

Vertinamieji tyrimai: gaires ir perspektyvos

Cory Heyman, Terry Salinger

KMUSR projektas – tai pedagogų profesinio tobulėjimo programa, kurios tikslas išmokyti dalyvius interaktyviųjų mokymo metodų: jų pagrindu mokiniai rengiami kaip atviros visuomenės piliečiai. Per pirmuosius trejus mokslo metus (1997-2000) tūkstančiai mokytojų dvidešimt keturiose Rytų Europos ir Vidurinės Azijos šalyse baigė KMUSR mokymo seminarus. Programos plėtros tvarka tokia: šaliai pareiškus norą įsitraukti į KMUSR projektą, Tarptautinė skaitymo asociacija (*International Reading Association, IRA*) parenka keturis savanorius, pasirengusius vyksti į šią šalį mokyti 20-40 mokytojų taikyti KMUSR metodus. Per 12-15 mėnesių jie organizuoja iš keturių seminarų susidedančių kursų ciklą; mokytojai išeina praktinių užsiėmimų programą ir pradeda taikyti KMUSR metodus savarankiškai. Tarp seminarų KMUSR dalyviai kartą per mėnesį susitinka su kolegomis ir aptaria padarytą pažangą bei KMUSR metodų taikymo problemas. Jie taip pat sužino savo bendraamžių bei KMUSR savanorių, kurie stebi jų mokymą per pamokas, nuomonę. Tikimasi, jog pirmųjų metų dalyviai, baigę visą KMUSR kursą, taps projekto mokytojų mokytojais.

Anksčiau ar vėliau kiekvienam, susijusiam su projekto plėtra, kyla klausimas: ar ši programa veiksminga? Kokie duomenys galėtų padėti suvokti mūsų laimėjimus ir sunkumus? Kokius metodus turėtume pasitelkti, kad surinktume naudingos informacijos?

2000 m. balandį KMUSR projekto įkūrėjas Atviros visuomenės institutas (*Open Society Institute, OSI*) užsakė Amerikos tyrimų institutams (*American Institutes for Research, AIR*) atlikti Kritinio mąstymo ugdymo skaitant ir rašant (KMUSR) projekto įvertinimą. Paprašyta ištirti tris klausimus:

1. Kiek KMUSR projekto mokytojai laikosi KMUSR modelio vedami pamokas?
2. Kiek mokiniai, kurių mokytojai dalyvauja KMUSR projekte, yra įgiję aukštesnio kritinio mąstymo įgūdžių negu mokiniai iš kontrolinių grupių, kurių mokytojai nesusiję su KMUSR projektu?
3. Kiek KMUSR projekto mokytojų ir mokinių požiūris į mokymą bei mokymąsi skiriasi nuo mokytojų ir mokinių požiūrio kontrolinėse grupėse?

Vertinimas buvo atliekamas keturiose šalyse: Čekijoje, Kirgizijoje, Latvijoje ir Makedonijoje. Norėta išsiaiškinti skirtumus tarp KMUSR projekto klasių mokinių ir mokytojų bei jų kolegų ir bendraamžių kontrolinėse grupėse. Vertinimo rezultatai parodė, kad dalyvavimas KMUSR projekte turėjo teigiamos ir pastebimos įtakos kiekvienam iš šių mokymo ir mokymosi aspektų:

Mokymo praktika

- KMUSR projekto mokytojai laikosi KMUSR modelio mažiau pamokos laiko skirdami paskaitų skaitymui ir daugiau – nedidelių grupių veiklai, labiau skatina mokinių vedamas diskusijas per pamokas;
- KMUSR projekto mokytojai daugiau laiko nei jų bendraamžiai iš kontrolinių grupių skiria „autentiškai pedagogikai“, kuri apima aukštesnio lygio mąstymą, diskusijas esminiais klausimais, sąsajas su išoriniu pasauliu ir įvairių požiūrių į problemos sprendimą vertinimą.

Mokytojo ir mokinių požiūriai

- emociniu požiūriu KMUSR mokytojai aktyviai remia projektą ir yra labiau patenkinti savo darbu nei kontrolinių grupių mokytojai;

- KMUSR mokiniai labiau nei kontrolinių grupių mokiniai patenkinti pamokų veikla, kurios metu jie gali aktyviai dalyvauti mokymosi procese.

Mokinių kritinio mąstymo įgūdžiai

- KMUSR mokiniai apskritai turi aukštesnio kritinio mąstymo įgūdžių negu mokiniai iš KMUSR projektui nepriklausančių kontrolinių grupių; skirtumai labiausiai susiję su tuo, kaip bendraujama klasėje ir kaip mokytojas taiko kritinio mąstymo metodus savo mokymo praktikoje.

Vertinimas: kodėl ir kaip?

Tyrimo metu AIR ekspertai dalijosi savo žiniomis ir metodais su šalies koordinatoriais ir duomenų rinkimo šalies viduje komandomis Čekijoje, Kirgizijoje, Latvijoje bei Makedonijoje. Be kita ko, jie taip pat gavo informacijos apie tarptautinio vertinimo tyrimus, atliktus Rytų Europoje ir Vidurinėje Azijoje.

Šio straipsnio tikslas – apibendrinti AIR panaudotus būdus ir aptarti, kas sužinota programos vertinimo metu. Tikimės, kad čia išdėstyti pastebėjimai ir rekomendacijos įkvėps pradėti tyrimus ir kitus; be to, viliamės, jog mūsų patirtis padės priimti reikiamus sprendimus.

Ataskaitoje nagrinėjami šeši pagrindiniai klausimai, susiję su programos vertinimu:

1. Programos vertinimas: kas tai?
2. Kam reikalingas vertinimas?
3. Kokius klausimus vertintojai turi apsvarstyti ruošdamiesi vertinimui?
4. Kaip nustatyti tyrimo kokybę?
5. Kaip tinkamai pateikti vertinimo rezultatus?
6. Kokius klausimus reikia apsvarstyti vykdant tarptautinį tyrimą?

Programos vertinimas: kas tai?

Programos vertinimas – tai tyrimas, kokią įtaką programa ar jos įgyvendinimas daro tikėtiniems rezultatams. Duomenys programoms įvertinti gali būti renkami keliais būdais, tarp jų interviu, stebėjimais, apžvalgomis bei vertinimais. Analizė gali apimti tiek kokybinius, tiek kiekybinius metodus.

Kam reikalingas vertinimas

Programos vertinimas programos vadovams naudingas keliais atžvilgiais. Pirmia, vertinimas, atliktas įvairiose programos įgyvendinimo stadijose (formuojamasis vertinimas), parodys, ar gerai programa veikia ir kaip ją galima tobulinti. Formuojamasis vertinimas padeda vadovams planuoti ateities veiksmus. Vertinimas, atliktas projekto pabaigoje (apibendrinamasis vertinimas), gali parodyti bendrą projekto sėkmę. Apibendrinamasis vertinimas paprastai nėra diagnozė – greičiau tai atsakas projekto rėmėjams ir šalininkams, padedantis:

- bendruomenėms nustatyti dalyvavimo naudą;
- esamiems rėmėjams nustatyti investicijų vertę;
- potencialiems rėmėjams numatyti programos plėtojimo (t. y. „proporcingo didinimo“) esant tam pačiam dalyvių skaičiui ar programos pristatymo naujiems dalyviams padarinius.

KMUSR projekto vertinimas yra apibendrinamasis. Atviros visuomenės institutui rūpėjo žinoti, ar mokytojai, kurie baigė visą KMUSR seminarų ciklą, tebenaudoja KMUSR metodus per savo pamokas praėjus dvejiems metams, ir, jei taip, kaip šie mokymo metodai padeda mokiniams pasiekti geresnių rezultatų, tokių kaip kritinio mąstymo įgūdžiai.

„KMUSR projekto vertinimo pamokos gali padėti koordinatoriams ir mokytojams planuoti būsimų KMUSR projektų veiklą. Pavyzdžiui, vertinimas parodė, kad bendravimo klasėje pavyzdžiai – ypač mokinių galimybės dirbti kartu – daug ką paaikškina apie mokytojų dalyvavimą KMUSR projekte ir jų mokinių kritinio mąstymo įvertinimo balus. Šis atradimas gali paskatinti mokytojus tęsti tyrimą, ar papildomas darbas su bendraamžiais ir toliau gerins mokinių kritinio mąstymo įgūdžius.

Kokius klausimus vertintojai turi apsvarstyti ruošdamiesi vertinimui

Prieš pradėdant plačią vertinimo programą, reikia apsvarstyti daug klausimų, tarp jų:

- vertinimo tikslą;
- laiką;
- išteklius;
- tyrimo objektą;
- respondentų skaičių;
- pagrindines sąvokas ir vertinimo kriterijus;
- būdus atskleisti programos sėkmę.

Vertinimo tikslo nustatymas

Pirmasis bet kurios vertinimo programos žingsnis yra vertinimo tikslo nustatymas. Kodėl vykdomas tyrimas? Kas vertins jo rezultatus? Kokie klausimai domina pirmiausia: ar programa yra „veiksminga“, o galbūt postūmis „iš šalies“ gali teigiamai paveikti tikėtinus rezultatus?

Programos vertinimo rezultatai gali pasirodyti labai platūs ir apgaulingi (pavyzdžiui, „programa sėkminga“) arba jie gali būti siauri (pavyzdžiui, ar laikomasi KMUSR modelio per pamokas). Tačiau kuo plačiau suvokiami rezultatai, tuo sunkiau juos pamatuoti. Pavyzdžiui, ką reiškia pasakymas „programa sėkminga“? Kas tai yra minimali sėkmė? Kiek ilgai ji trunka? Vertintojai privalo apmąstyti, kaip jie naudos realius duomenis rezultatams ištirti.

Laiko apmąstymas

Vertinimui skirtas laikas daro esminį poveikį mūsų gebėjimui atsakyti į tyrimo klausimus. Per trumpas laikas pasiruošti vertinimui ir jo atlikimui pastūmės vertintojus priimti pirmalaikius, nekompetentingus sprendimus, kurie turės neigiamos įtakos duomenų analizei. Per ilgas laikotarpis pridarys bėdų renkant duomenis iš tyrimo dalyvių. Pavyzdžiui, panašios mokytojų grupės į apklausos klausimus gali atsakyti skirtingai mokslo metų pradžioje ir pabaigoje.

Sprendimai dėl duomenų rinkimo laiko grafiko ne mažiau svarbūs. Pavyzdžiui, kuriuo programos įgyvendinimo metu vertintojai turėtų rinkti duomenis? Vertinimo rezultatai tikriausiai iš esmės skirsis, priklausomai nuo to, ar duomenys rinkti įgyvendinimo pradžioje, ar vėliau. Antra, ar duomenys bus renkami tik kartą, ar keletą kartų? Tik kartą renkami duomenys parodys, kaip programa veikia vienu tam tikru metu, bet nesuteiks išsamaus vaizdo apie programos poveikį apskritai. Be to, tyrimo klausimai apie *poveikį* reikalauja, kad vertintojai rinktų duomenis apie tuos pačius dalyvius tam tikrą laiką. Jeigu ilgalaikių duomenų surinkti negalima, vertintojai nagrinės tik *sąryšio* tarp programos įgyvendinimo ir rezultatų stiprumą.

Turimų išteklių vertinimas

Ištekliai, kaip ir laikas, gali paveikti programos vertinimo kokybę. Lėšų ir žmonių, prisidedančių prie tyrimo, kiekybinis veiksnys turės esminės įtakos sprendimams dėl paties vertinimo ir jo vykdymo. Pavyzdžiui, galimybė samdyti ir

mokyti duomenų rinkėjus sąlygos tai, ar vertinimas remsis anketomis paštu, asmeninėmis apklausomis, ar, sakykim, pamokų stebėjimu. Apskritai tyrimas iš esmės pagerėja naudojant įvairius duomenų rinkimo būdus, ypač nagrinėjant tokį kompleksinį procesą kaip KMUSR mokymas. Kitaip tariant, nuo turimų išteklių priklauso duomenų, kuriuos rinks rinkėjai, kiekis ir pobūdis bei pasitenkinimas darbo rezultatais.

Vertintojai turėtų būti dėmesingi išteklių šaltiniams, galintiems pagerinti vertinimo kokybę bei sumažinti jo kaštus. Pavyzdžiui, vadinamieji „ekstantiniai“ (turintys išliekamąją vertę) duomenys, kaip antai reklaminė medžiaga, ataskaitos ar straipsniai, tikriausiai suteiks informacijos apie programos istoriją, programos tikslus ir jos įgyvendinimo eigą dar prieš pradedant tyrimą. Vertintojai taip pat gali pasinaudoti esamais duomenų rinkimo būdais bei metodais, kurie panašiomis aplinkybėmis buvo naudingi. Jei tokios priemonės tinka naujam vertinimui, jų panaudojimas sumažins tyrimo išlaidas ir pagerins vertinimo rezultatus.

Tyrimo objektas

Sprendimas, ką tirti tyrimo metu, turėtų priklausyti nuo pateikiamų klausimų. Vienais atvejais vertintojus domina tik viena žmonių kaip respondentų grupė, kitais atvejais gali prireikti nustatyti skirtingas grupes. Kai vertinimas turi kelis tikslus, galbūt teks apibrėžti, ar tiriamos grupės yra vienodai svarbios, ar kai kurios iš jų yra svarbesnės už kitas.

Visų programos dalyvių tyrimas tikriausiai bus labai imlus laiko požiūriu ir brangiai kainuos, ir galiausiai gali paaiškėti, kad jis nesuteikia daugiau informacijos nei siauresnis, tačiau gerai apgalvotas pasirinkimas. Vis dėlto vertintojas, rinkdamasis tyrimo objektą, turi apsvarstyti ir kitus tyrimo klausimus. Pavyzdžiui, ar svarbu, kad apibendrinamieji rezultatai tiktų platesniam dalyvių ratui? Kokio dydžio turėtų būti tyrimo objektas? Kaip turėtų tyrimo dalyviai būti parenkami?

Kaip duomenų rinkimo priemonės įvairūs atrankos metodai turi skirtingų privalumų ir trūkumų (žr. 1 lentelę).

1 lentelė. Įvairių atrankos metodų privalumai ir trūkumai

	Apibrėžimas	Privalumai	Trūkumai
Visuotinis tyrimas	Pasirenkami visi tiriamos visuomenės dalies nariai	Suteikia išsamią informaciją apie tiriamą visuomenės dalį	Informacijos rinkimas gali brangiai kainuoti
Atsitiktinė atranka	Tiriamos visuomenės dalies atrinkimas naudojantis metodais, užtikrinančiais, kad kiekvienas tiriamos visuomenės dalies atstovas galės būti parinktas	Potencialiai leidžia apibendrinti duomenis apie tiriamą visuomenės dalį	Verčia vertintoją daryti statistines išvadas apie tai, ar atrinktas gerai atspindi tiriamą visuomenės dalį
Tikslinga atranka	Tarptautinė būdingų tyrimo dalyvių atranka iš visos tiriamos visuomenės dalies	Renka informaciją iš žmonių, priklausančių tiriamai visuomenės daliai, kurie turi ypatingų bruožų	Sunku apibendrinti, ar gerai surinkta informacija atspindi visą tiriamą visuomenės dalį

Pavyzdžiui, tyrimai, kuriuose sąmoningai pasirenkami tyrimo dalyviai, gali suteikti vertingos informacijos įvairiais klausimais. KMUSR projekto vertinimo atveju interviu su žmonėmis, susipažinusiais su KMUSR projektu, suteikė svarbios informacijos apie projekto poveikį ir įsitvirtinimą vertinamose šalyse. Tačiau surinktai informacijai įtakos turėjo kriterijai, kuriais vadovaudamiesi mes parinkome respondentus. Ji tinkama apibendrinimui tik tada, jei respondentai iš visos tiriamos grupės parenkami atsitiktinai. Kuo didesnė atranka, tuo tikslesnis tiriamųjų apibūdinimas.

Respondentų skaičius

Pasirinktinai atliekamo tyrimo dydis geriausiai nustatomas „jėgų analizės“ (*power analysis*) metodu. Šis metodas padeda įvertinti, kokio dydžio turėtų būti atranka, kad ji tiksliai atspindėtų projekto įgyvendinimo poveikį. Jei atrankos dydis per mažas, tikrasis poveikis tikriausiai liks nenustatytas. Per didelė atranka, atvirkščiai, gali brangiai kainuoti ir būti neveiksminga. Jėgų analizė reikalauja iš vertintojų atsakyti į du klausimus:

1. *Kokio lygio statistinis patikimumas būtinas, kad rezultatai būtų įtikinami?* Įprasta norma – 95% (efektyvumas pakankamas, jei pakartotinis tyrimas patvirtina gautus rezultatus 95 atvejais iš 100), tačiau tai gali keistis priklausomai nuo tyrimo. Kuo aukštesnė statistinio patikimumo norma reikalinga, tuo daugiau turi būti atrenkama respondentų.

2. *Kokio dydžio numatomas projekto poveikis?* Jei vertintojai tikisi palyginti didelių skirtumų tarp tiriamosios ir kontrolinės grupių, atrinkta tyrimui grupė neturėtų būti tokia didelė.

Atlikus KMUSR projekto vertinimo jėgų analizę paaiškėjo, kad būtina pasirinkti bent 36 KMUSR projekto mokytojus ir 36 kontrolinės grupės mokytojus kiekvienoje šalyje norint pa matuoti vidutinį KMUSR projekto poveikį mokymo praktikai.

Pagrindinių sąvokų ir vertinimo kriterijų nustatymas

Teorinių programos vertinimo sąvokų vertinimas paprastai prasideda nuo pagrindinių sąvokų apibrėžimo. Kuo konkretnesnės sąvokos, tuo lengviau pasirinkti tinkamus kriterijus. Kitas žingsnis – sugalvoti duomenų rinkimo klausimus, kurių atsakymai būtų glaudžiai susiję su galimais apibrėžimais. Kai kurie atsakymai gali būti lengvai ir tiesiogiai įvertinami. Pavyzdžiui, vertintojai paprastai pasitiki informacija, kai žmonės sakosi esą vyrai arba moterys. Tačiau kitus atsakymus reikia vertinti atidžiau.

Vertinant mokinių kritinio mąstymo įgūdžius, AIR pradeda nuo KMUSR projekto pateikto kritinio mąstymo apibrėžimo, apimančio šešis komponentus. Anot jo, kritiškai mąstantis mokiny:

- formuoja savo nuomonę;
- protingai renkasi iš kelių skirtingų nuomonių;
- sprendžia problemas;
- atsakingai diskutuoja;
- vertina bendradarbiavimą, kurio dėka priimamas bendras sprendimas;
- moka vertinti kito nuomonę ir pripažįsta, kad žmogaus suvokimas ir jo santykiai su supančiu pasauliu formuojami veikiant daugeliui veiksnių.

Paskui pateikiami trys kriterijai, kuriais remiantis vertinami įvairūs kritinio mąstymo sąvokos komponentai (su jais galima susipažinti KMUSR vertinimo ataskaitos A priede, publikuojamame programos tinklapyje¹).

¹ American Institutes for Research, *The 2000-2001 Evaluation of the Reading and Writing for Critical Thinking Project*, New York, NY: Open Society Institute.

Vertindami KMUSR, kritinio mąstymo įvertinimą mes naudojome kaip programos rezultatų matą. „Programos rezultatai“ buvo tik vienas iš keturių aspektų, kuriuos vertintojai turėjo įvertinti. Kiti tyrimo aspektai buvo „programa“, „programos įgyvendinimas“ ir „kontrolė“.

Programos ir rezultatai paprastai matomi jau pačiuose vertinamojo tyrimo klausimuose. Klausime „Kiek mokiniai, kurių mokytojai dalyvauja KMUSR projekte, yra įgiję aukštesnio kritinio mąstymo įgūdžių negu mokiniai iš kontrolinių grupių, kurių mokytojai nesusiję su KMUSR projektu?“, KMUSR projektas yra programa, o kritinio mąstymo įgūdžiai – rezultatai. Tačiau dažnai vertintojai taip pat suinteresuoti išnagrinėti mechanizmus, kurių padedamos programos paveikia rezultatus, ir todėl jiems reikia surinkti informaciją ir apie programos įgyvendinimą. Pagaliau vertintojai nori suprasti, ar sąsajoms tarp programų ir rezultatų tikrai turėjo įtakos programos įgyvendinimas, o ne kiti veiksniai. Todėl svarbu nustatyti ir įtraukti į tyrimą kiek galima daugiau potencialių svertų.

Vertinant KMUSR projektą, programa buvo vertinama atsižvelgiant į tai, ar mokytojai dalyvavo ir baigė KMUSR seminarų kursą. Programos įgyvendinimo matai buvo kritinio mąstymo įgūdžių integravimas į pamokos veiklą, bendravimas pabrėžiant mokinių darbą grupėse, mokinių diskusijos bei mokinių atliekami tyrimai. Tuo tarpu vertinimo rezultatai buvo nustatyti priklausomai nuo tiriamo klausimo¹. Pavyzdžiui, vertindami mokytojų pasirengimą, AIR nagrinėjo dydį, kiek mokytojai naudoja kritinio mąstymo įgūdžius savo pamokose (kriterijus, kurį mes taip pat naudojome analizuodami programos įgyvendinimo matą). Tirdami mokinių lygį, AIR kartu su koordinatoriais vertinamosiose šalyse ieškojo visoms šalims vienodų ir prasmingų kritinio mąstymo įgūdžių matų. Galų gale AIR nustatė daugybę kriterijų, apimančių mokinių išsilavinimą, mokytojo charakteristikas ir mokomąjį dalyką².

Nustatę pagrindinius kriterijus, vertintojai pasirenka priemones, padėsiančias gauti pageidaujamą informaciją. 2 lentelėje pateikiami švietimo tyrimuose naudojami duomenų rinkimo šaltiniai bei palyginami kai kurie jų privalumai ir trūkumai. Naudojantis įvairiais šaltiniais, daugumą trūkumų galima įveikti.

Duomenys iš daugelio šaltinių vieni kitus gali papildyti ir patvirtinti. Pavyzdžiui, vertindami KMUSR, AIR ekspertai pasitelkė duomenis, gautus stebint pamokas, ir mokytojo bei mokinių apžvalgas, kuriose buvo nagrinėjami klausimai apie mokymo praktiką. Apžvalgų klausimai suteikė išvalgų apie mokytojo ir mokinio požiūrį į mokymą bei informacijos apie esamą mokymo praktiką, kuri galėjo likti nepastebėta remiantis vien pamokų stebėjimu. Mokinių apžvalgose taip pat buvo nagrinėjami kritinio mąstymo įgūdžių vertinimai.

Nustatyti ir išmatuoti pagrindinius vertinimo kriterijus gana sunku viename lingvistiniame ir kultūriniame kontekste, o atlikdami tarptautinius tyrimus, tokius kaip KMUSR vertinimas, vertintojai turi būti ypač atsargūs. Viena vertus, KMUSR sukūrė norminę tvarką įgyvendinant programą nepriklausomai nuo nacionalinių sienų. Antra, kiekviena šalis KMUSR projektą pritaiko savo poreikiams savitomis kultūros aplinkybėmis. Šio, kaip ir bet kurio kito tarptautinio vertinimo uždavinys, - užtikrinti, kad sąvokos būtų vienodai vertinamos visose šalyse. Turi būti organizuojamos plačios diskusijos su vertinimo partneriais tam, kad visose vertinamosiose šalyse kriterijai būtų suprantami vienodai, o duomenys renkami ir analizuojami panašiai.

¹ Išsamiau apie KMUSR projekto vertinimo kriterijus žr.: *KMUSR vertinimo ataskaita*, p. 4-6.

² *Ibid.*, p.14-16.

2 lentelė. Skirtingų duomenų šaltinių naudojimo vertinant švietimą privalumai ir trūkumai

Duomenų šaltinis	Apibrėžimas	Privalumai	Trūkumai
Vieninteliai duomenys	Jau egzistuojanti informacija apie vertinamą programą ir jos įgyvendinimą	▪ Nebrangu rinkti ▪ Teikia konteksto informacijos	Informacija šališka ir priklauso nuo rašiusiojo
Apžvalgos	Apklausos, pateikiančios iš dalies ar visiškai struktūriškai apibrėžtą informaciją	▪ Gana nebrangu apklausti didelę žmonių grupę ▪ Palyginti lengva apibūdinti įvairius pavyzdžius ir kryptis	▪ Pateikiama informacija gali būti šališka ir atspindėti tai, ką vertintojas nori žinoti ▪ Apžvalgos teikia ribotus pavyzdžių ir krypčių paaiškinimus
Interviu	Diskusijos su respondentais siekiant gauti daugiau išsamios informacijos apie programas ir programos poveikį	▪ Pateikia detalios informacijos apie programas ir programos poveikį ▪ Gali pateikti pavyzdžių ir krypčių, neatskleistų renkant duomenis	▪ Pateikiama informacija gali būti šališka dėl vertintojo lūkesčių išgirsti tai, ką norima išgirsti ▪ Interviu apklausiant kiekvieną respondentą paprastai gerokai brangesni už apžvalgas ▪ Informacija iš nedidelio kiekio apklaustųjų gali neatspindėti vertinamos visuomenės dalies
Stebėjimas	Galimybė pamatyti programą ar jos įgyvendinimą praktikoje	Suteikia tiesioginės informacijos apie tai, kaip programa veikia ir koki poveikį daro tikėtiniems rezultatams	▪ Programos dinamika gali būti šališka priklausomai nuo to, ką vertintojas nori matyti ▪ Stebėjimo duomenis brangu rinkti ir atitinkamai užkoduoti ▪ Ribotas stebėjimų kiekis vienoje vietoje neapima viso veiksmų ir sąveikų spektro
Įvertinimas	Veikla, skirta nustatyti, kaip respondentas veikia lyginant su numatytu programos rezultatu	Idealiu atveju tai nešališka priemonė nustatyti, kaip sekasi įgyvendinti trokštamą programos rezultatą	Vertinimo paruošimas, vykdymas ir analizė brangiai kainuoja

KMUSR vertinimo atveju mokinių kritinis mąstymas buvo vertinamas pasitelkus priemones:

- sukurtas konsultuojantis su KMUSR koordinatoriais vertinamose šalyse;
- išverstas vienos vertėjų grupės į penkias kalbas ir paskui vėl išverstas į anglų kalbą kitos vertėjų grupės, kad tikrasis vertimas turėtų tą pačią *pažodinę* prasmę visose instancijose;
- eksperimento būdu vykdomas visose keturiose šalyse būti tikriems, kad priemonės nebuvo šališkos ir turėjo tą pačią *kultūrinę* prasmę visose instancijose.

Programos vertinimas

Programos, jos įgyvendinimo ir rezultatų nustatymas bei taikymas, apsisprendimas, ką reikėtų tirti, ir tinkamų priemonių informacijai rinkti įvertinimas – bene svarbiausi programos vertinimo žingsniai. Bet vertintojai privalo pagalvoti ir apie tai, kaip duomenys galiausiai bus analizuojami vertinant pačią programą, atsižvelgiant į vertinimo tikslus. Labai svarbu, kad vertintojai sukurtų analizės planą prieš pradėdami rinkti duomenis, norėdami būti tikri, jog neišleis iš akių jokių esminių duomenų elementų ir kad visa jų renkama informacija bus naudinga atliekant analizę. Vertintojai naudoja įvairius metodus siekdami nustatyti, ar programa turi norimą poveikį tikėtiniems rezultatams. Galbūt akivaizdžiausias metodas yra paklausti žmonių apie poveikį tiek apklausos, tiek apžvalgos forma. Žmonės, dirbę su programa, programos dalyviai ir programos stebėtojai – gynėjai ir kritikai, - paprastai pateikia daugybę įžvalgų apie programos sėkmę ir jos priežastis. Šio metodo trūkumas, kaip matyti iš 2 lentelės, yra tai, kad žmonių nuomonė šališka. Pavyzdžiui, vieni turi asmeninių priežasčių girti ar peikti programą. Kiti gali pasakyti vertintojams tai, ką, jų nuomone, vertintojai nori išgirsti. Dar kiti gali būti nuoširdūs reikšdami savo nuomonę apie programą ir nevisiškai suprasti tikrą programos ar jos įgyvendinimo poveikį tikėtiniems rezultatams.

Todėl vertintojai privalo rasti kitų būdų programos sėkmei įvertinti. Galbūt geriausias būdas parodyti, kad tam tikros programos dalies (tokios kaip KMUSR seminarai) įgyvendinimas daro realų poveikį tikėtiniems rezultatams (tokiems kaip mokinių kritinio mąstymo įgūdžiai) – imtis mokslinio metodo. Pagal jį vertintojas atsitiktinai iš tiramos visuomenės dalies (tokius kaip mokytojai ar mokiniai) pasirinktus žmones suskirsto į dvi grupes, kurių nariai turi panašų išsilavinimą ir patirtį; tada viena grupė dalyvauja projekte, kita ne. Vertintojas renka duomenis apie abi grupes dvejopu laiku – prieš pradėdamas tyrimą ir tyrimo pabaigoje. Šie duomenys vėliau naudojami nustatyti, ar vertinant rezultatus susijusi su programa grupė pasiekė ko nors, kas skirtųsi nuo grupės, kuri nepatyrė intervencijos. Šis atsitiktinio priskyrimo procesas yra vienintelis vertinimo metodas, kuris patenkina vertintojus, kai skirtumai tarp dviejų grupių atsiranda dėl programos įgyvendinimo, o ne dėl kitų priežasčių.

Tačiau mokslinio tyrimo metodo taikymas švietimo tyrimams paprastai nebūdingas. Tokie projektai kaip KMUSR sukurti tam, kad jais pasinaudotų kuo daugiau žmonių. Taigi mažai tikėtina, kad vertintojai turėtų galimybę tirti programą nuo pat jos pradžios ar stebėti poveikį dalyvaujančiai ir kontrolinei grupėms tam tikrą laiką. Tokie tyrimai bus brangūs ir atims daug laiko. Todėl vertintojai turi metodiškai nagrinėti tyrimo klausimus, prie mokslinio tyrimo bandydami priartėti kiek galima arčiau. Pavyzdžiui, vertintojai gali mėginti sukurti kontrolinę grupę, labai panašią į susijusią su projektu, ir ištirti skirtumus pagal vertinimo klausimus. Jie taip pat gali sugalvoti tyrimus, kaip išnagrinėti skirtumus tarp žmonių, dirbančių su projektu skirtingą laiko tarpą (pvz., KMUSR mokytojai su vienerių, dvejų ar trejų metų patirtimi), užuot tyrę žmonių patirtį ilgą laiką.

Kaip nustatyti tyrimo kokybę

Sprendimai apie vertinimo kokybę paprastai sąlygojami kelių veiksnių (kai kuriuos jų jau aptarėme):

- Ar vertinimo klausimai atskleidžia svarbiausius ryšius tarp programų ir tikėtinų rezultatų?
- Ar vertinimo klausimų įvertinimo kriterijai tiksliai atspindi teorines sąvokas?
- Ar pagrindinių sąvokų kriterijai buvo taikyti mažose respondentų grupėse?
- Ar pavyzdžių apibrėžimas ir duomenų rinkimo priemonės tinka tiksliai ir veiksmingai atsakyti į vertinimo klausimus?
- Ar gali vertintojai įrodyti, kad ryšys tarp programos įgyvendinimo ir tikėtinų rezultatų – kad ir kokie jie būtų – yra realus, o ne kito kokio nors veiksnio padarinys?

Kiti sprendimai dėl programos vertinimo kokybės susiję su duomenų rinkimo procesu. Dažniausiai iškyla trys klausimai, susiję su duomenų rinkimo kokybe:

- Ar iš visų respondentų buvo renkami reikalingi duomenys?
- Ar buvo imtasi priemonių išvengti duomenų rinkėjų ir respondentų šališkumo?
- Ar pakankamai buvo kontroliuojama tyrimo kokybė?

Nuoseklumas

Svarbiausias kiekvienos duomenų rinkimo veiklos tikslas yra užtikrinti, kad duomenų rinkimo procedūros būtų nuosekliai taikomos visų respondentų atžvilgiu – ypač lyginant tiriamą ir kontrolinę grupes (jei jos yra). Nuoseklumas svarbus todėl, kad duomenų rinkimas neiškreiptų mūsų supratimo apie tikrus respondentų, kurie tam tikrą laiką dalyvavo ir nedalyvavo projekte ar jį įgyvendinant, skirtumus. Šis nuoseklumo klausimas apima duomenų rinkimą panašiomis aplinkybėmis ir panašioje terpėje. Žmonės, kurie prižiūri duomenų rinkimo veiklą, taip pat turėtų pasirinktinai patikrinti respondentus, kad įsitikintų, jog duomenys visais atžvilgiais surinkti nuosekliai.

Šališkumo sumažinimas

Jeigu numatyta, kad tyrimas bus vykdomas lyginant žmones, kurie dalyvavo ir kurie nedalyvavo projekte ar jį įgyvendinant, duomenų rinkimas turėtų būti toks, kad duomenų rinkėjai ir respondentai kuo mažiau netyčia iškreiptų tyrimo rezultatus. Duomenų rinkėjai neturėtų žinoti, ar jie stebi dalyvaujančius projekte, ar nedalyvaujančius, o respondentai neturėtų žinoti, ar jie susiję su programos įgyvendinimu. Toks atsargumas apsaugo duomenų rinkėjus ir respondentus nuo to, kad jų nuomonė apie projektą neiškreiptų tyrimo rezultatų. Žinoma, vertinimo koordinatoriai neturėtų tyčia apgaudinėti respondentų apie vertinimo tikslus, bet tam tikras duomenų rinkimo tikslų neapibrėžtumas gali apsaugoti respondentus nuo tokio elgesio ar atsakymų, kokie, jų manymu, yra teisingi, užuot reagavus natūraliai.

Kokybės kontrolė

Vienas iš geriausių būdų padidinti informacijos nuoseklumą ir sumažinti šališkumą duomenų rinkimo metu – tai aiškos ir suprantamos kokybės kontrolės tvarkos nustatymas. Pavyzdžiui, vertintojai turėtų nuolat bendrauti su duomenų rinkėjais, kad įsitikintų, jog pastarieji laikosi duomenų rinkimo tvarkos ir neturi jokių keblumų. Duomenų rinkimo metu duomenų rinkėjams naudinga susirinkti į grupę pasidalyti patirtimi ir sugalvoti atitinkamus metodus, kaip reaguoti į nenumatytus nukrypimus nuo įprastos duomenų rinkimo tvarkos. Vertintojai taip pat gali susitikinėti su duomenų rinkėjais, kad įsitikintų, jog duomenų rinkimo priemonės vykdomos kruopščiai ir nuosekliai.

Taip pat svarbu, ypač atliekant didesnius tyrimus, pasirinktinai apklausti respondentus siekiant gauti papildomos informacijos apie duomenų rinkimo procesą. Ar duomenų rinkėjas atvyko numatytu laiku? Ar asmuo laikėsi nustatytos tvarkos? Pasirinktinės apklausos padeda vertintojams sužinoti, ar duomenų rinkėjai dirbo profesionaliai ir ar respondentams duomenų rinkimas nesudarė tam tikrų keblumų. Tyrimo dalyvių geranoriškumas – vienas iš būdų vertintojams dirbti ilgesnį laiką.

Kaip tinkamai pateikti vertinimo rezultatus

Pamokos, gautos atliekant vertinamąjį tyrimą, suteiks vertingos informacijos įvairioms auditorijoms, tarp jų projekto darbuotojams, pedagogams praktikams, politikams, programos rėmėjams ir galimiems finansuotojams. Todėl vertintojai supranta, kaip svarbu pasidalyti projekto rezultatais. Veiksmingi skleidimo metodai – tinkama auditorija ir sprendimas, kas šiai auditorijai svarbu. Vienas iš bendriausių dokumentų formų yra vertinimo ataskaita. Kiti metodai pranešti tyrimo rezultatus – tai skelbimai, vaizdo įrašai bei interneto svetainės. Nesvarbu kokia forma, produktas turėtų aiškiai ir tiksliai paaiškinti rezultatus, suteikti pakankamai informacijos apie tai, kokie duomenys rinkti ir kaip jie buvo renkami bei analizuojami.

Pavyzdžiui, „Vertinamasis KMUSR tyrimas“ (dokumentas publikuojamas adresu (www.rwct.org)) yra tradicinė vertinimo ataskaita raštu. Joje pateikiama išsami informacija apie tyrimo planus ir vertinimo rezultatus. Ataskaitos prieduose taip pat pateiktos duomenų rinkimo priemonės ir rubrikų įvertinimas, kad būtų galima geriau suprasti, kokie duomenys buvo renkami. Santraukoje kartu su lentelėmis ir schemomis glaustai pateikiamos vertinimo tezės, leidžiančios gana greit susipažinti su svarbiausiais rezultatais. Reziümė parašyta taip, kad ją būtų galima išversti į įvairias KMUSR projekto kalbas KMUSR projekto darbuotojų visame pasaulyje labai.

Apibendrinamosios išvados

Vertinimas – puiki priemonė nustatyti švietimo programą ir programos skleidimo rezultatus. Jis rekomenduotinas kaip svarbi veikla planuojant bet kokią programą. Kadangi programos vadovui paprastai sunku rasti resursų programos vertinimui atlikti, kompromisas galėtų būti tiesioginės paslaugos; informacija apie profesionalų vertinimą padės vadovams įgyti kompetencijos ir sutaupyti resursų. Tačiau kokybiškas ir veiksmingas vertinimas pareikalaus daug pastangų. Programos vadovai ir vertintojai privalo artimai bendradarbiauti, norėdami suformuluoti tinkamus vertinimo klausimus, kuriuos reikėtų nagrinėti kartu su kitais duomenimis. Be to, pasiruošimas vertinimui reikalauja išsamios esamų šaltinių ir galimų problemų analizės. Ir galiausiai ne mažiau svarbu parengti kontrolės tvarką, siekiant sumažinti šališkumą ir padidinti vertinimui surinktų duomenų patikimumą.

Cory Heyman – vyresnysis tyrimų analitikas;

Terry Salinger – tyrimų vadybininko padėjėja. Abu jie – Amerikos tyrimų institutų bendradarbiai.
