

बड़े पैमाने पर आकलन

अनुभव और नजरिया

वैजयन्ती शंकर

किसी भी शैक्षिक संस्था या व्यवस्था के लिए यह आवश्यक है कि वह बच्चों की प्रगति एवं सीखने से जुड़ी समस्याओं पर नजर रखे। इसे ध्यान में रखते हुए बच्चों का आकलन किया जाता है। इस लेख में एजुकेशनल इनिशिएटिव द्वारा किए गए व्यापक स्तर के आकलनों के अनुभवों और उनकी समस्याओं पर चर्चा की गई है।

मैं बड़े पैमाने पर किए जाने वाले आकलनों को बनाने से जुड़े अपने अनुभवों और तरीकों को आपके साथ साझा करूंगी* और यह बताऊंगी कि कैसे उनके बारे में सोच-विचार किया गया और इससे जुड़ी गलतफहमियां कौन-कौनसी थीं। मैं इस पर भी बात करूंगी, कि आकलन व्यापक शैक्षिक संदर्भ में किस तरह की भूमिका अदा करता है और हकीकत में व्यापक समाज को किस तरह से फायदा पहुंचाता है।

आकलन के प्रकार

एजुकेशनल इनिशिएटिव (आगे से संक्षेप में ईआई) के द्वारा बड़े पैमाने पर किए जाने वाले आकलनों को मोटे तौर पर तीन किस्मों में बांटा जा सकता है। पहली किस्म 'कई क्षेत्रों के एक साथ' अध्ययन की है। इन अध्ययनों को सभी जगहों पर एक साथ किसी तयशुदा वक्त पर किया जाता है, जैसे अलग-अलग राज्यों या जिलों में या इसी तरह से एक ही वक्त पर एक साथ खटाखट खींचे गए फोटो। उदाहरण के तौर पर, 'स्टूडेंट लर्निंग स्टडी' भारत के 19 राज्यों में एक साथ की गई थी। इसके दायरे में सरकारी स्कूलों के 1.6 लाख बच्चों को और 48 जिलों को शामिल गया था। यह अध्ययन सर्व शिक्षा अभियान के 12वें, 13वें और 14वें संयुक्त समीक्षा अभियान के दौरान किया गया था। इसी तरह से 'क्वालिटी एजुकेशन स्टडी' पिछले साल विप्रो के साथ की गई थी और इससे पहले 'स्टूडेंट लर्निंग इन मेट्रो' की गई थी जिसका प्रकाशन इंडिया टुडे पत्रिका की कवर स्टोरी के तौर पर किया गया था। इन्हें हम कई क्षेत्रों का एक साथ अध्ययन करना कह सकते हैं।

दूसरी किस्म के अध्ययनों को मैं 'दीर्घकालीन अध्ययन' का नाम दूंगी, जिसमें हम किसी प्रणाली और सुधार की प्रक्रिया का सालों-साल लगातार अध्ययन करते हैं। इसका एक उदाहरण भूटान में पिछले पांच सालों से किया जा रहा हमारा काम है। वहां पर प्रणालीगत बदलावों पर नजर रखने के लिए हर साल बच्चों का आकलन किया जाता है और उसके नतीजों को सुधार की प्रक्रिया में शामिल किया जाता है। इसी तरह की एक परियोजना 'गणोत्सव' के नाम से हम गुजरात में कर रहे हैं। गुजरात सरकार ने इसे पिछले तीन सालों

* यह 'विप्रो एप्लाइंग थोट इन स्कूल्स' के वर्ष 2012 के सालाना पार्टनर फोरम में दिया गया वक्तव्य है। यह लेख इस वक्तव्य का लिप्यान्तरित एवं अनुदित रूप है।

से शुरू किया है और हम इसमें पिछले एक साल से शामिल हैं। इस अध्ययन में सीखने के नतीजों पर ध्यान केन्द्रित करते हुए प्रणालीगत सुधारों पर नजर रखी जा रही है। मुंबई नगर निगम के साथ किया जा रहा 'स्कूल एक्सीलेंस प्रोग्राम' भी एक दीर्घकालीन अध्ययन है।

तीसरी किस्म 'प्रभाव के आकलन के अध्ययन' की है। मैं इसको एक अलग खाने में रखती हूँ क्योंकि ये किसी भी तरह के हस्तक्षेप के प्रभाव की जांच करने के लिए किए जाने वाले एक खास तरह के शुरुआती और आखिरी अध्ययन होते हैं। ये तब किए जाते हैं जब उस हस्तक्षेप को लागू करने वाली गैर-सरकारी संस्था या दूसरे कोई भी समूह यह जानना चाहते हैं कि क्या उनका कार्यक्रम ठीक से चल रहा है और इसीलिए हस्तक्षेप से पहले और कार्यक्रम को एक निश्चित समय तक लागू करने के बाद उस कार्यक्रम का आकलन करते हैं, जैसा कि शुरुआती-आखिरी (अंग्रेजी में बेसलाइन-एंडलाइन) अध्ययन के मॉडल में किया जाता है। आंध्र प्रदेश में किया गया अध्ययन (मनमाने तरीके से चुनाव करके किया गया आकलन) एक प्रभाव के आकलन का अध्ययन था और इसके साथ ही एक दीर्घकालीन अध्ययन भी था क्योंकि उसे छह सालों तक किया गया था। यह अध्ययन विश्व बैंक, आंध्र प्रदेश सरकार, अजीम प्रेमजी फाउंडेशन, ईआई और हार्वर्ड विश्वविद्यालय द्वारा समर्थित था। इस अध्ययन में मनमाने तरीके से इस बात का परीक्षण किया गया था कि किस हस्तक्षेप का शिक्षार्थियों के सीखने पर बेहतर प्रभाव (परीक्षण आंकड़ों से मापा गया) पड़ता है। शिक्षकों को प्रोत्साहन देने से, स्कूलों में निवेश बढ़ाने के लिए अनुदान देने से, अतिरिक्त शिक्षक लगाने से आदि। इसमें छह साल तक कुछ बच्चों, शिक्षकों और स्कूलों पर साल-दर-साल नजर रखी गई।

यहां पर इन तीन किस्मों का वर्णन करने के पीछे मेरा मकसद यह दर्शाना है कि अध्ययन की किस्म से यह तय होता है कि आकलन को किस तरह से गढ़ा जाएगा। हरेक का अलग मकसद है और उसकी डिजाइन या गढ़न्त अलग तरह की होगी। बड़े पैमाने पर किए जाने वाले आकलन की गढ़न्त में सबसे पहले अध्ययन के

मकसद पर ध्यान दिया जाता है। उदाहरण के लिए, हम किसी व्यक्ति पर आधारित आकलनों की गढ़न्त और मकसद के अलग-अलग प्रकारों को देखते हैं। दाखिले व नियुक्ति के लिए किए जाने वाले आकलनों में बेहतरीन को चुनने की जरूरत होती है, जबकि बोर्ड की परीक्षाओं में यह पक्का करने की जरूरत होती है कि बड़ी संख्या में शिक्षार्थियों का प्रमाणीकरण हो जाए। इनमें अलग-अलग तरह की डिजाइनों की जरूरत पड़ेगी। आपको इस बात की एक झलक देने के लिए एक उदाहरण देती हूँ, आईआईटी-जेईई की परीक्षा या किसी प्रवेश परीक्षा को ऊंचे दर्जे की कठिनाई स्तर पर डिजाइन किया जाएगा ताकि सिर्फ ऊंची काबिलियत वाली शिक्षार्थी ही उस परीक्षा में कामयाब हो सकें, जबकि एक बोर्ड परीक्षा को खास तौर पर कम कठिनाई स्तर पर गढ़ा या डिजाइन किया जाएगा क्योंकि तंत्र यह चाहता है कि बड़ी संख्या में शिक्षार्थी उस परीक्षा में कामयाबी हासिल करें। आकलन छानबीन के मकसद से भी डिजाइन किए जा सकते हैं, जिसमें कोई व्यक्ति यह जानने की कोशिश कर सकता है कि बच्चे की सीखने की क्षमताओं के मुताबिक उसके लिए कौनसी कक्षा उपयुक्त रहेगी। उदाहरण के लिए, अभी हमारे पास बहुत सारे बच्चे ऐसे हैं जो स्कूल से बाहर हैं, जिन्हें स्कूली तंत्र में जोड़े जाने की जरूरत है। असल में हम नहीं जानते कि सीखने के मामले में वे बच्चे किस कक्षा के स्तर के हैं। ऐसे मामलों में, आकलन को छानबीन के परीक्षण की तरह से डिजाइन किया जा सकता है और उसमें ऐसी चीजें हो सकती हैं जो अलग-अलग कक्षाओं तक फैली हों। निदान के मकसद से गढ़े गए आकलन, जैसे कि, मजबूतियों व कमजोरियों को विस्तार से बताने वाली एक रूपरेखा में दक्षताओं को व्यापक तौर पर समेटने की और सरल व कठिन सवालों के एक वाजिब मिश्रण की जरूरत पड़ेगी।

अब तक हमने व्यक्ति आधारित डिजाइनों को देखा है। जब हम किसी तंत्र पर आधारित आकलन की डिजाइनों को देखते हैं, तो फिर से हमें कई प्रकार नजर आते हैं। उदाहरण के लिए, आकलन का मकसद सिर्फ खतरे के लिए चेतावनी की घंटी बजाना हो सकता है, ताकि तंत्र को यह बताया जा सके कि कि चीजें दुरुस्त नहीं हैं। तब हमारा जोर इस बात पर रहता है कि कितनी फुर्ती और आसानी से उस परीक्षण को किया जा सकता है और ऐसे मामले में हम झटपट और खराब तस्वीरें उतारते हैं।

एक आकलन ऐसा भी हो सकता है जिसमें शिक्षक यह जानना चाहता हो कि क्या ठीक चल रहा है और क्या ठीक नहीं चल रहा है और तब उसमें हर बच्चे व हर सवाल पर बहुत बारीकी व विस्तार के साथ जानकारी इकट्ठी की जाएगी। या फिर कार्यक्रम

वैजयन्ती शंकर

अहमदाबाद की संस्था एजुकेशनल इनिशिएटिव में वर्ष 2003 से कार्यरत हैं। वर्तमान में एजुकेशनल इनिशिएटिव के स्ट्रेटेजिक रिलेशनशिप की उपाध्यक्ष हैं और संस्था द्वारा व्यापक स्तर पर किए गए आकलनों से जुड़ी रही हैं।

के असर का आकलन करने के मामले में डिजाइन अलग तरह की होगी और उसमें आकलन शुरुआती-आखिरी हालात के बारे में होगा। ऐसे मामले में आखिरी अध्ययन की डिजाइन शुरुआती अध्ययन की डिजाइन से बहुत अलग नहीं होनी चाहिए और उस परीक्षण में एकसी मर्दे (आइटम) या समरूप मर्दे रखने की जरूरत पड़ेगी। या आकलन को कुछ इस तरह का होना चाहिए कि जिससे बेंच मार्क बने, इस मामले में परीक्षण को इस तरह से गढ़ा जाएगा कि उससे सामान्य वितरण में जानकारी को इकट्ठा किया जा सके, यानी, कम क्षमता से काम करने वालों के साथ-साथ ऊंची क्षमता के साथ काम करने वालों और वे क्या करने के काबिल है, से जुड़ी जानकारी परीक्षण को मुहैया करवानी चाहिए। ये संकेतक सिर्फ यह दर्शाते हैं कि किस तरह से परीक्षण के मकसद यह तय करते हैं कि आकलन कैसे होगा और असल में वह उसकी बनावट कैसी होगी।

जब हम अलग-अलग समूहों के साथ काम करते हैं तो कुछ गलतफहमियां आम तौर पर हमारे सामने आती हैं, जैसे क्या हम जनगणना कर रहे हैं या एक नमूना ले रहे हैं? आम तौर पर जनगणना की तब तक जरूरत नहीं पड़ती जब तक कि बहुत ही निचले स्तर और हर बच्चे के स्तर तक जानकारी की जरूरत न हो। मुंबई नगर निगम के लिए 'स्कूल एक्सीलेंस प्रोग्राम' में किया गया आकलन असल में एक जनगणनीय अध्ययन था। मुंबई में कक्षा 3 व 4 के हरेक बच्चे का चार माध्यमों में परीक्षण किया गया था। मुंबई नगर निगम ने जनगणनीय अध्ययन करने का चुनाव किया क्योंकि वे सभी 24 वार्डों में फीडबैक देना चाहते थे। इसलिए हरेक वार्ड को और हरेक माध्यम को अध्ययन के दायरे में समेटना जरूरी था। उसका मकसद हरेक बच्चे और हरेक स्कूल के बारे में जानकारी मुहैया करवाना था। एक सबसे आम गलतफहमी लोगों में यह पाई जाती है कि उन्हें बच्चे के परीक्षण और तंत्र के परीक्षण के बीच फर्क को समझने में उलझन होती है। जब मकसद बच्चे की जांच या परीक्षण करना होता है, तब परीक्षण बनाने में सरोकार इस बात को जांचने पर रहता है कि सवाल बहुत ज्यादा कठिन या बहुत ज्यादा सरल तो नहीं हैं, आदि। जबकि तंत्र का परीक्षण करते वक्त, परीक्षण की डिजाइन में अलग-अलग कठिनाई के सवालों का मिश्रण लिया जाता है, ताकि कमजोर से लेकर बेहतरीन शिक्षार्थियों की क्षमताओं के सामान्य वितरण को अध्ययन के दायरे में समेटा जा सके।

किसी को भी एक परीक्षण में ही बहुत सारे मकसदों को हासिल करने की कोशिश नहीं करनी चाहिए। उदाहरण के लिए, शिक्षार्थियों की उपलब्धि की मजबूतियों व कमजोरियों को पहचानने के लिए

डिजाइन किए गए आकलन को शिक्षकों की तरक्की के लिए इस्तेमाल नहीं किया जाना चाहिए। तो डिजाइन हमेशा ही सीमाओं को तय करती है और वह हद भी तय करती है जहां तक किसी खास आकलन की खोजों को इस्तेमाल किया जा सकता है। एक दूसरी गलतफहमी यह है कि आकलन शिक्षार्थियों के सीखने को बेहतर करने की रामबाण दवा है। एक आम कहावत है, 'रोज-रोज तुलने से वजन अपने-आप नहीं घट जाता'। इसी तरह सिर्फ परीक्षण करने से शिक्षार्थियों के सीखने से जुड़े सभी मुद्दे हल नहीं हो सकते। हालांकि, इससे ऐसे बहुत से उपयोगी आंकड़े मिलते हैं जो शिक्षार्थियों के सीखने को बेहतर करने में लिए जाने वाले निर्णयों को लेने में मददगार हो सकते हैं।

एक दूसरा सवाल जो हमारे सामने आता है, वह यह है कि क्या अलग-अलग राज्यों, भाषाओं और सांस्कृतिक संदर्भों आदि में एक ही परीक्षण को लागू किया जा सकता है। अगर हम यह उम्मीद करते हैं कि सभी राज्यों के बच्चे एक खास उम्र में ज्ञान हासिल करें और कुछ खास चीजों को करना सीखें तो यह न्यायसंगत ही है कि उनकी जांच एक सामान्य परीक्षण के जरिए की जाए। बेशक अनुवाद, अनुकूलन और परीक्षण को अंतिम रूप देते समय विभिन्न सांस्कृतिक संदर्भों से जुड़े मुद्दों को ध्यान में रखा जाना चाहिए। हमारे ज्यादातर आकलन, उदाहरण के तौर पर, स्टूडेंट लर्निंग स्टडी 19 राज्यों में की गई थी। उसमें 15 भाषाएं और अलग-अलग सांस्कृतिक संदर्भ थे। अंतर्राष्ट्रीय आकलन जैसे कि टीआईएमएसएस (टिम्स) और पीआईएसए (पीसा) बहुत सारे देशों में एक साथ किए जाते हैं।

अक्सर हम यह पाते हैं कि बड़े पैमाने पर किए जाने वाले परीक्षणों और कक्षा परीक्षणों के बारे में कुछ गलतफहमियां होती हैं। इनमें से कोई भी दूसरे की जगह नहीं ले सकता। कक्षा परीक्षण इस बात के करीब रहेगा कि शिक्षक को क्या करने की जरूरत है। शिक्षक को इस बात की जरूरत होती है कि उसे तुरंत यह पता चले कि बच्चों ने पाठ पूरा करने के बाद क्या सीखा है, जो सिखाया गया है वह उसने समझा है या नहीं, आदि। तो कक्षा परीक्षण का मकसद अलग है। बड़े पैमाने पर परीक्षण किए गए हैं, सिर्फ इसी वजह से किसी को कक्षा परीक्षण छोड़ नहीं देने चाहिए और इसी तरह किसी ने कक्षा परीक्षण कर लिए हैं, तो वे नीतिगत स्तर पर सीखने के व्यापक पैटर्नों के बारे में जानकारी मुहैया करवाने के लिहाज से नाकाफी ही होते हैं।

आकलन में पेश आने वाली चुनौतियां

अब हम आकलनों को गढ़ते समय पेश आने वाली तंत्र संबंधी

चुनौतियों पर नजर डालते हैं। डिजाइन बनाते समय एक चुनौती हमारे सामने यह आती है कि अक्सर हमें यह नहीं पता होता कि हम उसे पाठ्यपुस्तक की बुनियाद पर बनाएं या जमीनी हकीकत की बुनियाद पर। उदाहरण के लिए, जब हमने आंध्र प्रदेश में 2004 में काम करना शुरू किया और उस वक्त यह पहली बार था कि ईआई सरकारी स्कूली तंत्र के साथ काम कर रही थी। हमने पाया कि पाठ्यपुस्तक तो एक स्तर पर है और बच्चों का स्तर उसकी तुलना में काफी कम है। तो हम क्या करें? क्या हम पाठ्यपुस्तक के साथ रहें या बच्चों के स्तर के साथ जाएं और इस बात को पक्का कैसे करें कि हमें सामान्य विवरण मिले? यहां पर अपने औजारों की शुरुआती जांच ने हमारी मदद की। इसी तरह से पाठ्यपुस्तकों में विसंगतियां मौजूद हो सकती हैं। उदाहरण के लिए, 'नगर निगम बेंच मार्क अध्ययन' के दौरान गणित की पाठ्यपुस्तकों के विश्लेषण के वक्त उत्तराखंड में यह पाया गया कि कुछ किताबों में अंग्रेजी के संख्यांक काम में लिए गए थे जबकि कुछ किताबों में देवनागरी के संख्यांक काम में लिए गए थे और तो और कुछ खास संख्यांकों के लिए उन्हीं किताबों के अलग-अलग पन्नों पर अलग-अलग तरह के फॉन्ट/लिपि का इस्तेमाल किया गया था।

एक मानकीकृत अध्ययन, जिसमें पूरे राज्य में एक ही परीक्षण पत्र काम में आने वाला हो, में यह तय करना होता है कि परीक्षण में संख्यांक किस तरह से दर्शाए जाएंगे। शिक्षकों के साथ किए गए साक्षात्कार से इस बात को सुलझाने में मदद मिली। विभिन्न क्षेत्रों के बीच भी पाठ्यचर्या में शामिल विषयवस्तु को लेकर अनपेक्षित फर्क पाए जा सकते हैं। उदाहरण के लिए, गुजरात में शुरुआती कक्षाओं में ज्यामिति को शामिल नहीं किया जाता है जबकि बाकी सभी राज्यों में शामिल किया जाता है। तो आप इसका क्या करेंगे? परीक्षण को बनाते समय इन क्षेत्रों के सवाल शामिल किए जाएं या नहीं, इस बारे में निर्णय लेना पड़ता है और उसके लिए वाजिब कारण देना पड़ता है। प्रश्न पत्र वाले परीक्षण में भाषा संबंधी कुछ दूसरी पेचीदगियां भी हमने देखीं। उदाहरण के लिए, कुछ क्षेत्रों जैसे आंध्र प्रदेश में हमने पाया कि वहनीय या वहन करने लायक निजी स्कूलों में भी बच्चों को चार से पांच भाषाएं सिखाई जाती हैं। वह अंग्रेजी माध्यम का स्कूल हो सकता है, लेकिन बच्चे हिंदी, उर्दू या तेलगु बोलने वाले होंगे और शिक्षक भी कक्षा में इन भाषाओं के मिश्रण की मदद से बच्चों के साथ संवाद करते नजर आएंगे। ऐसे मामलों में हमने जब पर्चे बनाए, तब उसी पर्चे के पहले खड़े खाने में सवाल को हिंदी में, दूसरे खड़े खाने में उर्दू में, तीसरे खड़े खाने में तेलुगु में दिया और बच्चों को इस बात की अनुमति दी कि वे जिस भाषा में आसानी महसूस करते हों, उसमें जवाब दें। एक दूसरी चुनौती यह होती है कि हम दक्षताओं को कैसे तय करेंगे?

क्या सीखा जाना है और कब सीखा जाना है? उदाहरण के लिए, आंध्र प्रदेश की कक्षा 3 की गणित की पाठ्यपुस्तक परीक्षण के दो दौर के बीच में संशोधित कर दिया गया।

शिक्षार्थियों संबंधी चुनौतियों के संदर्भ में कक्षा 1 व 2 परीक्षण के लिहाज से आदर्श उम्र नहीं है और शिक्षार्थियों के मूड के कारण मापन के दौरान कई तरह की समस्याएं पैदा होती हैं और आप सामान्यतः परीक्षण को नहीं कर सकते जब तक कि उसे करने की कोई बहुत बड़ी मजबूरी या अनिवार्यता न हो। विषयवस्तु संबंधी भी कई तरह के मुद्दे होते हैं, उदाहरण के लिए, पूरा शब्द लिखना बनाम छोड़ा गया वर्ण लिखना। हो सकता है कि बच्चा पूरा शब्द तो लिख ले लेकिन वह वर्णों को लिखने में नाकाम रहे क्योंकि वे उसके लिए अजनबी हों। हमने अक्सर पाया है कि जब हम सरकारी स्कूली तंत्र के साथ काम करते हैं तब वहां के बच्चे पढ़ने में नाकाम रहते हैं। कई बार बच्चे जानते हैं, उदाहरण के लिए, गणित के किसी सवाल को हल करना, लेकिन चूंकि वे निर्देश नहीं पढ़ पाते। नतीजन वे सवाल को भी नहीं पढ़ पाते। तो आप इस जानकारी को कैसे हासिल करेंगे। नतीजे में, परीक्षण में हमने एक नया घटक गढ़ा, जिसे हम 'समूह में मौखिक पढ़ना' कहते हैं। इसका मतलब यह है कि वहां एक आकलनकर्ता होता है जिसे सवाल को पढ़ने के लिए प्रशिक्षित किया गया होता है। आकलनकर्ता उस सवाल को मानकीकृत तरीके से एक बार पढ़ता है। बच्चे के लिए रुकता है, फिर उसे दोहराता है, फिर रुकता है और यह जांचता है कि सभी बच्चों ने जवाब दिया है या नहीं और उसके बाद अगले सवाल की ओर बढ़ जाता है। संयोगवश, राष्ट्रीय शैक्षिक अनुसंधान एवं प्रशिक्षण परिषद् ने आठवें सर्वे में इसकी शुरुआत की है, हालांकि वे इसे अंतर्क्रियात्मक विधि का नाम देते हैं। टीवी के कार्यक्रमों जैसे कौन बनेगा करोड़पति की वजह से मोटे तौर पर वस्तुनिष्ठ सवाल कम या ज्यादा समझ लिए जाते हैं। यहां तक कि दूर-दराज के इलाकों में भी कोई भी यह पता लगा सकता है कि बच्चे विकल्प (ऑप्शन) अ/ब/स जैसी चीजों को जानते हैं।

जैसा कि मैंने कहा था कि जब डिजाइन संबंधी चुनौतियों की बात आती है तो यह सवाल उठता है कि कौनसा हिस्सा लिखित में होना चाहिए और कौनसा मौखिक में रखा जाना चाहिए? व्यक्तिगत तौर पर मौखिक परीक्षण की लागत बहुत ज्यादा बैठती है। इसमें आकलनकर्ता अपनी बेहतरीन क्षमता के साथ काम करे तब भी तीन या चार घंटे में बहुत ही कम बच्चों की जांच कर सकता है। डिजाइन पर लिए जाने वाले निर्णयों को लेते समय इस बात पर ध्यान देने की जरूरत होती है कि मौखिक तौर पर ली जाने वाली जानकारी, लिखित तौर पर ली जाने वाली जानकारी की गुणवत्ता

को बढ़ाएगी या नहीं। इसी तरह, सवालों की किस्मों के मामलों में दो तरह की किनारे या सिरे देखने को मिलते हैं। कुछ सालों पहले हर कोई वस्तुनिष्ठ सवालों की खिलाफत करता पाया जाता था। अब यह मामला एक दूसरी किनारे या सिरे पर पहुंच गया है- बहुत सारे परीक्षणों में सिर्फ वस्तुनिष्ठ सवाल ही होते हैं। बड़े कार्यक्रमों में, कुछ ऐसी खुले किस्म के सवाल रखना बेहतर रहता है जिसमें बच्चा किसी समस्या को हल कर सके, चित्र बना सके, कुछ लाइनें लिख सके, आदि। डिजाइन बनाते समय यह बात ध्यान में रखनी चाहिए कि सवालों का मकसद सिर्फ याददाश्त के आधार पर याद करके दोहराना भर है या अवधारणात्मक समझ की जांच करना है और इन दोनों को किस अनुपात में परीक्षण में शामिल किया जाना चाहिए।

जैसा कि मैंने शुरुआत में कहा था, कि आप कठिनाई के स्तर को कैसे तय करेंगे, तो मनोमिति (साइकोमेट्रिक्स) के मामले में भी चुनौतियां होती हैं। सामान्य वक्र के बारे में हमें कुछ मान्यताएं तय करने की जरूरत पड़ती है और इसी तरह पर्व की विश्वसनीयता को भी देखने की जरूरत होती है। इसका मतलब यह है कि हमारे पास पर्याप्त संख्या में सवाल होने चाहिए ताकि आंकड़े विश्वसनीय हो सकें। परीक्षण को बनाने के दौरान सवालों को अंतिम रूप देने के बाद उसकी जांच यानी पायलट करने के वक्त हमें आधुनिक तकनीकों जैसे आईआरटी (आइटम रिस्पॉन्स थियरी) का भी इस्तेमाल करने की जरूरत पड़ती है। आईआरटी बुनियादी तौर पर सवाल की कठिनाई और सवाल विभेदीकरण की जांच करता है। उदाहरण के लिए, अगर दो बच्चे परीक्षा दे रहे हैं और मान लो कि दोनों को दस में से सात अंक मिले। एक ने सात अंक कठिन सवालों का जवाब देकर हासिल किए और दूसरे ने सात अंक कुछ सरल सवालों का जवाब देकर हासिल किए। लेकिन जब आप परीक्षण के आंकड़ों को देखेंगे तो आपको दोनों बराबर नजर आएंगे। तो बच्चों की क्षमताओं को तय करते वक्त आईआरटी न सिर्फ यह जांच करती है कि शिक्षार्थियों ने सवाल को सही किया है या नहीं, बल्कि यह भी जांच करती है कि उसने सरल सवाल को चुना है या कठिन सवाल को चुना है।

आकलन की प्रक्रिया

आम तौर पर, बड़े पैमाने पर किए जाने वाले परीक्षणों में तीन कदम उठाए जाते हैं। पहले कदम में परिस्थिति का विश्लेषण किया जाता है या एक शुरुआती परिकल्पना बनाई जाती है। हम पाठ्यपुस्तकें, पाठ्यक्रम, पाठ्यचर्या, आदि देखते हैं और इस स्तर पर सवालों को बनाना शुरू कर दिया जाता है। एक बार सवालों के बन जाने के बाद, दूसरे कदम में हम असल में जमीनी स्तर पर

जाकर कुछ शिक्षकों का साक्षात्कार करते हैं और यह जांच करते हैं कि सवाल ठीक है या नहीं। तब पूर्व परीक्षण किया जाता है। पूर्व परीक्षण सिर्फ सवालों का हो सकता है जिसे दो या तीन स्कूलों में किया जाता है और तब हम पर्वों को बनाते हैं। यही वक्त होता है जब खाके (ब्लू प्रिंट) को अंतिम रूप देते समय के मुद्दों को ध्यान में रखा जाता है।

तीसरे कदम में फील्ड में जाकर परीक्षणों की जांच (पायलट) करते हैं। परीक्षण की जांच में परीक्षण की पूरी प्रक्रिया की रिहर्सल की जाती है क्योंकि बड़े पैमाने पर किए जाने वाले आकलन को सिर्फ डिजाइन में और कागज में ही सही नहीं होना चाहिए, बल्कि उससे जमीनी स्तर पर मिलने वाले आंकड़े भी सही होने चाहिए, इसके लिए पूरी प्रक्रिया को करके देखा जाता है। पूरी प्रक्रिया में एक सिरे से दूसरे सिरे तक गुणवत्ता को मानक स्तर पर रखा जाता है। तो फील्ड में किया गया जांच परीक्षण हमें फील्ड के मुद्दों, कि क्या वह परीक्षण काम करेगा या नहीं, आदि, के बारे में जानकारी देता है। इसके बाद उसके दूसरे संस्करण बनाए जाते हैं, उनके अनुवाद किए जाते हैं, वैधता जांची जाती है आदि। हम जिन मुद्दों का सामना करते हैं उन्हें आप मेरी प्रस्तुति में देखेंगे।

किसी भी आकलन के केन्द्र में उस परीक्षण में काम लिए गए सवाल होते हैं। अच्छे सवालों को बनाते समय इन बातों का ध्यान रखा जाता है : 1. सीखने के जिस नतीजे को सवाल जांचने वाला है, क्या वह महत्वपूर्ण है?, 2. सवाल को जिस इरादे से बनाया गया है या सवाल जिस कौशल की जांच करता नजर आता है, क्या वही सवाल हकीकत में किसी दूसरे कौशल की जांच करता है?, 3. क्या सवाल शिक्षार्थी की उम्र या काबिलियत के लिहाज से उपयुक्त है?, 4. क्या सवाल में ऐसी कोई अस्पष्टता है जिससे कि शिक्षार्थी उसके एक से ज्यादा मतलब निकाल ले?, 5. क्या सवाल एक ही वक्त में तयशुदा के बजाय कई सारे नतीजों/कौशलों/दक्षताओं की जांच कर रहा है?

भारत जैसे देश में पूरे देश के स्तर पर, यहां तक कि राज्य के स्तर पर भी तुलनीय परीक्षण के लिए बहुभाषाई अध्ययन करने भी जरूरी है क्योंकि उसमें भी शिक्षण का माध्यम बहुभाषिक पाया जाता है। इसकी वजह से सामने आने वाली चुनौतियां कुछ इस तरह की हैं:

1. भारतीय भाषाओं में पाई जाने वाली विभिन्नताएं:

शब्दावली परीक्षण में, हिंदी के एक सादे से दो शब्दांशीय उल्लू में व्यंजनों का कोई गुच्छा या समूह नहीं है और लिपि भी सरल है,

जबकि यही शब्द तेलगु में गुडालगुब्बाह हो जाता है जिसमें चार शब्दांश होते हैं और संयुक्ताक्षर और पेचीदा लिपि भी होती है। इसलिए ये सवाल आपस में तुलनीय नहीं है।

तेलुगु भाषा के परीक्षण में शब्द पहचान के काम में, पतंग का चित्र (तेलुगु में गलीपाटम्म) लेने से तेलगु में एक से ज्यादा जवाब मिलते हैं जैसे हिंदी का शब्द पतंग या अंग्रेजी का शब्द काइट। इस तरह के जवाबों के लिए पेचीदा किस्म के आंकड़ों की प्रक्रिया का इस्तेमाल करने की जरूरत पड़ती है।

एक गुंथी हुई भाषा जैसे तेलगु में एक सरल-सा वाक्य जैसे हिंदी में 'उसने दौड़ना शुरू कर दिया' एक शब्द में कहा जाता है जैसे 'पारुगुल-अनकिंचुन्नासडु' (शुरू किया दौड़ना)। इससे सभी भाषाओं में व्याकरण वाले हिस्से में परीक्षण को बनाने में मुश्किल आती है।

2. अंग्रेजी के इस्तेमाल की वजह से पाई जाने वाली विभिन्नताएं:

गणित के एक परीक्षण में एक सवाल में 75 में से 43 को घटाना अंग्रेजी में लिखा था, जिसमें वियोजक (घटाई जाने वाली संख्या), वियोज्य (जिस संख्या में से वियोजक को घटाना है) से पहले लिखा गया था ताकि गणना करने से पहले संख्याओं को क्रम में रखने की जरूरत पड़े, जब इसका अनुवाद हिंदी और दूसरी भाषाओं में किया गया तो उस सवाल को बगैर क्रम बदले, सरल शब्दों में रख दिया गया जिसमें गणना करने की प्रक्रिया का सुराग मिलता था, जैसे 75 में से 43 घटाओ। इससे अलग-अलग भाषाओं के बीच कठिनाई के स्तर में तब्दीली आ गई।

अंग्रेजी के व्यापक इस्तेमाल ने कुछ शब्दांशों, जैसे ट्रैफिक सिग्नल, को तेलगु के असामान्य शब्द गमना-आगमना-गनता/घंटा-सूचिका की तुलना में ज्यादा आसान बना दिया है। दूसरे शब्दों में उसका अंग्रेजी समानार्थी स्थानीय भाषा के शब्द की तुलना में ज्यादा व्यापक तौर पर समझा जाता है। नतीजन, मदों को बनाते वक्त उन समानार्थी शब्दों को ध्यान में रखा जाता है जिनके परिचय व प्रसार का दायरा फैला हुआ हो।

3. सामाजिक, आर्थिक, सांस्कृतिक और भौगोलिक विविधताओं की वजह से विभिन्नताएं:

हालांकि अंग्रेजी के स्कूल शब्द का हिंदी में सही अनुवाद शाला किया गया, लेकिन मुमकिन है कि उसी राज्य के कुछ हिस्से इस शब्द से अपरिचित हों या कुछ क्षेत्रों में इसी शब्द का दूसरा रूप पाठशाला काम में लिया व समझा जाता हो। परीक्षण को बनाते

समय इस तरह के बोलीगत फर्कों को भी ध्यान में रखा जाना चाहिए।

बहुत मुमकिन है कि देश के उत्तर पश्चिम भाग, जैसे राजस्थान के बच्चे समुद्र और नारियल के पेड़ के विचार को समझ पाने में नाकाम रहें। इसी तरह, यह भी मुमकिन है कि देश की सांस्कृतिक विविधताओं के चलते जगहों, त्यौहारों, खानों, आदि की कहानियां भी अलग-अलग समुदायों में परिचित या अपरिचित हों। कामों के लिए प्रकरणों को चुनते वक्त इन विविधताओं को भी ध्यान में रखा जाना चाहिए।

आकलन में किए गए विश्लेषण और उससे बनने वाली रपटों में परीक्षण आंकड़ों से परे जाकर अंतर्दृष्टि भरी जानकारी देनी चाहिए ताकि सार्थक संवाद व बदलाव की राह बन सके। यह तभी मुमकिन है जब हमारा ध्यान सिर्फ प्रतिशत आंकड़ों को दर्ज करने से हटकर शिक्षार्थियों के सीखने में मजबूतियों व कमजोरियों के निदानात्मक विश्लेषण के श्रेणीकरण करने, गलतफहमियों के पैटर्नों को पहचानने, देश भर के अंदर और विभिन्न देशों की तुलना में समझ के मामले में दूसरे बच्चों की तुलना में हमारे बच्चे क्या जानते हैं और क्या करने के काबिल हैं- इसे जानने, आदि में लगे। आकलनों को न सिर्फ नीति-निर्माताओं को सीखने के पैटर्नों पर जानकारी मुहैया करवानी चाहिए बल्कि उससे शिक्षकों को भी काफी अंतर्दृष्टि मिलनी चाहिए ताकि वे कक्षा में अपने शिक्षाशास्त्र को और बेहतर बना सकें।

किसी तंत्र में शिक्षार्थियों के सीखने के मापन में आई परिपक्वता का नीचे लिखे स्तरों में समूहीकरण किया जा सकता है।

- 0 स्तर : सीखने पर बहुत कम ध्यान केन्द्रित किया गया।
- 1 स्तर : सीखने का जिक्र लेकिन औजारों पर बातचीत नहीं की गई।
- 2 स्तर : औजारों पर बातचीत, लेकिन मदवार आंकड़े नहीं।
- 3 स्तर : मदवार आंकड़े और विश्लेषण पर चर्चा की गई।

सिर्फ तीसरे स्तर पर आकर ही तंत्र हकीकत में यह समझना शुरू करता है कि बच्चे कैसे सीख रहे हैं। ◆

भाषान्तर : रविकांत