de prestatie van de student, krijgt deze na beantwoording van een vraag een volgende vraag van hetzelfde type of een vraag van het volgende type. Elke student krijgt dus een eigen test met vragen, gebaseerd op de volgende factoren:

- Maximale tijd (sectie 1): studenten zullen vaak al klaar zijn voordat de maximale tijd is bereikt; dit is afhankelijk van hun scores op de vragen. Dit is zeer gebruikelijk bij adaptieve testen.
- Naarmate het examen vordert, wordt er steeds meer informatie verzameld van het niveau van de student. Een student die de opdrachten structureel slecht of structureel goed

uitvoert, zal minder opdrachten krijgen en sneller klaar zijn dan een student die de verschillende opdrachten afwisselend goed/slecht maakt.

- Het maximale aantal opdrachten: het aantal opdrachten is gemaximeerd, zodat studenten niet door de test kunnen 'scrollen' en alleen de vragen beantwoorden waarop ze het antwoord weten.

In het eerste deel van het examen is de tijdsaanduiding indicatief. Studenten hoeven zich geen zorgen te maken over het aantal vragen dat ze niet hebben gemaakt. Als sectie 1 wordt afgesloten, omdat de tijd voorbij is, ondervinden studenten hier geen nadeel van. Het is zeer onwaarschijnlijk dat zij betere resultaten behaald zouden hebben als ze meer vragen hadden beantwoord (dit heeft te maken met het eerder omschreven principe dat het algoritme de antwoorden van de studenten interpreteert). Bij het tweede en derde gedeelte van het examen varieert dit per niveau.

2.1.1 Het testen van de spreekvaardigheid

Op veel scholen wordt het onderdeel spreekvaardigheid getest door middel van interviews en gesprekken die studenten voeren met een examiner. Deze manier van toetsen wordt vaak gebruikt vanuit de gedachte dat deze manier lijkt op spreken in het echte leven (authentieke situatie). Onderzoek wijst echter uit dat dit niet altijd een duidelijk of objectief beeld geeft. Studenten gebruiken technieken om hun zwakke punten te verhullen en de voortgang van het gesprek te vertragen. Examinatoren proberen soms ingangen te vinden om studenten te helpen. Zij zijn ook niet altijd objectief wanneer zij in korte tijd grote groepen studenten moeten beoordelen. Bovendien gaan andere factoren, zoals vermoeidheid, een rol spelen.

Met Pearson SEE mbo gebeurt het testen van spreekvaardigheid geautomatiseerd en op verschillende manieren. Studenten moeten bijvoorbeeld gesproken zinnen herhalen, korte antwoorden geven op vragen naar aanleiding van een verhaal, de onderdelen van een zin in de juiste volgorde plaatsen en een kort verhaal navertellen waarbij bepaalde informatie in hun verhaal moet terugkomen (omstandigheden, personen, afloop). Door middel van deze testvragen meet Pearson SEE mbo de psycholinguïstische aspecten van spreekvaardigheid. Concreet betekent dit dat onderstaande criteria gehanteerd worden om de vaardigheid (wijze en inhoud) te meten:

- spreektempo: praten ze in een gebruikelijk tempo, zoals we in het dagelijks leven doen?
- juiste en passende antwoorden: kunnen ze gepaste antwoorden geven op vragen?
- juist taalgebruik: maken ze gebruik van passend en juist taalgebruik (woordenschat en correcte en grammaticale constructies)?
- welbespraaktheid: kunnen ze zich vlot uitdrukken? Gebruiken ze de correcte uitspraak?
- complexiteit: hoe complex is het taalgebruik?
- responstijd: kunnen ze binnen een bepaalde tijd antwoord geven? Hoe snel wordt op een vraag gereageerd?

Op basis van deze competenties kan vastgesteld worden wat het niveau is van de studenten en hoe communicatief vaardig ze zijn. Zowel de validiteit als de betrouwbaarheid van deze manier van scoren met Pearson SEE mbo is zeer hoog (consistent en accuraat testen). Zie voor meer informatie over validiteit, Technical Report.

2.2 Hoe wordt het cijfer voor het instellingsexamen berekend?

Pearson SEE mbo beoordeelt Engelse taalvaardigheid volgens de Global Scale of English (GSE), een schaal die loopt van 10 tot 90. De schaal is een uitbreiding van het Common European Framework of Reference for languages (CEFR) of Europees Referentiekader voor talen (ERK), zodat het niveau van een student nauwkeuriger wordt aangetoond binnen een ERK-bandbreedte. Daarnaast wordt het studentenniveau weergegeven als cijfer op een schaal van 1 tot 10. Digitaal toetsen is een zeer objectieve manier van toetsen. Ook spellingsfouten worden automatisch fout gerekend door het systeem.

Pearson SEE mbo komt overeen met de generieke exameneisen voor Engels MBO 4 (dus niet met de beroepsspecifieke eisen). Pearson SEE A2 meet resultaten op de GSE 21 tot 55, wat overeenkomt met net iets onder en boven het ERK A2-niveau. Deze bandbreedte is gekozen voor een instellingsexamen op A2-niveau. Pearson SEE B1 meet resultaten op de GSE 29 tot 64, wat overeenkomt met net iets onder en boven het ERK B1-niveau. Deze bandbreedte is gekozen voor een instellingsexamen op B1-niveau.

Cijferberekening

In de ingebouwde cijferlijst in Pearson SEE mbo kunt u zowel de GSE-score, het ERK-level als het cijfer 1-10 per vaardigheid oproepen.

In Pearson SEE mbo krijgen de studenten voor het onderdeel 'Schrijven' een GSE-score, die wordt omgerekend naar een cijfer van 1-10. Voor de onderdelen 'Spreken' en 'Gesprekken voeren' krijgen de studenten één GSE-score, die wordt omgerekend naar een cijfer.

Waarom één cijfer voor Spreken en Gesprekken voeren?

Onder 'Speaking' in het Common European Framework (ERK) en de Global Scale of English vindt u Can Do's/Learning Objectives voor zowel Spoken Production als Interaction, net zoals bij Writing and Written Interaction de Can Do's/Learning Objectives onder één noemer Writing gecombineerd worden.

Het uiteindelijke cijfer voor de drie vaardigheden (Schrijven, Spreken en Gesprekken voeren) is dus het gemiddelde van het cijfer voor Schrijven en Spreken en Gesprekken voeren. De school kan zelf de weging bepalen van de verschillende onderdelen. We raden onderstaande cijferberekening aan. Zie ook: document¹

Voorbeeld:

Onderdeel	Cijfer	Uiteindelijke cijfer
Schrijven	5,5	

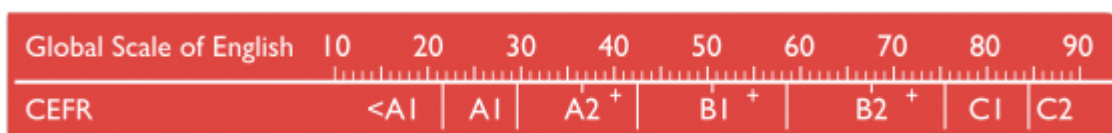
¹ Een mbo-instelling moet weliswaar al deze vaardigheden toetsen, maar hoeft slechts één eindcijfer te leveren voor het instellingsexamen voor de onderdelen Spreken, Gesprekken voeren en Schrijven.

Spreeken	7,3	(5,5 + 7,3): 2 = 6,4
Gesprekken voeren		

Binnen het ERK geldt het uitgangspunt dat een student een niveau behaalt wanneer hij/zij meer dan 50% van alle taken (*Can Do descriptors*) die op dat niveau zijn gedefinieerd, kan uitvoeren. Stijgt dit percentage tot 80%, hoe groter de kans dat studenten 50% van de descriptoren van het volgende niveau beheersen. Op basis van het ERK-uitgangspunt, kunnen we het volgende afleiden:

- Het cijfer 1 op de MBO-schaal betekent dat de student niet het onderliggende ERK behaalt (SEE A2: <A1 / SEE B1: A2-niveau).
- Het cijfer 5.5 betekent dat de student 50% van de taken die op A2 (SEE A2) / B1 (SEE B1) niveau gedefinieerd zijn, beheerst. De student bevindt zich dus aan de ondergrens van het A2 (SEE A2) / B1 (SEE B1) niveau.
- Van studenten die een score van 9 of hoger behalen, wordt verwacht dat zij reeds een hoger CEFR-niveau dan A2 (SEE A2) / B1 (SEE B1) beheersen.

De omreken tabel hieronder toont hoe een GSE-cijfer correspondeert met de 1-10-schaal.



MBO-omreken tabel 1-10-score voor A2-proficiency									
GSE		Cijfer				GSE		Cijfer	
< 21		1.0				35		7.3	
21		1.0				36		7.5	
22		1.6				37		7.8	
23		2.1				38		8.1	
24		2.7				39		8.3	
25		3.2				40		8.5	
26		3.7				41		8.7	
27		4.2				42		8.8	
28		4.6				43		9.0	
29		5.1				44		9.2	
30		5.5				45		9.4	
31		5.9				46		9.6	
32		6.3				47		9.8	

MBO-omrekeningstabel 1-10-score voor B1-proficiency					
GSE	Cijfer		GSE	Cijfer	
< 29	1.0		47	6.6	
29	1.0		48	6.8	
30	1.4		49	7.1	
31	1.7		50	7.3	
32	2.1		51	7.5	
33	2.4		52	7.7	
34	2.8		53	8.0	
35	3.1		54	8.2	
36	3.4		55	8.4	
37	3.8		56	8.6	
38	4.1		57	8.7	
39	4.4		58	8.9	
40	4.7		59	9.1	
41	5.0		60	9.2	
42	5.3		61	9.4	
43	5.5		62	9.6	
44	5.8		63	9.8	
45	6.1		64	10	
46	6.3		> 64	10	
33	6.6		48	10.0	
34	6.9		> 48	10.0	

2.3 Hoe wordt het eindcijfer (centraal examen en instellingsexamen) berekend?

Als een examenonderdeel bestaat uit een centraal examen én een instellingsexamen, dan wordt de waardering van het centraal examen (COE) en het instellingsexamen uitgedrukt in één cijfer (met één decimaal). De eindwaardering voor dat onderdeel wordt bepaald door het gemiddelde van beide examens te nemen.

Bijvoorbeeld:

Instellingsexamen: 6,2

Centraal examen: 5,0

Eindcijfer: 5,6

Zie ook: www.steunpunttaalenrekenenmbo.nl/steunpuntmbo/download/common/201_5-07-09-servicedocument-afronden-cijfers.pdf

2.4 Proefexamen en instellingsexamen

Pearson SEE mbo bevat twee delen: een proefexamen en het instellingsexamen. Het proefexamen is er voornamelijk op gericht de studenten vertrouwd te maken met de toetsindeling. Daarnaast kunt u het

proefexamen afnemen om inzicht te krijgen in het niveau van de student en te zien op welke punten er ruimte is voor verbetering. Hierdoor kunnen studenten nog worden bijgestuurd voordat ze het instellingsexamen maken.

Training

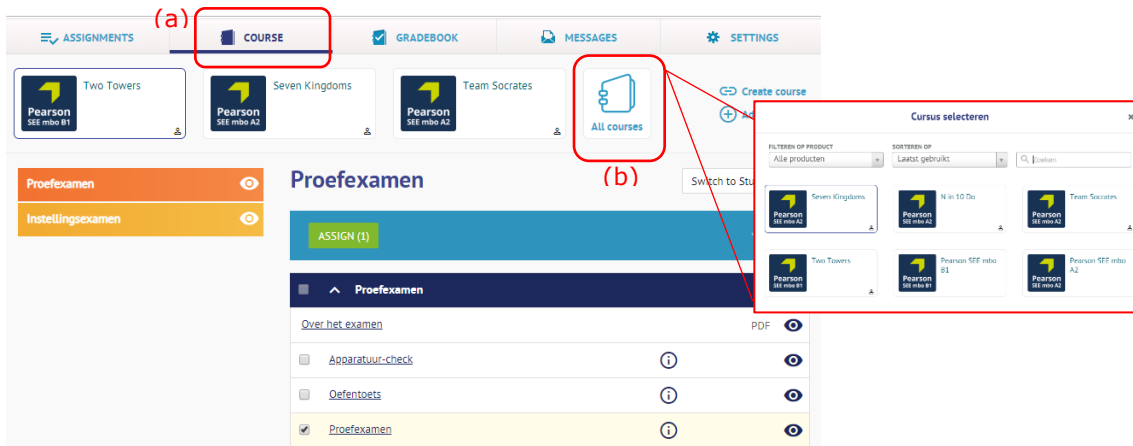
Voordat u uw studenten Pearson SEE mbo laat doen, volgt u een training van ons. Tijdens de training wordt u wegwijs gemaakt binnen het platform, zodat u vertrouwd raakt met het gebruik van Pearson SEE mbo. Bovendien laten we zien hoe u het proefexamen en instellingsexamen kunt klaarzetten voor uw studenten. In een tijdlijn geven we voor de verschillende betrokkenen (docent, medewerker Examenbureau, student, ICT) aan welke activiteiten in welke volgorde zou moeten plaatsvinden. Daarnaast komt de organisatie van Pearson See mbo in het platform ter sprake.

Pearson SEE hulplijn
voor ondersteuning & hulp tijdens Pearson SEE toetsmomenten.

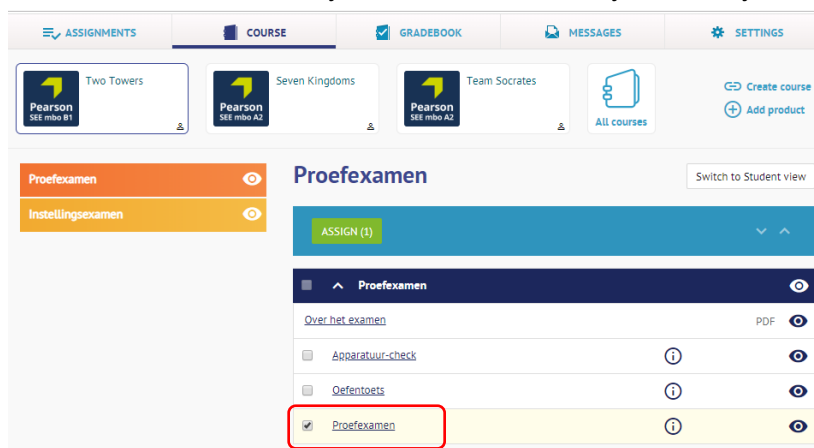
+31208089873

Klaarzetten van het examen

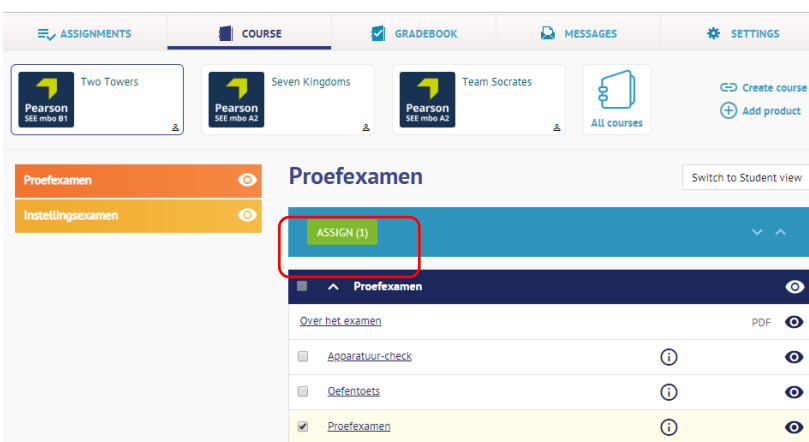
1. Ga naar de **Course/Cursus** pagina en selecteer de examengroep (a). Indien nodig, klik op **All courses/Alle cursussen** om de juiste examengroep te selecteren (b).



2. Plaats een vinkje voor het examen dat je wil toewijzen.



3. Klik vervolgens op **Assign/Toewijzen**.



4. Selecteer de studenten voor wie u het examen wilt klaarzetten.
Vink 'select all students'/'alle studenten selecteren' aan om de hele examengroep in 1 keer te selecteren (4a), of vink individuele studenten aan (4b)

1 Studenten selecteren

☒ Alle studenten selecteren (a)

☐ Stark, Ned	☐ Drogo, Khal	☐ of Myr, Thoros	☐ Higher, Jaqen
☐ Lannister, Jaime	☐ Tyrell, Margaery	☐ Royce, Yohn	☐ Trant, Meryn
☐ Baratheon, Robert	☐ Baratheon, Stannis	☐ Tyrell, Mace	☒ Cassel, Rodrik (b)
☐ Lannister, Cersei	☐ of Tarth, Brienne	☐ Marsh, Bowen	☐ Tynt, Janos
☐ Targaryen, Daenerys	☐ Bolton, Ramsay	☐ Targaryen, Viserys	☒ Lannister, Lancel
☐ Beelish, Petyr	☐ Sand, Ellaria	☐ Mormont, Jorah	☐ Luwin, Maester
☐ Snow, Jon	☐ Naharis, Daario	☐ Stark, Bran	☒ Lannister, Kevan
☐ Stark, Sansa	☐ Bolton, Roose	☐ Clegane, Sandor	☐ Selmy, Barristan
☐ Stark, Arya	☐ Baratheon, Tommen	☐ Tarly, Samwell	☐ Aemon, Maester
☐ Greyjoy, Theon	☐ Stark, Rickon	☐ of Asshai, Melisandre	☐ Tyrell, Loras
☐ Baratheon, Joffrey	☐ Lannister, Lancel	☐ Mormont, Jeor	☐ of Myr, Thoros
☐ Lannister, Tyrion	☐ Worm, Grey	☐ Maegyr, Talisa	
☐ Lannister, Tywin	☐ Tyrell, Olenna	☐ Giantsbane, Tormund	

5. Stel de inleverdatum in.
Klik op 'Custom Due date'/'Aangepaste inleverdatum' (5a) en vul bij 'date/hour'/'datum/uur' (5b) de eindtijd in. We raden aan de eindtijd zo'n 2 tot 3 later dan de starttijd te zetten.

2 Inleverdatum instellen

☐ Een week vanaf nu (10/1/2020, 10:45)
☐ Twee weken vanaf nu (17/1/2020, 10:45)
☒ Aangepaste inleverdatum (a)

Datum: 2020-01-03 (b)

Uur: 11 : 00

JANUARI 2020

ma	di	wo	do	vr	za	zo
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

11 : 00

instellen To

10 Pearson Educat

6. Activeer de tijd klok bij 'Enable timer'/'Timer inschakelen': wanneer deze is aangevinkt, moeten alle studenten binnen een vaste tijd het examen afronden. Dit resulteert in beter vergelijkbare resultaten en het voorkomt bedrog.

☒ **Timer inschakelen**

☐ **Onderbreken en hervatten toestaan**

☐ **Begintijd instellen**

7. Activeer 'Allow Pause and Resume'/'Onderbreken en hervatten toestaan'. Als de kans bestaat dat de internetverbinding verbroken wordt, zorgt deze optie ervoor dat de tijd klok ook pauzeert.

☒ Timer inschakelen
☒ **Onderbreken en hervatten toestaan**
☐ Begintijd instellen

8. Stel een starttijd in bij **Set start time/Begintijd instellen**, zodat alle studenten tegelijkertijd aan het examen beginnen.

Opmerking: Zet je het examen net voor afname klaar, dan hoef je geen starttijd in te vullen

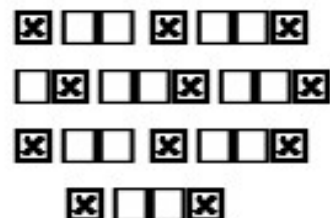
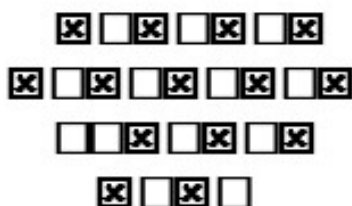
☒ Timer inschakelen
☒ Onderbreken en hervatten toestaan
☒ **Begintijd instellen**

Datum: 

Uur: :

Belangrijke informatie bij het afnemen van het examen:

- Zorg ervoor dat de studenten op de hoogte zijn van wat wel en niet is toegestaan bij het maken van het examen. We adviseren u de studenten een geprint exemplaar van de Handleiding voor studenten te geven. Dit document kunnen studenten ook downloaden vanuit hun eigen Pearson SEE mbo-account.
- Houd voldoende ruimte tussen studenten om de effecten van achtergrondgeluiden te reduceren bij de toetsscores. De volgende indeling van zitplaatsen biedt de beste resultaten:
 - één open zitplaats tussen elke twee studenten, zowel verticaal als horizontaal
 - twee open zitplaatsen tussen elke twee studenten in een horizontale rij, waarbij studenten niet in dezelfde verticale rij worden geplaatst. We raden aan om studenten niet direct naast elkaar te plaatsen.



- Geluidsonderdrukkende headsets en eenrichtingsmicrofoons zullen de interferentie van achtergrondgeluiden en afleidende ruis reduceren. Indien dit niet haalbaar is, dan is het aanbevolen om gebruik te maken van privacyschermen tussen de computerstations van 1,5m hoogte vanaf de grond gemeten, waarbij de toetsomgeving vrij moet zijn van extern geluid.
- De studenten krijgen de audio slechts één keer te horen. In Pearson SEE mbo zijn de opgaves waarbij de student zinnen moet herhalen of vragen moet beantwoorden vrij kort. Herhalen zou het examen eenvoudiger maken en deze zou dan aan validiteit verliezen. Bij de scores wordt ervan uitgegaan dat studenten slechts één keer luisteren, wat een beter beeld van het luistervermogen van de student geeft.
- Pearson SEE mbo is een adaptief examen, bedoeld om studenten te prikkelen een zo accuraat mogelijke score te behalen. Ze zullen sommige vragen moeilijk vinden. Details over het soort vragen dat op het examen gesteld wordt, zijn beschikbaar in het Test Information Booklet.

Browser lockdown

In Pearson SEE mbo is geen browser lockdown ingebouwd. De school is hier zelf verantwoordelijk voor. We raden scholen aan om browser lockdown-software te installeren bv. Safe Exam Browser.

Spellingcontrole

In Chrome staat vaak de spellingscontrole automatisch aangevinkt. Wanneer u deze wilt uitschakelen, moet dit manueel gebeuren. Dit is de verantwoordelijkheid van de school.

Advies Pearson SEE & dyslexie/leerbependingen

Er is niet één variant van dyslexie die voor iedereen met de diagnose geldt. Sommige dyslectici hebben vooral moeite met spellen terwijl anderen juist problemen ervaren bij het (hardop) lezen. Afhankelijk van de dyslexie variant, zal de ene student meer baat hebben bij de voorleesfunctie en anderen bij extra tijd of de automatische spell check.

1. Extra tijd

Aan studenten met dyslexie of andere leerbependingen kunt u een apart examen toewijzen. U kunt hun extra tijd geven door de timer van het examen uit te schakelen (disable timer aanvinken). Hierdoor kunnen zij langer over het examen doen. U kunt aangeven hoe lang zij over het examen mogen doen door een deadline in te stellen.

Voorbeeld

Stel u wilt een student bijvoorbeeld 50% extra tijd geven en het examen begint om 9.00 uur.

1. Selecteer de student voor wie u het examen wilt klaarzetten.

1 Studenten selecteren

☐ Alle studenten selecteren

☐ Stark, Ned	☐ Drogo, Khal	☐ of Myr, Thoros	☐ H'ghar, Jaqen
☐ Lannister, Jaime	☐ Tyrell, Margaery	☐ Royce, Yohn	☐ Trant, Mervyn
☐ Baratheon, Robert	☐ Baratheon, Stannis	☐ Tyrell, Mace	☒ Cassel, Rodrik
☐ Lannister, Cersei	☐ of Tarth, Brienne	☐ Marsh, Bowen	☐ Slynit, Janos
☐ Targaryen, Daenerys	☐ Bolton, Ramsay	☐ Targaryen, Viserys	☒ Lannister, Lancel
☐ Baelish, Petyr	☐ Sand, Ellaria	☐ Mormont, Jorah	☐ Luwin, Maester
☐ Snow, Jon	☐ Naharis, Dario	☐ Stark, Bran	☒ Lannister, Kevan
☐ Stark, Sansa	☐ Bolton, Roose	☐ Clegane, Sandor	☐ Selmy, Barristan
☐ Stark, Arya	☐ Baratheon, Tommen	☐ Tully, Samwell	☐ Aemon, Maester
☐ Greyjoy, Theon	☐ Stark, Rickon	☐ of Asshai, Melisandre	☐ Tyrell, Loras
☐ Baratheon, Joffrey	☐ Lannister, Lancel	☐ Mormont, Jeor	☐ of Myr, Thoros
☐ Lannister, Tyrion	☐ Worm, Grey	☐ Maegyr, Talisa	
☐ Lannister, Tywin	☐ Tyrell, Olenna	☐ Giantsbane, Tormund	

2. Stel de inleverdatum in.

Klik op 'Custom Due date/Aangepaste inleverdatum' (5a) en vul bij date/hour/ datum/uur (5b) de eindtijd in. In het geval van dit voorbeeld wordt de eindtijd 10u30 (anderhalf uur na de starttijd van 9u = 50% extra tijd).

2 Inleverdatum instellen

☐ Een week vanaf nu (10/1/2020, 10:45) ☐ Aanstaaende maandag (6/1/2020, 10:45) ☒ Aangepaste inleverdatum (a)

Datum: 2020-01-03 Uur: 11 : 00 (b)

Calendar view: JANUARI 2020

ma	di	wo	do	vr	za	zo
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

Time picker: 11 : 00

3. **Enable timer**/**Timer inschakelen** vinkt u uit.

☐ **Timer inschakelen**

☒ **Onderbreken en hervatten toestaan**

☒ **Begintijd instellen**

Datum: 2020-01-03

Uur: 11 : 00

-
4. Activeer **Allow Pause and Resume**/**Onderbreken en hervatten toestaan**. Als de kans bestaat dat de internetverbinding verbroken wordt, zorgt deze optie ervoor dat de tijd klok ook pauzeert.

☐ **Timer inschakelen**

☒ **Onderbreken en hervatten toestaan**

☒ **Begintijd instellen**

Datum: 2020-01-03

Uur: 11 : 00

-
5. Stel een starttijd in bij **Set start time**/**Begintijd instellen**. In het geval van dit voorbeeld is de begintijd 09:00 uur.

☐ **Timer inschakelen**

☒ **Onderbreken en hervatten toestaan**

☒ **Begintijd instellen**

Datum: 2020-01-03

Uur: 11 : 00

Bij het uitschakelen van de timer, is het belangrijk dat studenten zelf goed de tijd in het oog houden per toetsonderdeel.

Hieronder een inschatting van tijdsduur per onderdeel:

Onderdeel 1: 23 minuten (max. 30 minuten)

Onderdeel 2: 22 minuten

Onderdeel 3: 18 minuten

Bij het toewijzen van het examen kan je dus de timer uitschakelen en zelf een begin- en einduur kiezen (afh van de extra tijd die studenten krijgen).

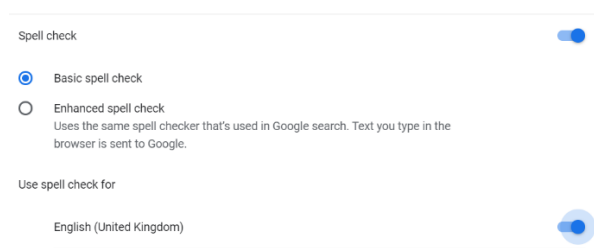
Echter, bepaalde opdrachten blijven getimed. Bv, bij een opdracht 'you have 25 seconds to prepare' blijft de 25 seconden gelden.

2. Voorleessoftware

Het is mogelijk om (externe) voorleessoftware (TTS – Text To Speech) met het examen te gebruiken, zoals bv. Claroread (www.clarosoftware.com) of Read Aloud (via de Chrome webstore). Dit valt onder de verantwoordelijkheid van de school.

3. Automatische Spell Check

Digitale toetsen zijn anders dan toetsen die met pen en papier worden afgenomen. Spellingsfouten die door AI worden beoordeeld, worden fout gerekend, terwijl deze door menselijke beoordelaars niet of in mindere mate aangerekend zouden kunnen worden. U kan ervoor kiezen om de automatische spell check die ingebouwd is in webbrowsers, in te schakelen als hulpmiddel voor spelling. Let er op dat deze voor de Engelse taal geactiveerd is in de settings.



2.5 Examenresultaten

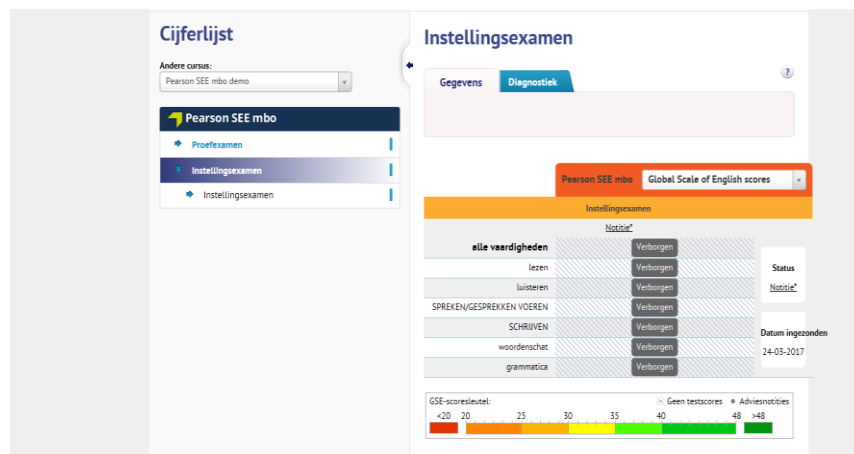
Gewoonlijk ontvangt u de examenresultaten in uw cijfermap, binnen een uur nadat het examen is ingeleverd en in vele gevallen zelfs veel sneller. Echter, lokale bandbreedtes, de bandbreedtes en internetverbinding van uw instelling, of periodes waarin uitzonderlijk grote hoeveelheden testen worden ingeleverd, kunnen er de oorzaak van zijn dat de wachttijd iets langer is.

Scores verbergen voor studenten

Het is mogelijk om de scores voor de studenten pas in een later stadium zichtbaar te maken. U kunt dan in het platform aangeven dat de scores onzichtbaar moeten zijn.



Studenten krijgen dan de melding dat de scores verborgen zijn.



Adviserende aantekeningen op scores

Adviserende aantekeningen worden in de cijfermap aangegeven met een asterisk naast de behaalde score, zoals in de afbeelding hieronder:

Pearson SEE mbo Global Scale of English scores				
Student	Proefexamens	Status	Instellingsexamen	Status
Rousseau, Jean-Ja...	32*	Submitted	22*	Submitted
Nietzsche, Friedr...	-	Re-test	-	Re-test
Marx, Karl	20	Submitted	-	Re-test
Voltaire, Francis...	140*	Submitted	22*	Submitted
Kant, Immanuel	32*	Submitted	140*	Submitted
Derrida, Jacques	140*	Submitted	22*	Submitted
Foucault, Michel	140*	Submitted	-	Re-test
Machiavelli, Nicc...	140*	Submitted	-	Re-test
Descartes, René	-	Re-test	-	Re-test
Freud, Sigmund	140*	Submitted	-	Re-test
de Montaigne, Mi...	32*	Submitted	-	Re-test
van Asselt, Tho...	-	Re-test	-	Re-test
Locke, John	140*	Submitted	-	Re-test
Sartre, Jean-Paul	140*	Submitted	-	Re-test

Een asterisk kan vanwege de volgende redenen worden geplaatst:



Advisory note

There are many factors that can influence a test score. For this student's test we have identified a problem with the score which may be lower or higher than expected due to one of the following possible reasons:

- The student's language level is too far below the level of this test or the course they are studying.
- The student's language level is too far above the level of this test or the course they are studying
- The student may have made little or no progress in their language level
- The student may have made progress but this cannot be detected by the test
- (One or both of) the tests the student has taken were taken under very different conditions
- The conditions the student took (one or both of) the tests under were unsuitable; for example there was a lot of noise affecting our ability to score the speaking test or there were problems with the computer
- The student had help from a friend on (one or both of) the tests, which has affected their score on the tests
- The student has not tried as hard on (one or both of) the tests, which has affected their score on the tests

We adviseren om voorzichtig om te gaan met het gebruik van scores met adviserende aantekeningen.

Status in het gradebook

Marsh, Bowen	-	Assigned
Maegyr, Talisa	28*	Submitted
Luwin, Maester	-	Assigned
Lannister, Tywin	-	1 Assigned
Lannister, Tyrion	-	Assigned
Lannister, Lancel	31*	2 Submitted
Lannister, Lancel	26	Submitted
Lannister, Kevan	-	3 Not Assigned
Lannister, Jaime	20*	Submitted
Lannister, Cersei	>48*	Submitted
H'ghar, Jagen	-	Not Assigned
Greyjoy, Theon	-	Assigned
Giantsbane, Torm...	-	Not Assigned
Drogo, Khal	-	Assigned
Clegane, Sandor	-	Not Assigned
Cassel, Rodrik	-	Assigned
Bolton, Roose	-	Not Assigned
Bolton, Ramsay	-	4 Re-test

1. Assigned
Het examen is toegewezen aan de student.

2. Submitted
De student heeft het examen gemaakt en ingediend.

3. Not Assigned
Het examen werd nog niet toegewezen aan de student.

4. Re-test
Het examen werd eerder aan de student toegewezen maar de student heeft het examen niet gemaakt (bijvoorbeeld, de student was door omstandigheden niet aanwezig op de testdag). De licentie is in dit geval niet verlopen. Het examen kan opnieuw toegewezen worden.

Pearson SEE mbo Score Challenge

Wanneer u het als docent niet eens bent met een score op de Pearson SEE-test van een of meer studenten, kunt u hier bezwaar tegen maken.

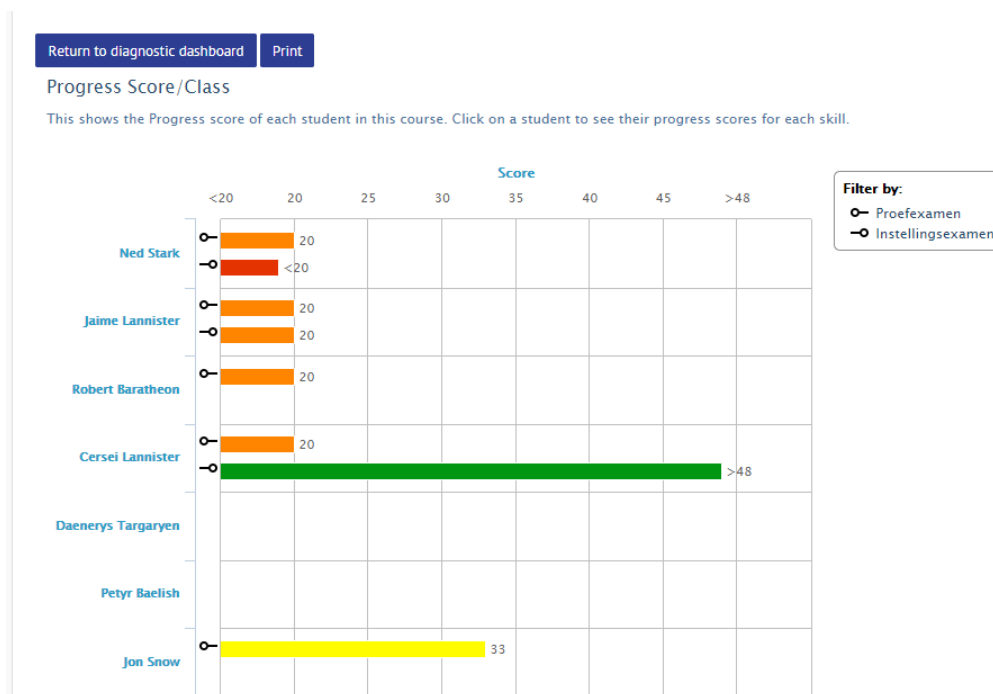
In bijlage 3 vindt u de Score Challenge voorwaarden.

Uw Pearson-consultant kan uw vragen over begeleidende notities en over de procedure verder beantwoorden.

Diagnostische tools

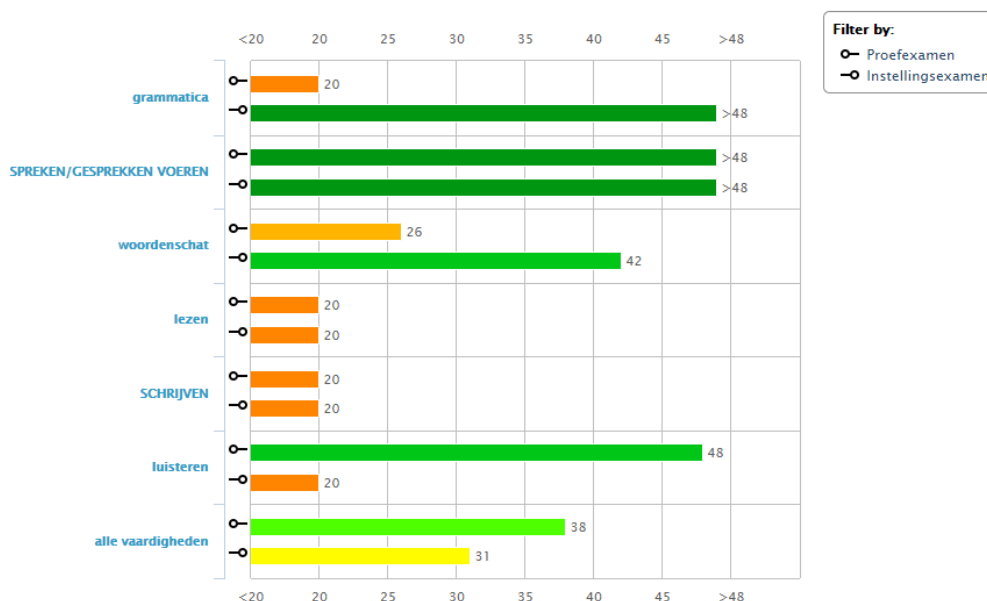
De diagnostische tools van het Pearson SEE mbo-platform kunt u gebruiken ter informatie bij het lesgeven en het leren en om studenten te helpen zich verder te verbeteren.

Met scorekaarten voor klassen en individuele studenten per examen en per vaardigheid kunt u met één oogopslag zien waar specifieke studenten meer hulp nodig hebben of specifieke vaardigheden meer aandacht nodig hebben.



Progress score/Skill/Student – Lannister, Lancel

This shows progress scores for each skill per student. Click on skill to see Learning Objectives this student needs to work on.



Per student zijn er zogenaamde can-do-indicatoren ('dit kan je'). Deze geven per vaardigheid specifieke gebieden aan waar de student aan moet werken en bieden praktische inzichten om de leerresultaten te verbeteren.

[Return to diagnostic dashboard](#)

[Return](#)

[Print](#)

Progress score/Skill/luisteren Can Do Statements

This shows a list of Learning Objectives this student needs to work on.

- Can distinguish between main ideas and supporting details in familiar, standard texts.
- Can follow most of a clearly structured presentation within their own field.
- Can follow most of an everyday conversation if speakers avoid very idiomatic usage.
- Can give clear, detailed descriptions on a wide range of familiar subjects.
- Can identify the main reasons for and against an argument or idea in a discussion delivered in clear standard speech.
- Can paraphrase in simpler terms what someone else has said.
- Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts.
- Can understand cause and effect relationships in informal conversation at natural speed.
- Can understand most TV news and current affairs programmes.
- Can understand scripted speech delivered quickly, if the accent is familiar.
- Can understand the details of extended and linguistically complex talks on a range of political, environmental, and social issues.
- Can understand the key points about a radio programme on a familiar topic.

Overige informatie

Voor overige informatie betreffende de validiteit van het examen, geautomatiseerde scores, onderzoek naar de toetsing van geïntegreerde vaardigheden of de conversie van de GSE-score naar andere schalen: <http://product.pearsonelt.com/progress/>

Technische ondersteuning: <https://support.pearsonelt.com/>

Details over het soort vragen in Pearson SEE mbo vindt u in de Test Information Booklet:

http://product.pearsonelt.com/progress/downloads/test_info_booklet.pdf

Zo geeft u toestemming aan de webbrowser voor toegang tot de microfoon:

<https://www.youtube.com/watch?v=s1PQR422sUc>

Wijzigen van de volumes van de microfoon en headset:

<https://www.youtube.com/watch?v=GZ8hx53JlyQ>

Pearson SEE hulplijn
voor ondersteuning & hulp tijdens Pearson SEE toetsmomenten.

+31208089873

2.6 Concrete invulling van het platform Person SEE mbo

Plan een afspraak in met uw consultant om de organisatie van Pearson SEE mbo te bespreken. Tijdens deze afspraak bespreken we in ieder geval de volgende zaken:

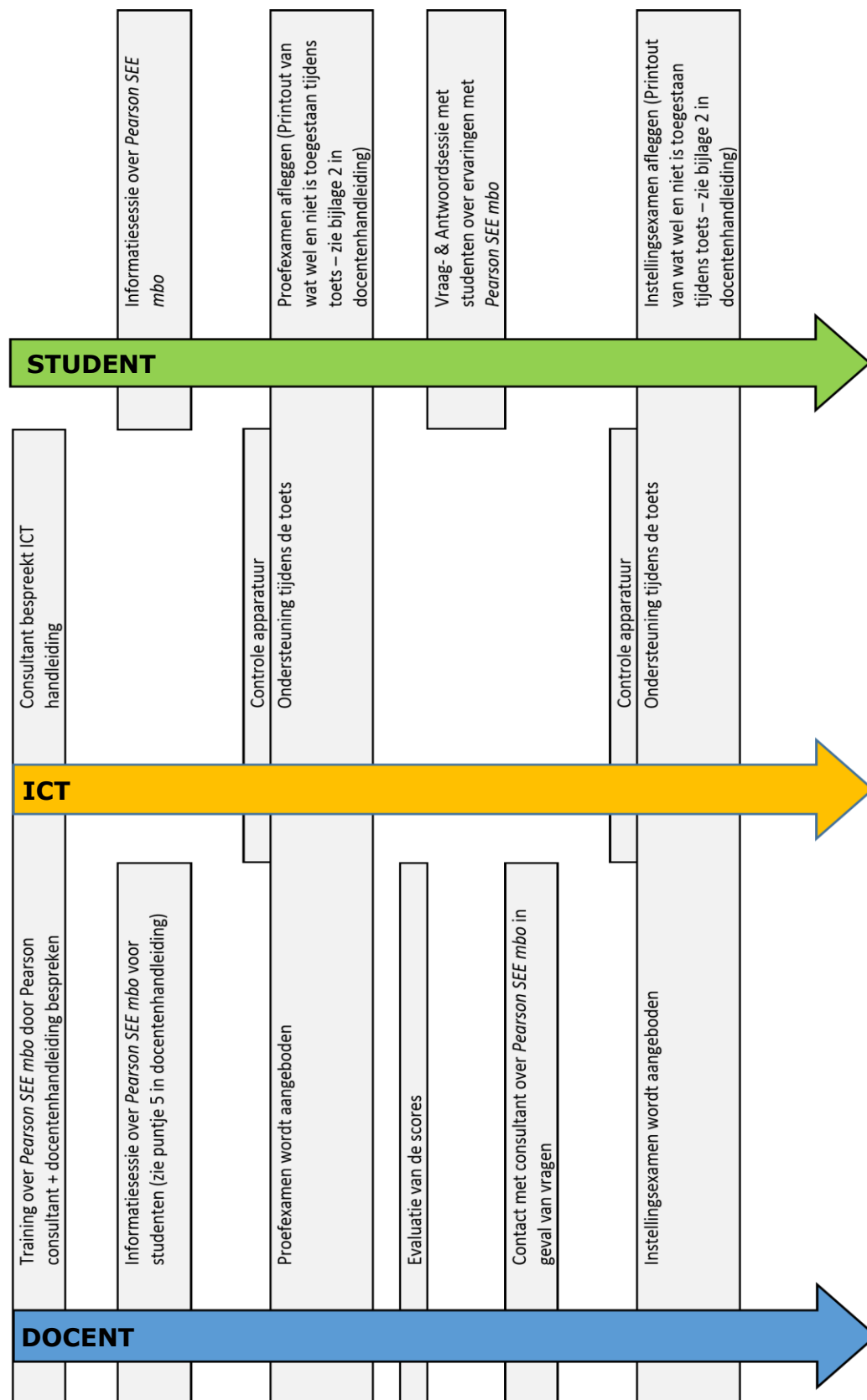
- de inrichting van het platform
- de toegang tot het platform
 - de eisen aan toegangsgegevens voor de studenten
 - de planning van het instellingsexamen

Contacteer je consultant, of mail elt.service@pearson.com

Pearson SEE hulplijn
voor ondersteuning & hulp tijdens Pearson SEE toetsmomenten.

+31208089873

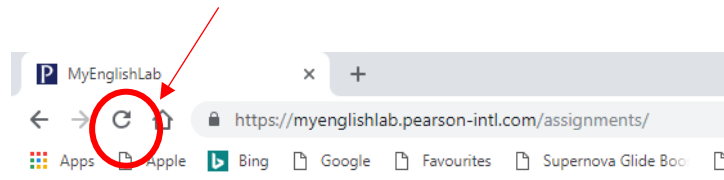
Bijlage 1: Tijdlijn



Bijlage 2: Over het examen

Wat mag je wel en niet doen tijdens het maken van het examen?

1. Sla geen vragen over.
2. Druk niet op de pijltjestoetsen om verder of terug te gaan wanneer je het examen maakt. Je kunt niet teruggaan naar de vorige vraag om je antwoord aan te passen. Doe je dit per ongeluk toch, dan krijg je een *Conflict ID error* te zien. Klik dan op *Pagina Vernieuwen (Reload this page)* of klik op de toets F5.



Pagina vernieuwen: F5 of 

3. Klik niet steeds op dezelfde knop, bijvoorbeeld op *Next (Volgende)*. Als er niets gebeurt nadat je op een knop hebt gedrukt, vernieuw dan de pagina.
4. Als een pagina van het examen blijft hangen, vernieuw dan de pagina.
5. Als de internetverbinding verbroken wordt, ga dan terug naar de inlogpagina en log weer in om verder te gaan met het examen. De antwoorden die je al had gegeven, zijn gewoon bewaard.
6. Het is niet toegestaan om hulpmiddelen te gebruiken tijdens het examen. Je mag dus tijdens het examen geen enkele andere software op de pc/laptop/tablet gebruiken.
7. Als je een schrijf- of spreekopdracht niet binnen de tijd af hebt, dan gaat het examen automatisch door naar de volgende opdracht. Je gegeven antwoord wordt dan wel gewoon bewaard en dus ook beoordeeld.
8. Het examen is getimed. Rechtsbovenaan in het scherm zie je een voortgangsbalkje die aangeeft hoever je ongeveer bent in het examen. Bepaalde opdrachten zijn individueel getimed. De tijd staat per opdracht in de instructies vermeld.
9. Wanneer je het examen hebt ingeleverd, is het resultaat definitief. Je kunt dus niet teruggaan en alsnog iets veranderen.

Bijlage 3:

Pearson SEE mbo score challenge terms en voorwaarden

Auteur: Marian Martin

Datum: 30 november 2018

Gebaseerd op PTE Academic terms

Als u het niet eens bent met de scores van een SEE mbo-test, dan kunt u ons verzoeken om de antwoorden te onderzoeken. U dient hierbij het volgende in ogenschouw te nemen:

- voorwaarden:
 - Student heeft het proefexamen afgelegd.
 - Onze aanbevelingen over optimale toetsomstandigheden werden nageleefd.Bij het niet naleven van de voorwaarden is score onderzoek altijd betalend.
- SEE mbo is een test die door de computer wordt nagekeken, waardoor de testresultaten niet zullen veranderen.
- We zullen het volgende onderzoeken:
 - of technische problemen de resultaten beïnvloed hebben.
 - of de opgenomen antwoorden op vragen duidelijk te horen zijn. Was het bijvoorbeeld mogelijk om de ingesproken antwoorden goed te verstaan terwijl er tevens achtergrondgeluiden te horen waren? Speelden technische vervormingen een rol? Bovendien wordt onderzocht of er problemen waren met de microfoon, of er te luid is gesproken en of de antwoorden in het Engels zijn gegeven.
 - of schrijfp opdrachten overgeslagen zijn of fouten bevatten.
- We zullen geen:
 - lees- en luisteroefeningen onderzoeken.
 - informatie over specifieke vragen geven.
- Alleen klanten van Pearson komen in aanmerking voor een onderzoek. De onderzoeksresultaten worden alleen naar de betreffende instelling gestuurd.
- De kosten voor het onderzoeken van een examen bedragen €150.00. Wanneer een technisch probleem de testresultaten beïnvloed heeft, worden er geen kosten in rekening gebracht. Pearson zal dan een (gratis) nieuwe test beschikbaar stellen.
- Aanvragen voor nader onderzoek moeten binnen dertig kalenderdagen na het afnemen van de test gedaan worden.

- De onderzoeksresultaten worden binnen vijftien werkdagen na het verzoek om onderzoek opgestuurd.

Uw Pearson-consultant kan uw vragen over deze begeleidende notities en over de procedure verder beantwoorden.

3. Handleiding voor ICT

Op basis van de ervaringen van gebruikers met Pearson SEE mbo hebben we een aantal (technische) zaken op een rijtje gezet die belangrijk zijn bij het afnemen van het examen. We adviseren u de procedures, vereisten en adviezen zo nauwkeurig mogelijk op te volgen voor een optimaal resultaat.

3.1 Gereedmaken voor Pearson SEE mbo

3.1.1 Bandbreedte

Een aantal factoren is van invloed op de prestaties van de Pearson SEE mbo-toets:

1. de beschikbare bandbreedte op de school
2. het aantal studenten dat de toets tegelijkertijd afneemt
3. datgene waarvoor de bandbreedte op datzelfde moment ook voor gebruikt wordt.

Bepaal met de onderstaande tabel het aanbevolen aantal gebruikers dat tegelijkertijd aan de Pearson SEE mbo-toets kan deelnemen, gebruikmakend van een gedeelde, dedicated internetverbinding met een gespecificeerde upload- en downloadsnelheid.

Concurrent Users	Upload speed	Download Speed
1	128 kbps (0.125 Mbps)	128 kbps (0.125 Mbps)
2	256 kbps (0.25 Mbps)	512 kbps (0.5 Mbps)
4	512 kbps (0.5 Mbps)	1024 kbps (1Mbps)
10	1024 kbps (1Mbps)	2048 kbps (2 Mbps)
20	2048 kbps (2 Mbps)	5130 kbps (5 Mbps)
50	5130 kbps (5 Mbps)	10 Mbps (10 Mbps)
100	10260 kbps (10 Mbps)	20520 kbps (20 Mbps)
200	20520 kbps (20 Mbps)	41040 kbps (40 Mbps)

With these recommendations all users should expect to be able to upload audio responses within 3-5 seconds and listen to audio without interruptions

Higher numbers are supported for a given speed but users could potentially have to longer wait times when submitting audio record activities.

3.1.2 Systeemvereisten

Alle Computers		
Desktop of laptop (tablets worden momenteel niet ondersteund) Moderne processor 1.33 GHz of sneller 32-bit (x86) Geheugen: 1024 MB RAM of meer Schermbresolutie: 1024 x 768 of hoger Plug-ins: Adobe® Reader 2015, Adobe® Flash Player 17 en/of hoger Internetverbinding: DSL, Kabel/Breedband, T1, of andere hogesnelheidsverbinding Headsets & microfoon (externe aanbevolen). Bij voorkeur geluidsonderdrukkende headsets Opmerking: Bij voorkeur de internetbrowser Google® Chrome		
PC	Macintosh	Linux
Besturingssysteem	Besturingssysteem	Besturingssysteem
Windows® XP, Windows Vista, Windows 7, 8, 10	OS X 10.10	Debian GNU of Ubuntu-basis
Webbrowser	Webbrowser	Webbrowser
Firefox® 35 (tot 37), Google® Chrome 40.x (tot 42), Microsoft Edge	Safari® 8.x, Firefox® 35 (tot 37), Google® Chrome 40.x (tot 42)	Firefox® 23-27.x, Google® Chrome 30-33.x

3.1.3 Whitelist URL's

Gebruikt uw instelling een firewall, dan moet u controleren of deze firewall de toegang tot de toets niet blokkeert. De open poorten zouden 80 en 443 moeten zijn.

Wij verzoeken u de volgende URL's en IP-adressen toe te voegen aan uw firewall 'whitelist':

<http://myenglishlab.pearson-intl.com/>
media.pearsoncmg.com
api.learnosity.com
learnositymediaprocessed.s3.amazonaws.com
apicdn.learnosity.com
ssl.google-analytics.com
ajax.googleapis.com
google-analytics.com
vjs.zencdn.net
54.75.238.173
54.228.253.251

3.1.4 Toetsomgeving en apparatuur

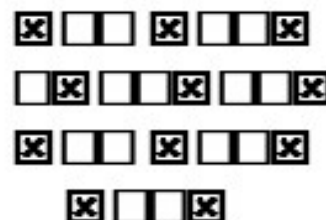
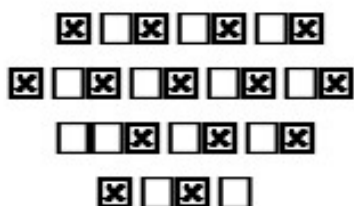
Handleiding voor de studenten

Zorg ervoor dat de studenten op de hoogte zijn van wat wel en niet is toegestaan bij het maken van de toets. We adviseren u de studenten een geprint exemplaar van de handleiding voor studenten te geven.

Opstelling tijdens het examen

Houd voldoende ruimte tussen studenten om de effecten van achtergrondgeluiden te reduceren bij de toetsscores. De volgende indeling van zitplaatsen biedt de beste resultaten:

- één open zitplaats tussen elke twee studenten, zowel verticaal als horizontaal.
- twee open zitplaatsen tussen elke twee studenten in een horizontale rij, waarbij studenten niet in dezelfde verticale rij worden geplaatst. We raden aan om studenten niet direct naast elkaar te plaatsen.



Achtergrondgeluiden

Geluidsonderdrukkende headsets en eenrichtingsmicrofoons zullen de interferentie van achtergrondgeluiden en afleidende ruis reduceren. Indien dit niet haalbaar is, dan is het aanbevolen om gebruik te maken van privacyschermen tussen de computerstations van 1,5m hoogte vanaf de grond gemeten, waarbij de toetsomgeving vrij moet zijn van extern geluid.

Browser lockdown

In Pearson SEE mbo is geen browser lockdown ingebouwd. De school is hier zelf verantwoordelijk voor. Wanneer u dit niet regelt, zouden studenten tijdens de toets bijvoorbeeld Google Translate kunnen raadplegen.

Dyslexie

Aan studenten met dyslexie kunt u een aparte toets toewijzen. U kunt hun extra tijd geven door de timer van de toets uit te schakelen (*disable timer* aanvinken). Hierdoor kunnen zij langer over de toets doen. U kunt aangeven hoe lang zij over de toets mogen doen door een deadline in te stellen. Dit is de enige tegemoetkoming voor dyslexie-studenten binnen Pearson SEE mbo.

Het is wél mogelijk om (externe) voorleessoftware aan de toets toe te voegen, zoals Claroread. Dit valt onder de verantwoordelijkheid van de school.

Digitale toetsen zijn anders dan toetsen die met pen en papier worden afgenomen. Spellingsfouten die op de computer worden gemaakt, worden fout gerekend, terwijl deze op papier niet of in mindere mate aangerekend zouden kunnen worden.

Spellingcontrole

In Chrome staat vaak de spellingscontrole automatisch aangevinkt. Wanneer u deze wilt uitschakelen, moet dit manueel gebeuren. Dit is de verantwoordelijkheid van de school.

Controle van de apparatuur

U dient de apparatuur te controleren voordat de toets wordt afgenomen om er zeker van te zijn dat alle apparatuur, bijvoorbeeld hoofdtelefoons, plug-ins en geluidskaarten, naar behoren functioneert. Na de check dient ervoor gezorgd te worden dat de hoofdtelefoons niet van de computers worden verwijderd. Mogelijk dient u toestemming te geven aan de webbrowser voor toegang tot de microfoon. Het is belangrijk dat de controle van de apparatuur plaatsvindt vóórdat de leerlingen aan de toets beginnen. Wij adviseren dat u de controle van de apparatuur uitvoert minstens een uur voordat de studenten arriveren en plaatsnemen.

U kan gebruik maken van onze Pearson SEE ICT test account voor de apparatuur-check. Inloggen via www.english.com/login met onderstaande inloggegevens.

Gebruikersnaam: ict@see

Password: pearson1

Klik op het Pearson SEE A2 icoon.



U vindt de Apparatuur-Check onder *Cursus* > *Proefexamen* of *Instellingsexamen*:

The screenshot shows the MyEnglishLab user interface. At the top, there's a navigation bar with 'Nederlands', 'Help', and 'Afmelden'. Below this is a secondary bar with 'OPDRACHTEN', 'CURSUS' (highlighted with a green box), 'CUFERLIJST', 'BERICHTEN', and 'INSTELLINGEN'. On the left, there's a sidebar with the Pearson SEE mbo A2 logo. The main content area is titled 'Proefexamen' and contains a list of links: 'Over het examen', 'Apparatuur-check' (highlighted with a green box), 'Oefentoets', and 'Proefexamen'. A button 'Overschakelen naar de Studentweergave' is also visible.

Elke student heeft een ingebouwde apparatuurcheck in de Pearson SEE mbo-toets. Wij adviseren dat u beschikbaar blijft terwijl de studenten deze check uitvoeren.

Overige informatie

Technische ondersteuning: <https://support.pearsonelt.com/>

U kunt ook contact opnemen met onze Helpdesk via:
<http://www.pearsoneltsupport.com> of +31707703682

Zo geeft u toestemming aan de webbrowser voor toegang tot de microfoon:

<https://www.youtube.com/watch?v=s1PQR422sUc>

Wijzigen van de volumes van de microfoon en headset:

<https://www.youtube.com/watch?v=GZ8hx53JLyQ>

3.2 Contactinformatie

Pearson SEE hulplijn
voor ondersteuning & hulp tijdens Pearson SEE toetsmomenten.

+31208089873

Tijdslijn

DOCENT

ICT

STUDENT

Training over Pearson SEE mbo door Pearson consultant + docentenhandleiding bespreken

Consultant bespreekt ICT Handleiding

Informatiesessie over *Pearson SEE mbo* voor studenten (zie puntje 5 in docentenhandleiding)

Informatiesessie over *Pearson SEE mbo*

Proefexamen wordt aangeboden

Controle apparatuur
Ondersteuning tijdens de toets

Proefexamen afleggen (Printout van wat wel en niet is toegestaan tijdens toets – zie bijlage 2 in de docentenhandleiding)

Evaluatie van de scores

Vraag- en antwoordsessie met studenten over ervaringen met Pearson SEE mbo

Contact met consultant over *Pearson SEE mbo* in geval van vragen

Controle van de apparatuur
Ondersteuning tijdens de toets

Instellingsexamen wordt aangeboden

Instellingsexamen afleggen (Printout van wat wel en niet is toegestaan tijdens toets – zie bijlage 2 in de docentenhandleiding)

APPENDIX A



Pearson SEE mbo

Global Scale of English



Global Scale of English Learning Objectives for Adult Learners

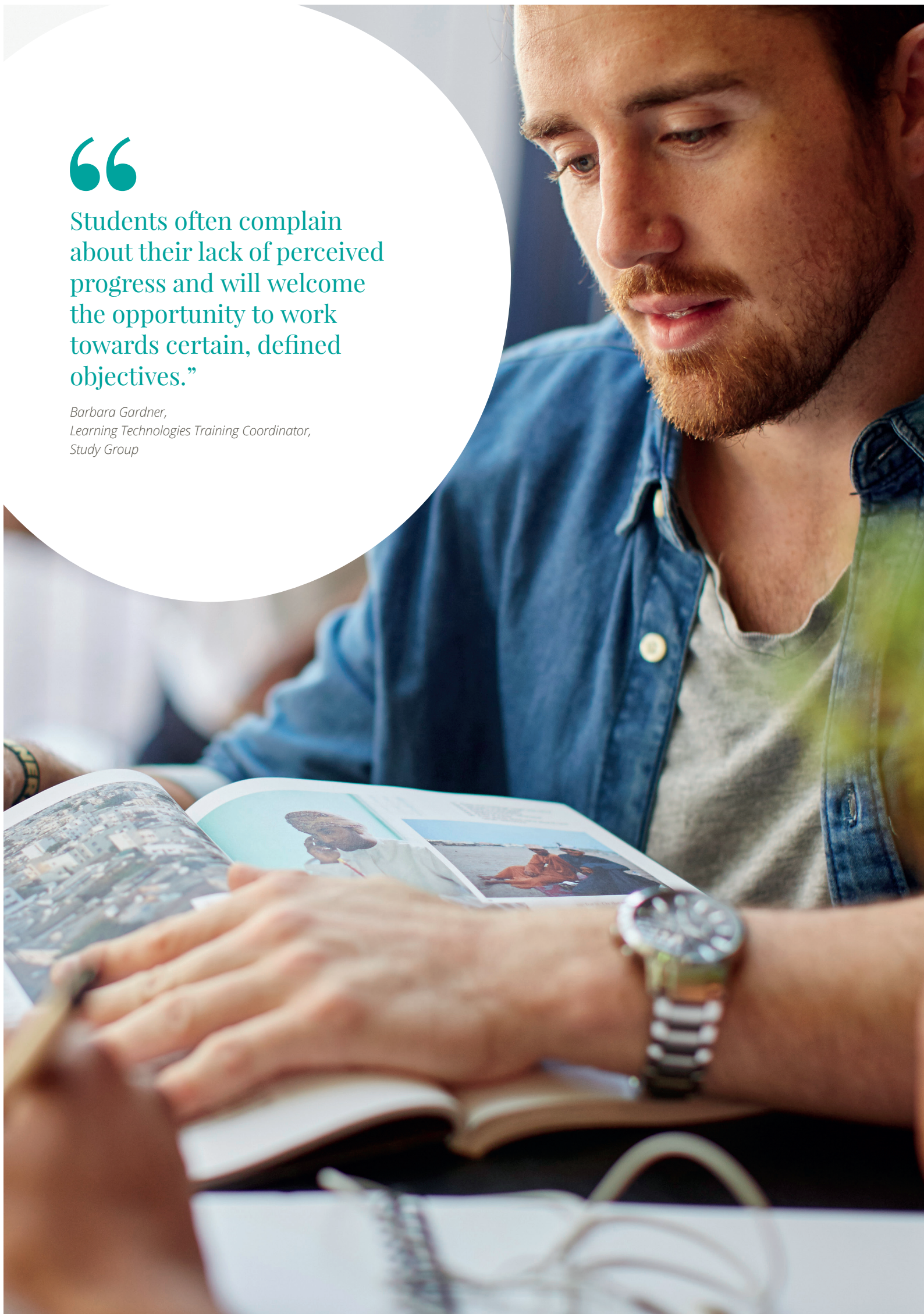


Published October 2015

“

Students often complain about their lack of perceived progress and will welcome the opportunity to work towards certain, defined objectives.”

*Barbara Gardner,
Learning Technologies Training Coordinator,
Study Group*



Global Scale of English Learning Objectives: making learning more effective

This document introduces the Global Scale of English (GSE) Learning Objectives. These Learning Objectives have been developed by Pearson English over a number of years in collaboration with teachers, ELT authors and language experts from around the world. They form the backbone of our mission to ensure that our products and services have a positive and measurable impact on learners' lives.

In order to learn English as effectively and efficiently as possible, a learner needs to know three things:

- What level is my English?
- Am I making progress?
- What should I aim to learn next to meet my long-term goals?

To answer these three questions, learners and teachers need the following as part of an overall English learning ecosystem:

- A clear and precise definition of what it means to be at a particular 'level' of proficiency
- English teaching and learning materials which are aligned to the 'level' definitions
- An assessment tool designed to profile a learner's proficiency across all four skills: reading, writing, speaking and listening



The Global Scale of English represents the most significant advance in performance-based approaches to language learning, teaching and assessment since the development of the Common European Framework of Reference.”

David Nunan Ph.D., Professor Emeritus of Applied Linguistics, University of Hong Kong

The GSE Learning Objectives form part of the first point above: providing an accurate definition of what it means to be at a particular 'level' on a scale of English language proficiency. Four sets of Learning Objectives have been developed, each tailored to meet the needs of specific audiences – Learners of Professional English, Academic English, Young Learners and General Adult learners.

For more information about our work in this field, please visit **[English.com/gse](https://www.english.com/gse)**.

The Global Scale of English

The Global Scale of English (GSE) is a standardised, granular scale which measures English language proficiency. Unlike some other frameworks which describe attainment in broad bands, the Global Scale of English identifies what a learner can do at each point on the scale across speaking, listening, reading and writing skills.

The scale is designed to motivate learners by accurately demonstrating step-by-step progress. Using the Global Scale of English, teachers can match a student to the right course materials to suit their exact level and learning goals.

The Global Scale of English serves as a standard against which English language courses and assessments worldwide can be benchmarked, offering a truly global and shared understanding of proficiency levels.

Visit **English.com/gse** to learn more.

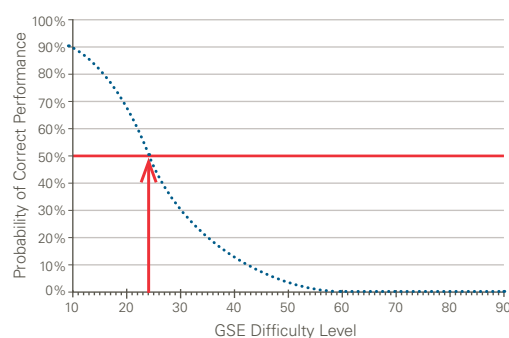
The GSE Learning Objectives are mapped to the Global Scale of English and describe what a learner can do at different levels of proficiency on the scale.

What it means to be at a level

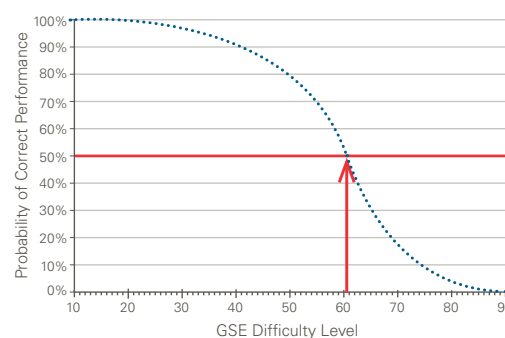
Learning a language is not like learning mathematics or electrical engineering, where each topic builds upon a previous one in a sequence. Language learning is not necessarily sequential, and a learner might be strong in one area, where they have had a lot of practice or a particular need or motivation, but quite weak in another.

For that reason, to say that a learner is 'at' a certain level on the Global Scale of English does not mean they have necessarily mastered every GSE Learning Objective for every skill up to that point. Neither does it mean that they have mastered none at a higher GSE value. If a student is assessed as being at 61 on the scale, it means s/he has a 50% probability of being able to perform Learning Objectives at that level, a greater probability of being able to perform Learning Objectives at a lower level, and a lower probability of being able to perform Learning Objectives at a higher level. Language learning is unique to every individual.

A learner at 25 on GSE



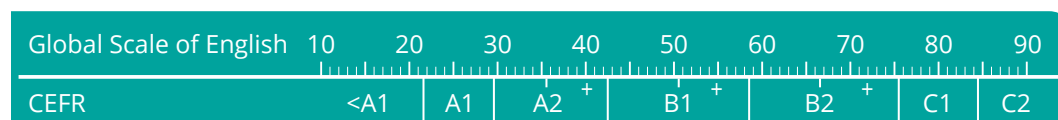
A learner at 61 on GSE



Extending the Common European Framework of Reference for Languages

The work to develop the GSE Learning Objectives builds upon the research carried out by Brian North and the Council of Europe in creating the Common European Framework of Reference for Languages (CEFR)¹.

The Global Scale of English itself has been psychometrically aligned to the CEFR.



The CEFR uses a six-level classification of learner proficiency from A1 (low basic) to C2 (fully proficient). The amount of instruction needed to progress learners from one level to the next varies widely according to level, context, native language, age, ability, and other factors, so it is difficult to quantify exactly. However, it has been observed that most people studying for three or four hours per week (as is the case for many language learners) may take two or more years to move from one CEFR level to the next – and as proficiency increases, it takes even longer to move to the next CEFR level. When learners spend two or more years studying without reaching a new CEFR level, it can leave them feeling that they are making little or no progress.

In developing the GSE Learning Objectives, we have extended the number and range of learning objectives – or ‘Can Do Statements’ – that are included in the CEFR, providing information to support a far more granular definition of language proficiency. Students are much more motivated when they can see every small step in their progression.

Personalisation of progress



For teachers, assessment specialists and content developers, the GSE ecosystem provides a detailed picture of language performance at different levels of proficiency and for individual skills. By combining course materials with assessment tools that are aligned to the Global Scale of English, teachers can:

- understand their students' levels of proficiency more precisely
- monitor students' progress at a granular level
- make more informed choices for each student or class

¹ These learning objectives extend and build on the principles of the Common European Framework of Reference for Languages (CEFR) in their intended use “...[for]...the planning of language learning programmes in terms of their assumptions regarding prior knowledge, ...their objectives, [and] their content” (Council of Europe, 2001, p6). Like CEFR, the GSE Learning Objectives are also intended for use in “the planning of language certification in terms of the content syllabus of examinations [and] assessment criteria, in terms of positive achievement.” (ibid)

Informing teaching and assessment materials

Authors, editors and teachers are increasingly aware of the need to be learner-focused and to create syllabuses and courses that reflect learner needs and expectations. The GSE Learning Objectives constitute a detailed and graded model of student target performance across a range of skills and domains.

Our authors and editors are using the GSE Learning Objectives as the starting point for the creation of all new teaching and learning materials. The following example from *SpeakOut* illustrates how the authors have developed content to support the GSE Learning Objective for Speaking:

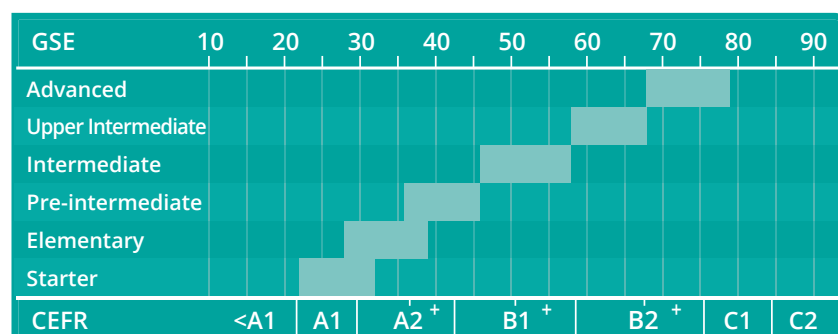
35 Can describe a travel experience with a few very basic stock phrases.

In the lesson, learners engage with a series of staged activities designed to support and guide them towards being able to describe a travel experience of their own.

- Speaking and vocabulary activities check learners' language knowledge and pre-teach key travel vocabulary, working with images and text
- Controlled writing provides practice of travel vocabulary in context
- Pre-reading prediction prepares learners for the reading passage describing a travel experience
- Follow-up activities raise awareness of key phrases used to describe a travel experience
- Speaking activities enable practice in a controlled context
- As a final activity, learners work in groups and prepare to describe a travel experience of their own, which they then present to the class

By the end of the lesson, learners have received the input required (skills, structures, vocabulary) and demonstrated their ability to perform the Learning Objective.

Many Pearson English courses for adult and young adult learners, such as *SpeakOut*, *Cutting Edge*, *Total English*, *Language Leader* and *Focus* are aligned to the Global Scale of English. You will find information on the back cover of the Student's Book that indicates the GSE range covered by each stage of a course:



The above chart indicates the target range on the Global Scale of English (and the CEFR). This indicates the range of GSE Learning Objectives that authors have written to for that level.

Assessment tools have also been created to report on the Global Scale of English. The newest assessment offering from Pearson English is *Progress* – a package of three tests taken at the beginning, in the middle and at the end of a course of study to measure progress. At each stage, a test report indicates a learner's score on the Global Scale of English – and identifies specific GSE Learning Objectives that require further work to improve. To learn more about *Progress*, go to [English.com/progress](https://www.english.com/progress).

GSE and the CEFR Levels

In the following we define how the Global Scale of English is related to the CEFR levels. To give an impression of what the levels mean, i.e., what learners at particular levels can do, we use the summary descriptors published as Table 1 in the CEFR (Council of Europe, 2001, p. 24).

GSE 10–21 (Below A1): Global assessment

The range on the Global Scale of English from 10 to 21 covers the area of measurable proficiency below the A1 level of the CEFR.

It includes the level which North (2000, p. 295) characterises as ‘Tourist’, corresponding to a range of 13–21 on the GSE, and a still lower ability which North (ibid.) labels ‘Smattering’. Neither of these was included in the CEFR, because A1 was “considered the lowest level of generative language use” (Council of Europe, 2001, p. 33) and ‘Tourist’ and ‘Smattering’ rely “purely on a very finite rehearsed, lexically organised repertoire of situation-specific phrases” (ibid.). GSE Learning Objectives in the range 10–21 have however been included, representing the key steps in learners’ progress towards A1.

GSE 22–29 (A1): Global assessment

The capabilities of learners at Level A1 have been summarised in the CEFR as follows:

Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help (Council of Europe, 2001, Table 1, p. 24).

GSE 30–35 (A2) and 36–42 (A2+): Global assessment

The capabilities of learners at Level A2 have been summarised in the CEFR as follows:

Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need (Council of Europe, 2001, Table 1, p. 24).

GSE 43–50 (B1) and 51–58 (B1+): Global assessment

The capabilities of learners at Level B1 have been summarised in the CEFR as follows:

Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans (Council of Europe, 2001, Table 1, p. 24).

GSE 59-66 (B2) and 67-75 (B2+): Global assessment

The capabilities of learners at Level B2 have been summarised in the CEFR as follows:

Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options (Council of Europe, 2001, Table 1, p. 24).

GSE 76–84 (C1): Global assessment

The capabilities of learners at Level C1 have been summarised in the CEFR as follows:

Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices (Council of Europe, 2001, Table 1, p. 24).

GSE 85-90 (C2): Global assessment

The capabilities of learners at Level C2 have been summarised in the CEFR as follows:

Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations (Council of Europe, 2001, Table 1, p. 24).

Sources and origins of the Learning Objectives

Many of the communicative descriptors contained in the Learning Objectives either come from the original CEFR document (Council of Europe, 2001) or are slightly modified versions of these. We have created the additional descriptors by consulting a variety of sources, including:

- The Council of Europe's *Breakthrough* specification (Trim, 2009)
- The EAQUALS/British Council Core Inventory (British Council/EAQUALS, 2010)
- British and American English course materials
- Syllabuses from various Ministries of Education.

The GSE Learning Objectives are coded for purposes of copyright and to show their origin:

(C)	Common European Framework descriptor, verbatim, © Council of Europe
(C _A)	Common European Framework descriptor, adapted or edited, © Council of Europe
(N2000)	North (2000) descriptor, verbatim
(N2000 _A)	North (2000) descriptor, adapted or edited
(N2007 _A)	North (2007) expanded set of C1 and C2 descriptors, adapted or edited
(P)	New Pearson descriptor

GSE Learning Objectives

Reading

GSE 10–21/Below A1: Reading

10	Can recognise cardinal numbers up to 10. (P)
	Can recognise the letters of the alphabet. (P)
12	Can read and understand simple prices. (P)
14	Can identify very common food and drink on a menu. (P)
	Can recognise basic plural forms of nouns (e.g. cars, books). (P)
15	Can recognise familiar names, words and very basic phrases on simple notices. (C _A)

GSE 22–29/A1: Reading

26	Can follow short, simple written directions (e.g. to go from X to Y). (C)
27	Can understand short written notices, signs and instructions with visual support. (P)

GSE 30–35/A2: Reading

31	Can understand simple questions in questionnaires on familiar topics. (P)
	Can find specific, predictable information in everyday materials (e.g. menus, timetables). (C _A)
	Can understand short, simple messages on postcards, emails and social networks. (C _A)
34	Can understand the general meaning of short, simple informational material and descriptions if there is visual support. (C _A)
	Can understand simple instructions on everyday equipment (e.g. cash machines). (C _A)

GSE 36–42/A2(+): Reading

37	Can understand rules and regulations (e.g. safety) if expressed in simple language. (C)
	Can identify specific information in simple letters, brochures and short articles. (C _A)
	Can understand short, simple personal emails and letters. (C _A)
	Can understand reports of travel delays and cancellations. (P)
38	Can make basic inferences from simple information in a short text. (P)
39	Can understand simple factual titles and headlines relating to common events. (P)
	Can identify specific information in a simple factual text. (P)
	Can understand the instructions to buy tickets online. (P)
42	Can understand the main information from simple diagrams (e.g. graphs, bar charts). (P)

GSE 43–50/B1: Reading

43	Can understand simple technical information (e.g. instructions for everyday equipment). (C _A)
	Can make basic inferences or predictions about text content from headings, titles or headlines. (P)
44	Can understand written instructions for taking medication. (P)

45	Can understand information in advertisements for jobs and services. (P)
	Can find and understand information in advertisements for sporting or cultural events. (P)
46	Can derive the probable meaning of simple unknown words from short, familiar contexts. (CA)
	Can understand clearly written, straightforward instructions on how to use a piece of equipment. (CA)
	Can understand basic types of standard letters and emails on familiar topics (e.g. enquiries, complaints). (CA)
	Can generally understand straightforward factual texts on familiar topics. (CA)
48	Can extract relevant details in everyday letters, brochures and short official documents. (CA)
49	Can identify the main topic and related ideas in a structured text. (P)
50	Can understand the relationship between a main point and an example in a structured text. (P)

GSE 51–58/B1(+): Reading

51	Can recognise the direct repetition of ideas as a simple cohesive device. (P)
	Can distinguish between fact and opinion in relation to common topics. (P)
	Can make simple inferences based on information given in a short article. (P)
	Can generally understand details of events, feelings and wishes in letters, emails and online postings. (CA)
52	Can follow chronological sequence in a formal structured text. (P)
54	Can recognise examples and their relation to the idea they support. (P)
55	Can understand written advice and instructions for resolving a problem with a product or piece of equipment. (P)
56	Can recognise significant points and arguments in straightforward newspaper articles on familiar topics. (CA)
57	Can understand cause and effect relationships in a structured text. (P)
	Can recognise the general line of a written argument though not necessarily all the details. (CA)
58	Can understand most correspondence relating to their field of interest. (CA)
	Can recognise the writer's point of view in a structured text. (P)
	Can understand problem and solution relationships in a structured text. (P)

GSE 59–66/B2: Reading

59	Can use a monolingual dictionary to check the meaning of words without needing to refer to a bilingual dictionary. (P)
	Can understand instructions for making financial transactions online. (P)
	Can identify the main conclusions in a text that presents and contrasts arguments in a clearly signalled way. (CA)
60	Can make inferences or predictions about the content of newspaper and magazine articles from headings, titles or headlines. (P)
61	Can follow the exchanges on the discussion board of a website. (P)
	Can distinguish supporting details from the main points in a text. (P)
62	Can interpret the main message from complex diagrams and visual information. (P)
	Can recognise the author's use of irony in a simple text, if guided by questions. (P)
	Can understand the author's purpose and intended audience. (P)

-
- 63 Can use reference materials to check factual information, when guidance on finding relevant sources of information is provided. (P)
-
- Can scan a long text or a set of related texts in order to find specific information. (C_A)
-
- 65 Can recognise the tone and intended audience of a structured text. (P)
-
- 66 Can summarise, comment on and discuss a wide range of factual and imaginative texts. (C_A)
-

GSE 67–75/B2(+): Reading

- 67 Can recognise the repetition of ideas expressed by substitution, paraphrasing, etc. (P)
-
- 68 Can recognise contrasting arguments in structured, discursive text. (P)
-
- Can understand complex technical information such as operating instructions, specifications for familiar products and services. (C)
-
- Can infer what will come next in an unstructured text by using contextual, grammatical and lexical cues. (P)
-
- Can understand the details of long complex instructions in their field, rereading as necessary. (C_A)
-
- 69 Can synthesise information from different sources in order to give a written or oral summary. (P)
-
- Can quickly scan long, complex texts for key information. (P)
-
- Can use a variety of reference materials to check factual information quickly and efficiently. (P)
-
- 70 Can understand inferred meaning in formal structured text. (P)
-
- Can understand complex, detailed correspondence, with occasional support from a dictionary. (C_A)
-
- 71 Can distinguish between fact and opinion in complex formal contexts. (P)
-
- Can understand complex questions in questionnaires designed to elicit opinions. (P)
-
- 73 Can identify specific information in a linguistically complex factual text. (P)
-
- Can recognise the author's use of irony in a text. (P)
-
- 74 Can understand the intended double meaning of a word in a written text. (P)
-

GSE 76–84/C1: Reading

- 76 Can understand correspondence containing idiomatic or non-standard language (C_A)
-
- 78 Can extract information, ideas and opinions from highly specialised sources within their field. (C_A)
-
- 79 Can understand complex arguments in newspaper articles. (P)
-
- 80 Can understand the details of long complex texts in their field without needing to reread. (C_A)
-

GSE 85–90/C2: Reading

- 87 Can understand highly colloquial language in unstructured texts that use complex structures. (P)
-

Listening

GSE 10–21/Below A1: Listening

10	Can recognise simple informal greetings. (P)
	Can understand the letters of the alphabet. (P)
	Can understand cardinal numbers from 1 to 20. (P)
13	Can recognise simple formal greetings. (P)
16	Can understand very basic common classroom instructions. (P)
18	Can understand the time of day when expressed in full hours. (P)
19	Can understand cardinal numbers from 21 to 100. (P)
20	Can understand ordinal numbers from 1 to 100. (P)
	Can understand simple language related to prices and quantities. (P)
21	Can understand basic personal details if given carefully and slowly. (P)
	Can understand basic questions about personal details if addressed slowly and clearly. (P)

GSE 22–29/A1: Listening

23	Can understand questions addressed carefully and slowly. (C _A)
	Can understand short, simple instructions addressed carefully and slowly. (C _A)
	Can understand the time of day when expressed to the quarter hour. (P)
	Can understand the time of day when expressed to within five minutes. (P)
24	Can distinguish between <i>can</i> and <i>can't</i> . (P)
	Can understand basic questions about people's likes and dislikes. (P)
25	Can understand basic information about someone's likes and dislikes. (P)
	Can follow speech which is very slow and carefully articulated, with long pauses. (C _A)
	Can understand cardinal numbers from 101 to 1000. (P)
27	Can understand a phone number from a recorded message. (P)
	Can understand basic questions about free time activities. (P)
	Can understand basic information about free time activities. (P)

GSE 30–35/A2: Listening

30	Can recognise phrases and content words related to basic personal and family information. (C _A)
	Can extract key factual information such as prices, times and dates from short clear, simple announcements. (P)
31	Can recognise phrases and content words related to familiar topics (e.g. shopping, local geography). (C)
33	Can understand simple, everyday conversations if conducted slowly and clearly. (C _A)
	Can follow short, simple social exchanges. (P)
35	Can understand who a phone call is intended for. (P)
	Can extract key factual information such as prices, times and dates from a recorded phone message. (P)

GSE 36–42/A2(+): Listening

- | | |
|----|---|
| 36 | Can recognise when speakers agree in a conversation conducted slowly and clearly. (P) |
| | Can recognise when speakers disagree in a conversation conducted slowly and clearly. (P) |
| 37 | Can understand enough to respond to direct requests expressed slowly and clearly. (CA) |
| 38 | Can understand standard speech on familiar matters, with some repetition or reformulation. (CA) |
| 39 | Can generally identify the topic of discussion around them when conducted slowly and clearly. (C) |
| 40 | Can understand enough to manage simple routine exchanges without undue effort. (C) |
| 41 | Can derive the probable meaning of simple, unknown words from short, familiar contexts. (P) |
| 42 | Can understand basic medical advice. (P) |

GSE 43–50/B1: Listening

- | | |
|----|--|
| 43 | Can listen to a short narrative and predict what will happen next. (N2000) |
| | Can understand instructions delivered at normal speed and accompanied by visual support. (P) |
| 44 | Can recognise when a speaker is checking that the listener has understood something in a conversation conducted slowly and clearly. (P) |
| | Can follow everyday conversation, with some repetition of particular words and phrases. (P) |
| 45 | Can follow the main points of short talks on familiar topics if delivered in clear standard speech. (CA) |
| | Can follow familiar topics if the speaker is clear and avoids idiomatic usage. (CA) |
| 46 | Can follow the main points of extended discussion around them if in standard speech. (CA) |
| 47 | Can understand the main points of narratives and conversations about familiar topics (e.g. work, leisure) delivered in clear standard speech. (CA) |
| 49 | Can understand simple technical instructions for everyday equipment. (CA) |
| | Can identify the main point of TV news items reporting events, accidents, etc. where the visual supports the commentary. (C) |
| | Can generally follow changes of topic in discussions related to their field if conducted slowly and clearly. (CA) |
| | Can follow most of an everyday conversation if speakers avoid very idiomatic usage. (CA) |



GSE 51–58/B1(+): Listening

-
- 51 Can distinguish between main ideas and supporting details in familiar, standard texts. (P)
 Can follow the main points in TV programmes on familiar topics if delivered in clear standard speech. (CA)
 Can follow recorded instructions and information given on a phone-delivered service. (P)
-
- 52 Can follow many films in which visuals and action carry much of the storyline. (CA)
 Can recognise that a joke has been made, even if the meaning is not fully understood. (P)
-
- 53 Can follow detailed directions. (C)
-
- 55 Can recognise examples and their relation to the idea they support. (P)
 Can understand advice and instructions for resolving a problem with a product or piece of equipment. (P)
-
- 57 Can extract the meaning of unknown words from context if the topic discussed is familiar. (CA)
 Can generally follow rapid or extended speech, but may require repetition or clarification. (CA)
 Can follow most of a clearly structured presentation within their own field. (CA)
 Can deduce the general meaning of a passage from context in a longer, structured text. (P)
-
- 58 Can understand a large part of many TV programmes on familiar topics. (CA)
 Can understand problem and solution relationships in informal conversation. (P)
-

GSE 59–66/B2: Listening

-
- 59 Can identify the main reasons for and against an argument or idea in a discussion delivered in clear standard speech. (P)
-
- 60 Can understand most of a radio programme about a familiar topic. (CA)
-
- 61 Can recognise when examples are being given in a structured presentation on an unfamiliar topic. (P)
 Can follow changes of topic in factual TV news items and form an idea of the main content. (C)
 Can understand scripted speech delivered quickly, if the accent is familiar. (CA)
-
- 62 Can follow a natural group discussion, but may find it difficult to participate effectively. (CA)
 Can understand detailed instructions well enough to be able to follow them without making mistakes. (C)
-
- 63 Can recognise the speaker's point of view in a structured presentation. (P)
-
- 64 Can understand TV documentaries, interviews, plays and most films in standard speech. (CA)
 Can understand main points and check comprehension by using contextual clues. (CA)
 Can understand unscripted speech delivered quickly, if the accent is familiar. (CA)
-
- 65 Can extract the main points from news items, etc. with opinions, arguments and discussion. (CA)
 Can understand cause and effect relationships in informal conversation at natural speed. (P)
 Can follow chronological sequences in extended informal speech at natural speed. (P)
-
- 66 Can understand the main ideas of complex technical discussions in their field. (CA)
 Can recognise the tone and intended audience of a formal presentation. (P)
 Can distinguish between fact and opinion in informal discussion at natural speed. (P)
-

GSE 67–75/B2(+): Listening

-
- | | |
|-------|---|
| 69 | Can extract specific details from poor quality public announcements, e.g., in a station, sports stadium, etc. (N2000A) |
| <hr/> | |
| 70 | Can follow a wide range of factual and creative texts and summarise themes and opinions. (CA) |
| <hr/> | |
| 71 | Can differentiate between rhetorical and genuine questions in informal discussion. (P) |
| <hr/> | |
| | Can infer attitude and mood in discussions by using contextual, grammatical and lexical cues. (CA) |
| <hr/> | |
| | Can understand the main points of complex and abstract presentations in their field. (CA) |
| <hr/> | |
| | Can understand when something is being said ironically in a casual conversation. (N2007A) |
| <hr/> | |
| 72 | Can understand the main points of complex academic/professional presentations. (CA) |
| <hr/> | |
| | Can understand most TV news and current affairs programmes. (C) |
| <hr/> | |
| | Can understand detailed medical advice. (P) |
| <hr/> | |
| 73 | Can follow extended speech expressing unstructured ideas and thoughts. (P) |
| <hr/> | |
| 74 | Can follow an animated conversation between two fluent speakers. (CA) |
| <hr/> | |
| | Can follow extended speech on abstract and complex topics outside their field, if able to ask for confirmation of details. (CA) |
-

GSE 76–84/C1: Listening

-
- | | |
|-------|---|
| 76 | Can follow a fast-paced conversation between fluent speakers well enough to be able to contribute. (CA) |
| <hr/> | |
| 77 | Can recognise coherence devices and follow complex arguments on unfamiliar topics. (P) |
| <hr/> | |
| 78 | Can understand the intended double meaning of a word used in a joke. (P) |
| <hr/> | |
| 79 | Can follow presentations on abstract and complex topics outside their field of interest. (P) |
| <hr/> | |
| 80 | Can follow films employing a considerable degree of slang and idiomatic usage. (C) |
| <hr/> | |
| 83 | Can recognise a wide range of idiomatic expressions and colloquialisms, appreciating register shifts. (N2000) |
-

GSE 85–90/C2: Listening

-
- | | |
|----|---|
| 88 | Can understand in detail discussions on abstract and complex topics among speakers with a variety of accents and dialects. (CA) |
|----|---|
-

Speaking

GSE 10–21/Below A1: Speaking

10	Can ask someone for their name. (P)
	Can say their name. (P)
11	Can make simple purchases by pointing or other gestures. (N2000A)
	Can say the letters of the alphabet. (P)
12	Can greet people using a few basic fixed expressions. (P)
	Can name a few very common everyday objects. (P)
	Can recognise and say the name of their own country, nationality and language. (P)
13	Can ask someone what their nationality is. (P)
	Can read out phone numbers. (P)
14	Can spell out their own name and address. (P)
15	Can say other people's nationalities. (P)
16	Can ask and answer basic requests for information with <i>What's this/that?</i> (P)
	Can tell the time of day in full hours. (P)
	Can name a few common jobs. (P)
17	Can use some very basic words to ask for food and drink. (P)
	Can name very common forms of transport. (P)
	Can ask for and give a phone number. (P)
18	Can say their own age and ask someone about their age. (P)
19	Can ask for and give the day and date. (N2000A)
	Can say what they do (e.g. name of their job, student). (P)
	Can ask about the price of something. (P)
	Can ask someone what their job is. (P)
	Can establish basic social contacts with simple, polite greetings and farewells. (CA)
20	Can say a range of basic numbers, quantities and prices. (CA)
	Can ask for and give a date of birth. (P)
	Can ask where other people are in a limited way. (P)
21	Can ask and answer simple questions about things they have in a limited way. (CA)
	Can ask very simply for repetition when they don't understand. (C)

GSE 22–29/A1: Speaking

-
- 22 Can ask for the spelling of a word, or for a word to be written down. (P)
 Can ask and answer basic questions about family and friends in a limited way. (P)
 Can tell the time of day to within five minutes. (P)
 Can say where they and other people are in a limited way. (P)
-
- 23 Can describe the position of something in a very basic way. (P)
 Can give basic information about the price of something. (P)
-
- 24 Can greet people, ask how they are and react to news. (CA)
 Can tell the time of day to the quarter hour. (P)
 Can ask for a drink or food in a limited way. (P)
-
- 25 Can ask and answer simple questions in areas of immediate need or on very familiar topics. (CA)
 Can ask for and give very basic information about the home. (P)
 Can indicate time by such phrases as *next week, last Friday, in November, three o'clock*. (C)
 Can accurately repeat clearly spoken words, phrases, and short sentences. (P)
-
- 26 Can describe where they live. (CA)
 Can check into a hotel using a few basic fixed expressions. (P)
-
- 27 Can express ability or lack of ability with regard to basic activities using *can* or *can't*. (P)
-
- 28 Can ask and answer simple questions about people they know in a limited way. (CA)
 Can make an introduction and use basic greeting and leave-taking expressions. (C)
 Can ask people for things and give people things. (C)
 Can express how they are feeling using very basic fixed expressions. (P)
 Can exchange personal details (e.g. where they live, things they have). (CA)
 Can express preferences about food and drink using basic fixed expressions. (P)
 Can use basic words to describe common weather conditions. (P)
 Can describe a person's likes and dislikes using simple language. (P)
 Can use brief, everyday expressions to describe wants and needs, and request information. (CA)
-
- 29 Can ask for attention. (C)
-

GSE 30–35/A2: Speaking

-
- 30 Can express basic intentions with simple time markers (e.g. *tomorrow*). (P)
 Can initiate and respond to simple statements on very familiar topics. (CA)
 Can talk about hotel accommodation using simple language. (P)
 Can give a short description of their home, family and job, given some help with vocabulary. (P)
 Can describe a person's hobbies and activities using simple language. (P)
-
- 31 Can use simple phrases to order a meal. (CA)
 Can make simple purchases by stating what is wanted and asking for the price. (C)
 Can make and accept a simple apology. (P)
 Can ask simple questions to find out about a subject. (P)
 Can describe what someone is wearing using a limited range of expressions. (P)
 Can talk about furniture and rooms using simple language. (P)
 Can ask for and provide things using simple phrases. (CA)
-
- 32 Can use brief, everyday expressions to ask for and give personal details. (CA)
 Can ask for simple directions from X to Y on foot or by public transport. (P)
 Can handle common everyday transactions (e.g. buying a ticket). (CA)
-
- 33 Can make simple transactions in shops, post offices and banks. (CA)
 Can describe their family, living conditions, education and present or most recent job. (C)
 Can introduce themselves on the phone and close a simple call. (P)
 Can describe basic activities or events that are happening at the time of speaking. (P)
 Can make simple references to the past using *was/were*. (P)
 Can describe skills and abilities using simple language. (P)
-
- 34 Can use simple language to describe people's appearance. (N2000)
 Can say what they like and dislike. (C)
 Can give simple directions from X to Y on foot or by public transport. (P)
 Can ask for basic advice using simple language. (P)
 Can ask simple questions in a face-to-face survey. (P)
 Can answer simple questions in a face-to-face survey. (P)
 Can describe people's everyday lives using a short series of simple phrases and sentences. (CA)
 Can ask someone to repeat a specific point or idea. (P)
 Can answer simple questions on the phone using fixed expressions. (P)
-
- 35 Can ask for and provide everyday goods and services. (C)
 Can ask and answer questions about what they do at work and in their free time. (C)
 Can describe a travel experience with a few very basic stock phrases. (P)
 Can use a limited range of fixed expressions to describe objects, possessions, or products. (P)
 Can ask for repetition or clarification on the phone in a simple way. (P)
-

GSE 36–42/A2(+): Speaking

- 36 Can communicate in routine tasks requiring simple, direct exchanges of information. (C_A)
 Can make and accept offers. (N2000)
 Can discuss what to do and where to go, and make arrangements to meet. (C)
 Can use simple, everyday polite forms of greeting and address. (C)
 Can ask and talk about very basic symptoms and ailments (e.g. cold, flu). (P)
 Can leave simple phone messages using fixed expressions. (P)
- 37 Can answer simple questions and respond to simple statements in an interview. (C)
 Can make simple, direct comparisons between two people or things using common adjectives. (P)
 Can describe what something is used for, using basic fixed expressions. (P)
- 38 Can give an extended description of everyday topics (e.g. people, places, experiences). (N2000_A)
 Can deal with practical everyday demands, exchanging straightforward factual information. (C_A)
 Can express how they feel in simple terms. (C_A)
 Can ask and answer questions about habits and routines. (C)
 Can describe habits and routines. (C_A)
 Can ask and answer questions about basic plans and intentions. (P)
 Can describe very basic events in the past using simple linking words (e.g. *then*, *next*). (P)
 Can take simple phone messages using fixed expressions. (P)
 Can read out a short, rehearsed statement (e.g. introduce a speaker, propose a toast). (C_A)
 Can make hotel, restaurant, or transport reservations on the phone. (P)
- 39 Can get information from a tourist office of a straightforward, non-specialised nature. (C_A)
 Can describe plans and arrangements. (C_A)
 Can initiate, maintain and close simple, restricted face-to-face conversations. (N2000)
 Can give basic advice using simple language. (P)
- 40 Can ask for and give or refuse permission. (N2000)
 Can ask for clarification about key words not understood, using fixed expressions. (C_A)
 Can ask and answer questions about past times and past activities. (C)
 Can discuss what to do in the evening or at the weekend. (C)
 Can explain what they like or dislike about something. (C)
 Can express enthusiasm and excitement in a limited way. (P)
 Can give a simple description of how to carry out an everyday process (e.g. a recipe). (P)
 Can tell a story or describe something in a simple list of points. (C)
 Can make a short rehearsed announcement on a familiar topic. (C_A)
- 41 Can talk about personal possessions, including household pets. (N2000_A)
 Can participate in short conversations in routine contexts on topics of interest. (C)
 Can make and respond to suggestions. (C)
 Can discuss what to do next using simple phrases. (C_A)
 Can cancel hotel, restaurant, or transport reservations on the phone. (P)

-
- 42 Can give a short, basic description of events and activities. (C)
-
- Can use some basic interjections to express understanding, surprise, disappointment, and excitement. (P)
-
- Can make simple future arrangements and plans with reference to a diary or schedule. (P)
-
- Can make an appointment on the phone. (P)
-

GSE 43–50/B1: Speaking

- 43 Can describe future plans and intentions using fixed expressions. (P)
-
- Can paraphrase a simple factual statement related to a familiar topic. (P)
-
- 44 Can ask for, follow and give detailed directions. (CA)
-
- Can summarise short written passages using the original wording and ordering. (CA)
-
- 45 Can convey simple information of immediate relevance and emphasise the main point. (CA)
-
- Can deal with less routine situations on public transport (e.g. asking where to get off). (N2000_A)
-
- Can carry out a prepared structured interview with some spontaneous follow-up questions. (CA)
-
- Can deal with common situations when making travel arrangements or travelling. (CA)
-
- Can express belief, opinion, agreement and disagreement politely. (C)
-
- Can convey simple relevant information emphasising the most important point. (CA)
-
- Can narrate a story. (C)
-
- Can express opinions using simple language. (P)
-
- Can explain the rules of a familiar game or sport using simple language. (P)
-
- Can give an opinion when asked directly, provided they can ask for repetition. (CA)
-
- Can use simple appropriate language to check that information has been understood on the phone. (P)
-
- Can ask for changes to hotel, restaurant, or transport reservations on the phone. (P)
-
- Can return a phone call, explaining who is calling and the reason for the call. (P)
-
- 46 Can initiate, maintain and close simple, face-to-face conversations on familiar topics. (CA)
-
- Can give or seek personal views and opinions in discussing topics of interest. (C)
-
- Can express preferences about food and drink in detail. (P)
-
- Can arrange a delivery time on the phone. (P)
-
- Can buy a ticket for an event on the phone. (P)
-
- 47 Can enter unprepared into conversation on familiar topics (e.g. family, hobbies, work). (CA)
-
- Can give straightforward descriptions on a variety of familiar subjects. (CA)
-
- Can begin to use a repertoire of common idiomatic phrases in routine situations. (CA)
-
- Can describe events, real or imagined. (C)
-
- Can express attitudes using simple language. (P)
-
- 48 Can describe dreams, hopes and ambitions. (C)
-
- Can introduce a conversation topic with the present perfect and provide details in the past. (P)
-
- 49 Can give detailed accounts of experiences, describing feelings and reactions. (C)
-
- Can discuss everyday, practical issues when the conversation is conducted slowly and clearly. (CA)
-
- Can respond in a simple way to verbal challenge. (P)
-

-
- 50 Can take some initiative in an interview, but is generally very dependent on interviewer. (CA)
 Can relate the plot of a book or film and describe their reactions. (C)
 Can ask someone to clarify or elaborate what they have just said. (C)
 Can negotiate prices and ask for discounts. (P)
 Can give simple reasons to justify a viewpoint on a familiar topic. (P)
 Can discuss films, books or plays in simple terms, using fixed expressions. (P)
 Can give detailed directions to a driver. (P)
-

GSE 51–58/B1(+): Speaking

- 51 Can express opinions as regards possible solutions, giving brief reasons and explanations. (CA)
 Can express and respond to feelings (e.g. surprise, happiness, interest, indifference). (C)
 Can express opinions and react to practical suggestions of where to go, what to do, etc. (CA)
 Can make a complaint. (C)
 Can briefly give reasons and explanations for opinions, plans and actions. (C)
 Can report the opinions of others, using simple language. (P)
- 52 Can repeat back what is said to confirm understanding and keep a discussion on course. (CA)
 Can use a suitable phrase to invite others into a discussion. (CA)
 Can speak in general terms about environmental problems. (P)
 Can express opinions and attitudes using a range of basic expressions and sentences. (CA)
- 53 Can compare and contrast alternatives about what to do, where to go, etc. (CA)
 Can use a basic repertoire of conversation strategies to maintain a discussion. (CA)
 Can define the features of something concrete for which they can't remember the word. (C)
 Can develop an argument using common fixed expressions. (P)
 Can give a short, rehearsed talk or presentation on a familiar topic. (CA)
 Can re-tell a familiar story using their own words. (P)
 Can signal that they wish to bring a conversation to an end. (P)
 Can ask someone to paraphrase a specific point or idea. (P)
- 54 Can describe basic symptoms to a doctor, but with limited precision. (CA)
 Can relate the basic details of unpredictable occurrences (e.g. an accident). (CA)
 Can leave phone messages containing detailed information. (P)
- 55 Can use synonyms to describe or gloss an unknown word. (CA)
 Can explain the main points in an idea or problem with reasonable precision. (C)
 Can express their thoughts in some detail on cultural topics (e.g. music, films). (CA)
 Can explain why something is a problem. (C)
 Can respond to ideas and suggestions in informal discussions. (CA)
 Can generally follow most of what is said and repeat back details to confirm understanding. (CA)
-

- 56 Can summarise and comment on a short story or article and answer questions in detail. (CA)
 Can give brief comments on the views of others. (C)
 Can summarise and give opinions on issues and stories and answer questions in detail. (CA)
 Can give an opinion on practical problems, with support when necessary. (CA)
 Can express and comment on ideas and suggestions in informal discussions. (CA)
 Can ask for confirmation of understanding during a live discussion or presentation. (P)
- 57 Can carry out a prepared interview, checking and confirming information as necessary. (CA)
 Can collate information from several written sources and summarise the ideas orally. (CA)
 Can ask for advice on a wide range of subjects. (P)
 Can reasonably fluently relate a straightforward narrative or description as a linear sequence of points. (CA)
- 58 Can ask a question in a different way if misunderstood. (N2007A)
 Can report the opinions of others. (P)
 Can express disagreement in a manner that shows they were actively listening to the other person. (P)
 Can express support in a manner that shows they were actively listening to the other person. (P)

GSE 59–66/B2: Speaking

- 59 Can deal with less common situations in a shop, post office (e.g. returning an unsatisfactory purchase). (CA)
 Can exchange information on a wide range of topics within their field with some confidence. (CA)
 Can describe objects, possessions and products in detail, including their characteristics and special features. (P)
 Can give basic technical instructions in their field of specialisation. (P)
- 60 Can justify a viewpoint on a topical issue by discussing pros and cons of various options. (CA)
 Can correct mistakes if they have led to misunderstandings. (N2000)
 Can justify and sustain views clearly by providing relevant explanations and arguments. (CA)
 Can give the advantages and disadvantages of various options on a topical issue. (CA)
 Can pass on a detailed piece of information reliably. (CA)
 Can take part in routine formal discussions conducted in clear standard speech in which factual information is exchanged. (CA)
 Can describe future plans and intentions in detail, giving degrees of probability. (P)
 Can paraphrase in simpler terms what someone else has said. (P)
- 61 Can use a limited number of cohesive devices with some 'jumpiness' in a long contribution. (CA)
 Can engage in extended conversation in a clearly participatory fashion on most general topics. (CA)
 Can respond to clearly expressed questions on a presentation they have given. (CA)
 Can give detailed answers to questions in a face-to-face survey. (P)
 Can express feelings (e.g. sympathy, surprise, interest) with confidence, using a range of expressions. (P)

-
- 62 Can make a note of favourite mistakes and consciously monitor speech for them. (C)
 Can construct a chain of reasoned argument. (C)
 Can describe how to do something, giving detailed instructions. (C)
 Can encourage discussion by inviting others to join in, say what they think, etc. (CA)
-
- 63 Can develop an argument giving reasons in support of or against a particular point of view. (N2000)
 Can give a clear, detailed spoken description of how to carry out a procedure. (C)
 Can describe the personal significance of events and experiences in detail. (CA)
 Can accurately describe a problem with a product or piece of equipment. (P)
-
- 64 Can express views clearly and evaluate hypothetical proposals in informal discussions. (CA)
 Can explain a problem and demand what action should be taken in an appropriate way. (CA)
 Can summarise orally the plot and sequence of events in an extract from a film or play. (CA)
 Can speculate about causes, consequences, hypothetical situations. (N2000)
 Can use stock phrases to gain time and keep the turn whilst formulating what to say. (CA)
 Can plan what is to be said and the means to say it, considering the effect on the recipient. (CA)
 Can make spontaneous announcements clearly and fluently. (CA)
 Can fluently substitute an equivalent term for a word they can't recall. (CA)
-
- 65 Can manage discussion on familiar topics confirming comprehension, inviting others in, etc. (CA)
 Can describe goals using a range of expressions. (P)
-
- 66 Can summarise a wide range of texts, discussing contrasting points and main themes. (CA)
 Can develop a clear argument with supporting subsidiary points and relevant examples. (CA)
 Can give clear, detailed descriptions on a wide range of familiar subjects. (CA)
 Can contribute to a conversation fluently and naturally, provided the topic is not too abstract or complex. (P)
 Can develop an argument well enough to be followed without difficulty most of the time. (C)
 Can give advice on a wide range of subjects. (P)
 Can outline an issue or problem clearly. (CA)
-



GSE 67–75/B2(+): Speaking

-
- 67 Can use a suitable phrase to intervene in a discussion on a familiar topic. (C_A)
 Can adjust to the changes of direction, style and emphasis normally found in conversation. (C)
 Can speculate about the causes of an issue or problem. (C_A)
 Can narrate a story in detail, giving relevant information about feelings and reactions. (P)
-
- 68 Can initiate, maintain and end discourse naturally with effective turn-taking. (C_A)
-
- 69 Can politely avoid answering a question without making it obvious to the listener. (N2007_A)
-
- 70 Can present their ideas with precision and respond to complex lines of argument convincingly. (C_A)
 Can state clearly the limits to a concession. (N2000)
 Can give clear presentations highlighting significant points with relevant supporting detail. (C_A)
 Can lead a discussion in an interview, expanding and developing ideas with little help from the interviewer. (C_A)
 Can compare the advantages and disadvantages of possible approaches and solutions to an issue or problem. (C_A)
-
- 71 Can use a variety of linking words efficiently to mark clearly the relationships between ideas. (N2000)
 Can give well-structured, detailed presentations on a wide range of familiar subjects. (C_A)
-
- 72 Can carry out an effective, fluent interview, spontaneously following up on interesting replies. (C_A)
-
- 73 Can negotiate a solution to a dispute (e.g. an undeserved traffic ticket, blame for an accident). (C_A)
 Can relate their own contribution skilfully to those of other speakers. (C)
 Can use circumlocution and paraphrase to cover gaps in vocabulary and structure. (C)
 Can give detailed technical instructions in their field of specialisation. (P)
-
- 74 Can exchange complex information on a wide range of matters related to their work. (C_A)
 Can make a clear strong argument during a formal discussion. (C_A)
 Can give a detailed response to a counter-argument presented by someone else during a discussion. (N2007_A)
-
- 75 Can express attitudes using linguistically complex language. (P)
 Can depart from and return fluently to a prepared speech in order to answer audience questions. (C_A)
-

GSE 76–84/C1: Speaking

-
- 76 Can substitute an equivalent term for a word they can't recall so smoothly that it isn't noticeable. (N2000_A)
 Can contribute to group discussions even when speech is fast and colloquial. (C_A)
 Can answer questions about abstract topics clearly and in detail. (C_A)
-
- 77 Can rephrase controversial statements into more neutral language. (P)
-
- 78 Can summarise clearly and precisely the arguments and event descriptions from a complex text. (N2007_A)
-
- 79 Can contribute fluently and naturally to a conversation about a complex or abstract topic. (P)
-
- 80 Can participate in a fast-paced conversation with fluent speakers. (C_A)
-
- 81 Can join a conversation already in progress between fluent speakers on complex topics. (C_A)
-

GSE 85–90/C2: Speaking

-
- 87 Can convey finer shades of meaning precisely by accurately using a wide range of modification devices. (N2000_A)
-

Writing

GSE 10–21/Below A1: Writing

- | | |
|----|--|
| 10 | Can write the letters of the alphabet in upper and lower case. (P) |
| | Can write their name, address and nationality. (CA) |
| 18 | Can copy familiar words and short phrases about everyday objects and set phrases. (CA) |
| 20 | Can write consistently with joined-up letters. (P) |

GSE 22–29/A1: Writing

- | | |
|----|---|
| 23 | Can complete simple forms with basic personal details. (CA) |
| 24 | Can copy short sentences on everyday subjects (e.g. directions how to get somewhere). (C) |
| 25 | Can write simple sentences about things that they and other people have. (P) |
| 26 | Can use basic punctuation (e.g. commas, full stops, question marks). (P) |
| 27 | Can spell a range of common names. (P) |
| | Can write simple sentences about their family and where they live. (CA) |
| | Can write simple sentences about personal interests. (P) |
| | Can complete a simple form requiring travel information (e.g. landing card, customs declaration). (P) |
| 28 | Can write short, simple notes, emails and postings to friends. (N2000A) |
| | Can spell a range of common greetings. (P) |
| | Can spell a range of common jobs. (P) |
| | Can write dates using both digits and words. (P) |
| | Can write simple sentences about someone's life and routines. (P) |
| 29 | Can write times using both digits and words. (P) |

GSE 30–35/A2: Writing

- | | |
|----|--|
| 30 | Can write simple sentences about what they and other people do. (CA) |
| 31 | Can write simple sentences about someone's work and duties. (P) |
| | Can use very basic connectors like <i>and</i> , <i>but</i> , <i>so</i> and <i>then</i> . (CA) |
| | Can write a simple description of a room, house or apartment. (P) |
| | Can give personal details in written form in a limited way. (CA) |
| 32 | Can write very short, basic directions. (P) |
| 33 | Can write simple sentences about personal skills. (P) |
| 34 | Can give an example of something in a very simple text using <i>like</i> or <i>for example</i> . (P) |

GSE 36–42/A2(+): Writing

-
- 36 Can make simple comparisons between people, places or things. (P)
Can ask for personal details in written form in a limited way. (CA)
-
- 38 Can write short, simple notes, emails and messages relating to everyday matters. (CA)
Can write short, basic descriptions of places, people or things. (P)
Can write a description of a simple everyday process (e.g. a recipe). (P)
Can write very simple personal emails/letters expressing thanks and apology. (CA)
-
- 39 Can signal contrast in simple text with a limited range of language. (P)
Can write basic instructions with a simple list of points. (P)
Can take a short, simple message if they are able to ask the speaker to repeat and reformulate it. (CA)
Can write short basic descriptions of past events and activities. (P)
-
- 40 Can use appropriate openings and endings in simple informal emails. (P)
Can write a brief summary of their own personal details, e.g. for a business card. (P)
-
- 41 Can write a basic informal email/letter of invitation with simple, key details. (P)
Can write a basic email/letter accepting or declining an invitation. (P)
Can write about everyday things (e.g. people, places, job, study) in linked sentences. (CA)
Can write simple sentences about their educational background and present or past job. (CA)
-

GSE 43–50/B1: Writing

-
- 44 Can clearly signal the end of a simple narrative or description. (P)
Can write a basic description of experiences, feelings and reactions, given a model. (P)
-
- 45 Can make simple, logical paragraph breaks in a longer text. (P)
Can write a short, simple description of a familiar device or product. (P)
Can write an email/letter congratulating someone on something. (P)
-
- 46 Can show a basic direct relationship between a simple problem and a solution. (P)
Can write a basic formal email/letter requesting information. (P)
Can write a description of a future event or activity. (P)
Can write the concluding sentence or sentences of a basic paragraph, given a model. (P)
Can write simple instructions on how to use a device or product, given a model. (P)
Can write simple informal emails/letters and online postings giving news or opinions. (P)
Can write short, simple essays with basic structure on familiar topics. (P)
-
- 47 Can write about personal interests in some detail. (CA)
Can write descriptions of past events, activities, or personal experiences. (P)
Can write an everyday connected text using a set of short elements or facts and building them into a sequence. (CA)
-
- 48 Can write personal emails/letters giving some details of events, experiences and feelings. (CA)
Can write a basic letter of application with limited supporting details. (P)
Can complete a detailed form requiring travel information (e.g. visa application). (P)
-
- 49 Can use simple cohesive devices such as repetition in a structured text. (P)
-

-
- 50 Can write fractions using both digits and words. (P)
-
- Can write about experiences, feelings and reactions in a simple connected text. (CA)
-
- Can write a basic summary of a simple text using the original wording and paragraph order. (CA)
-
- Can write personal emails/letters giving advice. (P)
-

GSE 51–58/B1(+): Writing

- 51 Can use limited discourse devices to link sentences smoothly into connected discourse. (CA)
-
- Can write a basic email/letter of complaint requesting action. (P)
-
- Can use common connectors to tell a story or describe an event in writing. (CA)
-
- Can write short, simple biographies about real or imaginary people. (CA)
-
- Can complete a form requiring educational information (e.g. applying for a course of study). (P)
-
- Can write an email/letter sending a message of sympathy. (P)
-
- 52 Can summarise the main message from simple diagrams (e.g. graphs, bar charts). (P)
-
- Can clearly signal chronological sequence in narrative text. (P)
-
- Can write a simple review of a film, book or TV programme using a limited range of language. (P)
-
- 53 Can take messages, communicate enquiries and explain problems. (CA)
-
- Can write emails/letters exchanging information, emphasising the most important point. (CA)
-
- Can give someone clear, detailed directions on how to get somewhere in a letter, email, or online posting. (P)
-
- Can write a description of a real or imagined event (e.g. a recent trip). (C)
-
- 54 Can write a formal email/letter requesting information. (P)
-
- Can complete a form requiring health information. (P)
-
- Can respond to and comment on other people's personal updates on a social media website. (P)
-
- 55 Can write a formal email/letter accepting or declining an invitation. (P)
-
- Can take notes of key points during a talk on a familiar topic, if delivered clearly. (CA)
-
- Can complete a form requiring financial information (e.g. application for a bank account or credit agreement). (P)
-
- Can write personal updates on a social media website using an appropriate style. (P)
-
- Can post comments on the discussion board of a website. (P)
-
- Can vary the formality of greetings in emails/letters based on intended recipients. (P)
-
- 56 Can introduce a counter-argument in a simple discursive text using *however*. (P)
-
- Can write instructions on how to use a device or product. (P)
-
- 57 Can summarise factual information within their field of interest. (CA)
-
- Can write a description of items for sale on a trading website. (P)
-
- Can write emails/letters of complaint with supporting details. (P)
-
- 58 Can write instructions on how to look after an object, device or product. (P)
-
- Can write emails/letters responding to personal news and views in detail. (CA)
-
- Can collate short pieces of information and summarise them for somebody else. (CA)
-
- Can write a description of a problem with a product or piece of equipment. (P)
-
- Can write comments and complaints about products and services. (P)
-

GSE 59–66/B2: Writing

-
- 59 Can write detailed descriptions of real or imaginary people. (P)

 Can demonstrate understanding of formality and conventions in standard letters. (P)

 Can comment on factual information within their field of interest. (C_A)

 Can write a formal email/letter of thanks or apology with appropriate conventions. (P)

 Can support ideas with relevant examples. (P)

 Can write a detailed description of an object, device or product. (P)
-
- 60 Can write a formal email/letter of invitation with appropriate register and conventions. (P)

 Can clearly signal cause and effect relationships in structured text. (P)

 Can write personal emails/letters giving and commenting on news in detail. (C_A)
-
- 61 Can reformulate an idea in different words to emphasise or explain a point. (P)

 Can end a discursive argument with a clear conclusion and opinion. (P)

 Can write instructions on how to repair an object, device or product. (P)

 Can clearly signal the difference between fact and opinion in structured text. (P)

 Can write a brief standard report conveying factual information, stating reasons for actions. (C_A)
-
- 62 Can signal additional information in a formal structured text with a range of language. (P)

 Can write a structured text clearly signalling main points and supporting details. (P)

 Can clearly signal problem and solution relationships in structured text. (P)

 Can write informal emails or letters using a range of idiomatic phrases. (P)
-
- 63 Can write detailed descriptions of real or imaginary places. (P)
-
- 64 Can write a concise summary of the main ideas of a longer structured text. (P)

 Can write personal emails/letters about abstract or cultural topics (e.g. music, films). (C_A)
-
- 65 Can systematically evaluate the advantages and disadvantages of various options. (P)

 Can structure longer texts in clear, logical paragraphs. (P)

 Can give a structured written explanation of a problem. (C_A)

 Can write a letter of complaint with appropriate register, structure and conventions. (P)

 Can develop a clear written description or narrative with relevant supporting detail and examples. (C_A)

 Can write a letter of application with appropriate register, conventions and supporting detail. (P)

 Can express news and views effectively in writing and relate to those of others. (C)
-
- 66 Can show the relationship between an opinion and a counter argument in a discursive text. (P)

 Can compare and evaluate ideas in a structured and logical text. (P)

 Can adopt a level of formality appropriate to the circumstances. (N2000)

 Can write an accurate summary of an essay or article on a familiar topic. (P)

 Can write engaging headlines or titles to capture a reader's attention. (P)
-

GSE 67–75/B2(+): Writing

-
- | | |
|-------|---|
| 67 | Can write relevant subheadings to structure longer more complex texts. (P) |
| | Can systematically develop an argument giving the reasons for or against a point of view. (P) |
| | Can write about feelings and the personal significance of experiences in detail. (CA) |
| | Can write clear, detailed descriptions on a variety of subjects related to their field of interest. (C) |
| | Can write a forceful but polite letter of complaint, including supporting details and a statement of the desired outcome. (P) |
| | Can synthesise and evaluate familiar information and arguments from a number of sources. (CA) |
| <hr/> | |
| 68 | Can seek confirmation of information and ideas on abstract and concrete topics. (CA) |
| | Can respond in writing to other people's arguments in an appropriate style. (N2007A) |
| | Can write clear and precise emails intended to create rapport and put the addressee at ease. (CA) |
| <hr/> | |
| 69 | Can write a structured review of a film, book or play with some references and examples. (CA) |
| <hr/> | |
| 70 | Can express the same idea with a different level of formality appropriate to the audience. (P) |
| | Can convey information and ideas on abstract and concrete topics. (CA) |
| <hr/> | |
| 71 | Can demonstrate understanding of structure and conventions of different written genres. (CA) |
| <hr/> | |
| 72 | Can check and correct spelling, punctuation and grammar mistakes in long written texts. (P) |
| | Can write a detailed, reasoned argument for or against a case. (N2007A) |
| <hr/> | |
| 74 | Can structure longer complex texts using a range of cohesive devices. (P) |
| | Can write a detailed description of a complex process. (P) |
| <hr/> | |
| 75 | Can develop an argument with appropriate highlighting of significant points and relevant supporting detail. (CA) |
-

GSE 76–84/C1: Writing

-
- | | |
|-------|---|
| 76 | Can write a clear summary of a complex factual text, maintaining its original tone and message. (P) |
| | Can make jokes in writing using words with similar spelling but different meanings. (P) |
| <hr/> | |
| 77 | Can write an accurate summary of a complex, discursive text. (P) |
| | Can write a detailed critical review of cultural events (e.g. plays, films, concerts) or literary works. (P) |
| | Can express themselves fluently in writing, adapting the level of formality to the context. (P) |
| <hr/> | |
| 81 | Can confidently argue a case in writing, specifying needs and objectives precisely and justifying them as necessary. (N2007A) |
| <hr/> | |
| 82 | Can smoothly switch between a range of writing styles to address specific audiences and topics in a personal way. (P) |
-

GSE 85–90/C2: Writing

-
- | | |
|----|---|
| 85 | Can create well-structured complex texts with underlying inferred meaning. (CA) |
|----|---|
-

References

- British Council/EAQUALS (2010) *Core Inventory for General English*. London, British Council and EAQUALS. <http://www.teachingenglish.org.uk/article/british-council-eaquals-core-inventory-general-english-0>
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Council of Europe (2009) *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR): A Manual*. Strasbourg: Council of Europe; Language Policy Division. http://www.coe.int/t/dg4/linguistic/Source/ManualRevision-proofread-FINAL_en.doc.
- North, B. (2000) *The Development of a Common Framework Scale of Language Proficiency*. New York: Peter Lang.
- North, B. (Ed) (2007) *Expanded set of C1 and C2 descriptors*. Strasbourg: Council of Europe; Language Policy Division. http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key_reference/C1C2descriptors_EN.pdf
- Trim, J. (2009) *Breakthrough*. Strasbourg, Council of Europe. Available on www.ealta.eu.org

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the Publishers.

©Pearson Education Limited 2015

© The copyright of the descriptive scales and the illustrative scales (in all languages) reproduced in this document belongs to the Council of Europe.

All users and publishers must ask formal and written permission prior to using these by writing to the Language Policy Unit of the Council of Europe (language.policy@coe.int)

Photography

Christof Van Der Walt: Cover, p2, p13

Somenath Mukhopadhyay: p23

Personalised learning

We have developed learning objectives tailored to meet the needs of four specific audiences: Adult Learners of General English, Learners of Academic English, Learners of Professional English and Young Learners (aged 6-14). These sets of learning objectives take account of the different goals across all four skills for each group of learners.

Visit **English.com/gse** for more information about the Global Scale of English and to download the four sets of GSE Learning Objectives.



Interested in taking part in our research?

The research project to create and calibrate learning objectives for different groups of learners – Adults, Young Learners, learners of Professional and Academic English – is ongoing. We would love to hear from experienced teachers who are interested in getting involved.

Do you have a couple of hours a month to rate the difficulty level of new learning objectives? If yes, then please contact us for more information at: **rating@pearson.com**

We look forward to hearing from you!

The Pearson English Team

APPENDIX B

Pearson SEE mbo Progress Technical Report

Pearson SEE mbo is een gecustomiseerde versie van Progress. Alle informatie in het Progress Technical Report en Progress Test Information Booklet is van toepassing op Pearson SEE mbo.



Technical Report

September 2014

V.1.0

Global Scale of English 10-58

Preface: Purpose of this report

The Progress technical report is written with the purpose of describing to stakeholders the background, development and technical aspects of the test. The report also provides preliminary validation evidence gathered during field testing.

It is presumed that the intended readership will have a basic knowledge of statistics and testing terminology, although a glossary is included for reference. For other stakeholders, a range of other information about the test - including the test information booklet and a sample test - is available on our website <http://product.pearsonelt.com/progress>.

Test validation is an ongoing process and the studies described in this document are not regarded as final. Test validation continues with a growing programme of research carried out each year, both by Pearson staff and independent researchers/teachers. For more information about our research programme, please visit our website.

Contents

1. Introduction	5
2. Test Description	7
3. Content Design and Development	12
4. Score Reporting	15
5. Field Testing	19
6. Psychometric Analysis and Validation Studies	29
7. Summary	36
8. Glossary	37
9. References	38

Executive Summary

Progress is a four skills test of English Language proficiency designed to support English language learners' learning by measuring their learning progress on the Global Scale of English (GSE) scale of 10 to 90. The test and the reporting scale are aligned to the Common European Framework of Reference for languages (CEFR). *Progress* uniquely combines a computer-adaptive part and non-adaptive part in one test. Furthermore, *Progress* is a fully automatically scored test including productive speaking and writing skills. This ensures that the student's proficiency level and progress is measured accurately and efficiently. *Progress'* score reporting includes the Overall score as well as six sub-scores: Reading, Writing, Speaking, Listening, Grammar and Vocabulary.

A total of 4,327 learners with 63 different first languages participated in field testing a prototype version of the test. Their performance data were used for calibrating items and for developing automated scoring systems. During the development phase, a set of 299 test-takers' data were set aside for conducting independent validation analyses and a series of preliminary analyses were run. The test reliability is estimated at 0.91, suggesting that the test is very reliable. A correlation coefficient of 0.94 was observed for the Overall score between the automatically generated scores and the skilled human raters. Similarly, for the sub-scores, correlation coefficients of 0.78 to 0.99 were observed. The data, therefore, suggest that *Progress* test scores are highly correlated with human scores and can be an effective tool for assessing learners' English proficiency.

1. Introduction

1.1 Global Scale of English

The Global Scale of English (GSE) is a granular scale from 10 to 90 that is designed to report English language proficiency. It has been empirically validated and enhances the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001) by showing finer gradations of a learner's level and progress within a CEFR band.

The Global Scale of English is currently used to report scores on the internationally recognised English language test, PTE Academic™. Validation evidence has been collected to demonstrate the alignment claim to the CEFR and the correlation to other test score scales such as Versant, TOEFL® iBT, TOEIC® and IELTS.



Figure 1 The Global Scale of English (GSE)

1.2 Progress Product Concept

1.2.1 Test description

Progress is an online test of English language proficiency designed to measure progress for individuals and groups of learners. The test is delivered through the MyEnglishLab platform and is normally combined with a learning program, to be taken at the start, middle and end of a course. It assesses reading, writing, listening, speaking, vocabulary and grammar. The test results include an overall score and language skill scores that are reported on the GSE from 10 to 90. The initial three levels, available from 2014, are designed for students in the range of A1 to mid B1 which is in the region of 22 to 50 on the GSE.

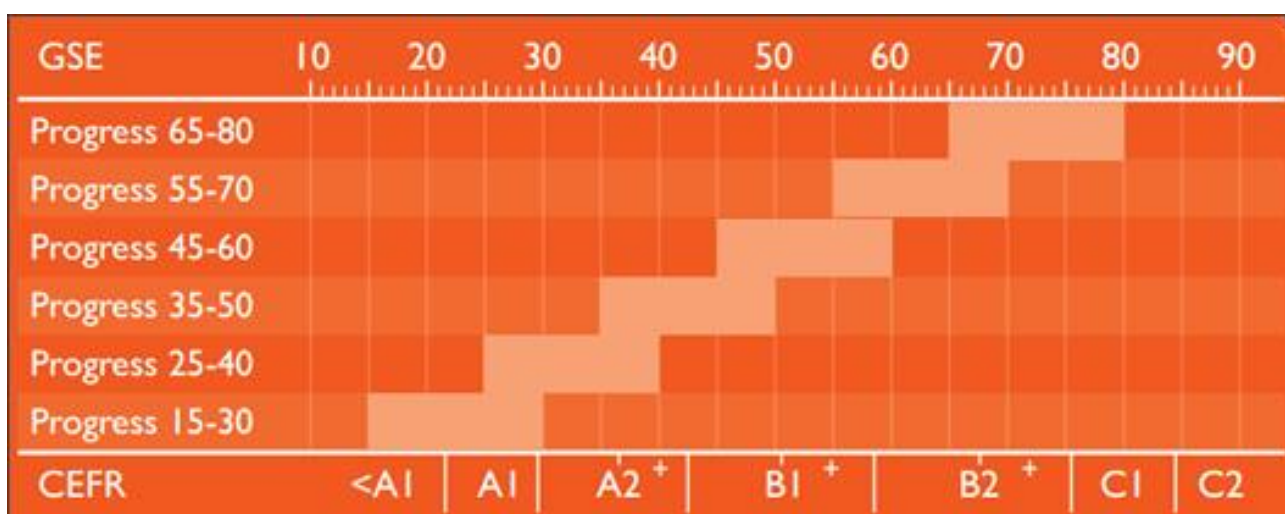


Figure 2 Progress levels mapped to the GSE and CEFR.

Progress is an automatically scored test, comprised of two parts. The first part of the test is computer adaptive to obtain an initial estimate of the learner's language ability. This estimate is then used to choose the second part of the test from among many fixed forms that have been tailored at different GSE scores. This partially adaptive approach allows for selecting a form that is appropriate for the learner's proficiency level, enabling a fine-grained measurement necessary for tracking the learner's progress. The adaptive design of the test makes measurement more efficient in terms of covering a range of ability and keeping the total test time short.

The first part of the test utilises mostly multiple choice item types and the second part of the test assesses productive language skills using item types that integrate multiple language skills (e.g., listen to an item and then respond verbally).

Items from both the adaptive part of the test and the linear section contribute to the learner's overall score. The scores for particular skills are based on the questions which are designed to assess those particular skills and knowledge.

1.2.2 Theoretical Rationale of the Intended Test Purpose

The conceptual framework for *Progress* is Assessment for Learning – the idea that testing should inform and support teaching, and also be informed and supported by teaching, in a continuous feedback loop. There is considerable academic evidence for the potential of Assessment for Learning to not only demonstrate outcomes but actually improve them. For example,

- In 1998 the Nuffield Foundation commissioned Professors Paul Black and Dylan William to evaluate the evidence from more than 250 studies linking assessment and learning. Their analysis was definitive - *initiatives designed to enhance effectiveness of the way assessment is used to promote learning in the classroom can raise pupil achievement. Learning outcomes demonstrated that the linking of assessment with instruction can deliver improved student achievement at a greater pace than many other interventions.* (Black and William, 1999)
- Teachers who are supported to collect and analyse data in order to reflect on their practice are more likely to make improvements as they learn new skills and practice them in the classroom. Through the evaluation process, teachers learn to examine their teaching, reflect on practice and evaluate their results based on student achievement. (Speck and Knipe, 2001)
- Teachers can adapt instruction on the basis of evidence, reflect and make changes that will yield immediate benefits to student learning. (Stiggins, Arter, Chappuis and Chappuis, 2006)

Following the Assessment for Learning concept, *Progress* is intended to be used in the context of learning in conjunction with learning materials and formative assessment tasks. The test content is, therefore, designed to relate to the common elements which are found in learning materials. Content is task-based with a focus on the assessment of the development of reading, listening, speaking and writing skills, and the structural elements of language, namely grammar and vocabulary. The test content was guided by Pearson's structured syllabus. The syllabus provided a detailed description of learning objectives at each level and these were supplemented by documents which described functional language areas, language exponents and appropriate lexis for each CEFR level.

2. Test Description

2.1 Test Design

Each administration of *Progress* involves a two-stage test:

- **Stage 1** is the adaptive part of the test, directed at the general level of the student's course of study. Stage 1 is designed to collect data for scores on vocabulary, grammar and reading. The items are mostly multiple choice items with a specific assessment focus on grammatical structures, lexis and reading content appropriate for the level of the student.
- **Stage 2** consists of items to test listening, speaking and writing as well as integrated skills items which test more than one skill. This part of the test is not adaptive; however, responses including productive language skills such as speaking and writing are automatically scored by Pearson's computerised automated scoring system. There are a number of equivalent forms across the levels to ensure that learners will not repeat exactly the same test content when they take *Progress* for a second or third time.

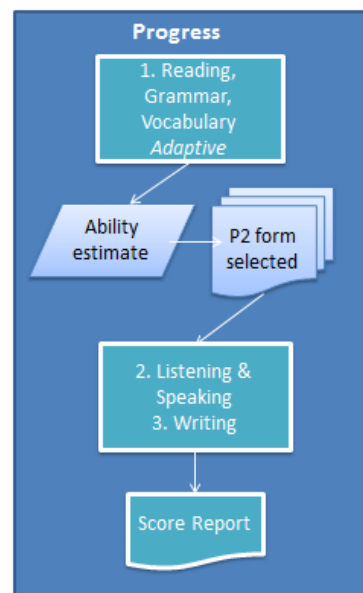


Figure 3 Progress test design

The estimated ability from Stage 1 is used to choose the most appropriate Stage 2 form. This means that the test as a whole is well targeted to the learner's level.

An overview of the item types in each stage is provided in Table 1 below. The overall score is calculated from all the items on the test. The skill scores and scores for grammar and vocabulary are calculated based on the items which focus on that skill or knowledge. The scores are calculated independently; they are not averages.

Table 1 Overview of Progress Item Types

Stage	Item Type	Description	Skills Measured	Testing Focus
Stage 1	Fill in the Table	Complete missing information in a table by typing into the gaps.	Vocabulary	Lexical sets, lexical knowledge including accurate production of the word form.
Stage 1	Choose the Right Word or Phrase	Choose the correct answer from a drop-down list to fill gaps in sentences.	Vocabulary or Grammar	a) Vocabulary knowledge in a written context. The ability to use a word or phrase appropriately. It also tests vocabulary range. b) Range and accuracy of grammatical knowledge in a written context.
Stage 1	Choose the Right Word or Phrase. You may choose more than one.	Choose the correct answer(s) to multiple choice questions.	Grammar	Words which are related to each other in that they have similar meanings or grammatical uses. It tests grammatical knowledge in a written context.
Stage 1	Complete the Dialogue	Complete the dialogue by dragging words into gaps.	Vocabulary	Vocabulary in a spoken context. Knowledge and understanding of particular vocabulary in the context of a conversation. It also tests vocabulary range.
Stage 1	Drag and Drop	Drag and drop words to re-order a set of 3 scrambled sentences.	Grammar	Word order, connectors and discourse markers. It tests grammatical knowledge in a written context at the sentence level.
Stage 1	Choose the Right Word or Phrase	Read a short text and answer multiple choice questions about the text.	Reading	Global understanding of short messages, notes and short pieces of writing.
Stage 1	Choose the right picture	Read a short text and answer multiple choice questions about the text by selecting the correct image.	Reading	Global understanding of short messages, notes and short pieces of writing.
Stage 1	Short Answer	Read a short text and answer 4 short answer questions about the text in writing.	Reading	Reading comprehension. It tests understanding of specific information included in the text.

Stage 2	Read and then Speak	Read a short passage aloud.	Reading, Speaking	Oral reading fluency. It tests the ability to decode and articulate words and sentences accurately and fluently.
Stage 2	Listen and then Speak	Listen to a sentence and repeat it.	Speaking, Listening	Listening comprehension at the word and sentence level. It tests pronunciation and fluency. It tests if the words heard are understood and repeated accurately.
Stage 2	Describe Image	Look at an image and describe it out loud.	Speaking, Vocabulary	The content of the talk as well as pronunciation, stress and fluency. It tests the use of appropriate words to describe the photograph or picture.
Stage 2	Listen to the Conversation and Answer	Listen to a conversation between two speakers followed by a question and give a verbal response to the question in one or a few words.	Listening	Understanding of the meaning, context and/or detail in a short dialogue.
Stage 2	Listen and Read	Listen to an audio recording and see its transcript. Click on the words in the transcript that do not match what is heard.	Listening, Reading	Reading and listening comprehension. It tests the ability to recognise individual words in a text.
Stage 2	Listen and then Write	Listen to a sentence and type it.	Listening, Writing	Listening comprehension at the word and sentence level. It tests the ability to write accurately and understand sentence structure, word order and connectors.
Stage 2	Read and then Write	Read a short passage for 30 seconds. After 30 seconds, the passage disappears. Reconstruct the passage in writing.	Reading, Writing, Vocabulary, Grammar	Reading comprehension. It also tests the ability to write accurately and understand sentence structure, word order and connectors.
Stage 2	Short Essay	Look at an image and describe it in writing OR Read a short prompt and write a response to the prompt.	Writing, Vocabulary, Grammar <i>(to be enabled after successful evaluation in BETA)</i>	Overall writing skills. Paragraph and sentence structure, the range and accuracy of language, the ability to structure an argument or discussion in a written context. It tests grammar and vocabulary as an essential part of writing.

2.2 Test Administration

The test is administered on the MyEnglishLab platform either as a standalone test which can be used with any course or bundled with a specific Pearson course. The test is web based and can be taken with or without a proctor. The course leader (in most cases the classroom teacher) has a number of options to control how and when the test is taken. For example, the course leader can choose whether the test should be timed or untimed, so that they can arrange for different levels of supervision depending on students' needs.

The test is presented to students in three sections. If enabled by the course leader, a timer is displayed for each section. The first section is equivalent to Stage 1 (described above) and comprises the adaptive part of the test. Sections 2 and 3 correspond to Stage 2 (described above). Section 2 contains the listening and speaking items. Section 3 contains the writing items.

2.3 Test Construct

Language proficiency is seen as the ability to use language as a social tool to achieve certain ends. It is communicative and interactive. The approach which underpins the development of this test embraces the concept of communicative competence following Widdowson (1978) who distinguishes between use and usage. The terms used by to refer to two aspects of communicative performance:

- a) the ability to produce correct sentences, or manifestations of the linguistic system = USAGE
- b) the ability to use the knowledge of the rules for effective communication = USE

The concept of use is directly related to communicative language teaching and an action orientated framework such as the CEFR. Usage relates to the underlying structure of the language which can be codified in different ways, for example grammar and vocabulary. *Progress* items which have been written with specific skills in mind are coded with appropriate 'can do' statements from the CEFR and Pearson's syllabus. This maintains the theoretical link with the framework and the underlying communicative approach. Vocabulary and grammar items have been coded using functional and notional descriptions of language which are associated with the CEFR and as such helps define appropriate exponents at different levels of communicative performance. Through its development *Progress* has drawn extensively from descriptive frameworks of the Common European Framework of Reference for Languages; Learning, teaching, assessment (2001) and related publications such as Threshold (1990a) and Waystage (1990b).

The model recognises the advances in understanding of language learning in that it is interactive and iterative; that learners learn primarily by doing activities either in the classroom or in the real world. In this respect the model recognises the task-based nature of language learning in developing language proficiency and this is reflected in what learners are asked to do in the test.

In *Progress* learners are presented with a number of tasks (items). These are in the main familiar items that learners will have met in the classroom. They are based on similar exercises which learners are asked to complete as part of their learning and in review and remedial studies. There are multiple task types for grammar, vocabulary, listening, reading speaking and writing. These tasks measure the skill directly or through specific traits which are related to that skill. As such the test construct recognises that

there are multiple skills or traits which contribute to language proficiency and that these can be represented as different composites which can be represented by different scores.

In *Progress* a number of the test items are integrated. This means they are designed to test multiple traits or skills.

3. Content Design and Development

3.1 Alignment to the CEFR & GSE

Progress is designed to be aligned to the CEFR and GSE. The Common European Framework of Reference for Languages (CEFR) has gained currency in Europe and beyond as a standard for defining, comparing and equating levels of language competence. There is no central accreditation scheme which test developers can invoke in order to seek endorsement of their claims to be aligned to the CEFR. However, the Council of Europe has published a detailed manual (Council of Europe, 2009) which recommends procedures test developers can apply in order to substantiate such claims. These procedures are grouped into five stages:

- Familiarisation: establishing procedures to ensure that personnel engaged in the linking process are well informed about the CEFR and have a well developed awareness of its levels
- Specification: assessing the coverage of the test as set out in test specifications in relation to the descriptions of language competence found in the CEFR
- Standardisation training and benchmarking: ensuring that personnel engaged in determining the level of benchmark samples responses have a sound common understanding of the CEFR reference levels
- Standard setting: determining the cut scores which determine how test takers are allocated to CEFR levels on the basis of their test performance
- Validation: collecting empirical evidence, such as statistical data, as to how far test scores reflect true performance at the CEFR levels.

The above procedures were implemented at the appropriate junctures in the process of developing *Progress*. During the item development phase, several steps were taken:

- Item writers were asked to consider the CEFR as the construct model for the test design
- Item Writer Guidelines specified in detail the characteristics of each item and gave item writers rules and checklists to ensure that a high proportion of their items were fit for purpose and suitable for inclusion in the item bank
- Item writers received training in understanding, interpreting and using the CEFR and GSE (detailed below)
- Item writers were asked to provide a target CEFR level of A1 to B2 for each item submitted. These pre-estimates were stored in the item bank. Their predictions of level were validated or updated after the items were field-tested
- Item writers were asked to assign one or more learning objectives to each item they submitted using 'can do' statements taken from the CEFR and Pearson syllabus
- The topics, vocabulary and grammar used in test items is based on the CEFR Functions and Notions documents (van Ek & Trim 1990a, 1990b) and represent the ranges of functions and notions that learners should be able to use to function effectively and efficiently at a particular CEFR level. They include themes, specific and general notions and language functions for different CEFR levels.

3.2 Item Development

Quality assurance measures were adopted at each stage of item development starting with initial item writer recruitment and training, including several rounds of review stages by Test Development staff and external experts and ending in feedback and evaluation. In general, the language structures used in the test were designed to reflect those that are common in spoken and written English. The items were designed to be independent of social and cultural nuance, and high-cognitive functions.

Item writer recruitment and training

Item writers were recruited from the USA, the UK, Hong Kong and Australia to ensure an international perspective. They were required to meet minimum qualification requirements: L1 English speaker or L1-like proficiency in English, and a bachelor's or higher degree in applied linguistics, English language and literature, education, or a closely related field. All Progress item writers had previous experience in item writing and in teaching ESL or EFL.

Successful applicants attended a training event run by Pearson test developers. The training covered familiarisation with the *Progress* test specification, item writer guidelines and authoring tools, and the CEFR.

Item Banking

An authoring tool was used, allowing texts, questions and media to be uploaded to the item bank automatically using standardised templates and formats. All the stages of item development were controlled and monitored; all access to the item bank was audited and all comments and decisions on items were tracked and stored. Information from the item bank software was also fundamental to the feedback given to item writers, described below.

Item Writing and Review Process

Item writers were commissioned to write sets of items. The item writers followed detailed Item Writer Guidelines and checklists to ensure that a high proportion of items were accepted.

The item content review process is considered another vital part of validation. Although the item writer guidelines highlight the concerns of construct irrelevance and construct under-representation, both may still exist as a result of, for example, the writers' different cultural backgrounds, genders, ages, or religious beliefs. The target test-taking population of *Progress* is heterogeneous, may speak a wide variety of first languages, and have a wide variety of cultural and social backgrounds. If the content of a task upsets certain groups of test takers, it may affect their test performance. Validity checks by external content reviewers and Pearson test developers were conducted to evaluate the appropriateness of *Progress* item content and to eliminate potential bias. On occasion, items were rejected when the following issues were flagged and could not be resolved:

- *Content*: when items demonstrated bias and sensitivity issues or when they failed to meet test specifications;
- *Knowledge*: When answers could be obtained based on general knowledge without context, required special knowledge, or were of an inappropriate genre;
- *CEFR*: when stimulus or questions were out of test range or unsuitable for that level

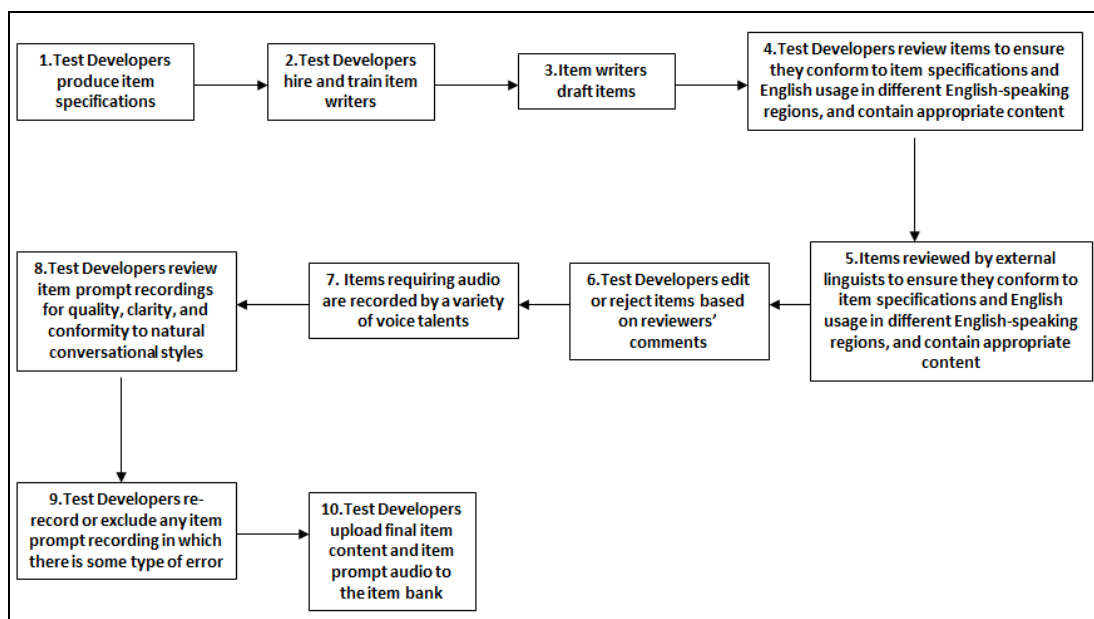


Figure 4 The item development process

Quality control for item writers continued throughout the process. Individual and team performance for both item writing and item reviewing was closely monitored by test development staff. Detailed findings of these performance reviews were fed back to writers at the end of the writing session and used to improve training and development.

Item Prompt Recording and Review

A total of fourteen native speakers (7 men and 7 women) representing the US, UK, and Australia were selected for recording the spoken prompt materials. Of the spoken prompt materials, 40% were recorded by the US speakers, 40% by the UK speakers, and 20% by the Australian speakers.

The comprehension questions were recorded by a professional US voice talent whose voice was distinct from the voices used on the prompt. The prompts were recorded by L1 English speakers who were not professional voice talents; this was to ensure that their speaking styles represented a range of natural speaking patterns that learners would encounter in conversational contexts in English.

The speakers were instructed to record items in the same way as they normally speak in English. That is, the speakers were not asked to change their pronunciation (e.g., no special enunciation or over-articulation) or their style of delivery. This helped to ensure that the types of speech the test takers encounter during the test are representative of characteristics in real-life conversation.

To ensure intelligibility and acceptability of the speakers, all recordings were vetted by Pearson test developers for quality, clarity, and conformity to natural conversational styles. Any recording in which reviewers noted some type of error was either re-recorded or excluded.

4. Score Reporting

4.1 The Progress Score Report

The *Progress* Score Report consists of an overall score, and sub-scores for communicative and enabling skills. All scores are reported on the Global Scale of English (GSE) which ranges from 10 to 90. Scores can also be displayed on the CEFR.

Overall score: This reflects a test taker's overall English language ability. The score is based on performance on all items in the test. Each test consists on average of between 60-65 items across 17 item types.

Skill scores (listening, reading, speaking, writing, vocabulary and grammar): These scores are based on test items that assess those particular skills, either as a single skill or integrated skill tasks. For vocabulary, the score relates to knowledge and use of appropriate choice of words used to express meaning, as well as lexical range. For grammar, the score relates to knowledge and use of appropriate language with respect to word form and word order.

Progress is normally combined with a learning program, to be taken at the start, middle and end of a course. Scores from each of the three test administrations are displayed alongside each other to quickly see in which areas students have made progress and where their strengths and weaknesses lie.

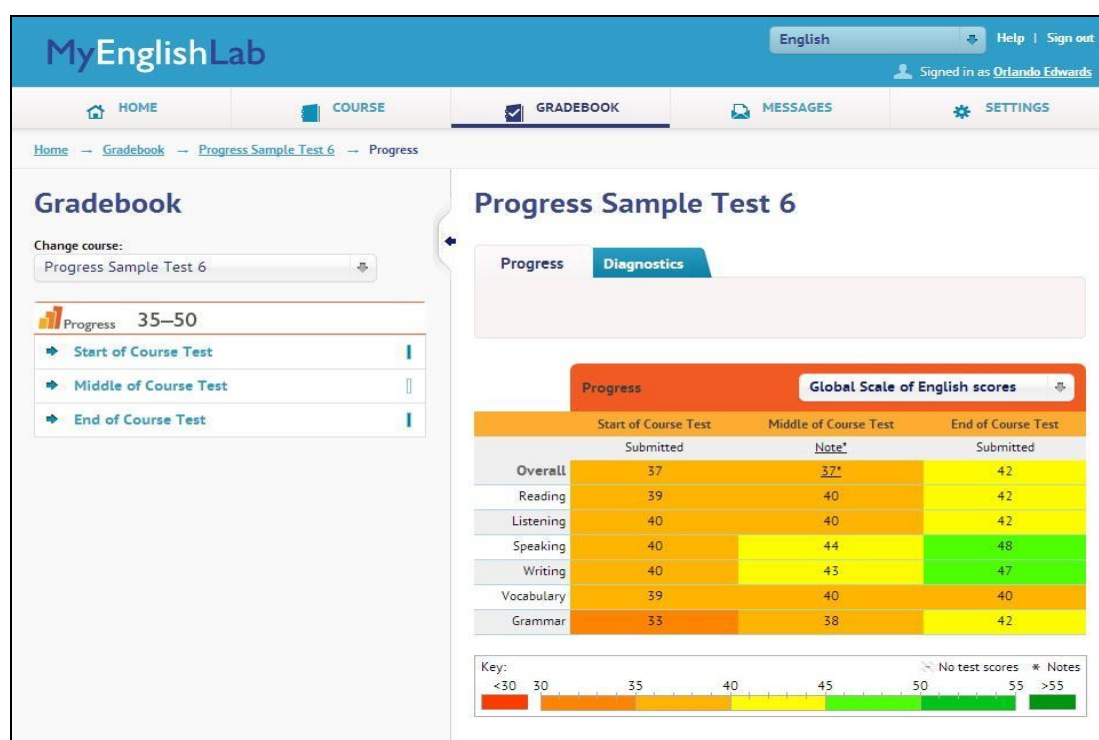


Figure 5 Sample Progress Score Report for a single student

The item types contributing to each of the reported scores are shown in Table 2.

Table 2 Item types contributing to each of the Progress reported scores

Item Type	Overall Score	Skill Scores	Enabling Scores
Listen to the Conversation and Answer*	Y	Listening	
Describe Image*	Y	Speaking	Vocabulary
Short Essay* (scoring to be enabled after successful evaluation in BETA)	Y	Writing	Vocabulary, Grammar
Read and then Speak*	Y	Speaking Reading	
Listen and then Write	Y	Listening Writing	
Listen and then Speak*	Y	Listening Speaking	
Read and then Write*	Y	Reading Writing	Vocabulary Grammar
Listen and Read	Y	Listening Reading	
Choose the Right Picture	Y	Reading	
Choose the Right Word or Phrase	Y	Reading	
Short Answer	Y	Reading	
Fill in the Table	Y		Vocabulary
Choose the Right Word or Phrase	Y		Vocabulary
Complete the Dialogue	Y		Vocabulary
Choose the Right Word or Phrase	Y		Grammar
Choose the Right Word or Phrase. You may choose more than one.	Y		Grammar
Drag and Drop	Y		Grammar

* automated scoring of speaking/writing

Progress uses automated scoring for six of the speaking and writing item types (marked with an asterisk in Table 2). The automated scoring engines were trained using large

samples of learners' and English L1 speakers' answers which were collected in the Field Test and marked by experienced and well-trained examiners. Scores produced by the automated scoring engines were further validated by using separate sets of learners' responses that were not used to train the system. The process used to train and validate the scoring engines is described in detail in Section 5.

4.2 Score Use

Progress is a formative assessment instrument and it is not expected to be taken at a secure test center. As such, the intended score use is not for certification of the student's English proficiency level; it is for supporting the student's learning.

Progress is designed as a test to be taken at the start, middle and end of a course of study. It is designed to accurately measure the progress made by a student over the course duration in order to focus teaching and learning. Once a candidate has completed a test, scores are reported in the gradebook. Teachers and students can then view their scores from the most recent test along with scores from their earlier sittings.

Scores from the test given at the start and middle of a course can be used to understand the strengths and weaknesses of a student or class in order to better target their individual learning needs. Tests at the middle and end of a course demonstrate how much progression is being made.

The overall score is the key indicator of a student's English proficiency. The skill scores give an indication of how well the student did on questions aimed at that particular skill. The skill scores can be used by the teacher in conjunction with the 'Can-Do dashboard' (see Figure 6 below). This diagnostic tool identifies particular learning objectives which the student could work on. Teachers will also use other information they have about the student, performance in class etc., to further hone and make relevant the information from the dashboards.

The screenshot displays the 'MyEnglishLab' interface. At the top, there's a blue header with the logo, a language dropdown set to 'English', and links for 'Help' and 'Sign out'. Below the header is a navigation bar with icons and labels for 'HOME', 'COURSE', 'GRADEBOOK' (which is active), 'MESSAGES', and 'SETTINGS'. A breadcrumb trail shows the path: 'Home → Gradebook → 15-30 → Diagnostics → Progress Scores'. The main content area is titled '15-30' and includes three buttons: 'Return to diagnostic dashboard', 'Return', and 'Print'. Below these is the heading 'Progress score/Skill/Overall Can Do Statements' with a subtext: 'This shows a list of Learning Objectives you need to work on.' A list of ten learning objectives follows, each preceded by a radio button. To the right of the list is a 'Filter by:' box with three radio button options: 'End of Course Test' (selected), 'Middle of Course Test', and 'Start of Course Test'.

Figure 6 Illustration of the 'Can-Do dashboard' from the student gradebook

4.3 Score Interpretation

The table below offers a reference to interpret GSE scores in terms of the CEFR level descriptors¹.

Table 3 Main features of a learner's ability at different GSE/CEFR levels.

Global Scale of English (GSE)	Common European Framework (CEFR) Level	Main features of a learner's language ability at this level (<i>Council of Europe, 2001</i>)
22 - 29	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.
30 - 42	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
43 - 58	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst in an area where the language is spoken. Can produce simple connected text on topics, which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.

¹ © 2001 The copyright of the level descriptors reproduced in this document belongs to the Council of Europe.

5. Field Testing

As part of the test development process, a large field test was carried out between June and September 2013. The purpose of the field test was to ascertain the appropriateness of the pool of items to serve as a source for constructing individual test forms which would allow reliable predictions of students' ability in English.

The goals were to:

- Gather the necessary data to calibrate the *Progress* item bank on the GSE;
- Gather the necessary data required to finalise the adaptive scoring algorithm for Stage 1 and the linear design for Stage 2;
- Gather the necessary data to train the automated scoring engines;
- Finalise the operational test format, e.g. number of items, timings, instructions;
- Finalise the item-level and test-level scoring requirements and score report design;
- Gather additional data for research and validation purposes.

5.1 Field Test Design

The field test was divided into four 'tiers', with each tier targeting slightly above and below a CEFR level.

A total of 64 test forms were created with an overlapping model whereby each item appeared on at least two forms. Some items appeared across the four tiers to form a linking design. In addition, anchor items (calibrated items with a known difficulty value on the GSE) were placed across the forms and tiers to create an additional link for the psychometric analysis.

Students were assigned to a tier by their teacher/institution according to their estimated ability level (A1 to B2). Most of the students (96%) were recruited from Wall Street English² (WSE), whose course levels are aligned to the CEFR (Lee, 2006). WSE participants were primarily recruited from *Progress*' target-market regions in South America, Asia (excluding mainland China), Europe and Middle East. Of the non-WSE students, these came from 8 institutions currently using Pearson learning materials.

An overlapping student model was created by distributing students on WSE courses at the top and bottom of a CEFR level onto two adjacent tiers.

² Wall Street English (WSE) is part of Pearson. It has over 450 centres in 28 countries/territories in Africa, Asia, Europe, Latin America, and the Middle East. It uses a blended learning method. For more information, visit: <http://www.changingfutures.com/>

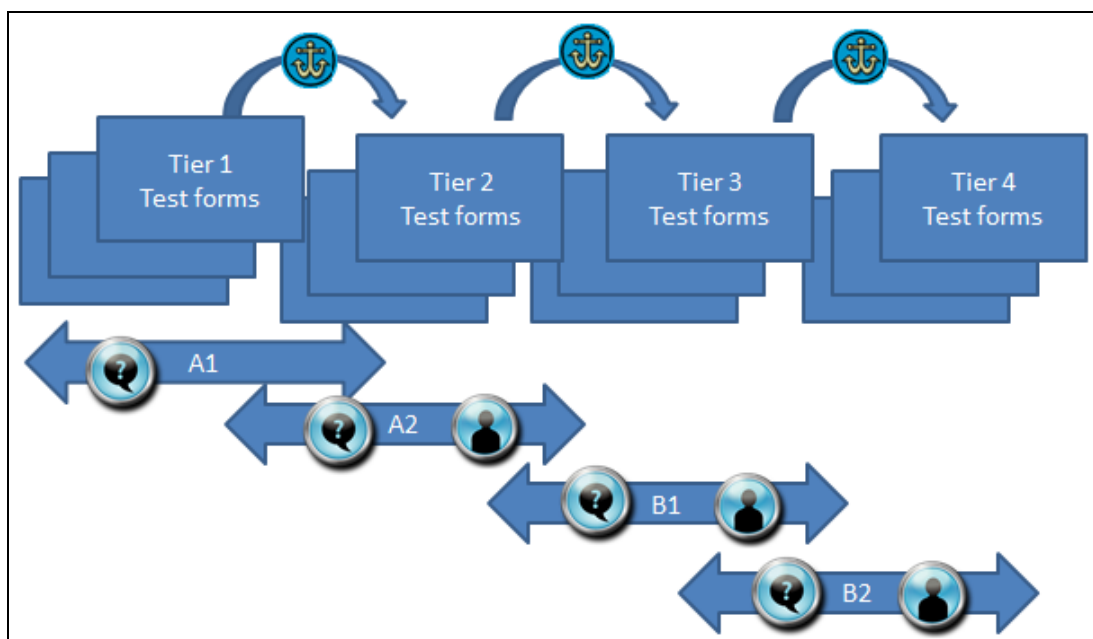


Figure 7 Field Test Model showing overlap between anchors, item level and student ability across tiers and test forms

The test was administered by the participating institutions/WSE centres and proctored by the teacher. For WSE students, the test was partially embedded into the structure of their course, so that when they finished a particular WSE level they were offered the test as part of their study experience. However, it was not compulsory.

As an incentive to participate and take the test seriously, students were provided with an online subscription to Longman Dictionaries. Participants were also provided with a percentage score for Part 1 (reading, vocabulary and grammar). This was based on the raw scores from uncalibrated, computer-scored items only.

L1 English speakers were also recruited to take the *Progress* speaking items in order to build the scoring models. These tests were administered via web-access and were not proctored.

5.2 Participant information

4,327 participants completed *Progress* tests during the field test phase. An additional 93 participants were L1 English speakers who took the 'native speaker' form, which was limited to only a few item types. Some basic biographical information on participants was gathered in a pre-test survey and it was completed by 4,245 participants.

Age: The demographic for the operational test will be upper secondary and young adult. Table 4 shows a breakdown of the participants' ages. The majority (72%) were aged 16-35 ($M = 29.7$, $SD = 11.6$).

Table 4 Field test participant ages

Age Group	Count	% of Total
Below 16	86	2.0%
16 to 20	920	21.3%
21 to 25	979	22.6%
26 to 30	743	17.2%
31 to 35	506	11.7%
36 to 40	336	7.8%
41 to 45	247	5.7%
46 to 50	193	4.5%
Above 50	317	7.2%
Grand Total	4327	100%

Gender: There was a fairly even divide between female (54.5%) and male (45.5%) participants.

Country of Birth: Table 5 presents the distribution of participants over country of birth. 109 countries were reported, of which 18 account for about 93% of the total number of participants. These 18, where N>50, are reported in the table.

Table 5 Field test participant country of birth

Country	Count	% of Total
Indonesia	464	10.7%
France	309	7.1%
Russia	305	7.0%
Italy	299	6.9%
Portugal	284	6.6%
Colombia	274	6.3%
Argentina	264	6.1%
Hong Kong SAR China	262	6.1%
Germany	230	5.3%
Saudi Arabia	225	5.2%
Thailand	224	5.2%
Turkey	196	4.5%
Ecuador	179	4.1%
South Korea	159	3.7%
China	109	2.5%
Mexico	93	2.1%
Brazil	90	2.1%
Spain	50	1.2%
Other countries	311	7.3%

First Language: Test takers selected their L1 from a drop-down list. In total, participants reported 63 different L1s. The languages with over 100 people are reported in Table 6.

Table 6 Field test participant first languages

Language	Count	% of Total
Spanish	887	20.5%
Indonesian	465	10.7%
Portuguese	391	9.0%
Chinese	379	8.8%
French	336	7.8%
Russian	333	7.7%
Italian	303	7.0%
Arabic	272	6.3%
German	238	5.5%
Thai	221	5.1%
Turkish	192	4.4%
Korean	160	3.7%
Other languages	150	1.2%
Total	4327	100%

5.3 Transcription

As a result of the field test, over 290,000 responses were collected. All response data was stored in a secure database. Spoken responses were transcribed by a team of trained transcribers. The purpose of transcribing spoken responses was to transform the audio recorded test-taker responses into annotated orthographic text. The audio and orthographic transcriptions were then used to develop and validate automated scoring systems based on a large sample of candidates at various levels and from various first language backgrounds.

Responses were transcribed by a group of twenty educated L1 English speakers located in the United States. They all underwent rigorous training, which included understanding the purpose of transcriptions and learning a specific set of rules and annotation symbols. Subsequently, they completed a series of practice sets. Transcribers who did not perform well on the practice sets did not continue. During the real transcription process, the quality of their transcriptions was closely monitored by the test development team and the transcribers were given feedback throughout the process. As an additional quality

check, when two transcriptions for the same response did not match, the response was automatically sent to a third transcriber for adjudication.

The actual transcription process was carried out with an online interface. Transcribers could listen to each response as many times as they wished in order to understand the response. Audio was presented from a stack semi-randomly, so that a single test taker's set of responses would be spread among many different transcribers. The process is illustrated in Figure 8.

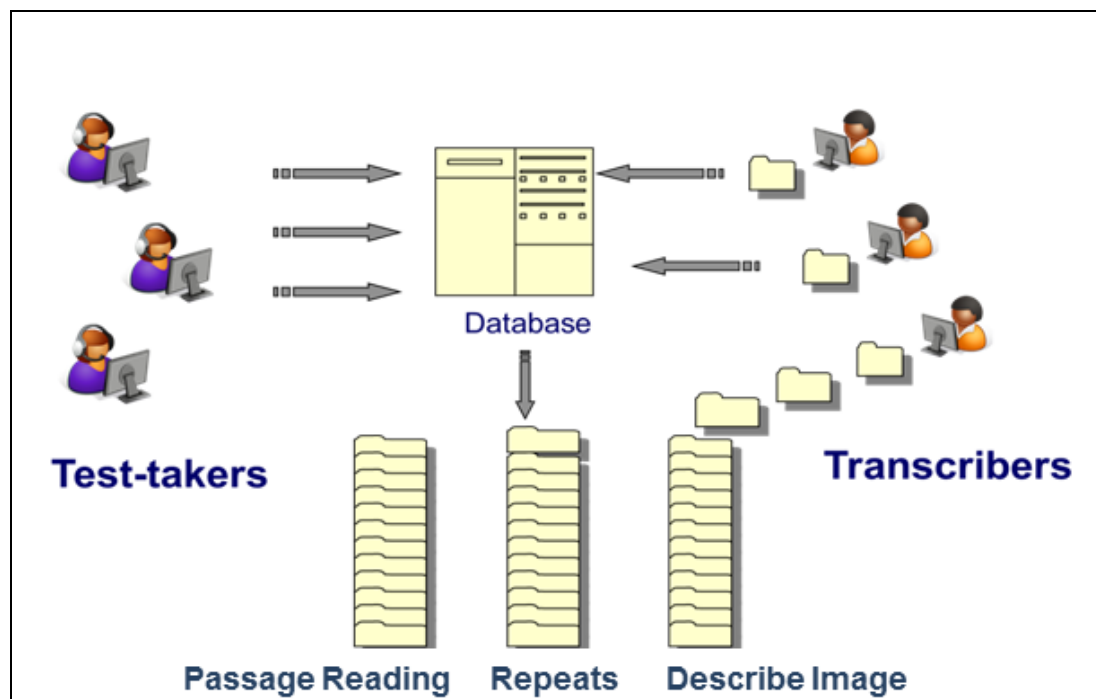


Figure 8 Transcription process in Progress Test development

There were two steps to the transcription task. The first step involved producing transcriptions to develop automated scoring systems. For this step, a total of 24,614 transcriptions were produced (4,357 transcriptions for native responses and 20,257 transcriptions for learner responses.) The second step was to produce transcriptions to conduct a series of validation analyses. An additional 14,059 transcriptions were produced for the validation analysis purpose.

5.4 Human Rating

Rater recruitment and training

Progress raters were required to meet minimum qualification requirements: L1 English speakers or L1-like proficiency in English, have a bachelor's or higher degree in humanities, education, arts or social sciences, and EFL teaching experience. A recognised qualification in ESL or EFL was strongly preferred.

Successful applicants attended a standardisation event run by Pearson test developers to ensure that rating was carried out consistently. The standardisation training covered familiarisation with the *Progress* test specification, item types and target population, the CEFR and the scoring rubrics. A standardisation guide was developed by Pearson staff using a selection of field test test-taker responses. The selected responses represented

each item-trait combination and the full range of possible score points as well as a rationale for the assigned score referenced to the *Progress* scoring rubric.

At the end of the standardisation event, a rater qualification exam was administered. Participants had to achieve 80% adjacent agreement in order to continue. All 22 raters passed the minimum requirement.

Rating process

Both spoken and written responses from the field test were rated by twenty-four expert raters in order to provide a criterion for the machine scoring. The raters judged for Narrative Clarity and Accuracy, Content, Opinion Writing, Writing Conventions, Vocabulary, Pronunciation, and Fluency. For each one of the seven traits, a set of rubrics was developed based on the CEFR and, prior to rating actual responses, the experts received training in how to evaluate responses according to the specific rating criteria. The traits, trait descriptions, and tasks used for trait ratings are summarised in Table 7.

Table 7 Trait descriptions used for human scoring

Trait	Trait Description	Tasks rated
Narrative Clarity and Accuracy	Ability to reproduce writing in clear and accurate English. Ability to convey fully the situation, characters, actions and ending of a text.	Read and then Write
Content	Ability to address all aspects and elements of the image.	Short Essay (picture) Describe Image
Opinion Writing	Ability to state his/her opinion on the topic and provide supporting ideas and/or examples.	Short Essay
Writing Conventions	Ability to adhere to the standard rules of spelling, punctuation, capitalisation, sentence formation and grammar.	Short Essay
Vocabulary	Range of vocabulary and the accuracy and appropriacy of its use.	Short Essay
Pronunciation	Ability to produce consonants, vowels, and stress in a native-like manner in sentence context; also, knowledge of the phonological form (phonemes and lexical stress) of everyday words.	Read and then Speak Describe Image Listen and then Speak
Fluency	Smooth, rapid speaking; evidenced by appropriate rhythm, phrasing, and pausing in continuous speech.	Read and then Speak Describe Image Listen and then Speak

The actual rating process involved was carried out in the same online interface used by the transcribers. For the rating of Pronunciation and Fluency, raters could listen to each response as many times as they wished. Audio was presented from a stack semi-randomly, so that a single test taker's set of responses would be spread among many different raters in a way to the transcription process (see Figure 8 above).

For Narrative Clarity and Accuracy, Content (Short Essay), Opinion Writing, Writing Conventions, and Vocabulary, the ratings were judged based on the test taker's written responses. For Content (Describe Image), the ratings were judged based on transcriptions of responses. The raters read the transcribed responses on screen and rated them for Content. This was done intentionally, so that judgments were not influenced by the test taker's speech characteristics such as pronunciation or fluency.

Raters were only able to rate for the specific trait they were assigned and trained to rate. Separating the judgment of different traits is intended to minimise the transfer of judgments from pronunciation to fluency or to vocabulary and content judgments, by having the raters focus on only one trait at a time. This is illustrated in Figure 9. For Content (Short Essay), Opinion Writing, Writing Conventions, Vocabulary, Pronunciation, and Fluency, rating stopped when each response had been judged by two independent raters; with the Narrative Clarity and Accuracy and Content (Describe Image) rating, each response was judged by three independent raters.

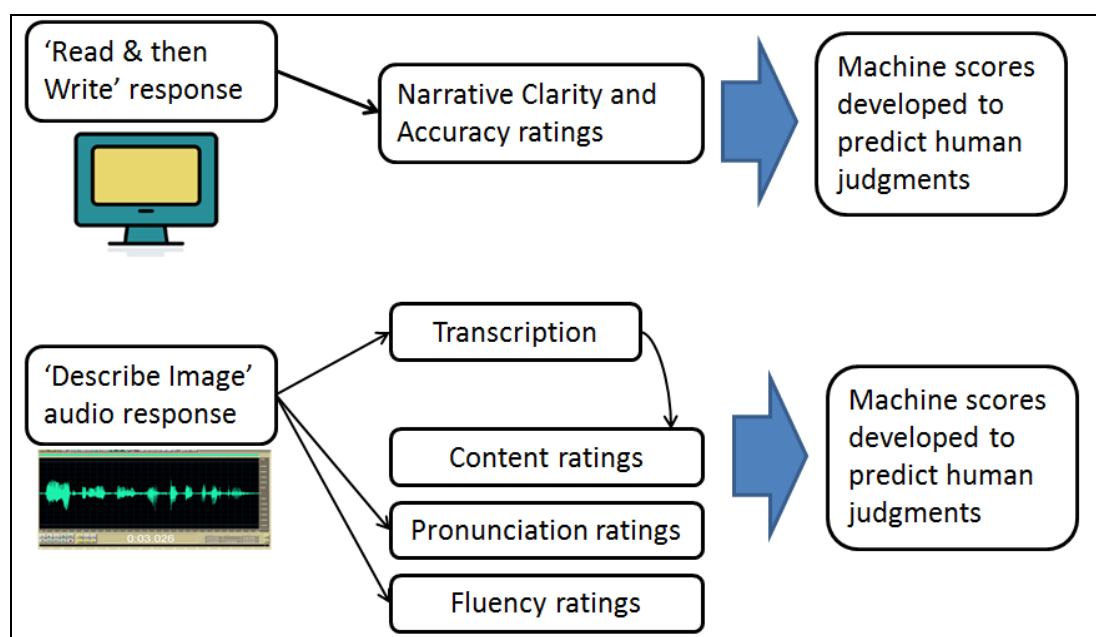


Figure 9 Human rating process in Progress test development

As in the transcription task, the rating task also consisted of two steps. The first part was to collect expert judgments to develop automated scoring systems and the second step was to conduct a series of validation analyses. The experts produced a total of 89,743 ratings for the development of automated scoring systems and a total of 28,200 ratings for the validation analysis purpose. Inter-rater reliabilities were computed for all rating tasks as in Table 8. The reliability ranges in the table are based on a 95% confidence interval.

Table 8 Inter-rater reliability estimates and 95% confidence intervals for development and validation purposes

Trait	Inter-rater Reliability Estimate (Development)	95% CI (Development)	Number of Rated Responses (Development)	Inter-rater Reliability Estimate (Validation)	95% CI (Validation)	Number of Rated Responses (Validation)
Narrative Clarity & Accuracy	0.79	0.78-0.80	8,159	0.83	0.80-0.85	600
Content (Describe Image)	0.61	0.60-0.62	13,394	0.65	0.62-0.68	1200
Content (Short Essay - picture)	0.56	0.52-0.60	1,212	0.61	0.43-0.74	67
Opinion Writing	0.59	0.57-0.61	2,919	0.53	0.43-0.62	225
Writing Conventions	0.67	0.65-0.69	4,131	0.61	0.53-0.68	292
Vocabulary	0.55	0.53-0.57	4,131	0.41	0.31-0.50	292
Pronunciation	NA	NA	NA	0.70	0.69-0.71	5,702
Fluency	NA	NA	NA	0.72	0.71-0.73	5,702

5.5. Machine Scoring

Automated scoring methods are used for both the spoken and written constructed responses in *Progress*. The responses collected on each item during the field test were subjected to human ratings on various aspects of language skills (traits) as described above. These human ratings were then used to train an artificial intelligence engine. The end result is a set of models able to produce ratings that are predictive of those that expert human judges would give. The model can then be used to rate new responses on the same prompts in operational testing. This approach creates time, effort, and cost savings in the scoring of new test responses, because human ratings of new data are no longer needed. For an overview of automated scoring technology, see Bernstein, Van Moere, & Chen (2010) for automated spoken response scoring and Foltz, Streeter, Lochbaum, & Landauer (2013) for written response scoring.

Trained scoring models are generally successful at reproducing the human ratings they have been trained on. However, in operational testing, the scoring system will be scoring new, unseen responses. To ensure that the machine's scores are comparable to scores given by skilled human raters for data the model has never before encountered, Pearson conducted preliminary validation analyses in which a new set of test taker responses was

scored by both human raters and by the automated scoring system. The relationship between the two sets of ratings was investigated to demonstrate validity evidence. The findings of these studies are reported in sections 6.3 and 6.4.

6. Psychometric Analysis and Validation Studies

Field test response data were extracted from databases for analysis. Response times in milliseconds were also extracted and associated with each candidate's response data.

6.1 Data cleaning and item scoring

Methods for scoring each item were developed and adjusted over time as data were collected and analysed. In the case of computer-scored items, these methods involved taking scored values (e.g., raw number of words placed in the correct position), converting to a percentage, binning into a small number of score categories, and rounding to an integer for entry into the psychometrics software. In the case of machine-scored items (see Section 5.5) item scores on each of the traits were generated by Pearson's system, and were then combined to comprise an overall item score made up of one or more traits. These aggregated overall item scores were truncated and rounded to integers to adhere to the values in the scoring rubrics originally presented to expert raters. These "binned" overall item scores were then included in the main dataset and subjected to analysis in a psychometric software package.

During field testing, test forms were presented in two stages to examinees. The first stage consisted mainly of computer-scored, multiple choice or short-answer items, and the second stage mostly of constructed responses to later be scored by trained automatic scoring systems. In order to best represent the operational test experience, timing on the field test was controlled at the section level. Candidates moved at their own pace through the section by clicking the next button when they had completed an item. A time/item remaining counter was displayed onscreen. If the section time expired before they reached the final item, candidates were moved on automatically to the next section. As a result, in some cases test takers did not complete all the items in a section. Through analysis of response content and timing data for each response, data were screened out of the final psychometric analysis if it could be determined that the examinee never saw the item presentation. If the response time was zero milliseconds, that response was not included in the final dataset. However, if a response time was greater than zero, and no response was given by the examinee, it was determined that the student may have chosen to skip the item, and a lowest obtainable scale score was assigned.

An initial analysis of the raw responses was conducted to screen data. Items were identified and not included in subsequent psychometric analyses if they met any of the following criteria:

- total number of responses less than 50;
- item facility values at or below .10;
- item facility values at or above .90;
- Rasch item-to-measure statistics below .30.

ConQuest 3.0 (Adams, R. J., Wu, M. L., & Wilson, M. R., 2014) was used to analyse responses in a Partial Credit Model. The model was run a number of times to identify and remove items and score categories that the software found problematic. A final model

was settled on and run. This model provided item parameter estimate values for the remaining items, and Expected A Posteriori (EAP) score estimates for all candidates using the calibrated item parameters. All calibrated items, therefore, contributed to the Overall score. A frequency distribution of the Overall scores is shown in Figure 10. These Overall scores are entirely based on machine scoring of responses.

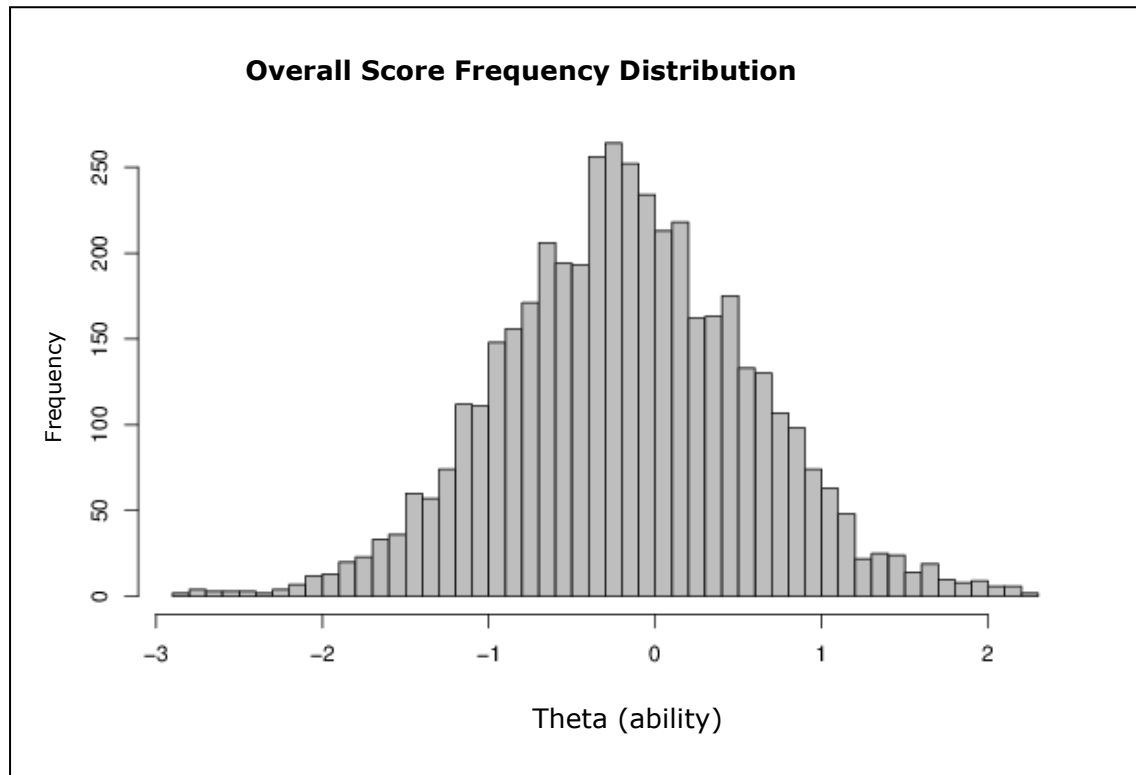


Figure 10 Frequency distribution of overall scores on theta (ability) scale

Theta on the X-axis of the figure denotes the test-taker's ability. The lower a theta value is, the lower the test-taker's ability is. In Item Response Theory (IRT), the typical theta range is -3 to +3. The ability distribution in Figure 10 looks very normally distributed. With the large sample size of more than 4,000 test-takers in field testing, a normal distribution of ability is expected.

In calculating skill scores (Reading, Writing, Speaking, Listening, Vocabulary, Grammar), the same item parameter calibrations were used from the Overall-score calibration. EAP scores were simply re-estimated for each candidate using the subset of items (all items of a particular item type) identified as contributing to each skill, as shown in Table 2. Frequency distributions of those scores are shown in Figure 11. Again, with the large sample size, data is expected to normally distribute as in Figure 10 for each skill.

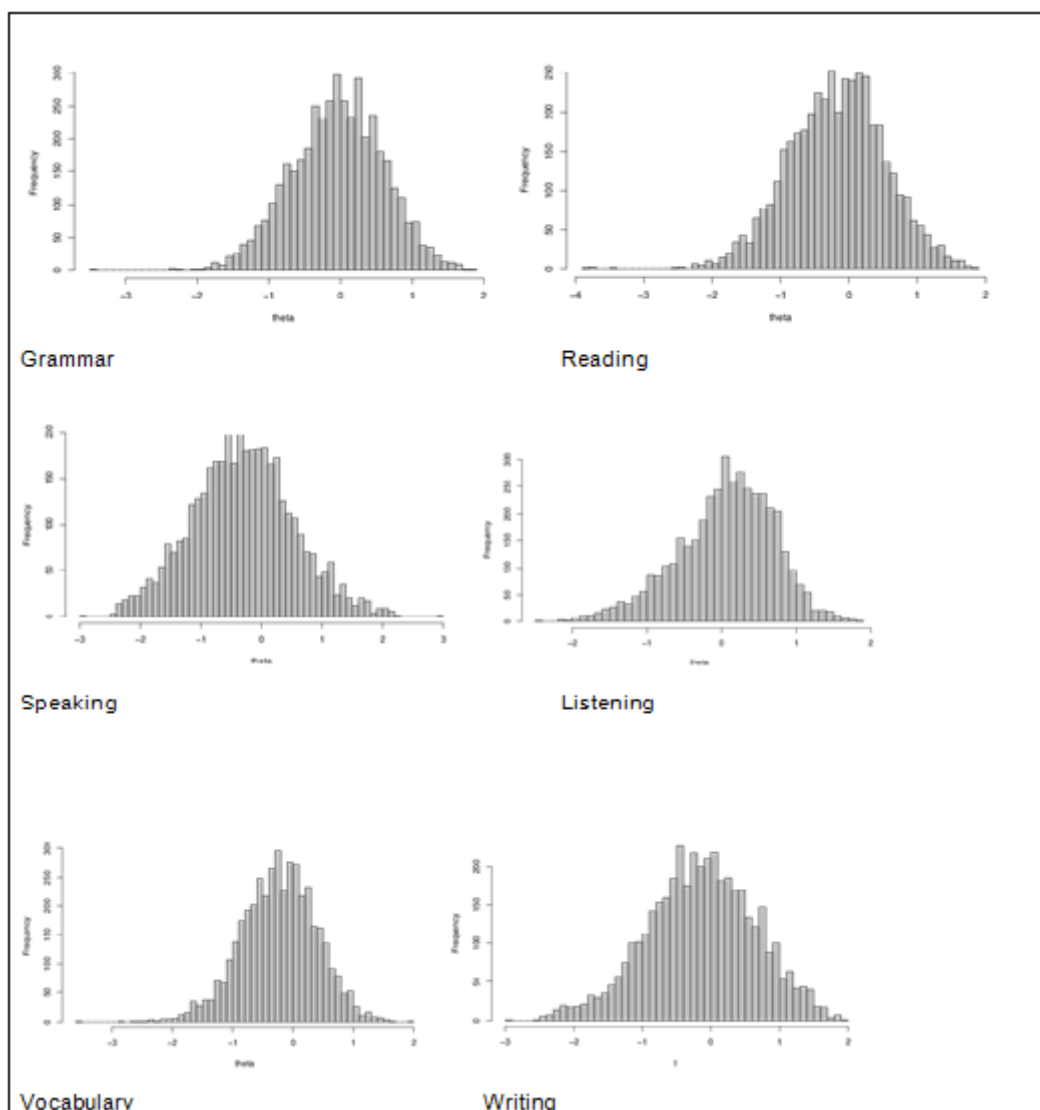


Figure 11 Frequency distributions for Progress skill scores on theta (ability) scale

6.3 Validity Study Design

A series of analyses were performed to demonstrate the reliability and validity of test scores of *Progress*. The preliminary validation analyses examined a few key aspects of the *Progress* scores:

1. Internal or structural reliability: the extent to which the *Progress* provides consistent and repeatable scores.
2. Accuracy of machine scores in relation to human judgments: the extent to which the *Progress* scores accurately reflect the scores that human listeners and raters would assign to the test-takers' performances.

In order to address these questions, a sample of field test participants was kept separate from the development data, and the data were used as an independent validation set. This is an important validation process for an automatically scored test because, once the test becomes operational, the automated scoring system needs to deal with test-takers and responses that the machine has never encountered before. Having a separate validation dataset simulates the real-life operational scoring situation. If machine scores

show close correspondences with human ratings, it will serve as a piece of validity evidence for the accuracy of *Progress*' machine scoring.

6.4 Internal Validity

To understand the consistency and accuracy of *Progress* Overall scores and the distinctness of the sub-scores, the following indicators were examined:

- the reliability of *Progress*;
- the correlations between the *Progress* Overall score and the *Progress* sub-scores, and between pairs of *Progress* sub-scores;
- comparison of machine-generated *Progress* scores with listener-judge scores of the same *Progress* tests.

These qualities of consistency and accuracy of the test scores are the foundation of any valid test (Bachman & Palmer, 1996).

Score Reliability

The reliability of overall scores is estimated directly within the model for all candidates. A scatterplot of score reliability for all candidates is shown in Figure 12. The average reliability across all candidates is .91, indicating that the test is highly reliable.

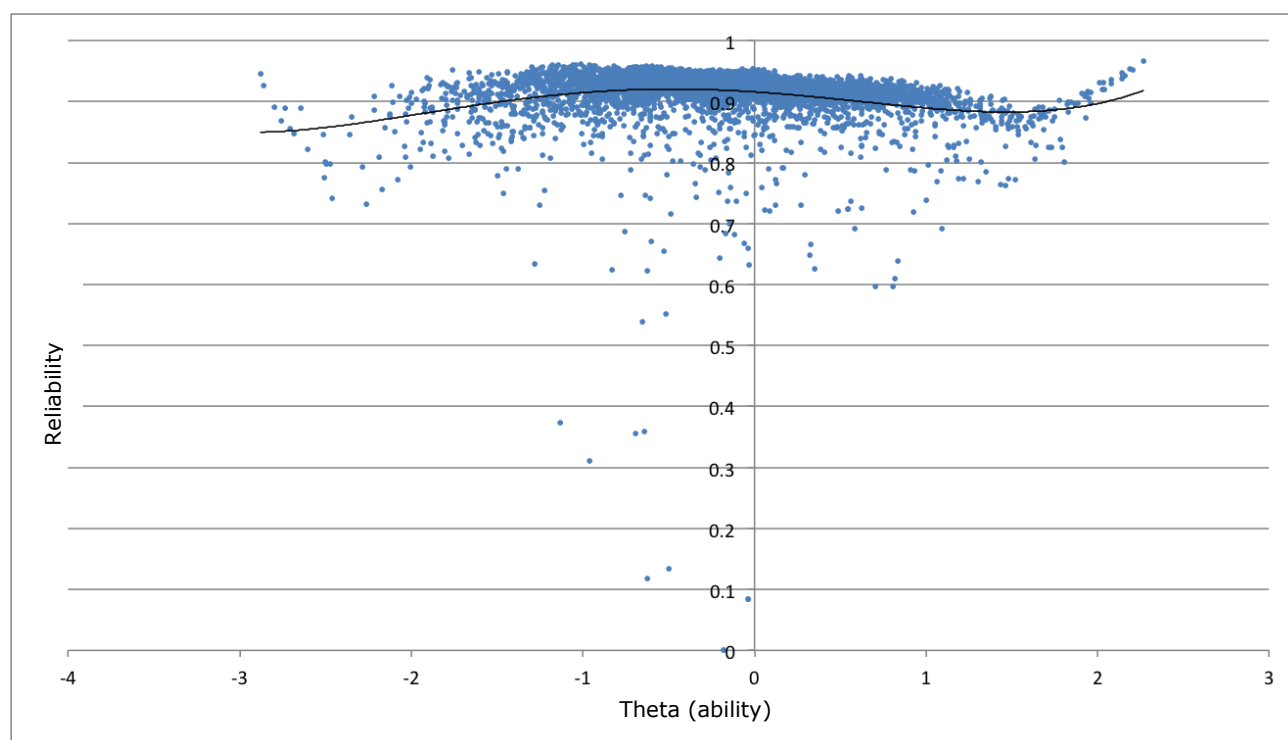


Figure 12 Reliability of candidate overall scores with trendline

Correlation between Sub-scores

Table 9 presents the correlations between pairs of *Progress* sub-scores as well as between the sub-scores and the Overall score for the sample of 4,327 tests. These scores are EAP estimates generated by the IRT model.

Table 9 Correlations between Progress scores for non-native test taker sample (n=4,327)

	Overall	Speaking	Grammar	Listening	Writing	Vocabulary	Reading
Overall	1						
Speaking	.765	1					
Grammar	.739	.423	1				
Listening	.820	.692	.491	1			
Writing	.844	.558	.609	.820	1		
Vocabulary	.826	.612	.596	.562	.644	1	
Reading	.791	.690	.557	.614	.621	.597	1

As expected, the correlations between the Overall score and each sub-skill score are higher than the correlations between sub-skills. Sub-skill scores correlate each other to some extent by virtue of presumed general covariance within the candidate population between different component elements of language skills. The correlations between the sub-skill scores are, however, significantly below unity, which indicates that the different scores measure different aspects of the test construct, using different measurement methods, and different sets of responses.

Validation Sample

As described in Section 5.5 of this document, some item types on this test were scored by machine algorithms that were trained using human ratings of performances. In order to check on the performance of the machine scoring method, 299 participants' data were set aside for a series of validation analyses. This study compared Overall and sub-skills scores that were generated both by the machine method, and by human ratings. Care was taken to ensure that the training dataset and validation dataset did not overlap – that is, the performance samples provided by the validation test takers were excluded from the datasets used for developing the automated scoring systems.

The mean age of the validation subjects was 30.5 years old ($SD = 12.3$). The gender was split between 45% male and 53% female.

Table 10 provides a summary of first languages amongst the group, with Spanish as the most reported L1 (21.7%).

Six (2%) of the validation subjects did not report their age, gender, or first language.

Table 10 Summary validation subjects' first language backgrounds

First Language	Percent
Spanish	21.7%
Portuguese	17.7%
Indonesian	11.0%
French	9.3%
Italian	7.0%
Russian	6.0%
Chinese	5.3%
Arabic	4.7%
Thai	4.3%
German	4.0%
Korean	2.3%
Turkish	2.0%
Romanian, Polish	0.7% each
Kyrgyz, Kabyle, Japanese, Greek	0.3% each

Correlations between the Progress Test Machine and Human Scores

Table 11 presents Pearson product-moment correlations between raw (i.e., non-IRT-based) machine scores and human scores, when both methods are applied to the same performances on the same *Progress* responses. At both Overall and sub-skill levels, total scores are computed within the item types that were identified as contributing to those skill areas, as shown in Table 2. In the case of item types where scoring was objective (and thus human ratings-based scores were not produced), the computer-scored responses are included in totals. As explained earlier, the human ratings used in this analysis were not included in the data to develop the automated scoring systems. The correlations presented in Table **11** suggest that *Progress* automated scoring will yield scores that closely correspond with human scoring of the same material.

Table 11 Correlation coefficients between human and machine scoring of the *Progress* performances (n=299).

Score	Correlation
Overall	.94
Grammar	.98
Vocabulary	.93
Reading	.98
Writing	.99
Speaking	.78
Listening	.83

As can be seen in Figure 13, at the Overall score level, the *Progress* automated scoring is highly compatible with scoring based on careful human transcriptions and multiple independent human judgments. That is, at the Overall score level, machine scoring does not introduce substantial error in the scoring as compared with human ratings even for the test takers that the machine has never encountered.

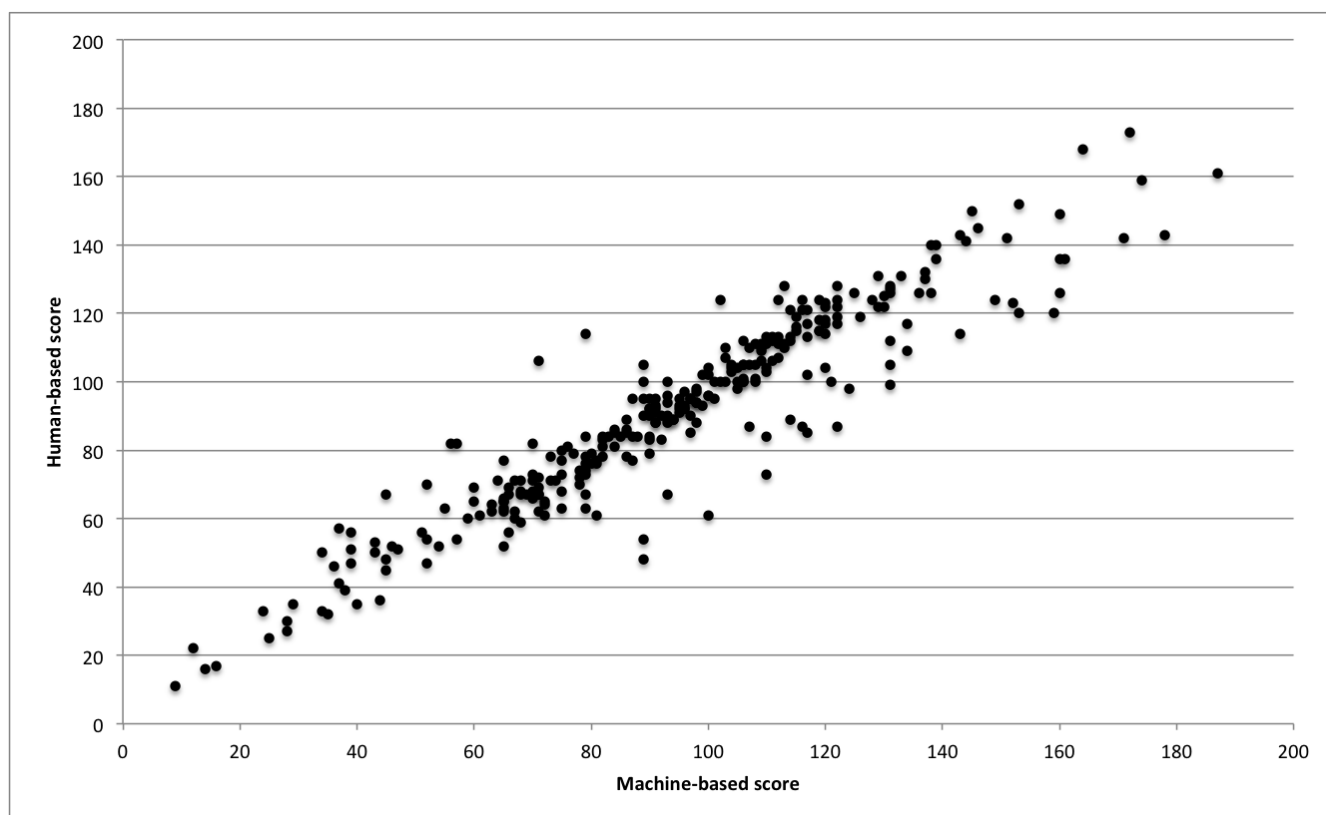


Figure 13 Scatterplot of automatically-generated Overall scores and human-based Overall scores ($r=.94$, $n=299$)

7. Summary

This report describes the initial stages of validation of *Progress* through a large-scale field test. The field test was designed to calibrate the item bank as well as help to provide evidence for alignment to the CEFR and the Global Scale of English (GSE). The results are in general positive with a large number of items calibrated and a strong relationship established between the item bank and the GSE.

The operational design of *Progress* utilises both adaptive and non-adaptive approaches to ensure the test is as accurate as possible. The adaptive stage of the test uses computer scored items which help estimate the learner's ability on the fly, selecting items which are targeted to the level. The adaptive algorithm also selects questions across different item types and different skills to ensure that there is a balance in terms of items responded to and the performance of the learner on the different skills. Stage 2 of the test is linear and uses computer scored items including automatically scored speaking and writing questions. All items have been calibrated based on the responses in the field test. The scores across all items are used to calculate the overall score.

An initial set of validity evidence was presented based on the field test forms and data, indicating a high IRT-based reliability ($r=0.91$) and a correlation coefficient of 0.94 between machine-generated Overall scores and human-generated Overall scores. Since the field test forms were slightly different from the operational test forms, further validation studies are planned in the Beta phase and later to provide further validity evidence on the operational versions of the test forms.

One of the key characteristics of *Progress* is that it is user friendly and learning focussed. *Progress* is primarily designed to be used by teachers and learners as part of a learning program supporting learning through assessment. The more granular nature of the GSE, as compared with CEFR levels, gives both learners and teachers a clearer idea of the progress being made and the 'drill down' menus in the Gradebook give useful and practical information which inform and support learning, including remediation.

The process of establishing validity for any test is an on-going endeavour. *Progress* will continue to be developed and a number of further validation and efficacy studies are planned in order to build a body of validity evidence. The results from these studies will be reported on the *Progress* website www.pearsonelt.com/progress

8. Glossary

Term	Explanation
Computer Adaptive	An approach for testing. An adaptive test uses an algorithm to select an item that is the closest to the test-taker's level based on the test-taker's responses to earlier items on the test
Calibrated item (anchor)	Items with known difficulty values – they have been analysed using IRT and placed onto the Global Scale of English
Common European Framework of Reference for Languages (CEFR)	A framework to describe language proficiency, developed by the Council of Europe. It uses 'can-do' statements to describe 6 levels of proficiency from A1 to C2. See: http://www.coe.int/t/dg4/linguistic/cadre1_en.asp
ConQuest	A psychometric software package used for item response analysis and modelling. See: shop.acer.edu.au/acer-shop/group/CON3
Correlation	A statistic from -1 to 1 indicating the strength of the relationship between two variables. 'Pearson product-moment' is a particular method for evaluating a correlation.
Expected A Posteriori (EAP) score	A particular statistical method used in IRT to derive an overall score from a set of items
Global Scale of English (GSE)	Global Scale of English: a granular scale from 10 to 90 used to report on English language proficiency
Gradebook	The area on the My English Lab platform where Progress scores and diagnostic reports are displayed
Item facility value	A statistic measuring the degree of difficulty of a test item. The higher the value to 1.0 (100%) the easier the item is.
Item Response Theory (IRT)	A measurement modelling approach. It models both the difficulty of items and test-taker abilities on the same scale
Reliability	Estimate of how consistent test scores are over time.
Standard Deviation (SD)	A measure of the dispersion of a set of data from its mean
Theta	The latent trait as estimated in IRT modelling. In Progress, it is the test-taker's ability.
Validity	The extent to which evidence and theory support the score interpretations for the intended purposes of the test.

9. References

- Adams, R. J., Wu, M. L., and Wilson, M. R. (2014). *ConQuest (3.0.1)* [Computer software]. Camberwell, Australia.
- Bachman, L.F. & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bernstein, J., Van Moere, A., & Cheng, J. (2010). Validating automated speaking tests. *Language Testing*, 27(3), 355-377.
- Black, P. and William, D. (1999) *Assessment for learning: Beyond the black box*. University of Cambridge School of Education.
- Council of Europe (2001). *Common European Framework of Reference for Language; Learning, teaching, assessment*. Cambridge University Press
- Council of Europe (2009). *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*. Council of Europe.
- van Ek, J.A., & Trim, J.L.M., (1990a) *Threshold Level*, Cambridge University Press
- van Ek, J.A., & Trim, J.L.M., (1990b) *Waystage Level*, Cambridge University Press
- Foltz, P. W., Streeter, L. A., Lochbaum, K. E., & Landauer, T. K (2013). Implementation and applications of the Intelligent Essay Assessor. *Handbook of Automated Essay Evaluation*, M. Shermis & J. Burstein, (Eds.). Pp. 68-88. Routledge, NY. NY.
- Lee, T. (2006) *Report on WSI CEFR Alignment Study Phase One (Alignment with BULATS Levels)*. Retrieved 3rd April from http://www.wallstreetinstitute.com/pdf/Microsoft%20Word%20-%20Alignment_Report220606_FINAL.pdf
- Speck, M., and Knipe, C. (2001). *Why can't we get it right? Professional development in our schools*. Thousand Oaks, California: Corwin Press, Inc.
- Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for learning*. Portland, OR: Educational Testing Service.
- Widdowson, H. G. (1978) *Teaching Language as Communication*. Oxford University Press.

ALWAYS
LEARNING

APPENDIX C

Pearson SEE mbo Progress Test Information

Pearson SEE mbo is een gecustomiseerde versie van Progress. Alle informatie in het Progress Technical Report en Progress Test Information Booklet is van toepassing op Pearson SEE mbo.



Test Information Booklet

Sep 2015

V. 0.6

Global Scale of English 10-85

Contents

Introduction	4
Test Validation	5
Test Coverage	9
Test Questions	10
Score Reporting	17

Introduction

Progress: An Overview

Progress is a part adaptive computer-based test of English language proficiency designed to be used on a number of occasions through a course of study to accurately show learners' progress.

The Purpose of Test

The test is normally used in a learning context in conjunction with relevant materials and formative assessment tasks, but it can be used independently. It provides detailed information to a teacher about a group or class of learners who are studying any course, and information about each learner in the group. The information allows the teacher to make decisions about adapting learning material to suit the level of both group and individual learners and providing extension activities where the group or particular individuals are weak. It also allows the teacher to tailor the learning program to particular learners, giving extra support and input where required.

Who is it for?

The test is designed for adult learners who are 16 or older. Progress can be used alongside any adult or upper secondary course. It is intended to be used with comprehensive integrated skills courses not short or partial courses.

Why take an integrated skills test?

Some of the questions Progress uses test a single skill such as speaking or writing. When assessing these skills we also test traits such as pronunciation and fluency, the ability to argue as well as written conventions along with grammar and vocabulary. A number of the questions on the test are integrated skills questions. These questions test more than one skill at the same time.

Using integrated skills questions means that Progress is a better test of a learner's English. In real life and in the classroom learners use more than one skill to complete communicative tasks. To order something in a restaurant we need to listen and speak, to take notes in a classroom we need to listen and write. Integrated skills questions test how well learners can use the skills they have learnt and practised in the classroom and used in real life.

Test Validation

Test Design

Progress is designed specifically to measure progress in language proficiency. It employs a part adaptive method. Part of the test uses an adaptive algorithm which takes a learner's answers to a previous question to select the most suitable question to present next. Progress selects these items from a large item bank making each learner's experience different.

The adaptive nature of the test allows Progress to quickly and accurately estimate a learner's English proficiency. This estimate is then used to choose further questions which are fine-tuned to the learner's level allowing a very accurate measure of their proficiency. By taking two or three tests over a period of study the learner, and the teacher, can see how much progress is being made.

Test Development

The questions in Progress have been developed by international teams of writers who are very experienced in writing assessment questions. Teams are based in the UK, Australia, the USA and Hong Kong. All questions have been tagged with a Global Scale of English (GSE) level and linked to a 'can do' statement.

Once written, all questions are reviewed by the teams in the different countries. Comments and suggestions for improvement are stored with the test questions on a secure database. The questions then go through a further review by an expert panel and decisions are made on the quality of the questions; which to keep and which to reject. All questions are then thoroughly checked by Pearson staff and images and high quality recordings are added to complete the questions before they go forward to be calibrated in a large scale field test.

After the field testing, further checks are made on item quality based on the measurement characteristics of the questions. Questions are eliminated from the item pool if they are too easy or too difficult, if weaker learners get them right but stronger learners get them wrong, or if they show any bias. These checks then result in a bank of the best quality questions. Questions are selected from this bank to go into the final tests.

Field Testing

As part of the test development process, a large field test, conducted in two phases, was carried out to ascertain the appropriateness of the pool of items and to serve as a source for constructing individual test forms which would allow reliable predictions of students' ability in English. A portion of the data collected was transcribed and rated which was used to train automated scoring systems.

Field test forms were created using a linking approach. That is, the forms were linked together with sets of items that appeared on all forms. Also, during the second phase of data collection, since most candidates took two tests, the field test forms were also linked through candidates.

Learners and L1 English speakers were recruited to participate in the field test. A total of 13,073 tests were submitted during the two field test phases. The demographic for Progress is upper secondary and young adult. The majority of participants were aged 16-35. Participants were from 96 countries. The countries with the largest number of participants included; Saudi Arabia, Poland, Panama, Ecuador, The Netherlands, Argentina, Brazil, Spain, Guatemala, Japan and Thailand. As an incentive to participate, students received a one year free access to the Longman Dictionary of Contemporary English Online (LDOCE). L1 English speakers were offered an Amazon voucher.

Validity Evidence

Test Reliability

Reliability is one aspect of validity - if a candidate took a test on multiple occasions, would that person get a similar score each time? During field testing, a large number of candidates took two tests in a short period of time. The two tests were made up of different items. Presumably, little or no learning occurred between these test administrations, so the correlation of the scores from these two tests should provide a good estimate of test reliability, known as test-retest reliability. The higher the observed correlation between the two test administrations, the more reliable the test scores are. In the observed field test data, after removing test data from candidates who either did not answer a sufficient number of items, or who got extreme scores outside of the normal GSE range, the test-retest correlation was .861 ($n=2,141$). This observed correlation demonstrates a high level of consistency of measurement of Progress test administrations.

The psychometric analysis tool, called Winsteps, also yielded another measure of test reliability estimate as part of item calibration. The reliability estimate is 0.90 ($n=11,908$). From these two estimates, it is clear that test reliability is high.

Automated scoring validation process

From the field test data, 300 candidates were randomly selected as the validation data set. A validation data set is a group of candidates whose data are segregated out prior to psychometric analysis in order to independently test how well automated scoring models work, once they are complete. Additionally, these candidates' data were not included in the psychometric item calibration, or in the scaling onto the GSE. If the test scores for these candidates as calculated by both automated and human scoring models are highly correlated, this provides evidence that the automated scoring models will work as expected for other new candidates in the operational setting.

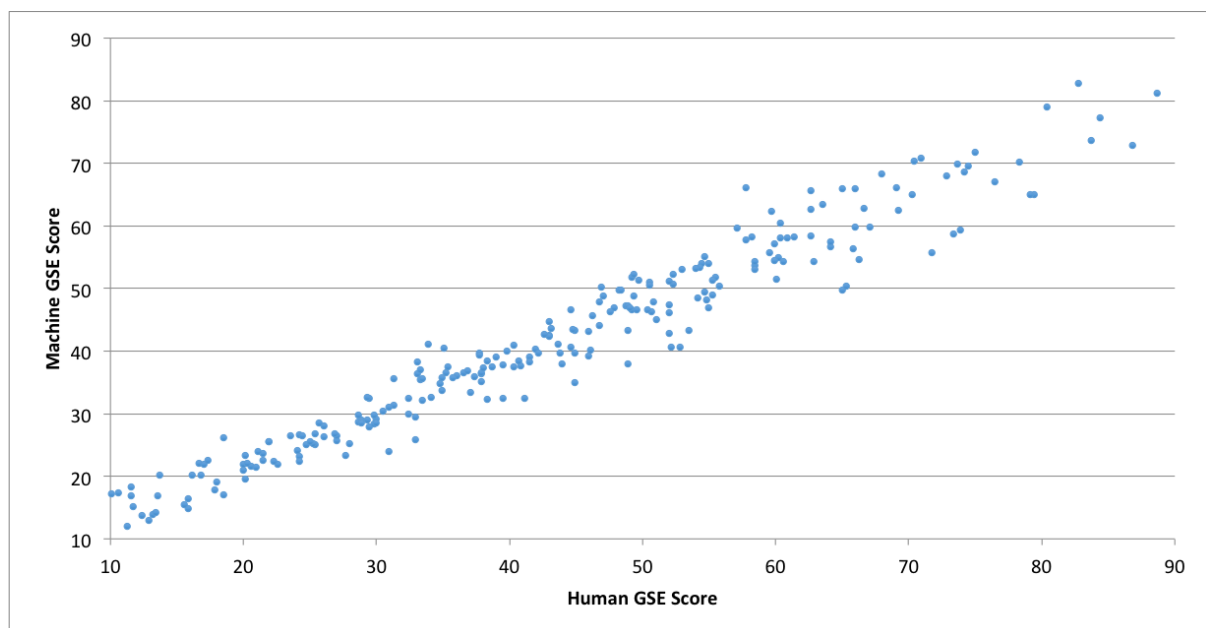
Once the automated scoring system was developed, the responses from the validation set were run through the same psychometric model to produce an Overall and six skill scores for each candidate. Those human and machine scores were then correlated to compare how similar those two kinds of scores are for each person. When candidates were identified as having extreme scores (i.e., well outside the reported score range of the GSE and not well estimated), or when they had fewer than five responses which were able to be scored in a particular skill area, their scores were excluded from the analyses. This reduced the n-count for the Overall score correlation to 288 candidates. The relationship between machine and human Overall scores was found to be a very strong one with a correlation of .97 (see Table 1).

Table 1. Correlations between Progress scores using machine and human scoring methods for Overall and skill reporting areas.

Score Type	Correlation
Overall	.97
Listening	.93
Speaking	.83
Reading	.90
Writing	.99
Grammar	.97
Vocabulary	.93

Machine scoring produces Progress scores that are nearly identical to those that a careful human rating process for many item types would require (see Figure 1).

Figure 1. Scatter plot of GSE scaled scores for validation set candidates using human and machine scoring methods.



Conclusion

Progress is a four-skill English language proficiency test that is delivered online and is scored completely automatically by automated scoring systems. The test consists of a computer adaptive part and a linear form part for an effective assessment of the learner's progress in English language proficiency. The validation analysis demonstrated that the test is highly reliable (i.e., the test-retest reliability of 0.861) and the scores from the automated scoring systems closely correspond to the scores from careful human raters (i.e., a correlation of 0.97 at the Overall score level).

Test Coverage

The test covers all four language skills; reading, listening, speaking and writing as well as knowledge of grammar and vocabulary.

Skills or Knowledge	Test Focus
Reading	To demonstrate reading skills, learners will be asked to: • read and understand the main points from signs, newspapers and magazines • understand the detail of short texts • understand the detail in longer texts
Listening	To demonstrate listening skills, learners will be asked to: • listen for specific information in listening texts • show understanding of meaning in context and the detail of short dialogues • follow and understand short texts and show understanding by writing down or repeating accurately what was said
Speaking	To demonstrate speaking skills, learners will be asked to: • speak clearly using appropriate stress and intonation • pronounce words so that they can be understood • describe pictures or other visual material connecting ideas together accurately and with a range of language
Writing	To demonstrate writing skills, learners will be asked to: • describe a scene or picture accurately using appropriate vocabulary • write a story accurately linking ideas and using appropriate vocabulary • write a short essay giving different points of view, linking ideas and using accurate and appropriate language
Grammar	To demonstrate knowledge of grammar , learners will be asked to: • choose the right word or phrase to make an accurate sentence • understand the difference between different grammatical tenses and other structures • put words in the right order to make grammatical sense
Vocabulary	To demonstrate knowledge of vocabulary , learners will be asked to: • produce words which relate to common themes and topics such as family, work and social situations • use appropriate words in different contexts • show an understanding of the different meaning of words and how they relate to other words

Test Questions

What kinds of questions are in the test and what do they measure?

The test has a number of different question types. This gives learners a chance to demonstrate their English skills in different ways. There are questions where learners choose the correct option or where they write the answer into an open question. There are questions where the learner repeats or copies what has been said as well as questions where learners describe something or write a short essay. The questions are similar to the questions and tasks learners will have done in the classroom as part of their learning and so should be familiar.

Because Progress is part adaptive, different learners will see different questions and may not be presented with all the questions described below.

Vocabulary Questions

There are three vocabulary question types. Vocabulary is also tested as part of *Describe Image*, *Short Essay* and *Read and then Write* which are Integrated Skills questions.

Question	What do the learners have to do?	What is being tested?
Fill in the Table	This question asks the learner to complete a set of vocabulary items with appropriate words. The words are presented as a table of related words.	This question tests the vocabulary knowledge of the learner. It tests the words the learner knows and the accuracy of the form of the word. It tests the learner's knowledge of word families and related sets of words that they may have met in the classroom or when learning English.
Choose the Right Word or Phrase	This question asks the learner to choose the correct word to complete a number of sentences. The sentences are related by a similar theme.	This question tests the vocabulary knowledge of the learner in a written context. It tests the vocabulary the learner knows and whether they can understand the use of the vocabulary in the context of a sentence. It tests the range of vocabulary the learner knows.
Complete the Dialogue	This question asks the learner to select words from a word bank to complete a dialogue.	This question tests the vocabulary of the learner in a spoken context. It tests the vocabulary the learner knows and whether they can understand the use of the vocabulary in the context of a conversation. It tests the range of vocabulary the learner knows.

Grammar Questions

There are four grammar question types. Grammar is also tested as part of *Short Essay*, and *Read and then Write* which are Integrated Skills questions.

Question	What do the learners have to do?	What is being tested?
Choose the Right Word or Phrase	This question asks the learner to choose the correct word to complete a number of sentences. The sentences are related by a similar theme.	This question tests the knowledge of grammar of the learner. It tests the range of grammatical knowledge as well as the accuracy of grammar in a written context.
Choose the Right Word or Phrase. You may choose more than one.	This question asks the learner to choose from a number of options. They may choose one or more than one answer. The sentences are related by the grammatical structure which is being tested.	This question tests the grammatical knowledge of the learner. It tests words which are related to each other in that they have similar meanings or grammatical uses. It tests grammatical knowledge in a written context.
Drag and Drop	This question asks the learner to re-order a sentence correctly.	This question tests the grammatical knowledge of the learner at sentence level. It tests word order, connectors and discourse markers. It tests grammatical knowledge in a written context.
Error correction	This question asks the learner to select one of the available options to correct the mistake in the sentence.	This question tests knowledge of grammatical rules in use.

Reading Questions

There are four reading question types. Reading is also tested as part of *Read and then Speak*, *Read and then Write*, and *Listen and Read* which are all Integrated Skills questions.

Question	What do the learners have to do?	What is being tested?
Choose the Right Picture	This question asks learners to read a short text and select the best picture to match with the text.	This question tests the global understanding of short messages, notes and short pieces of writing.
Choose the Right Word or Phrase	This question asks learners to read a short text and select the best word or phrase to complete the text.	This question tests the global understanding of short messages, notes and short pieces of writing.
Short Answer	This question asks the learner to read a longer text and answer questions on the text.	This question tests the reading comprehension of the learner. It tests specific information included in the text.
Drag and Drop	This question asks the learner to read a text and select the word or phrase that best completes each gap.	This question tests the global understanding of a sentence and short pieces of writing.

Listening Questions

There is one listening question type which tests only listening. Listening is also tested as part of *Listen and then Write*, *Listen and then Speak*, and *Listen and Read* which are all Integrated Skills questions.

Question	What do the learners have to do?	What is being tested?
Listen to the Conversation and Answer	This question asks the learner to listen to a short conversation and then answer a question about the conversation.	This question tests listening comprehension. It tests the accuracy of the listening comprehension of the learner.

Speaking Questions

There is one speaking question type which tests speaking and grammar. Speaking is also tested as part of *Read and then Speak* and *Listen and then Speak* which are Integrated Skills questions.

Question	What do the learners have to do?	What is being tested?
Describe Image	This question asks the learner to look at a photograph or picture and describe what they see.	This question tests the learner's ability to speak in an extended way linking concepts and ideas. It tests the accuracy of speech including accurate grammar, pronunciation and stress as well as the fluency of the speech. It tests the use of appropriate words to describe the photograph or picture.

Writing Questions

There is one writing question type which tests only writing. Writing is also tested as part of *Listen and then Write* and *Read and then Write* which are Integrated Skills questions.

Question	What do the learners have to do?	What is being tested?
Short Essay	This question asks the learner to write a short essay in response to a prompt.	This question tests global writing skills. It tests paragraph and sentence structure, the range and accuracy of the language used, the ability to structure an argument or discussion in a written context. It tests grammar and vocabulary as an essential part of writing.

Integrated Skills Questions

There are seven questions types which measure more than one skill at the same time. These are called Integrated Skills Questions.

Question	What do the learners have to do?	What is being tested?
Read and then Speak	This question asks the learner to read aloud a sentence or short text.	This question tests accurate pronunciation and how fluent the learner is at speaking. It tests if the words in the text are understood and repeated accurately.
Listen and then Write	This question asks the learner to listen to a sentence or short text and write what they have heard.	This question tests listening comprehension at the word and sentence level. It tests the ability to write accurately and understand sentence structure, word order and connectors.
Listen and then Speak	This question asks the learner to listen to a sentence or short text and then repeat it.	This question tests listening comprehension at the word and sentence level. It tests pronunciation and fluency. It tests if the words heard are understood and repeated accurately.
Read and then Write	This question asks the learners to read a short story or short piece of factual text. The text then disappears and the learner has to reconstruct the text.	This question tests reading comprehension. It tests the ability to write accurately and understand sentence structure, word order and connectors.
Listen and Read	This question asks the learner to read a text and at the same time listen to the text. The learner has to find the differences between the written text and the spoken text.	This question tests reading and listening comprehension. It tests the ability to recognise individual words in a text.

Listen and Speak	This questions asks the learner to listen to a short narrative and then retell the narrative using their own words.	This questions tests listening and speaking. It assesses understanding of a short narrative.
Passage Comprehension	This question asks the learner to listen to a short passage and answer three comprehension questions with a short spoken response.	This question tests listening comprehension. It tests understanding of a short passage.

Test Familiarity

Learners can take the sample test at any time to familiarise themselves with the question types in the test. Teachers can also assign this test if they want to ensure learners do this as a homework activity.

Question Type and Level

Most questions are used across the levels but some questions are more appropriate for students at A1 or for students with a high proficiency level. The table over shows how questions are distributed in relation to Common European Framework (CEFR) levels.

		Progress Level					
Item Type	Skill	15-30	25-40	35-50	45-60	55-70	65-80
Fill in the table	Vocabulary						
Choose the right word or phrase	Vocabulary						
Complete the Dialogue	Vocabulary						
Choose the right word or phrase	Grammar						
Choose the right word or phrase. You may choose more than one.	Grammar						
Drag and Drop	Grammar						
Error Correction	Grammar						
Choose the right picture	Reading						
Choose the right word or phrase	Reading						
Short Answer	Reading						
Drag and Drop	Reading						
Listen to the Conversation	Listening						
Describe image	Speaking						
Short Essay	Writing						
Read and then Speak	Speaking & Reading						
Listen and then Write	Listening & Writing						
Listen and then Speak	Listening & Speaking						
Read and then Write	Reading & Writing						
Listen and Read	Listening & Reading						
Listen and Speak	Listening & Speaking						
Passage Comprehension	Listening & Speaking						

Score Reporting

Score Reporting

The Progress Score Report consists of an overall score, and skill scores for communicative and enabling skills. All scores are reported on the Global Scale of English (GSE) which ranges from 10 to 90. Scores are also displayed on the Common European Framework of Reference (CEFR).

Overall score: This reflects a candidate's overall English language ability. The score is based on performance on all items in the test. On average, each test presents between 60-65 items across 17 item types.

Skill scores (listening, reading, speaking, writing, vocabulary and grammar): These scores are based on test items that assess those particular skills, either as a single skill or integrated skill tasks. For vocabulary, the score relates to knowledge and use of appropriate choice of words used to express meaning, as well as lexical range. For grammar, the score relates to knowledge and use of appropriate language with respect to word form and word order.

Progress is a formative assessment instrument and it is not expected to be taken at a secure test centre. As such, the intended score use is not for certification of the student's English proficiency level; it is for supporting the student's learning.

Progress is designed as a test to be taken at the start, middle and end of a course of study. It accurately measures the progress made by a student over the course duration in order to focus teaching and learning. Once a candidate has completed a test, scores are reported in the gradebook. Teachers and students can then view their scores from the most recent test, along with scores from their earlier sittings.

Scores from the test given at the start and middle of a course can be used to understand the strengths and weaknesses of a student or class in order to better target their learning needs. Tests at the middle and end of a course demonstrate how much progress has been made.

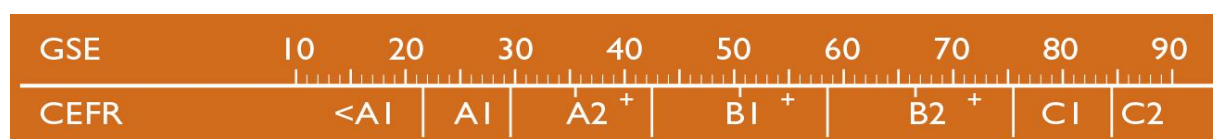
The overall score is the key indicator of a student's English proficiency. The skill scores give an indication of how well the student did on questions aimed at that particular skill. The skill scores can be used by the teacher in conjunction with the 'Can-Do dashboard.' This diagnostic tool identifies particular learning objectives which the student could work on. Teachers will also use other information they have about the student, performance in class etc., to further utilize and make relevant the information from the dashboards.

The Global Scale of English

The test result provides scores on the Global Scale of English which ranges from 10 to 90. The scores consist of an overall score, and sub scores for listening, reading, speaking, writing, vocabulary and grammar. The test also reports Common European Framework levels.

The Global Scale of English is a numeric, granular scale from 10 to 90 which measures English language proficiency. It enhances the Common European Framework of Reference (CEFR) by showing finer gradations of a learner's level within a CEFR band, and can therefore demonstrate smaller and more precise improvements in a learner's English level.

The Global Scale of English is currently used to report scores on the internationally recognised English language test, PTE Academic™. It is empirically aligned to the CEFR, as described in the *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Council of Europe, 2001), and correlated to other test score scales such as TOEFL® iBT, TOEIC® and IELTS.



Global Scale of English and the Common European Framework levels

In the following tables we define how the Global Scale of English is related to the CEFR levels. To give an impression of what the levels mean, i.e., what learners at particular levels can do, we use the summary descriptors published in the CEFR (Council of Europe, 2001, p. 24).

Global Scale of English	Link to the CEFR levels
GSE 10–21 Global assessment The range on the Global Scale of English from 10 to 21 covers the area of measurable proficiency below the A1 level of the CEFR. It includes the level which North (2000, p. 295) characterises as 'Tourist', corresponding to a range of 13–21 on the GSE, and a still lower ability which North (ibid.) labels 'Smattering'. Neither of these was included in the CEFR, because A1 was "considered the lowest level of generative language use" (Council of Europe, 2001, p. 33) and 'Tourist' and 'Smattering' rely "purely on a very finite rehearsed, lexically organised repertoire of situation-specific phrases" (ibid.). A few descriptors in the range from 10 to 21 have however been included, representing the key steps in learners' progress towards A1.	

GSE 22–29 Global assessment The range on the Global Scale of English from 22 to 29 corresponds to the A1 level of the CEFR. The capabilities of learners at Level A1 have been summarised in the CEFR (Council of Europe, 2001, Table 1, p. 24) as follows:	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.
GSE 30-35 and 36-42 Global assessment The interval on the Global Scale of English from 30 to 35 corresponds to the lower part of the A2 level of the CEFR, while the interval from 36 to 42 corresponds to the upper part of the A2 level, which is also sometimes referred to as the A2+ level. The capabilities of learners at Level A2 have been summarised in the CEFR (Council of Europe, 2001, Table 1, p. 24) as follows:	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

GSE 36–42 and 43–58 Global assessment The interval on the Global Scale of English from 36 to 42 corresponds to the lower part of the B1 level of the CEFR, while the interval from 43 to 58 corresponds to the upper part of the B1 level, which is also sometimes referred to as the B1+ level. The capabilities of learners at Level B1 have been summarised in the CEFR (Council of Europe, 2001, Table 1, p. 24) as follows:	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
GSE 59–66 and 67–75 Global assessment The interval on the Global Scale of English from 59 to 66 corresponds to the lower part of the B2 level of the CEFR, while the interval from 67 to 75 corresponds to the upper part of the B2 level, which is also sometimes referred to as the B2+ level. The capabilities of learners at Level B2 have been summarised in the CEFR (Council of Europe, 2001, Table 1, p. 24) as follows:	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and Independent disadvantages of various options.

GSE 76–84**Global assessment**

The interval on the Global Scale of English from 76 to 84 corresponds to the C1 level of the CEFR. The capabilities of learners at Level C1 have been summarised in the CEFR (Council of Europe, 2001, Table 1, p. 24) as follows:

Can understand a wide range of demanding, longer texts, and recognise implicit meaning.

Can express him/herself fluently and spontaneously without much obvious searching for expressions.

Can use language flexibly and effectively for social, academic and professional purposes.

Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.

**ALWAYS
LEARNING**