

**Давтан хэрэглэгдсэн даалгаврын чанарын улирал дамнасан шинжилгээ: Хүндрэлийн зэрэг, ялгах чадвар, буруу хариултын хувилбарын ажиллах байдал**

Ш.Сайнбилэг, Ч.Нямгэрэл, Н.Амгалан, П.Манлайбаатар

Өгүүллийн мэдээлэл**Түүх:**

Хүлээн авсан: 2026.04.02

Засаж сайжруулсан: 2026.04.07

Хэвлэхийг зөвшөөрсөн: 2026.05.19

Түлхүүр үг:

Ерөнхий хими II

Давтан хэрэглэгдсэн даалгавар

Даалгаврын хүндрэлийн зэрэг

Даалгаврын ялгах чадвар

Буруу хариултын хувилбарын

ажиллах байдал

Сонгодог тестийн онол

Хичээлд суурилсан үнэлгээ

ХУРААНГУЙ

Энэхүү судалгаагаар Ерөнхий хими II хичээлийн явцын шалгалтад гурван улирал дараалан хэрэглэсэн олон-сонголтот 41 даалгаврын чанарыг үнэлэв. Судалгаанд 2024-2025 оны намар ($n = 28$), хавар ($n = 52$), 2025-2026 оны намар ($n = 44$)-ын улирлын өгөгдлийг ашиглав. Даалгаврын чанарыг сонгодог тестийн онолын хүрээнд хүндрэлийг зэрэг, ялгах чадвар, буруу хариултын хувилбарын ажиллах байдлаар шинжилж, өөрчлөлт орсон 13 даалгаварт дагалдах хайгуул шинжилгээ хийв. Үр дүнгээс харахад даалгаврын хүндрэлийн зэрэг гурван улирлын хооронд статистик ач холбогдолтой өөрчлөгдсөн бол ялгах чадвар сорилгын түвшинд харьцангуй тогтвортой байв. Засвар орсон даалгавруудын ялгах чадвар эерэг чиглэлтэй өөрчлөгдсөн ч энэ үр дүнг түүврийн хэмжээ, бүлгийн ялгааг харгалзан болгоомжтой тайлбарлах шаардлагатай. Даалгаврын шинжилгээ нь чанарын өөрчлөлт бүх даалгаварт жигд бус, харин даалгавар тус бүрийн онцлогоор илэрч байгааг харуулав. Иймээс олон-сонголтот даалгаврын чанарыг зөвхөн нэг удаагийн гүйцэтгэлээр дүгнэхгүй, харин давтан хэрэглэгдсэн даалгаврын мөрдөлтийг буруу хариултын шинжилгээтэй хамтатган авч үзэх нь нотолгоонд тулгуурласан даалгаврын сан боловсруулахад дэмжлэг болохыг харууллаа.



Энэхүү өгүүлэл нь *Creative Commons Attribution-NonCommercial-ShareAlike 4.0* Олон улсын лицензээр хамгаалагдсан болно

Холбоо барих зохиогч:

Чойжилсүрэнгийн Нямгэрэл

Монгол улсын их сургууль, Шинжлэх ухааны сургууль, Байгалийн ухааны салбар, Химийн тэнхим

Email: nyamgerel@num.edu.mn

A Cross-Semester Analysis of the Quality of Repeatedly Items: Item Difficulty Discrimination, and Distractor Functioning

Sh.Sainbileg, Ch.Nyamgerel, N.Amgalan, P.Manlaibaatar

Article information**Article history:**

Received: 04.02.2026

Revised: 04.07.2026

Accepted: 19.05.2026

Keywords:

General Chemistry II

Repeatedly used items

Item difficulty

Item discrimination

Distractor functioning

Classical test theory

Classroom-based assessment

ABSTRACT

This study evaluated the quality of 41 multiple-choice items used across three consecutive semesters in the midterm assessments of General Chemistry II. Data were collected from the Fall 2024-2025 ($n = 28$), Spring 2024-2025 ($n = 52$), and Fall 2025-2026 ($n = 44$). Item quality was analyzed within the framework of Classical Test Theory using item difficulty, item discrimination, and distractor functioning. Exploratory follow-up analyses were conducted for 13 modified items. The results showed that item difficulty differed significantly across semesters, whereas item discrimination remained relatively stable at the test level. Modified items showed a more positive change in discrimination, but this finding should be interpreted cautiously because of the small number of items and cohort differences. Overall, changes in item quality were not uniform but item-specific. These findings suggest that multiple-choice item quality should not be judged from a single administration alone; rather, tracking repeatedly used items together with distractor analysis can support evidence-informed item-banks improvement.



This is an open access article under the *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License*.

Corresponding author:

Nyamgerel Chojilsuren

Department of Chemistry, School of Science, National University of Mongolia

Email: nyamgerel@num.edu.mn

УДИРТГАЛ

Дээд боловсролын хичээлд суурилсан үнэлгээ нь оюутны мэдлэг, чадварыг хэмжих хэрэгсэл төдийгүй, тухайн хичээлийн суралцахуйн үр дүн хэрхэн хэрэгжиж буйг нотолгоонд тулгуурлан үнэлэх үндэс болдог. Ялангуяа олон-сонголтот даалгавар нь том бүлгээр зохион байгуулах шалгалтад хэрэглэхэд тохиромжтой, оноожуулалтын зардал бага, найдварт чанар өндөр байх боломжтой тул стандартчилагдсан сорилгод өргөн хэрэглэдэг. Гэвч даалгаврын хэмжилзүйн чанар, буруу хариултын хувилбарын ажиллах байдал сул байх нь сорилгын найдвартай байдал, хүчинтэй байдлын талаарх нотолгоо суларч, улмаар үнэлгээний шийдвэрийн чанарт сөргөөр нөлөөлөх эрсдэлтэй. Иймд даалгаврын болон буруу хариултын шинжилгээ нь хичээлийн агуулга, арга зүй, сорилгын чанарыг нотолгоонд тулгуурлан сайжруулах чиглэлийг оновчтой тогтоох боломж олгодог.

Олон улсын судалгаанд олон-сонголтот даалгаварт тавигдах шаардлага, буруу хариултыг боловсруулах зарчим, химийн ойлголтыг оношлох хэрэгслийн чанар, даалгаврын хэмжилзүйн шинжилгээний асуудлууд өргөн судлагдсан байдаг. Харин хичээлд суурилсан үнэлгээнд ашигласан даалгаврыг улирал дамнан давтан мөрдөж, даалгаврын хэмжилзүйн чанарыг системтэй шинжилсэн судалгаа харьцангуй ховор байна. Энэ нь нэг удаагийн шалгалтын хүрээнд “сайн” гэж үнэлэгдсэн даалгавар дараагийн улиралд ч ижил хэмжилзүйн чанартай байх эсэхийг тодруулах хэрэгцээ байгааг харуулж байна. Улмаар багшийн боловсруулж хэрэглэж буй сорилго болон даалгавар нь оюутны мэдлэг, чадварыг зөв хэмжиж, хичээлийн суралцахуйн үр дүнг зөв үнэлж чадаж байна уу? гэдэгт нотолгоонд суурилан хариулах боломжтой болно.

Иймд бид Ерөнхий хими II хичээлийн явцын шалгалтад ашигласан даалгавруудыг гурван улирал дараалан мөрдөж, даалгаврын хэмжилзүйн үзүүлэлтүүдийн өөрчлөлтийн хэв шинжийг тодорхойлохыг зорив. Энэхүү судалгаанд дараах асуултуудад хариу эрэв.

СА1. Гурван улирал дараалан давтан хэрэглэгдсэн олон-сонголтот даалгавруудын хүндрэлийн зэрэг, ялгах чадварын өөрчлөлтийн хэв шинж хэрхэн илэрсэн бэ?

СА2. Өөрчлөлт дагуулсан даалгавар, өөрчлөлтгүй хэрэглэгдсэн даалгаврын хүндрэлийн зэрэг, ялгах чадварын өөрчлөлтийн хэв шинж хоорондоо ялгаатай байсан уу?

СА3. Өөрчлөлт дагуулсан даалгавруудын хүндрэлийн зэрэг, ялгах чадвар улирал хооронд хэрхэн өөрчлөгдсөн бэ?

СА4. Өөрчлөлт дагуулсан даалгавар тус бүрийн ялгах чадвар, буруу хариултын хувилбарын ажиллах байдал хэрхэн өөрчлөгдсөн бэ?

СЭДВИЙН СУДЛАГДСАН БАЙДАЛ

Боловсролын үнэлгээний орчин үеийн үзэл баримтлалд сорилгыг зөвхөн оноо гаргах хэрэгсэл бус, харин суралцагчийн мэдлэг, чадварын талаар нотолгоо цуглуулах, тайлбарлах, шийдвэр гаргахад ашиглах тогтолцоо гэж үздэг (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014; National Research Council, 2001).

Сонгодог тестийн онол (Classical Test Theory - CTT)-оор сорилгын оноо нь бодит оноо болон хэмжилтийн алдааны нийлбэр гэж үзэгддэг бөгөөд даалгаврын хүндрэлийн зэрэг, ялгах чадвар, сорилгын дотоод нийцэл зэрэг үзүүлэлтүүд нь сорилгын чанарын талаарх үндсэн нотолгоо болдог (Fan, 1998; Novick, 1966). Олон-сонголтот даалгаварт суурилсан сорилгын дотоод нийцлийг илэрхийлэхэд Кудер-Ричардсоны 20 коэффициент (Kuder-Richardson 20 coefficient KR-20) болон Кронбахын альфа коэффициент өргөн хэрэглэгддэг. Гэхдээ найдвартай байдлын коэффициент өндөр байх нь даалгавар бүр сайн ажиллаж байна гэсэн үг биш (Barbera & VandenPlas, 2011; Downing, 2004). Сорилго дахь даалгаврын хүндрэлийн зэрэг, ялгах чадвар нь анхан шатны шинжилгээ, сорилгын чанарын хяналт, даалгаврын сангийн сайжруулалтад өргөн ашиглагддаг (Fan, 1998; Novick, 1966). Олон-сонголтот даалгавар боловсруулах удирдамжийг нэгтгэсэн судалгаанууд даалгаврын чанарыг зөвхөн зөв хариултаар бус, харин түүний өгүүлбэр, буруу хариултуудын зохиомж, хэмжилзүйн үзүүлэлтүүдийн уялдаагаар үнэлэх ёстойг харуулсан (Haladyna et al., 2002). Химийн боловсролын салбарт энэ чиглэлийн судалгаа сүүлийн жилүүдэд эрчимтэй өргөжсөн байна.

Breakall (2019) нар ерөнхий химийн контекстэд олон-сонголтот даалгавар боловсруулахад гардаг доголдлыг системтэй үнэлэх хэрэгсэл боловсруулсан нь зөвхөн даалгаврын үр дүнг биш, түүний зохиомжийн чанарыг давхар авч үзэх шаардлагатайг тодруулжээ.

Ерөнхий хими хичээлийн шалгалтын чанарыг судалсан бүтээлүүдэд даалгаврын хэмжилзүйн үзүүлэлтийг суралцахуйн үр дүн, агуулгын хамрах хүрээ, багшийн үнэлгээний практиктай холбон авч үзэх хандлага хүчтэй байна (Clark et al., 2022; Hartman & Lin, 2011; Towns, 2014).

Sorenson, Hanson (2021) нар СТТ болон Рашийн загварыг хослуулан ерөнхий хими хичээлийн шалгалтын даалгаврыг давтан шинжилж, улирлаас улиралд шалгалтын дундаж оноо өөрчлөгдөх нь заавал оюутны чадварын өөрчлөлтийг илтгэхгүй, даалгаврын хүндрэлийн зэрэг өөрчлөгдсөнтэй холбоотой байж болохыг харуулсан. Хоёр болон гурван удаагийн давтан ашиглалтын дараа даалгаврын болон сорилгын чанар бараг бүх тоон үзүүлэлтээр мэдэгдэхүйц сайжирсан нь ажиглагдсан хэмээжээ. Энэ судалгаа нь багш өөрийн хичээлд ашигладаг сорилгын даалгаварт шинжилгээ хийж, сорилго болон хичээл сургалтын үйл ажиллагаагаа сайжруулж болдгийг баталсан судалгаа болжээ. Мөн даалгаврыг дахин хэрэглэх үед хүндрэлийн зэрэг, ялгах чадвар, буруу хариултын үр нөлөө өөрчлөгдөж болохыг хэд хэдэн судалгаа тэмдэглэсэн байна (Demircioğlu & Demircioğlu, 2016; Rezigalla et al., 2024).

Буруу хариултын шинжилгээ нь олон-сонголтот даалгаврын чанарын судалгаанд чухал байр суурь эзэлж байна. Буруу хариултуудын ажиллах байдал нь зөвхөн даалгаврын хүндрэлийн зэрэг, ялгах чадварт нөлөөлөөд зогсохгүй, суралцагчдын түгээмэл алдаа, ташаа ойлголтыг илрүүлэх оношилгооны үнэ цэнэтэй үзүүлэлт юм (Gierl et al., 2017). Зарим судалгаа нийт шалгуулагчдын маш бага хувь сонгосон буруу хариулт (Non-Functioning Distractor - NFD) түгээмэл тохиолддогийг, тэдгээрийг засах эсвэл солиход сорилгын чанар сайжрах боломжтойг харуулсан байна (Ali et al., 2016; Tarrant et al., 2009). Schaubert нар (2010) үүрэггүй буруу хариултыг хасах эсвэл засварлах нь даалгаврын хүндрэлийн зэрэгт нөлөөлж болохыг харуулсан бол Rezigalla нар (2024) үүрэггүй буруу хариултад даалгаврын зохиомж, агуулгын зураглал, давтан хэрэглээ зэрэг хэд хэдэн хүчин зүйл нөлөөлдгийг тэмдэглэсэн. Мөн олон-сонголтот даалгаварт буруу хариултын хувилбарын тоо чанарын баталгаа болдоггүй, харин буруу хариулт үнэхээр ажиллаж, зөв хариулттай өрсөлдөж чадаж байгаа эсэх нь илүү чухал болохыг Rodriguez (2005)-ийн мета-анализ тодруулжээ.

Азийн болон бусад орнуудын ижил төстэй судалгаа СТТ-д суурилсан даалгаврын шинжилгээ өргөн хэрэглэгдэж байгааг харуулдаг ч тэдгээрийн дийлэнх нь нэг удаагийн шалгалт, хөндлөн огтлол (cross-sectional)-ын өгөгдөлд төвлөрсөн байна (Muangkhoua, 2017; Sulistyono et al., 2020). Зарим судалгаанд багшийн боловсруулсан шалгалтад даалгаврын хэмжилзүйн үзүүлэлтийн доголдол түгээмэл байгааг тогтоосон нь хичээлд суурилсан үнэлгээнд даалгаврын шинжилгээг системтэй хийх шаардлагатайг баталж байна (Adegoke et al., 2018). Нөгөө талаас Зүүн Азийн шалгалт төвтэй соёлын нөхцөлд багш нарын үнэлгээний практик нь явцын үнэлгээг суралцахуйг дэмжих хэрэгсэл гэхээс илүү амжилтыг баталгаажуулах хэрэгсэл гэж ойлгох хандлагатай байж болохыг зарим судалгаанд тэмдэглэжээ (Xiaoqiong & Xianghong, 2022). Монголын их сургуулийн орчинд ч энэ хандлага илэрч болох бөгөөд даалгаврын шинжилгээг хийх нь энэ хандлагыг бууруулах нэг хүчин зүйл байх боломжтой.

Дээрх судалгаануудыг нэгтгэн үзвэл, одоогийн судалгааны хоосон орон зайг дараах гурван түвшинд тодорхойлж болно. Нэгдүгээрт, ерөнхий хими хичээлийн олон-сонголтот даалгаврын чанарыг дараалсан улирлуудад давтан мөрдсөн судалгаа ховор байна. Хоёрдугаарт, засварлагдсан болон засварлагдаагүй даалгаврын хүндрэлийн зэрэг, ялгах чадварыг буруу хариултын хувилбарын ажиллах байдалтай хамтатган харьцуулсан судалгаа цөөн байна. Гуравдугаарт, СТТ-ийн дээж хамааралт шинжийг хүлээн зөвшөөрсөн нөхцөлд, бодит хичээлийн өгөгдөл дээр сорилгын болон даалгаврын нотолгоог нэгтгэн тайлбарласан химийн боловсролын судалгаа Монголын дээд боловсролын орчинд бараг ажиглагдахгүй байна.

Иймээс энэ судалгаа нь хичээлд суурилсан үнэлгээний практикт чухал боловч судалгаа хангалттай биш буй давтан хэрэглэгдсэн даалгаврын мөрдөлт гэсэн асуудлыг химийн боловсролын бодит өгөгдөл дээр нотолгоонд суурилсан байдлаар авч үзэж буйгаараа онцлогтой.

ОНОЛЫН ҮНДЭС

СТТ-ийн даалгаврын хүндрэлийн зэрэг, ялгах чадвар зэрэг хэмжилзүйн үзүүлэлтүүд тухайн шалгуулагчдын бүлгээс хамаардаг буюу дээж хамааралт шинжтэй (Fan, 1998; Zenisky et al., 2013).

Даалгаварт хариулах онол (Item Response Theory - IRT) нь дээж хамааралт байдлыг бууруулах боломжтой боловч тогтвортой параметр тооцоолоход илүү их хэмжээний өгөгдөл шаарддаг тул хичээлд суурилсан жижиг болон дунд хэмжээний дээжтэй судалгаанд СТТ илүү бодитой сонголт хэвээр байна (Fan, 1998; Krishnan, 2010; Mango, 2009). Иймд СТТ-д суурилсан урт хугацааны мөрдөлт нь арга зүйн хязгаарлалттай боловч хичээлд суурилсан үнэлгээний практикт өндөр ач холбогдолтой чиглэл гэж үзэж болно. Гэхдээ улирал хоорондын өөрчлөлтийг тайлбарлахдаа даалгаврын засвар, хичээл заалтын өөрчлөлт, оюутны бүлгийн ялгаа, сорилгын бүтэц зэрэг хүчин зүйлсийг хамтатган авч үзэх шаардлагатай.

Давтан хэрэглэгдсэн даалгаврын мөрдөлтийг бодит хичээлийн өгөгдөл дээр шинжлэхэд СТТ-ийг практикийн хувьд хамгийн тохиромжтой хүрээ гэж үзсэн тул сорилгын найдвартай байдлын нотолгоо болгон KR-20 коэффициентийг, даалгаврын чанарын нотолгоо болгон p индекс, d индекс, буруу хариултын хувилбарын ажиллах байдлыг ашиглав. Мөн СТТ-ийн үзүүлэлт дээж хамааралт шинжтэй гэдгийг онолын түвшинд хүлээн зөвшөөрсөн. Өөрөөр хэлбэл, нэг даалгаврын хүндрэлийн зэрэг, ялгах чадвар нь зөвхөн тухайн улиралд шалгалт өгсөн оюутны бүлэгтэй холбоотой утгатай тул улирал хоорондын ялгааг тогтмол шинж чанарын өөрчлөлт гэж шууд тайлбарлах боломжгүй. Иймээс энэ судалгаанд улирал хоорондын өөрчлөлтийг шалтгаант нөлөө гэж бус, харин сорилгын болон даалгаврын түвшинд ажиглагдсан психометрийн хэв шинжийн өөрчлөлт гэж үзэн тайлбарлав.

АРГА ЗҮЙ

Судалгааны дизайн

Судалгаа нь хичээлд суурилсан үнэлгээний өгөгдөлд тулгуурласан, давтан хэрэглэгдсэн даалгаврыг шинжилгээний үндсэн нэгж болгон авч үзсэн, урт хугацааны харьцуулсан тоон судалгаа юм. Судалгаагаар ижил даалгаврыг өөр өөр улиралд хэрэглэсэн үед тэдгээрийн хэмжилзүйн үзүүлэлтүүд хэрхэн өөрчлөгдөж буйг тодорхойлж байгаагаараа онцлогтой.

Судалгааны орчин ба оролцогчид

Судалгаанд МУИС-ийн Шинжлэх Ухааны Сургууль (ШУС) болон Инженер Технологийн Сургууль (ИТС)-ийн хөтөлбөрөөр суралцаж, Ерөнхий хими II хичээлийг судалж буй оюутнуудын явцын шалгалтын өгөгдлийг ашигласан. Улирал бүрийн шалгалтад өөр өөр оюутны бүлэг хамрагдсан.

Хүснэгт 1. Ерөнхий хими II хичээлийг сонгон судалсан, шалгалтад оролцсон оюутны тоо, улирлаар

Улирал	Сонгон судалсан	Явцын шалгалт өгсөн	Оролцоо, %
2024-2025 намар	34	28	82.3
2024-2025 хавар	59	52	88.1
2025-2026 намар	47	44	93.6

Тайлбар. Оролцоо (%) нь явцын шалгалт өгсөн оюутны эзлэх хувийг илэрхийлнэ.

Судалгааны материал

Судалгаанд Ерөнхий хими II хичээлийн 1-8 дугаар долоо хоногийн агуулгыг хамарсан явцын шалгалтын өгөгдлийг ашиглав. Явцын шалгалтыг олон-сонголтот даалгавар бүхий сорилго хэлбэрээр авсан бөгөөд сорилго нь 2024-2025 оны намрын улиралд 25 даалгаврын өгөгдөл, нийт 44 даалгавартай, 2024-2025 оны хавар болон 2025-2026 оны намрын улиралд тус бүр 25 даалгаврын өгөгдөл, нийт 41 даалгавартай байв. Эхний улирлын шинжилгээний дараа 3 даалгаврыг хассан тул гурван улиралд нийт 41 давтан хэрэглэсэн даалгаврыг үндсэн шинжилгээнд ашиглав.

Хэмжилтийн үзүүлэлтүүд

Энэхүү судалгаанд дараах хэмжилтийн үзүүлэлтүүдийг ашиглав.

Даалгаврын хүндрэлийн зэрэг (p индекс). Тухайн даалгаварт зөв хариулсан шалгуулагчдын хувийг илэрхийлнэ.

Даалгаврын ялгах чадвар (d индекс). Дээд болон доод бүлгийн зөв хариултын хувийн ялгааг илэрхийлнэ. d индекс өндөр байх тусам тухайн даалгавар өндөр ба бага гүйцэтгэлтэй шалгуулагчдыг илүү сайн ялгаж байна гэж үзэв.

Буруу хариултын хувилбарын ажиллах байдал. Хариултын сонголтын тархалт болон тухайн буруу хариултыг сонгосон шалгуулагчдын хувь дээр үндэслэн үнэлэв. Судалгаанд нийт шалгуулагчдын 5-аас бага хувь сонгосон хариултыг үүрэггүй буруу хариулт гэж үзэв.

Даалгаврын ангилал

Улирал бүрийн даалгаврын шинжилгээний үр дүнд багш нарын гаргасан шийдвэрийг үндэслэн даалгаврыг хэвээр үлдээх, засварлах (даалгаврын агуулга засах, хариулт засах), хасах, хичээлийн агуулга-арга зүйн оролцоо хийх, эсвэл даалгаврын засвар ба агуулга-арга зүйн оролцоог хамтад нь хэрэгжүүлэх гэсэн төлөвөөр ангилсан. Үүний дараа даалгавруудыг өөрчлөлт дагуулсан даалгавар (ӨДД), өөрчлөлтгүй хэрэглэсэн даалгавар (ӨГД) гэсэн хоёр бүлэгт нэгтгэн харьцуулсан.

Судалгааны үйл явц

Нэгдүгээрт, Ерөнхий хими II хичээлийн суралцахуйн үр дүн, агуулгад нийцсэн олон-сонголтот даалгаврыг боловсруулсан. Даалгаврыг хичээлийн хамтлагийн багш нар хамтран хэлэлцэж, техникийн, агуулгын, хэл найруулгын алдааг засаж сайжруулсан.

Хоёрдугаарт, боловсруулсан сорилгыг гурван улирал дараалан явцын шалгалтад ашиглаж, оюутны хариултын өгөгдлийг цуглуулсан.

Гуравдугаарт, улирал бүрийн даалгаврын шинжилгээг хийж, р индекс, d индекс, буруу хариултын хувилбарын ажиллах байдлыг тооцоолсон. Үр дүнд үндэслэн зарим даалгаврыг хэвээр үлдээх, заримыг засварлах, хасах, эсвэл тухайн агуулга дээр хичээл заалтын өөрчлөлт хийх шийдвэр гаргасан.

Дөрөвдүгээрт, давтан хэрэглэгдсэн даалгавруудын улирал дамнасан өгөгдлийг нэгтгэн, харьцуулсан болон даалгаврын түвшний шинжилгээ хийсэн.

Олон сонголттой даалгавар боловсруулах үе шат

Хичээлийн хамтлагийн багш бүр Ерөнхий хими II хичээлийн тодорхой суралцахуйн үр дүнг хэрэгжүүлэхэд чиглэсэн, 1-8 дугаар долоо хоногт судлах агуулгыг хамарсан олон-сонголтот даалгавар зохиож, хичээлийн хамтлагийн багш нарын хурлаар хоёр талт хяналтын аргаар шүүн тунгааж, техникийн, хэл найруулгын, онолын алдаагүй байдлыг бататгасан хэлэлцүүлгээр эцэслэн боловсруулсан. Блумын таксономид үндэслэн доод эрэмбийн (санах, ойлгох, хэрэглэх), дээд эрэмбийн (задлан шинжлэх, үнэлэх) чадварыг үнэлэх даалгаврын тэнцвэрийг хангасан. Мөн олон-сонголтот даалгавар боловсруулахад тавигддаг шаардлагын дагуу оюутнуудын дунд түгээмэл тохиолддог ташаа ойлголтуудыг илрүүлэхүйц буруу хариултыг төлөвлөж даалгаврын чанарыг анхаарч боловсруулсан.

Өгөгдөл цуглуулах үе шат

Боловсруулсан олон-сонголтот даалгавар бүхий сорилгыг 2024-2025 оны намрын улирлаас эхлэн 3 улирал дамнасан Ерөнхий хими II хичээлийн явцын шалгалт авахад давтан ашиглаж судалгааны өгөгдлийг цуглуулсан. Улирал бүрд шалгалтыг нэг ижил удирдамжийн дагуу хичээлийн хамтлагийн багш нарын хяналт дор авсан.

Өгөгдлийн боловсруулалт ба статистикийн шинжилгээ хийх үе шат

Улирал бүрийн даалгаврын хэмжилзүйн үндсэн үзүүлэлтүүд (хүндрэлийн зэрэг, даалгаврын ялгаа чадвар, буруу хариултын хувилбарын ажиллах байдал)-ийг DEAnalysis программ ашиглан тооцоолсон.

Статистикийн шинжилгээг jamovi 2.6.44 программ ашиглан хийв. jamovi нь R-д суурилсан нээлттэй эхийн статистикийн программ бөгөөд ерөнхий шугаман болон ерөнхийшүүлсэн шугаман загварын шинжилгээнд GAMLj, нэмэлт боловсруулалтад Rj модуль ашиглав.

Судалгааны асуулт 1-ийн хүрээнд гурван улиралд давтан хэрэглэгдсэн 41 даалгаврын р индекс болон d индексийн улирал хоорондын ялгааг Фридманы шинжилгээгээр шалгав. Статистикийн хувьд ач холбогдолтой ялгаа илэрсэн тохиолдолд улирал хоорондын хос харьцуулалтыг Дёрбин–Коновер аргаар хийв.

Судалгааны асуулт 2-ын хүрээнд “Өөрчлөлт дагуулсан даалгавар”, “Өөрчлөлтгүй хэрэглэгдсэн даалгавар”-ын бүлгүүдийн р болон d индексийн өөрчлөлтийн хэмжээг харьцуулахад Манн–

Уитнигийн U шинжилгээг үндсэн арга болгон, хараат бус хоёр бүлгийн t-шинжилгээг нэмэлт шалгалтын арга болгон ашиглав.

Судалгааны асуулт 3-ын хүрээнд ӨДД-ын үзүүлэлтүүдийн өөрчлөлтийг даалгавар тус бүрээр харгалзуулан авч үзсэн. Даалгавар тус бүрийн хүндрэлийн зэргийн улирал хоорондын өөрчлөлтийг хосолсон t-шинжилгээ болон Вилкоксоны тэмдэгт зэрэглэлийн шинжилгээгээр үнэлэв. Даалгавар тус бүрийн ялгах чадварын улирал хоорондын өөрчлөлтийг илүү нарийвчлан судлахын тулд биномиаль ерөнхийшүүлсэн шугаман загвар (Binomial generalized linear model) тохируулж, зөв хариулсан эсэхийг хамаарах хувьсагч, улирал, дээд/доод бүлэг, улирал×бүлгийн харилцан үйлчлэлийг тайлбарлагч хувьсагчаар оруулав. Улирал×бүлгийн харилцан үйлчлэлийн коэффициентийг тухайн даалгаврын ялгах чадвар улирал хооронд өөрчлөгдсөн эсэхийн гол нотолгоо болгон авч үзэв. Даалгавар бүрийн буруу хариултын хувилбарын ажиллах байдлын улирал хоорондын өөрчлөлт буюу сонголтын тархалт улиралтай хамааралтай эсэхийг Пирсоны хи-квадрат шинжилгээгээр шалгаж, нөлөөний хэмжээг Крамерийн V коэффициентээр тайлбарлав.

Бүх шинжилгээнд статистикийн ач холбогдлын түвшнийг $p < .05$ гэж үзэв. Даалгаврын түвшний олон харьцуулалтыг хайгуул шинжтэй нотолгоо гэж үзсэн тул даалгаврын түвшний p -утгуудыг түүврийн онцлог, статистикийн хязгаарлалтыг харгалзан тайлбарлав.

Ёс зүйн асуудал

Судалгаанд ашигласан өгөгдлийг хичээлийн хэвийн заалт, үнэлгээний үйл явцын хүрээнд цуглуулсан. Өгүүлэлд хувь хүнийг таних боломжтой мэдээлэл оруулаагүй бөгөөд өгөгдлийг нэгтгэсэн болон даалгавар тус бүрийн түвшинд боловсруулсан.

ҮР ДҮН

Сорилгын ерөнхий үзүүлэлт ба дотоод нийцэл

Явцын шалгалтын сорилгын ерөнхий үзүүлэлт болон дотоод нийцлийг улирал тус бүрээр тооцоолсон үр дүнг Хүснэгт 2-т үзүүлэв. KR-20 коэффициент 2024-2025 оны намарт .790, 2024-2025 оны хаварт .829, 2025-2026 оны намарт .826 байсан нь сорилгын дотоод нийцэл улирал бүрд хангалттай түвшинд байсныг харуулж байна. Дундаж оноо харгалзан 27.86/44, 25.10/41, 27.89/41 байсан бөгөөд нийт боломжит оноонд харьцуулсан дундаж хувь нь тус тус 63.3%, 61.2%, 68.0% байв. Хэмжилтийн стандарт алдаа улирал хооронд бага зэрэг буурсан нь сорилгын онооны тогтвортой байдал тодорхой хэмжээнд сайжирсныг илтгэж болно. Гэвч сорилгын даалгаврын тоо болон шалгуулагчдын бүлгийн бүрэлдэхүүн улирал хооронд ялгаатай байсан тул эдгээр үзүүлэлтийг түүврийн онцлог, статистикийн хязгаарлалтыг харгалзан тайлбарлах шаардлагатай.

Хүснэгт 2. Явцын шалгалтын сорилгын ерөнхий үзүүлэлт ба дотоод нийцэл

Үзүүлэлт	2024-2025 намар	2024-2025 хавар	2025-2026 намар
Оюутны тоо (n)	28	52	43
Даалгаврын өгөгдөл	25	24	24
Даалгаврын тоо	44	41	41
KR-20	.790	.829	.826
Дундаж оноо	27.86	25.10	27.89
Дундаж, %	63.3	61.2	68.0
Стандарт хазайлт (SD)	6.21	6.68	6.25
Хамгийн бага оноо/хамгийн их оноо	16/38	3/39	13/39
Хэмжилтийн стандарт алдаа (SEM)	2.84	2.76	2.61

Тайлбар. KR-20 = Kuder-Richardson Formula 20; SD = стандарт хазайлт; SEM = хэмжилтийн стандарт алдаа. SEM-ийг $SD \times \sqrt{(1 - KR20)}$ томъёогоор тооцоолов. Эхний улирал болон дараагийн хоёр улиралд сорилгын нийт даалгаврын тоо ялгаатай байсан тул улирал хоорондын төв хандлагыг тайлбарлахдаа дундаж хувийг хамтатган авч үзэв.

Даалгаврын шинжилгээний үр дүн

Эхний улирлын шинжилгээний үр дүнд сорилгоос 3 даалгаврыг хассан тул гурван улиралд давтан хэрэглэгдсэн 41 даалгаврыг үндсэн шинжилгээнд ашиглав. Эдгээрээс 13 нь өөрчлөлт дагуулсан даалгавар байв. Өөрчлөлт дагуулсан даалгавруудын р индекс, d индекс, буруу хариултын хувилбарын ажиллах байдал болон багшийн гаргасан шийдвэрийн мэдээллийг Хүснэгт 3-д нэгтгэн үзүүлэв.

Даалгаврыг төлөвөөр нь ангилсан байдал (Багш нарын тэмдэглэлээс жишээлэн харуулав)

Засварлахаар шийдвэрлэсэн даалгавар (ЗД)-ын жишээ

Даалгавар 29. 2024-2025 оны хичээлийн жилийн хаврын улиралд даалгаврын В хариулт үүргийн сааталтай гарсан нь ОСДТШ-ын 8.7 биелэгдээгүй (сургалтын явцад багшийн хийх ажиглалт чухал) болохыг илтгэж байна. Мөн В хариулт ганцаараа 10-ын зэрэг эерэг байгаа нь нөлөөлсөн байж болох юм. Иймд хариултын хувилбарыг $1 \cdot 10^{-14} / 1.8 \cdot 10^{-5} = 5.5 \cdot 10^{-5}$ гэж өөрчлөх.

Хүснэгт 3. Өөрчлөлт дагуулсан даалгаврын р индекс, d индекс, буруу хариултын хувилбарын ажиллах байдал, даалгаврын төлөв

Д	2024-2025 намар				2024-2025 хавар				2025-2026 намар			
	р	d	БХ	төлөв	р	d	БХ	төлөв	р	d	БХ	
Д7	64.29	100.00	A (<5%)	ЗД	63.46	57.14	A (<5%)	ХҮД	81.82	36.36	A, D (<5%)	
Д8	75.00	57.14	A (<5%)	ЗД	61.54	42.86	A (<5%)	ХҮД	68.18	18.18	A, B (<5%)	
Д9	89.29	28.57	B, C, D (<5%)	ЗД	75.00	35.71	-	ХҮД	84.09	0.00	B, C (<5%)	
Д29	25.00	28.57	D (<5%)	ЗД	38.46	21.43	--	ЗД	38.64	63.64	-	
Д31	32.14	71.43	B (<5%)	ЗДОД	25.00	42.86	-	ХҮД	45.45	90.91	-	
Д32	60.71	-28.57	C (<5%) В эерэг	ЗДОД	53.85	42.86	-	ХҮД	47.73	90.91	-	
Д33	42.86	71.43	таасан	ЗД	63.46	21.43	D эерэг	ЗД	56.82	63.64	-	
Д35	57.14	85.71	A (<5%)	ХҮД	38.46	7.14	A (<5%)	ЗД	43.18	54.55	A (<5%)	
Д15	42.86	71.43	-	ХҮД	40.38	78.57	Математик тооцоо анхаарах	ОД	65.91	72.73	-	
Д37	46.43	42.86	таасан	ОД	48.08	42.86	таасан	ОД	54.55	45.45	-	
Д38	46.43	14.29	таасан	ОД	53.85	7.14	Дүгнэх боломжгүй	ХҮД	72.73	27.27	-	
Д40	46.43	42.86	C (<5%)	ЗД	34.62	21.43	B (<5%)	ХҮД	54.55	36.36	B (<5%)	
Д4	25.00	28.57	-	ЗДОД	38.46	7.14	A, D эерэг	ЗДОД	38.64	72.73	D (<5%)	

Тайлбар. р = даалгаврын хүндрэлийн зэрэг; d = даалгаврын ялгах чадвар; БХ = үүрэггүй буруу хариулт. ХҮД = хэвээр үлдээсэн даалгавар; ЗД = засварласан даалгавар; ОД = хичээлийн агуулга-арга зүйн оролцоо хийсэн даалгавар; ЗДОД = засвар болон агуулга-арга зүйн оролцоог хамтад нь хэрэгжүүлсэн даалгавар. БХ баганад буруу хариултыг тэмдэглэв.

Хичээлийн оролцоог бий болгосон даалгавар (ОД)-ын жишээ

Даалгавар 15. 2024-2025 оны хичээлийн жилийн хаврын улиралд А-7.1%, В-14.3%, С-35.7% байсан бол хаврын улиралд В, С хариултын хувилбарыг бараг тэнцүү сонгосон байна. Үүнийг оюутны чадварт ялгаа байсан гэж дүгнэх үндэсгүй юм. Зэрэгтэй тоон дээр үйлдэл хийх математик чадвар (тухайлбал, зэрэгт илэрхийллээс үл мэдэгдэх суурийг тооцоолох)-аас түлхүү хамаарч байж магадгүй юм. Иймд урвалын хурдад температурын үзүүлэх нөлөөг Вант-Гоффын тэгшитгэлээр тооцоолох дасгал лекцийн хичээлд тусгах.

Засварласан болон хичээлийн оролцоог бий болгосон даалгавар (ЗДОД)-ын жишээ

Даалгавар 32. 2024-2025 оны хичээлийн жилийн намрын улиралд даалгаврын В үүргийн сааталтай (таасан байх магадлалтай), С ялгах хүч маш сул гарсан. С-д эх бодисуудыг хамтад нь, бүтээгдэхүүн бодисуудыг хамтад нь хувирах хос болгон сонгосон тул сонгогдоогүй байна. В-д нэг эх бодисыг 1 бүтээгдэхүүн бодистой, мөн нөгөө эх бодистой хослуулсан тул сонгоогүй. Иймд А, С хариултын хувилбарыг хасав. D хувилбараас эхний хосыг, В хувилбарыг 2 тусдаа хариулт болгов. Брэнстед Лоурийн онол сэдвийн лекцийн агуулгад жишээ, дасгал нэмэх.

Статистикийн шинжилгээний үр дүн

Гурван улиралд давтан хэрэглэгдсэн нийт 41 даалгаврыг хэмжилзүйн мөрдөлтийн үндсэн нэгж болгон авч үзэв.

Судалгааны асуулт 1. Гурван улиралд давтан хэрэглэгдсэн олон-сонголтот даалгавруудын хүндрэлийн зэрэг, ялгах чадвар чадварын өөрчлөлтийн хэв шинж хэрхэн илэрсэн бэ?

Гурван улиралд давтан хэрэглэгдсэн даалгавруудын хүндрэлийн зэрэгт Фридманы шинжилгээ хийхэд ерөнхий ялгаа статистикийн хувьд ач холбогдолтой ялгаа илэрсэн, $\chi^2(2) = 18.3, p < .001$. Дескриптив үзүүлэлтээс харахад р индексийн дундаж 2024-2025 оны хаврын улиралд хамгийн бага, 2025-2026 оны намрын улиралд хамгийн өндөр байв. Үр дүнг Хүснэгт 4а, б-д харуулав.

Хүснэгт 4а. Давтан хэрэглэгдсэн даалгавруудын р индексийн дескриптив үзүүлэлт

Дескриптив үзүүлэлт		
Улирал	Дундаж	Медиан
2024-2025 намар	64.1	64.3
2024-2025 хавар	60.7	63.5
2025-2026 намар	68	68.2

Тайлбар. р индекс нь тухайн даалгаварт зөв хариулсан шалгуулагчдын хувийг илэрхийлнэ.

Хүснэгт 4б. р индексийн Фридманы шинжилгээний үр дүн

Фридманы шинжилгээ		
χ^2	df	p
18.3	2	<.001

Тайлбар. Фридманы шинжилгээг гурван улиралд давтан хэрэглэгдсэн 41 даалгаврын р индексийг сорилгын түвшинд харьцуулахад ашиглав.

Улирал хоорондын ялгааг тооцоход Дёрбин-Коновер хос харьцуулалтаар бүх хослолд статистикийн хувьд ач холбогдолтой ялгаа илэрсэн (Хүснэгт 5).

Хүснэгт 5. Давтан хэрэглэгдсэн даалгавруудын р индексийн улирал хоорондын хос харьцуулалтын үр дүн

Улирал	Статистик	p
2024-2025 намар - 2024-2025 хавар	2.67	0.009
2024-2025 хавар - 2025-2026 намар	4.78	<.001
2024-2025 намар - 2025-2026 намар	2.11	0.038

Тайлбар. Хос харьцуулалтыг Дёрбин-Коновер аргаар хийв. р-утгыг олон дахин харьцуулалтын нөхцөлд хязгаарлалттай тайлбарлав.

Харин даалгаврын ялгах чадварт Фридманы шинжилгээ хийхэд улирал хооронд статистикийн хувьд ач холбогдолтой ялгаа илрээгүй, $\chi^2(2) = 0.642, p = .725$. Үр дүнг Хүснэгт 6а, б-д харуулав.

Хүснэгт 6а. Давтан хэрэглэгдсэн даалгавруудын d индексийн дескриптив үзүүлэлт

Дескриптив үзүүлэлт		
Улирал	Дундаж	Медиан
2024-2025 намар	38.3	42.9
2024-2025 хавар	39.2	42.9
2025-2026 намар	38.8	36.4

Тайлбар. d индекс нь даалгаврын ялгах чадварыг илэрхийлнэ.

Хүснэгт 6б. р индексийн Фридманы шинжилгээний үр дүн

Фридманы шинжилгээ		
χ^2	df	p
0.642	2	0.725

Тайлбар. Фридманы шинжилгээг гурван улиралд давтан хэрэглэгдсэн 41 даалгаврын d индексийг сорилгын түвшинд харьцуулахад ашиглав.

Судалгааны асуулт 2. Өөрчлөлт дагуулсан, өөрчлөлтгүй хэрэглэгдсэн даалгавруудын хүндрэлийн зэрэг, ялгах чадварын өөрчлөлтийн хэв шинж хоорондоо ялгаатай байсан уу?

Өөрчлөлт дагуулсан даалгавар (ӨДД), өөрчлөлтгүй хэрэглэгдсэн даалгавар (ӨГД)-ын р болон d индексийн өөрчлөлтийн хэмжээг харьцуулахад 2024-2025 оны хавраас 2025-2026 оны намрын улирал хооронд d индексийн өөрчлөлтөд статистикийн хувьд ач холбогдолтой ялгаа илэрсэн. Манн-Уитнигийн U шинжилгээ болон хараат бус хоёр бүлгийн t-шинжилгээ аль аль нь энэ үр дүнг дэмжсэн. Өөрөөр хэлбэл, энэ хугацаанд ӨДД нь ӨГД-аас илүү эерэг өөрчлөлт бүхий ялгах чадвар үзүүлэв. Харин бусад улирал хооронд р болон d индексийн өөрчлөлтийн хэмжээнд статистикийн хувьд ач холбогдолтой ялгаа илрээгүй. (Хүснэгт 7а, б).

Хүснэгт 7а. Өөрчлөлт дагуулсан болон өөрчлөлтгүй хэрэглэгдсэн даалгавруудын өөрчлөлтийн дескриптив үзүүлэлт

Харьцуулсан улирал	Бүлэг	N	р өөрчлөлтийн дундаж	р өөрчлөлтийн медиан	р өөрчлөлтийн SD
2024–2025 намар - 2024–2025 хавар	ӨГД	31	-4.17	-3.30	9.71
2024–2025 намар - 2024–2025 хавар	ӨДД	10	-0.91	-1.65	10.0
2024–2025 хавар - 2025–2026 намар	ӨГД	35	7.27	7.87	9.84
2024–2025 хавар - 2025–2026 намар	ӨДД	6	7.31	2.45	12.6
2024–2025 намар - 2025–2026 намар	ӨГД	10	4.12	10.7	12.4
2024–2025 намар - 2025–2026 намар	ӨДД	31	3.83	5.19	11

Харьцуулсан улирал	Бүлэг	N	d өөрчлөлтийн дундаж	d өөрчлөлтийн медиан	d өөрчлөлтийн SD
2024–2025 намар - 2024–2025 хавар	ӨГД	31	2.30	7.14	29.2
2024–2025 намар - 2024–2025 хавар	ӨДД	10	-3.57	-7.14	30.7
2024–2025 хавар - 2025–2026 намар	ӨГД	35	-6.36	-4.55	25.6
2024–2025 хавар - 2025–2026 намар	ӨДД	6	34.40	42.20	25.5
2024–2025 намар - 2025–2026 намар	ӨГД	10	-0.712	-5.19	22.2
2024–2025 намар - 2025–2026 намар	ӨДД	31	4.16	-7.14	52.8

Тайлбар. ӨДД = өөрчлөлт дагуулсан даалгавар; ӨГД = өөрчлөлтгүй хэрэглэгдсэн даалгавар; SD = стандарт хазайлт. N нь тухайн харьцуулалтад багтсан даалгаврын тоог илэрхийлнэ.

Хүснэгт 7б. ӨГД ба ӨДД бүлэг хоорондын харьцуулалтын үр дүн

Харьцуулсан улирал	Шинжилгээ	р өөрчлөлт	р	d өөрчлөлт	р
2024–2025 намар - 2024–2025 хавар	Хараат бус t	-0.916	.365	0.546	.588
2024–2025 намар - 2024–2025 хавар	Манн-Уитни U	129.0	.439	117.0	.248
2024–2025 хавар - 2025–2026 намар	Хараат бус t	-0.010	.992	-3.61	< .001
2024–2025 хавар - 2025–2026 намар	Манн-Уитни U	95.0	.726	30.5	.006
2024–2025 намар - 2025–2026 намар	Хараат бус t	0.0705	0.944	0.418 ^a	0.678
2024–2025 намар - 2025–2026 намар	Манн-Уитни U	146	0.785	149	0.855

Тайлбар. Манн-Уитнигийн U шинжилгээг үндсэн арга, хараат бус хоёр бүлгийн t-шинжилгээг давтан шалгалтын арга болгон ашиглав.

Судалгааны асуулт 3. Өөрчлөлт дагуулсан даалгавруудын хүндрэлийн зэрэг болон ялгах чадвар улирал хооронд хэрхэн өөрчлөгдсөн бэ?

ӨДД-ыг даалгавар тус бүрээр харгалзуулан шинжлэхэд 2024-2025 оны хавар, 2025-2026 оны намрын улирлын хооронд d индексийн өөрчлөлт статистикийн хувьд ач холбогдолтой байна. Хосолсон t-шинжилгээ энэ өсөлтийг дэмжсэн бол Вилкоксоны тэмдэгт зэрэглэлийн шинжилгээнд ижил чиглэлийн үр дүн ажиглагдав. Харин бусад улирал хооронд р болон d индексийн өөрчлөлтөд статистикийн хувьд ач холбогдолтой ялгаа илрээгүй (Хүснэгт 8).

Хүснэгт 8. ӨДД-ын хосолсон харьцуулалтын үр дүн

Харьцуулсан улирал	Шинжилгээ	р индекс	р	d индекс	р
2024–2025 намар - 2024–2025 хавар	Хосолсон t	0.286	.781	0.368	.722
2024–2025 намар - 2024–2025 хавар	Вилкоксон	30.0	.838	31.0	.343
2024–2025 хавар - 2025–2026 намар	Хосолсон t	-1.42	.214	-3.30	.021
2024–2025 хавар - 2025–2026 намар	Вилкоксон			1.0	.058

2024–2025 намар -2025–2026 намар	Хосолсон t	-1.05	0.32	-0.249	0.809
2024–2025 намар -2025–2026 намар	Вилкоксон	15	0.221	28	1

Тайлбар. Улирал хоорондын дотоод өөрчлөлтийг хосолсон t-шинжилгээ болон Вилкоксоны тэмдэгт зэрэглэлийн шинжилгээгээр үнэлэв.

Даалгаврын түвшний дагалдах хайгуул шинжилгээ

Судалгааны асуулт 4. Өөрчлөлт дагуулсан даалгавар тус бүрийн ялгах чадвар, буруу хариултын хувилбарын ажиллах байдал хэрхэн өөрчлөгдсөн бэ?

Ялгах чадварын өөрчлөлт

ӨДД тус бүрийн өөрчлөлтийг илүү нарийвчлан шинжлэхийн тулд биномиаль ерөнхий шугаман загвараар шалгахад үндсэн 5 хэв шинж ажиглагдсан (Хүснэгт 9).

Хэв шинж 1: ӨДД 4, 32, 35-ын хувьд загварын нийт ач холбогдол статистикийн хувьд батлагдсан. Хүндрэлийн зэрэг улирал хооронд өөрчлөгдөөгүй боловч дээд, доод бүлгийг ялгаж байсан. Ялгах чадвар улирал хооронд өөрчлөгдсөн. Энэ нь тухайн даалгаврын засвар, агуулга-арга зүйн оролцоо, эсвэл шалгуулагчдын бүлгийн ялгаатай байдалтай холбоотой байж болно.

Хэв шинж 2: ӨДД 33-ын хувьд загварын нийт ач холбогдол статистикийн хувьд батлагдсан. Хүндрэлийн зэрэг өөрчлөгдөөгүй. Дээд, доод бүлгийг ялгаж байсан, ялгах чадвар улирал хооронд өөрчлөгдөх хандлагатай байв.

Хэв шинж 3: ӨДД 7, 8, 9, 15, 29, 31, 37-ын хувьд загварын нийт ач холбогдол статистикийн хувьд батлагдсан. Эдгээр даалгаврын хүндрэлийн зэрэг өөрчлөгдөөгүй, дээд болон доод бүлгийг бүх улиралд тогтвортой ялгаж байсан боловч ялгах чадвар улирал хооронд мэдэгдэхүйц өөрчлөгдөөгүй.

Хэв шинж 4: ӨДД 40-ийн хувьд загварын нийт ач холбогдол статистикийн хувьд батлагдаагүй. Хүндрэлийн зэрэг болон ялгах чадвар улирал хооронд өөрчлөгдөөгүй боловч дээд, доод бүлгийн ялгаа ажиглагдсан.

Хэв шинж 5: ӨДД 38-ын хувьд загварын нийт ач холбогдол статистикийн хувьд батлагдаагүй. Хүндрэлийн зэрэг өөрчлөгдсөн боловч дээд, доод бүлгийг хангалттай ялгаж байгаа эсэх нь тодорхой бус, ялгах чадвар улирал хооронд өөрчлөгдөөгүй.

Хүснэгт 9. ӨДД тус бүрийн ялгах чадварын улирал хоорондын өөрчлөлт: Биномиаль ерөнхийшүүлсэн шугаман загварын үр дүн

ӨДД	Загварын нийт ач холбогдлын p	Улирлын үндсэн нөлөө, p	Дээд, доод бүлгийн үндсэн нөлөө, p	Улирал × бүлгийн харилцан үйлчлэл, p	Хэв шинж
4	.003	.380	< .001	.021	
32	< .001	.764	< .001	< .001	1
35	.004	.248	< .001	.039	
33	.001	.476	< .001	.079	2
7	< .001	.118	< .001	.333	
8	.001	.154	< .001	.183	
9	.020	.415	.013	.364	
15	< .001	.395	< .001	.961	3
29	.012	.383	.001	.198	
31	< .001	.075	< .001	.444	
37	.040	.461	.002	.587	
40	.119	.239	.016	.889	4
38	.051	.041	.309	.101	5

Тайлбар. Загварт зөв хариулсан эсэхийг хамаарах хувьсагч, улирал, дээд/доод бүлэг, улирал × бүлгийн харилцан үйлчлэлийг тайлбарлагч хувьсагчаар оруулсан.

Дээд/доод бүлгийн нөлөө нь тухайн даалгаврын ялгах чадварыг, улирлын нөлөө нь бүлэг болон улирал × бүлгийн харилцан үйлчлэлийг харгалзан үзсэний дараа улирал хооронд зөв хариулах магадлал (хүндрэлийн зэрэг) тогтвортой өөрчлөгдсөн эсэхийг, харин улирал × бүлгийн харилцан үйлчлэл нь ялгах чадварын улирал хоорондын өөрчлөлтийг тус тус илэрхийлнэ.

Буруу хариултын хувилбарын ажиллах байдлын өөрчлөлт.

ӨДД тус бүрийн буруу хариултын хувилбарын ажиллах байдлыг улирал хооронд харьцуулах зорилгоор хариултын сонголтын тархалтад Пирсоны хи-квадрат шинжилгээ хийв. 2024-2025 оны намар болон 2024-2025 оны хаврын улирлын хооронд хийсэн харьцуулалтаар 13 ӨДД-оос Д7, Д31,

Д32 дээр сонголтын тархалт статистикийн хувьд ач холбогдолтой өөрчлөгдсөн байна. Тухайлбал, Д7 дээр $\chi^2(3) = 21.00, p < .001, \text{Cramer's } V = .512$; Д31 дээр $\chi^2(5) = 30.90, p < .001, \text{Cramer's } V = .622$; Д32 дээр $\chi^2(4) = 52.50, p < .001, \text{Cramer's } V = .810$ байсан нь эдгээр даалгаврын буруу хариултын хувилбарын ажиллах байдлын хэв шинж их хэмжээгээр өөрчлөгдсөнийг харуулж байна. Эдгээр гурван даалгаврын нийтлэг онцлог нь зөв хариултын байрлал тухайн хоёр улирлын хооронд өөрчлөгдсөн, тухайн хугацаанд буруу хариултын засвар хийгдсэн, хичээлийн агуулга, арга зүйд өөрчлөлт оруулсан явдал байв. Иймээс энэ хугацаанд ажиглагдсан өөрчлөлтийг буруу хариултын засвар, хичээл заалтын өөрчлөлт болон даалгаврын бүтцийн шинжтэй өөрчлөлт гэж тайлбарлаж болно.

Харин 2024-2025 оны хавар болон 2025-2026 оны намрын улирлын хооронд хийсэн харьцуулалтаар 13 ӨДД-оос зөвхөн Д8 болон Д35 дээр сонголтын тархалт статистикийн хувьд ач холбогдолтой өөрчлөгдсөн байна. Д8-д $\chi^2(3) = 8.40, p = .038, \text{Cramer's } V = .294$, Д35 дээр $\chi^2(4) = 12.50, p = .014, \text{Cramer's } V = .358$ байсан нь буруу хариултын хувилбарын ажиллах байдал дунд түвшинд өөрчлөгдсөнийг харуулж байна. Онцлог нь эдгээр хоёр даалгаврын зөв хариу тухайн хоёр улиралд ижил хэвээр байсан, буруу хариултын засвар хийгдсэн явдал юм. Иймээс эдгээр даалгаварт ажиглагдсан өөрчлөлтийг буруу хариултын хувилбарын татах хүч, сонголтын өрсөлдөөн, буруу хариултын засвар, хичээл заалтын нөхцөл, эсвэл оюутны ойлголтын ялгаатай байдалтай холбон тайлбарлах нь илүү үндэслэлтэй байна.

Үүнээс гадна Д7 болон Д15-д 2024-2025 оны хавар болон 2025-2026 оны намрын улирлын хооронд сонголтын тархалтын өөрчлөлт хил заагийн түвшинд ажиглагдсан боловч уламжлалт $\alpha = .05$ босгоор статистикийн хувьд ач холбогдолтой гэж үзэхэд хангалтгүй байв. Харин бусад ӨДД дээр буруу хариултын хувилбарын ажиллах байдлын хэв шинж улирал хооронд статистикийн хувьд ач холбогдолтой өөрчлөлт илрээгүй нь буруу хариултын хувилбарын ажиллах байдал харьцангуй тогтвортой байсныг илтгэж байна.

Нийт дүнгээр авч үзвэл, ӨДД-ын буруу хариултын хувилбарын ажиллах байдал бүх даалгаварт жигд өөрчлөгдөөгүй, харин даалгавар тус бүрийн онцлогоор илэрсэн байна. Энэ нь даалгаврын чанарын өөрчлөлтийг зөвхөн хүндрэлийн зэрэг, ялгах чадвараар бус, буруу хариултын хувилбарын ажиллах байдалтай хамтатган тайлбарлах шаардлагатайг харуулж байна.

Хүснэгт 10. ӨДД-ын буруу хариултын хувилбарын ажиллах байдлын өөрчлөлт

Даалгавар	Харьцуулсан улирал	χ^2	df	p	Cramer's V	Тайлбар
Д7	2024-2025 намар ↓ 2024-2025 хавар	21.00	3	< .001	.512	Сонголтын тархалт хүчтэй өөрчлөгдсөн; зөв хариу В → С. Буруу хариултын хувилбарт засвар орсон.
Д31	2024-2025 намар ↓ 2024-2025 хавар	30.90	5	< .001	.622	Сонголтын тархалт маш хүчтэй өөрчлөгдсөн; зөв хариу С → В. Буруу хариултын хувилбарт засвар орсон. Хичээл заалтын нөхцөлд өөрчлөлт орсон.
Д32	2024-2025 намар ↓ 2024-2025 хавар	52.50	4	< .001	.810	Сонголтын тархалт маш хүчтэй өөрчлөгдсөн; зөв хариу D → С. Буруу хариултын хувилбарт засвар орсон. Хичээл заалтын нөхцөлд өөрчлөлт орсон.
Д8	2024-2025 хавар ↓ 2025-2026 намар	8.40	3	.038	.294	Сонголтын тархалт өөрчлөгдсөн; зөв хариу өөрчлөгдөөгүй (С → С). Буруу хариултад хувилбарт засвар орсон.
Д35	2024-2025 хавар ↓ 2025-2026 намар	12.50	4	.014	.358	Сонголтын тархалт өөрчлөгдсөн; зөв хариу өөрчлөгдөөгүй (В → В). Буруу хариултын хувилбарт засвар орсон.

Тайлбар. Пирсоны хи-квадрат шинжилгээ нь тухайн даалгаврын сонголтын тархалт улирал хооронд өөрчлөгдсөн эсэхийг, Cramer's V нь өөрчлөлтийн хүчийг илэрхийлнэ. $p < .05$ бол тухайн даалгаврын буруу хариултын хувилбарын ажиллах байдал статистикийн хувьд мэдэгдэхүйц өөрчлөгдсөн гэж үзэв. Улирал бүрд өөр өөр оюутны бүлэг хамрагдсан тул энэхүү шинжилгээний үр дүнг шалтгаант нөлөөний баттай нотолгоо бус, харин сонголтын тархалтын өөрчлөлтийн хайгуул шинжтэй нотолгоо гэж тайлбарлав.

ХЭЛЭЛЦҮҮЛЭГ

Дараалсан гурван улирлын сорилгын KR-20 коэффициент .790, .829, .826; SEM-ийн утга 2.84, 2.76, 2.61 байсан нь Sorenson ба Hanson (2021) нарын хоёр болон гурван удаагийн давтан ашиглалтын дараа даалгаврын болон сорилгын чанар бараг бүх тоон үзүүлэлтээр мэдэгдэхүйц сайжирсан гэсэн үр дүнтэй нийцэж байна.

Нийт 41 даалгавраас 13-г нь голчлон буруу хариултын хувилбарын засвар хийгдсэн. Өөрчлөлт дагуулсан болон өөрчлөлтгүй хэрэглэгдсэн даалгавруудын харьцуулахад даалгаврын засвар болон хичээлийн оролцоонд өөрчлөлт илүү хийгдсэн 2024-2025 оны хавраас 2025-2026 оны намрын хооронд өөрчлөлт дагуулсан даалгавруудын ялгах чадвар илүү эерэг чиглэлтэй өөрчлөгдсөн. Энэ нь олон-сонголтот даалгаварт үүрэггүй буруу хариулт түгээмэл тохиолддог бөгөөд түүнийг засах эсвэл солиход сорилгын чанар сайжрах боломжтой (Ali et al., 2016; Tarrant et al., 2009) гэсэн өмнөх судалгааны үр дүнтэй нийцэж байна.

СУДАЛГААНЫ ХЯЗГААРЛАЛТ

Энэхүү судалгаанд дараах хязгаарлалт байна.

- Улирал бүрийн шалгалтад өөр өөр оюутны бүлэг оролцсон тул даалгаврын үзүүлэлтүүдийн өөрчлөлтийг зөвхөн даалгаврын засвар, эсвэл хичээлийн заалтын өөрчлөлтийн шууд үр нөлөө гэж тайлбарлах боломжгүй. Оюутны бүлгийн чадварын ялгаа нөлөөлсөн байж болзошгүй.
- Зарим харьцуулалтад ӨДД-ын тоо цөөн байсан тул статистикийн хүч хязгаарлагдмал байж болно. Иймээс даалгавар тус бүрийн шинжилгээний үр дүнг хайгуул шинжтэй нотолгоо гэж үзэх нь зүйтэй.
- Энэхүү судалгааны хүрээнд хичээлийн суралцахуйн үр дүнтэй уялдсан зарим нэмэлт шинжилгээ хийсэн боловч энэ өгүүлэл нь давтан хэрэглэгдсэн даалгаврын чанарын улирал хоорондын өөрчлөлтөд төвлөрсөн тул тэдгээр үр дүнг дэлгэрүүлэн оруулаагүй. Иймээс психометрийн үзүүлэлтүүд болон суралцахуйн үр дүнгийн биелэлтийн хоорондын уялдааг хамтатган авч үзэх нь цаашдын судалгаанд чухал чиглэл болно.

ДҮГНЭЛТ

Дараалсан гурван улирлын сорилгын KR-20 коэффициент .790, .829, .826; SEM-ийн утга 2.84, 2.76, 2.61 байв. Энэ үр дүн нь дараалсан улирал хооронд сорилгын санамсаргүй хэмжилтийн алдаа буурч, сорилгын дотоод нийцэл сайжирсан ерөнхий дүр зургийг харуулж байна.

Гурван улирлын өгөгдөлд үндэслэсэн шинжилгээгээр даалгаврын хүндрэлийн зэрэг статистикийн хувьд өөрчлөгдсөн боловч ялгах чадвар харьцангуй тогтвортой байв. Энэ үр дүн нь даалгаврын чанарын өөрчлөлт бүх даалгаварт жигд илрээгүй, харин тодорхой даалгаврууд дээр төвлөрөн илэрсэн байж болохыг харуулж байна.

Өөрчлөлт дагуулсан болон өөрчлөлтгүй хэрэглэгдсэн даалгавруудын харьцуулалтаас 2024-2025 оны хавраас 2025-2026 оны намрын хооронд буюу гол өөрчлөлтүүд хийгдсэн улиралд өөрчлөлт дагуулсан даалгавруудын ялгах чадвар илүү сайжирсан нь ажиглагдсан. Энэ үр дүн нь даалгаврын засвар, агуулга-арга зүйн оролцоо зэрэг сайжруулалтын үйл ажиллагаа нь зарим даалгаврын психометрийн чанартай холбоотой байж болохыг харуулж байна.

Даалгавар тус бүрийн дагалдах хайгуул шинжилгээ нь өөрчлөлтийн хэв шинж бүх даалгаварт ижил бус, даалгавар тус бүрийн онцлогоор илэрч байгааг нарийвчлан тодруулав. Буруу хариултын засвар, хичээл заалтын өөрчлөлтийн нөлөө харилцан адилгүй байна.

Олон-сонголтот даалгаврын чанарыг үнэлэхдээ r индекс, d индексээр хязгаарлахгүй, буруу хариултын шинжилгээ, даалгавар тус бүрийн мөрдөлт, багшийн хийсэн сайжруулалтын мэдээллийг хамтатган авч үзэх нь ач холбогдолтой болохыг харуулж байна. Давтан мөрдөж даалгаврын чанарын мэдээллийг харьцуулан судлах ийм хандлага нь хичээлд суурилсан үнэлгээний чанарыг сайжруулах, даалгаврын санг нотолгоонд суурилан боловсронгуй болгох, багшийн үнэлгээний шийдвэрийг илүү үндэслэлтэй болгоход ач холбогдолтой байх боломжтой. Цаашид даалгаврын психометрийн үзүүлэлтүүд болон суралцахуйн үр дүнгийн биелэлтийн уялдааг хамтатган судлах нь цаашдын судалгаанд чухал чиглэл болно.

ТАЛАРХАЛ

Энэхүү өгүүлийг боловсруулах явцад үнэтэй санал, зөвлөгөө өгсөн хамтран ажиллагчдаа талархал илэрхийлье.

ЗОХИОГЧДЫН ОРОЛЦОО

Онолын үндэс, санаа: Ш.Сайнбилэг, Ч.Нямгэрэл; **Арга зүй:** Ш.Сайнбилэг; **Мэдээлэл цуглуулалт:** Ч.Нямгэрэл, Н.Амгалан, Ш.Сайнбилэг, П.Манлайбаатар; **Бичилт (эхний хувилбар):** Ш.Сайнбилэг; **Бичилт (засвар):** Ш.Сайнбилэг, Ч.Нямгэрэл

НОМ ЗҮЙ

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Adegoke, B. A., Nwosu, A. A., & Okoye, N. N. S. (2018). Item analysis of university-wide multiple-choice objective examinations: The experience of a Nigerian private university. *SpringerPlus*, 7(1), 1–10. <https://doi.org/10.1186/s40064-018-2006-5>

Ali, S. H., Carr, P. A., & Ruit, K. G. (2016). Validity and reliability of scores obtained on multiple-choice questions: Why functioning distractors matter. *Journal of the Scholarship of Teaching and Learning*, 16(1), 1–14. <https://doi.org/10.14434/josotl.v16i1.19106>

Barbera, J., & VandenPlas, J. R. (2011). Clarity on Cronbach's alpha use. *Journal of Chemical Education*, 88(12), 1571–1572. <https://doi.org/10.1021/ed3004353>

Breakall, J., Randles, C., & Tasker, R. (2019). Development and use of a multiple-choice item writing flaws evaluation instrument in the context of general chemistry. *Chemistry Education Research and Practice*, 20(2), 369–382. <https://doi.org/10.1039/C8RP00262B>

Clark, T. M., Turner, D. A., & Rostam, D. C. (2022). Evaluating and improving questions on an unproctored online general chemistry exam. *Journal of Chemical Education*, 99(10), 3510–3521. <https://doi.org/10.1021/acs.jchemed.2c00603>

Demircioğlu, H., & Demircioğlu, G. (2016). Analysis of the difficulty and discrimination indices of multiple-choice questions. *Eurasian Journal of Educational Research*, 16(64), 1–20. <https://doi.org/10.14689/ejer.2016.64.1>

Downing, S. M. (2004). Reliability: On the reproducibility of assessment data. *Medical Education*, 38(9), 1006–1012. <https://doi.org/10.1111/j.1365-2929.2004.01932.x>

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, 58(3), 357–381. <https://doi.org/10.1177/0013164498058003001>

Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A comprehensive review. *Review of Educational Research*, 87(6), 1082–1116. <https://doi.org/10.3102/0034654317726529>


Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309–333. https://doi.org/10.1207/S15324818AME1503_5


Hartman, J. R., & Lin, S. (2011). Analysis of student performance on multiple-choice questions in general chemistry. *Journal of Chemical Education*, 88(9), 1223–1230. <https://doi.org/10.1021/ed100133v>


Krishnan, V. (2010). *An item analysis using Classical Test Theory (CTT) on Alberta's data*. University of Alberta, Community-University Partnership.


- Mango, C. (2009). Demonstrating the difference between Classical Test Theory and Item Response Theory using chemistry test data. *International Journal of Educational and Psychological Assessment*, 1(1), 1–16.
- Muangkhousa, S. (2017). Test analysis with Classical Test Theory (CTT): Finding item difficulty and item discrimination using SPSS program. *Vajira Medical Journal: Journal of Urban Medicine*, 61(6), 477–486. <https://doi.org/10.14456/vmj.2017.46>
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press. <https://doi.org/10.17226/10019>
- Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1), 1–18. [https://doi.org/10.1016/0022-2496\(66\)90002-2](https://doi.org/10.1016/0022-2496(66)90002-2)
- Rezigalla, A. A., Eleragi, S. A., Elhussein, A. B., Alfaifi, J., AlGhamdi, M. A., Al Ameer, A. Y., & Adam, M. I. E. (2024). Item analysis: The impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(1), Article 445. <https://doi.org/10.1186/s12909-024-05433-y>
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3–13. <https://doi.org/10.1111/j.1745-3992.2005.00006.x>
- Schauber, S. K., Hecht, M., & Nouns, Z. M. (2010). Rarely selected distractors in high-stakes medical multiple-choice examinations. *BMC Medical Education*, 10(1), Article 85. <https://doi.org/10.1186/1472-6920-10-85>
- Sorenson, B., & Hanson, K. (2021). Using classical test theory and Rasch modeling to improve general chemistry exams on a per instructor basis. *Journal of Chemical Education*, 98(5), 1529–1538. <https://doi.org/10.1021/acs.jchemed.1c00164>
- Sulistyo, G. H., Mukminatien, N., & Saukah, A. (2020). Item analysis of English final semester test. *Indonesian Journal of EFL and Linguistics*, 5(2), 519–536. <https://doi.org/10.21462/ijefl.v5i2.302>
- Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: A descriptive analysis. *BMC Medical Education*, 9, Article 40. <https://doi.org/10.1186/1472-6920-9-40>
- Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, 91(9), 1426–1431. <https://doi.org/10.1021/ed500076x>
- Xiaoqiong, H., & Xianghong, T. (2022). Formative assessment and exam culture in East Asian Confucian-heritage contexts. *Frontiers in Psychology*, 13, Article 990196. <https://doi.org/10.3389/fpsyg.2022.990196>
- Zenisky, A. L., Hambleton, R. K., & Sireci, S. G. (2013). Individual score reports on NAEP: Design, reporting, and interpretability. *Educational Measurement: Issues and Practice*, 32(4), 17–27. <https://doi.org/10.1111/emip.12017>

Зохиогчдын мэдээлэл

Овог, нэр: Шагдарсүрэнгийн Сайнбилэг
 <https://orcid.org/0009-0004-9130-6968>
 Байгууллагын нэр: МУИС
 Хаяг: СБД, 6-р хороо, Их сургуулийн гудамж, МУИС хичээлийн төв байр, 3-р давхар
 Улс: Монгол
 e-mail: Sainbileg@num.edu.mn

Овог, нэр: Чойжилсүрэнгийн Нямгэрэл
 <https://orcid.org/0000-0001-6040-0776>
 Байгууллагын нэр: МУИС
 Хаяг: СБД, 6-р хороо, Их сургуулийн гудамж, МУИС хичээлийн төв байр, 3-р давхар
 Улс: Монгол
 Nyamgerel@num.edu.mn

Овог, нэр: Нацагдаржийн Амгалан
 <https://orcid.org/0000-0002-7398-3971>
 Байгууллагын нэр: МУИС

Овог, нэр: Пүрэвсүрэнгийн Манлайбаатар
 <https://orcid.org/0009-0002-3660-1206>
 Байгууллагын нэр: МУИС

Хаяг: СБД, 6-р хороо, Их сургуулийн гудамж, МУИС хичээлийн төв байр, 3-р давхар Улс: Монгол e-mail: n_amgalan@num.edu.mn	Хаяг: СБД, 6-р хороо, Их сургуулийн гудамж, МУИС хичээлийн төв байр, 3-р давхар Улс: Монгол e-mail: manlaibaatar.p@num.edu.mn
---	---
