

Modrý test — nástroj kriminologické prognózy

MIROSLAV VEVERKA

I. Problematika kriminologické prognózy

Míra účinného zasahování do sociální reality je závislá na schopnosti předvídat. Každý člověk se musí dnes a denně pokoušet o pohledy do budoucnosti, odhadovat, jak se za určitých podmínek budou chovat jeho partneři a podle těchto odhadů volit plány svého vlastního chování. Správný odhad chování druhého a předvídaní jeho pravděpodobné reakce na určitou situaci může např. zabránit konfliktu, nesprávný odhad může naopak konflikt vyvolat.

Prognóza kolektivního chování společenských skupin a vrstev má pro život společnosti snad ještě větší význam, než je tomu v životě jedince. Intuice a subjektivní odhad na podkladě zkušenosti, s nimiž při tvorbě svých plánů chování vystačí jedinec, jsou ovšem slabou základnou pro odhad chování lidských skupin a vrstev. Zprostředkovanost informací a jejich ohromný rozsah brání v případě kolektivního chování tomu, aby se uplatnila lidská intuice alespoň na tom stupni spolehlivosti a platnosti jako při prognóze individuálního chování.

Pokusy o uvědomělé, či neuvědomělé předvídaní individuálního i kolektivního chování jsou staré jako lidstvo samo. Avšak teprve ve dvacátém století se objevuje snaha o systematický a vědecký přístup k tomuto dávnému problému. Úspěchy volebních předpovědí, jakož i prognózy populačních a ekonomických trendů ilustrují pokrok v předvídaní kolektivního chování. Pokud jde o předpovědi individuálního chování, bylo dosaženo zřetelných úspěchů v předvídaní prospěchu ve škole a úspěchu v povolání. První slibné výsledky byly získány i v oblasti prognózy adaptace v manželství a prognózy trestné činnosti. Na vědecké predikci lidského chování není nic tajemného. Je pouze třeba uspořádat data vycházející z lidské zkušenosti do stálých a spolehli-

vých kategorií a využít jich k pravděpodobnostním soudům.

Problematiku kriminologické prognózy, která je předmětem této stati, je třeba chápat v kontextu širšího světového trendu predikce individuálního i kolektivního chování.

Soudci i ostatní pracovníci kriminalistické fronty se snažili vědomě, či nevědomě předvídat budoucí chování delikventa již od té doby, kdy v trestním právu převládlo nápravné a preventivní pojetí funkce trestu nad původní funkcí represivní. Ale teprve v r. 1928, resp. 1929 se objevily první pokusy o prediktivní techniky, které jsou spojeny se jmény E. W. Burgesse a manželů Sh. a E. Glueckových. Uplynulých čtyřicet let přineslo celou řadu více nebo méně úspěšných pokusů o předvídaní delikventního chování, z nichž většina je orientována na předpověď recidivy u pachatelů propouštěných z výkonu trestu. Také v socialistických zemích je v posledních letech věnována zvýšená pozornost kriminologické prognóze. V Polsku byly již ukončeny první úspěšné výzkumy [17]. Značný zájem o problematiku kriminologické prognózy se projevuje v SSSR [7; 12; 14; 18; 19; 20 aj.].

Zakrzewski [25], který přehledně uvádí dosavadní výzkumy, rozlišuje tři hlavní typy kriminologických prognóz: 1. intuitivní; 2. analyticko-subjektivní; 3. výpočetní (v literatuře též nazývanou statistickou, matematickou, objektivní apod.).

V *intuitivní prognóze* delikventního chování mají hlavní slovo dostatečně nedefinované kategorie, jako je např. osobní dojem, a metodicky nespecifikovaný způsob, jímž se dochází k závěrům. V praktickém životě je aplikován zejména sociálními pracovníky, v domovech výchovy mládeže a v nápravně výchovných ústavech, kde je podkladem různých osobních posudků. *Analyticko-subjektivní prognóza* vyplývá z rozboru a zhodnocení dílčích údajů, při nichž hodnotící využívá svých

odborných vědomostí a poznatků vědy. Do určité míry se ovšem i zde uplatňuje intuice, je však pouze jedním z četných činitelů posuzovacího procesu. Příkladem může být posudek psychiatra nebo rozsudek dobře kvalifikovaného soudce. *Výpočetní prognóza* naproti tomu spočívá ve vyčíslení stupně pravděpodobnosti příštího delikventního chování jedince a svou specifikovanou a jednoznačnou metodikou je do značné míry nezávislá na osobě hodnotitele.

Různé typy výpočetních prognóz trestné činnosti jsou založeny na stejném myšlenkovém postupu. U zkoumaného jedince se nejprve zjišťuje přítomnost, či nepřítomnost určitých kritických znaků. Zjištěným znakům je přisouzena váha úměrná jejich souvislosti s trestnou činností. Ze součtu těchto vah se zjistí tzv. hrubé skóre. Přitom musí být k dispozici tabulka hrubých skóre pro měřenou populaci. Hrubé skóre zkoumaného jedince se vyhledá v uvedené tabulce a z tabulky se vyčte stupeň pravděpodobnosti delikventního chování.

Podle metody, již byly získány váhy pro jednotlivé kritické znaky a tabulka hrubých skóre, lze rozeznávat dva druhy výpočetních prognóz: 1. intuitivně vážené; 2. statisticky vážené. *Intuitivně vážené* postupy využívají odborných zkušeností autorů nebo nenáročných statistických šetření a používají málo diferencovaných vah, zpravidla v podobě několika celých čísel. *Statisticky vážené* postupy naproti tomu odvozují jemně diferencované váhy zcela empiricky z kriminologických výzkumů na podkladě postupů matematické statistiky. Do první kategorie patří řada psychologických testů použitých v pokusech o kriminalologickou prognózu, jakož i prediktivní techniky Burgesse, Freye, Gercka, Schiedta a dalších. Do druhé kategorie lze zařadit známé tabulky Glueckových [24]. Je zřejmé, že lepší predikční výsledky lze očekávat od postupů statisticky vážených. Jejich širšímu uplatnění však v minulosti bránil velký rozsah výpočetních prací, kterých je zapotřebí jak při vývoji těchto technik, tak při jejich masové aplikaci.

Skupina odborníků ČSLA, v níž spolu s autorem této stati pracovali JUDr. Václav Šedivý, MUDr. Zdeněk Heinzl, PhDr. Naďa Janoušková a ing. Vladimír Klimeš,

se pokusila o řešení tohoto problému za širokého využití samočinného počítače.

II. Vývoj Modrého testu

Problematika kriminologické prognózy byla pracovním týmem řešena v širším rámci výzkumu trestné činnosti a hrubé nekázně v armádě, o němž referujeme na jiném místě [23].

Uvedme nejprve základní myšlenku, s níž jsme přistupovali k projektování celého výzkumu. Předpokládali jsme, že podobně jako chování vůbec, i trestná činnost a další kázeňsky postižované formy nekonformního chování v armádě jsou výslednicí dvou rozsáhlých skupin faktorů, které se navzájem podmiňují a ovlivňují. Jsou to:

1. *Faktory vnitřní (endogenní)*, spočívající v psychice, v osobnosti jedince. Jsou víceméně dány jako relativně stabilizované znaky již před spácháním trestného činu.

2. *Faktory vnější (exogenní)*, dané aktuální situací a celkovým prostředím, v němž a pod jehož vlivem dochází ke spáchání trestného činu (tzv. situační proměnné).

Vlastní výzkum se pak pokoušel kvantifikovat jednak faktory osobnostní a faktory prostředí, jednak změřit kvalitu výsledného chování. Z takto kvantitativně vyjádřeného modelu byly potom samočinným počítačem izolovány faktory významné pro genezi trestné činnosti a současně byl zjišťován stupeň jejich závažnosti.

Projekt experimentu se opíral o jeden nezbytný předpoklad. Hned na začátku výzkumu bylo nutno s dostatečnou přesností změřit ony rysy osobnosti zkoumaných jedinců, které hypoteticky souvisejí s trestnou činností. Za tímto účelem jsme prostudovali zahraniční metody a techniky, které pro náš experiment přicházely v úvahu. Byly to především některé psychologické testy dotazníkového typu a některé pomůcky, speciálně konstruované k predikci delikvence.

Z psychologických testů nekriminologické povahy bylo relativně nejlepších výsledků dosaženo v analýze a prognóze delikvence pomocí Minnesotského dotazníku — MMPI [10]. Tento test dotazníkového typu je však příliš rozsáhlý a jen menší část z jeho 550 položek ukázala určitý vztah k trestné činnosti. Rovněž pomocí

Cattellova testu 16 P. F. byly zjištěny významné difference v osobnostním profilu normální a delikventní populace [4]. Tento test se pokouší v šestnácti navzájem nezávislých faktorech vyjádřit podstatnou část povahových rozdílů mezi lidmi. Již z tohoto důvodu byl pro naše účely pomůckou příliš obecnou. Na obecné a z kriminologického hlediska málo diferencované úrovni zůstává také Eysenckův dotazník MPI [6], konstatující u delikventní populace rozdíly v introverzi a extroverzi. Projektivní test Rorschachův, který v rukou specialisty může být dobrou diagnostickou pomůckou k diferenciaci některých delikventních typů [8] pak pro nás nebyl použitelný jednak pro zdlouhavé individuální snímání, jednak pro odborně náročnou a nedostatečně objektivní interpretaci. Společným nedostatkem všech uvedených pomůcek dále bylo, že dosud nebyly standardizovány na naši populaci a že tudíž nejsou pro nás jako exaktní měrný nástroj zatím dobře použitelné.

Z vysloveně kriminologických diagnostických pomůcek jsme se zaměřili především na ty, které jsou metodicky relativně nejlépe propracovány. Vedle již zmíněných predikčních tabulek manželů Glueckových to byla především Kvaraceova KD-škála [3, 79–80] a Haggerty-Olson-Wickmanovy BR-seznamy [2]. Neuspokojivé nebo vůbec chybějící údaje o reliabilitě a validitě uvedených pomůcek nás však varovaly před tím, založit tak rozsáhlý experiment pouze na nich.

Pokud jsme tedy chtěli zamýšlený výzkum vůbec realizovat, byli jsme postaveni před nutnost vyvinout nové, původní pomůcky k prognóze delikventního chování, samozřejmě při využití všech poznatků, které byly na tomto poli již shromážděny.

Z nejširšího hlediska je možný dvojí přístup ke zjišťování vnitřních faktorů příznačných pro delikventní dispozice — přímý a nepřímý. Přímá metoda spočívá ve zjišťování aktuálního profilu, v měření rysů osobnosti. Touto cestou směřuje použití uvedených psychologických testů.

Nepřímá metoda zachycuje profil osobnosti zprostředkovaně, přes kritické životní situace, jimiž byly delikventní sklonky vytvářeny nebo v nichž se manifestovaly. Jejím prototyp lze spatřovat v sociální predikční tabulce Glueckových. Tato

druhá metoda, vycházející z anamnézy, nezachycuje sice aktuální profil osobnosti tak přesně jako psychologický test, ale má jiné výhody. Dotaz na časově vzdálené a emocionálně již neutralizované kritické životní situace zpravidla nezraňuje respondenta tak, jako mnohé na současnost mířící otázky psychologického testu. Většina kritických situací je sama o sobě jasně definovatelná a bez obtíží identifikovatelná, takže jednoznačná odpověď zpravidla nečiní obtíže ani respondentu s nižším intelektem. Kritické životní situace ulpívají v paměti člověka dlouhou dobu a proto představují znaky s vysokou reliabilitou. Málo hodnoticích momentů, skrytý význam těchto zprostředkovaných znaků a možnost jejich objektivního ověření v případech potřeby vyvažují nedostatky této nepřímé metody.

Z důvodů vzájemné kontroly jsme zvolili k měření endogenních faktorů obě zmíněné metody. První byla realizována v podobě Zeleného testu, druhá se stala náplní Modrého testu. O Zeleném testu bude pojednáno na jiném místě. Modrý test, který svou vnější podobou velmi připomíná běžný sociologický dotazník, je vlastním předmětem tohoto pojednání.

Test má podobu sešitu s deskami z modrého kartonu. Listy sešitu jsou postupně kráceny tak, že každý následující list je užší než list předchozí. Do kapsy na zadní desce sešitu se zasouvá vyměnitelná vložka pro vyznačování odpovědí. Vložka je současně formulářem pro přenos údajů na děrný štítek. Test v dnešní podobě obsahuje 46 otázek. Jeho znění je k dispozici v českém, slovenském a maďarském jazyce. Respondent odpovídá tím způsobem, že u každé otázky najde předtištěnou odpověď, která u něho platí, a její číslo zaškrtně na viditelné části vložky při pravém okraji sešitu. Po obrácení stránky se objeví nový sloupec vložky pro vyznačování dalších odpovědí. Je-li respondent s odpověďmi hotov, popsaná vložka se vyjme a předá se k vyděrování na děrný štítek. Do kapsy sešitu se zasune vložka čistá. Tím je test připraven k novému snímání. Tato úprava dotazníku vede ke značné úspoře papíru při masovém snímání testu.

Modrý test je v podstatě dotazníkem zaměřeným na sociální zázemí respondenta. Jeho otázky byly vybrány z několika

základních oblastí, charakterizujících sociální prostředí, v němž jedinec vyrůstal. Především je v dotazníku podrobně zpracován úsek rodinných vztahů. Týká se to zejména úplnosti rodiny, vztahů mezi rodiči, vztahů rodičů k respondentu, postavení respondenta mezi sourozenci, alkoholu a jiné zátěže v blízkém příbuzenstvu a řady dalších faktorů, z nichž je možno usuzovat na celkovou atmosféru rodiny a na kvalitu plnění její společenské funkce. Zařazena byla i otázka týkající se jeslí, členství v dětské a mládežnické organizaci aj.

V dotazníku jsou dále zpracovány položky spojené s prospěchem a chováním ve škole, s dosaženou odbornou kvalifikací, s fluktuací v zaměstnání apod. Určitý indikační význam má i způsob trávení volného času, návštěva hostinců, tanečních zábav, trampingu apod.

Dotazník zachycuje též některé faktory, jež naznačují souvislost s duševní poruchou a celou řadou dalších znaků, které podle výsledků výzkumu souvisejí s pozdějšími sklony k trestné činnosti, nebo naopak s konformním chováním. Do Modrého testu byly zařazovány vesměs položky objektivně prověřitelné. Položky, u nichž tato možnost prověření není, byly zařazeny do Zeleného testu.

Je třeba poznamenat, že nepřímá souvislost uvedených otázek se stupněm delikventních dispozic není laikům na první pohled zřejmá, což má vliv na značnou spolehlivost výsledků.

Základem pro výběr položek Modrého testu byl předvýzkum na souboru 300 kriminálních recidivistů a 300 ukázněných vojáků, realizovaný autorem statí v r. 1964 až 1965 [23, 192—195]. Do původní makety Modrého testu byla také zařazena všechna kritéria sociální predikční tabulky Glueckových, což přineslo velmi užitečné poznatky [22].

Forma dotazníku nebyla zvolena náhodně. V předvýzkumu byl učiněn pokus zjišťovat faktory sociálního zázemí zkoumaných osob z různých úředních dokladů, ze zpráv o pověsti, z posudku zaměstnavatele, z posudků velitelů, ze záznamů o odměnách a trestech během vojenské služby, z výpovědi pachatelů o jejich osobních a rodinných poměrech, excerpcí dřívějších trestních spisů, psychiatrických posudků, jakož i pomocí dotazníků, vyplňo-

vaných rodiči zkoumaných osob. Výsledky jednoznačně potvrdily, že pro účely dostatečně přesného kvantitativního měření nejsou tyto techniky použitelné. Údaje z trestních spisů a ostatních úředních dokumentů byly často neúplné, vycházely z různých kritérií a nebyly srovnatelné, nebyly k dispozici ve všech případech a velmi často nebyly kvantifikovatelné. Také informace získané od rodičů nemohly zajistit úplnost potřebných dat. Údaje rozvedeného otce a matky se zpravidla navzájem velmi lišily. Kromě toho bylo nemožné touto cestou získat údaje z rodin, kde rodiče již nežili nebo kde odmítli spolupracovat. Z toho vyplynul úmysl získávat potřebné informace přímo od zkoumaných osob.

Modrý test byl od počátku koncipován jako plně automatizovaná diagnostická pomůcka. Úloha lidského činitele se omezuje na 15—20 minut, potřebných k jeho sejmutí. Sejmутí se provádí hromadně podle přesné instrukce a pod dohledem zcvičeného asistenta, který musí dbát na dodržení předepsané testovací atmosféry. Veškeré další úkony až po vytištění seznamů samočinným počítačem jsou plně mechanizovány a automatizovány. Použitím samočinného počítače při vyhodnocování testu jsme vyřešili několik překážek, které se stavěly do cesty širšímu využití prediktivních technik.

1. Mohli jsme zvolit složitý a vnitřně diferencovaný způsob vyhodnocení testu, jímž se pochopitelně ztrácí méně prvotní informace.
2. Vyřešili jsme problém chronického nedostatku odborných psychologů, kterých by jinak bylo zapotřebí k hromadnému použití testu.
3. Zajistili jsme objektivnost vyhodnocovacího procesu, který nezávisí na lidském činiteli.
4. Uspořili jsme mnoho lidské práce nutné k vyhodnocování masově používaných prediktivních technik.
5. Zkrátili jsme časové rozpětí mezi sejmутím testu a vyhodnocením výsledků na minimum, technicky dané přenosem testové vložky na nosné médium počítače, a na minimální strojový čas, potřebný k realizaci standardního programu.

Položky Modrého testu nepředstavují nějaký katalog kauzálních faktorů trestné činnosti. Vedle znaků, které mají více či méně kauzální povahu, zachycuje test řadu znaků symptomatického či indikačního

charakteru, jež jsou často spíše důsledkem nebo zprostředkovaným projevem tendencí obecně podporujících delikventní chování. Představují v podstatě pomůcku usnadňující rozhodování problémů, které zde existovaly již dříve, a umožňující poskytnout kvalitnější informace pro toto rozhodování. Techniky kriminologické prognózy tedy nemají nahradit rozhodování odpovědných pracovníků, ale mají doplnit jejich zkušenosti a poskytnout jim širší podklady pro rozhodnutí.

Praktický smysl úsilí o kriminologickou prognózu je stejný jako smysl moderní lékařské preventivní péče: po vyhledání (depistáži) ohrožených jedinců se zvýšenými předpoklady k trestné činnosti na ně soustředit výchovné a preventivní úsilí, které se při jiných postupech neefektivně rozptýluje do šířky. Stanovení prognózy umožňuje, aby zvlášť disponovaní jedinci byli pokud možno zařazeni do takového prostředí, které by jejich dispozice k delikventnímu chování nepodporovalo, ale naopak tlumilo. Umožňuje dále, aby takové dispozice mohly být cílevědomě oslabovány výchovným a převýchovným procesem.

Závěry kriminologické prognózy vycházející z analýzy vnitřních faktorů jsou v životě korigovány vlivem vnějších faktorů — celkového prostředí a situace, v níž se jedinec ocitne. Proto je také lépe hovořit o prognóze než o predikci delikventního chování. Mnoho jedinců se zřetelnými delikventními předpoklady a s relativně vysokou pravděpodobností delikventního chování se nikdy neocitne v situaci, která by tyto latentní předpoklady stimulovala. Jiné osoby s relativně nižší pravděpodobností podobného chování a s nižším skóre testu se naopak mohou dostat do tak těsnivé a provokující situace, že u nich k protizákonným projevům skutečně dojde. Situační faktory tedy snižují platnost kriminologické prognózy oproti hodnotám vypočteným z osobnostních faktorů.

Této okolnosti, která snad na první pohled může snižovat hodnotu popisovaných prognostických technik, lze však záměrně využít.

Vnější faktory jsou zpravidla mnohem proměnlivější, méně stabilní a tudíž také lépe přístupné kontrole, záměrnému výběru a usměrnění, než je tomu u relativně

stabilizovaných rysů osobnosti. Přes kontrolu vnějších faktorů lze tedy provádět prevenci trestné činnosti s menším množstvím vynaložené společenské energie, než je tomu při pokusech o přímou přestavbu faktorů vnitřních. To je praktická stránka problému. Za předpokladu dostatečně přesného změření vnitřních faktorů a skutečného chování lze vypočíst hodnotu znamenající vnějších faktorů, které svým vlivem vstoupily mezi prognózu a výsledek. To je teoretická stránka problému.

Možnosti Modrého testu byly vyzkoušeny v obou zmíněných směrech. První slibné výsledky se zdají potvrzovat výchozí hypotézu, jež tu byla naznačena.

III. Statistická analýza testu

K tomu, aby byl vytvořen skutečně účinný nástroj kriminologické prognózy, jsme využili zkušenosti soudobé psychologie s tvorbou a analýzou psychologických testů. Lienert [16, 13] shrnuje požadavky na dobrý test do tří hlavních a čtyř vedlejších kritérií. Dobrý test musí splňovat především tyto tři požadavky: 1. musí být objektivní; 2. musí být spolehlivý (reliabilní); 3. musí být platný (validní). Dále má být: 4. adekvátní; 5. srovnatelný; 6. hospodárný; 7. užitečný.

1. Test je *objektivní* tehdy, jestliže měří znaky osobnosti jednoznačně, tj. jestliže výsledky testu nejsou závislé na osobě vyhodnocovatele. Různí vyhodnocovatelé musí tedy dospět na základě stejného případu k totožným výsledkům. Tento požadavek není např. dostatečně zajištěn u známého Rorschachova testu (test inkoustových skvrn), kde se výsledky různých zpracovatelů od sebe zpravidla dosti liší. Automatickým zpracováním výsledků Modrého testu na samočinném počítači jsme požadavku objektivnosti testu v uvedeném smyslu maximálně vyhověli, neboť jsme lidskou složku z vyhodnocovacího procesu zcela vyloučili.

2. *Spolehlivost* (reliabilita) testu znamená, že příslušné znaky osobnosti jsou testem měřeny exaktně, tj. že při opakování testu za stejných podmínek dává test stejné výsledky (tzv. retestová reliabilita). Stupeň reliability testu tedy vypovídá jednak o rozsahu skutečného rozptylu, zachyceného testovými skóre, jednak o riziku chyby způsobené nepřesným měřením.

Nedostatečná spolehlivost testu může mít tři příčiny: a) nedostatečnou objektivitu testu; b) nedostatečnou stabilitu testu; c) nestálost měřeného znaku. Protože jsme předem odstranili možný zdroj nespolehlivosti sub a) a výběrem otázek minimalizovali zdroj sub c), museli jsme se v další etapě soustředit na zajištění dostatečně vysoké stability testu.

Stabilita testu, jak o tom ještě bude řeč, je zvláštní aspekt spolehlivosti testu. Stabilita charakterizuje spolehlivost testu bez zřetele na skutečné konkrétné provedené měření. Kdyby testem měřené znaky osobnosti byly dostatečně stálé a kdyby nekolísaly v čase, pak by stabilita testu měla být rovna jeho spolehlivosti. U mnohých psychologických testů jsou však vedle údajů o reliabilitě uváděna též data o stabilitě testu, a to z toho důvodu, aby se vlastnosti testu nesměšovaly s vlastnostmi měřených subjektů.

Reliabilita testu se udává korelačním koeficientem. U testů dotazníkového typu se obecně vyžaduje spolehlivost vyšší než 0,7, u zvlášť kvalifikovaných testů je žádoucí spolehlivost 0,9 [16, 14].

3. Test je *platný* (validní) tehdy, jestliže skutečně měří ten znak osobnosti, k jehož měření je určen, a nikoliv znak jiný, jemu podobný. Stupeň validity testu vypovídá tedy jednak o tom, jak test skutečně plní funkci, k níž je určen, jednak dává představu o riziku chyby způsobené hodnocením testových výsledků. Validita testu bývá definována jako stupeň korelace testových výsledků výběrového souboru s kritériem validity měřeným na témže souboru. Nemůžeme tedy mluvit o validitě testu vůbec, ale vždy jen o validitě vzhledem k určitému definovanému kritériu platnosti. Za testy s vysokou validitou se považují ty testy, jejichž platnost je vyjádřena korelačním koeficientem $r > 0,6$; při hodnotě $0,6 > r > 0,4$ jde o testy se střední validitou, nízkou platnost pak mají testy, jejichž validita vyjádřená korelačním koeficientem je nižší než 0,4 [16, 15].

Nižší platnost testu může být způsobena několika faktory: a) nedostatečnou spolehlivostí testu; b) nedostatečnou adekvátnos-

tečný fakt soudního potrestání vyšetřené osoby, pak je zřejmé, že v tomto verifikačním kritériu se uplatňují tři skupiny faktorů: 1. osobnostní (vnitřní) faktory měřené testem; 2. situační (vnější) faktory, testem nepodchycované; 3. faktory spojené s rysy subjektivismu v úsudcích orgánů trestního řízení, zejména v případech mezního typu. Činitelé sub 2. a 3. nejsou testem zachytitelné, ale přesto se v kritériu platnosti projevují. V důsledku toho snižují spolehlivost samotného kritéria validity, takže platnost Modrého testu, vyjadřovaná ve vztahu k osobám skutečně odsouzeným, nikdy nemůže nabýt hodnoty 1,0 nebo hodnot blízkých této hranici.

Zatímco údaje o reliabilitě musí být zásadně uvedeny u všech testů, údaje o validitě se v některých případech nevyžadují. Bývá to zejména tam, kde kritérium platnosti je méně spolehlivé než test sám. Testové výsledky pak mají povahu prvotních informací o znaku, který není jinak dostatečně přístupný měření.

4. *Adekvátní* test je takový test, který skutečně měří ten znak osobnosti, který má měřit, přičemž se ponechává stranou spolehlivost měření. V tomto smyslu se hovoří též o diagnostické hodnotě testu. Míní se tím stupeň reprezentativnosti, s nímž test zachycuje požadovaný rys osobnosti. Adekvátnost testu nelze vyčíslit koeficientem, jako je tomu u reliability a validity testu. Stupeň adekvátnosti lze však posoudit na základě rozdílu mezi koeficientem platnosti a koeficientem spolehlivosti: je-li tento rozdíl velký, je adekvátnost testu nízká a naopak. Již z toho, co bylo řečeno, vyplývá, že u Modrého testu lze očekávat spíše střední než vysokou adekvátnost. Test byl koncipován jako diagnostická pomůcka s vysokou adekvátností vzhledem k vnitřním (endogenním) faktorům kriminality; vzhledem ke skutečnému výsledku — ke skutečně realizovaným vnitřním faktorům v podobě trestné činnosti — však může dosáhnout jen adekvátnosti omezené.

5. Test je *srovnatelný* tehdy, jestliže jsou buď k dispozici dvě paralelní formy téhož testu, nebo jestliže máme k dispozici jiný

dispozic byl u obou těchto testů různý, kritérium validity bylo stejné. Oba testy měřené na téže souboru vykazují vzájemnou korelaci $r = 0,508$. Tím je umožněna individuální kontrola platnosti testových výsledků u této jedince.

6. O *hospodárném* testu můžeme hovořit tehdy, jestliže test a) klade malé časové nároky na snímání; b) spotřebuje málo materiálu; c) je jednoduše aplikovatelný; d) je přizpůsoben hromadné aplikaci; e) je rychle a pohodlně vyhodnotitelný. Všem těmto požadavkům jsme se pokusili vyhovět způsobem, který byl již popsán.

7. *Užitečný* test v Lienertově smyslu je takový test, který měří ty znaky osobnosti, jejichž sledování je důležité z hlediska praktické potřeby. Vysokou užitečnost má tedy takový test, který ve své funkci nemůže být zastoupen jiným a naopak, test má jen malou užitečnost v případě, že zjišťuje takové znaky osobnosti, jež by stejně dobře mohly být měřeny řadou jiných testů.

Z toho, co bylo až dosud řečeno, vyplývá, že o kvalitě testu rozhoduje v prvé řadě jeho spolehlivost a jeho platnost. Vysoce spolehlivý a platný test je možno vytvořit ovšem jen tehdy, jestliže všechny jeho položky (otázky) budou mít co nejvyšší spolehlivost a platnost. K dosažení tohoto cíle směřovala celá etapa tvorby testu nazývaná *položkovou analýzou*. Položková analýza testu představuje přijatelnou aproximaci k teoreticky žádoucí, ale prakticky zatím obtížně realizovatelné mnohonásobné korelaci. V rámci položkové analýzy byla každá položka testové makety podrobena statisticko-analytickému zkoumání tak, aby se zjistila hodnota, s níž přispívá žádoucím charakteristikám testu. Podle výsledků položkové analýzy byly pak některé položky vyloučeny nebo upraveny.

Výběrový soubor pro položkovou analýzu byl vytvořen tak, aby zachytil okraje předpokládaného kontinua. Toto hypotetické kontinuum je představováno stupni společenské přizpůsobivosti, resp. nepřizpůsobivosti. Proto delikventi představují jedince, kteří jsou na extrémním okraji sociální maladaptace. Aby bylo dosaženo maximálního informačního rozpětí souboru, byl jako jejich protiklad zvolen stejně velký soubor s vysokým stupněm sociální adaptace (elita). Celkový výběrový soubor pro statistickou analýzu testu se tak sklá-

dal z 500 delikventů a 500 příslušníků elity [23, 195].

Soubor delikventů byl vybírán z vojáků základní služby, kteří se dopustili trestných činů během vojenské služby a byli za ně odsouzeni k nepodmíněným trestům odnětí svobody. Již tímto způsobem byli tedy ze souboru vyloučeni jedinci, kteří se dopustili méně závažných činů a dostali jiné tresty než tresty nepodmíněné.

Soubor „elity“ byl sestaven z příslušníků Pohraniční stráže, čímž bylo využito hodnocení prováděné při výběru jedinců do tohoto útvaru. Z vojáků byli pak do výběrového souboru zařazeni ti, kteří podle dosavadního pozorování velitelů měli předpoklady pro ukázněné chování během vojenské služby. Tímto dvojím výběrem bylo zajištěno, aby se do kontrolního souboru dostali pokud možno jen ti vojáci, kteří podle dosavadních poznatků jsou skutečně dobře sociálně adaptováni.

Prvním krokem položkové analýzy bylo zjišťování *statistické významnosti* odpovědi pomocí testu χ^2 . Odpovědi statisticky nevýznamné ($P > 0,05$) byly ihned vyloučeny. Další položky pak odpadly v průběhu položkové analýzy, takže konečná forma testu obsahuje pouze položky vysoce statisticky významné ($P < 0,001$).

Dalším krokem bylo přisouzení vah jednotlivým odpovědím, přičemž jsme u Modrého testu použili postupu manželů Glueckových, jak je o tom referováno na jiném místě [22, 43]. V námi upravené podobě byla *váha K* jednotlivých typů odpovědi vypočtena podle vzorce:

$$K = \frac{d}{d + e}$$

d = relativní četnost dané odpovědi v souboru delikventů

e = relativní četnost dané odpovědi v souboru elity.

Aritmetický průměr těchto vah, přisouzený odpovědím konkrétního respondenta, byl pak označován jako *hrubé skóre*. Hrubé skóre bylo v této etapě souhrnnou kvantitativní charakteristikou testových výsledků respondenta.

Pro každou položku bylo pak vypočteno několik indexů charakterizujících její vlastnosti.

Index statistické obtížnosti vyjadřuje v procentech, jak často se ve zkoumaném

souboru vyskytuje kriminogenní odpověď ($k = K > 0,5$).

Byl počítán podle vzorce:

$$I_0 = 100 \cdot \frac{N_k}{N}$$

N_k = počet respondentů, kteří na danou otázku dali kriminogenní odpověď

N = počet všech respondentů

V celkové stavbě testu musí být respektováno určité rozložení položek podle jejich statistické obtížnosti. Index obtížnosti byl kromě toho použit při výpočtu indexu stability a kritériově stabilitního indexu položky.

Pokud jde o validitu položky, projevuje se tím, že platná položka je pozitivně zodpovídána velkým počtem respondentů s vyšším stupněm měřeného znaku a odpovědi negativní dávají naopak respondenti s nízkým zastoupením měřeného rysu. Z praktických důvodů se platnost položky nepřezkoušuje přímo na měřeném rysu osobnosti, ale na souhrnných skóre celého testu. Získaný index se nazývá *indexem rozlišnosti* a charakterizuje položku z toho hlediska, jak přispívá k rozlišení respondentů, v našem případě delikventů od elity. Index rozlišnosti je roven koeficientu korelace mezi dvojím typem odpovědi (kriminogenní a nekriminogenní) a hrubým skóre. Korelace dvou proměnných, z nichž jedna (odpovědi) má alternativní povahu a druhá (hrubá skóre) má spojitý charakter, se vyjadřuje pomocí tzv. bodově biseřální korelace. Tímto způsobem byl také index rozlišnosti vypočten:

$$I_r = \left(\frac{\sum X_k}{N_k} - \frac{\sum X}{N} \right) \sqrt{\frac{N_k}{N - N_k}} \cdot \frac{1}{\sqrt{N \sum X^2 - (\sum X)^2}}$$

X = hrubá skóre respondentů, kteří dali kriminogenní odpověď

X = hrubá skóre všech respondentů

Vzhledem k tomu, že jsme v našem případě měli k dispozici jasně definované vnější kritérium (delikventi — elita), vypočetli jsme kromě indexu rozlišnosti též přímo *index validity* položky. Obě proměnné (odpovědi i kritérium) měly v tomto případě alternativní charakter; proto byl k jejich výpočtu zvolen koeficient te-

trachorické korelace, počítaný ze čtyřpolní tabulky aproximací:

$$I_v = \cos \left(\frac{180^\circ}{1 + \sqrt{ad/bc}} \right)$$

Z podobné čtyřpolní tabulky byl vypočten též *index spolehlivosti* položky. Podkladem k jeho výpočtu byly jednak odpovědi při prvním snímání testu, jednak odpovědi týchž respondentů po osmi měsících.

Na základě takto získaných údajů o každé položce bylo přistoupeno k výběru těch nejvhodnějších otázek, které nejvíce přispívají k požadovaným kvalitám testu. Bylo k tomu využito postupu navrženého Gulliksenem. Z indexu rozlišnosti a indexu validity byl nejprve vypočten *index stability* (I_s) a *kritériově stabilitní index* (I_k) tím způsobem, že první indexy byly vynášeny rozptylem položky (σ_i).

$$I_s = I_r \cdot \sigma_i \quad I_k = I_v \cdot \sigma_i$$

$$\sigma_i = \sqrt{p \cdot q} \quad p = \frac{I_o}{100} \quad q = 1 - p$$

Oba takto vypočtené indexy se potom u každé položky vynesly do systému souřadnic a grafickým způsobem, který popisuje Gulliksen [9] i Lienert [16], byly postupně vylučovány nejméně hodnotné položky. Tento postup byl na obdobných souřadnicích opakován s indexem spolehlivosti a kritériově stabilním indexem, jakož i s indexem spolehlivosti a indexem stability.

Z původních 59 položek Modrého testu bylo ponecháno 42 položek, které nejlépe přispívaly k potřebným charakteristikám testu (zbývající 4 položky jsou v testu z jiných důvodů).

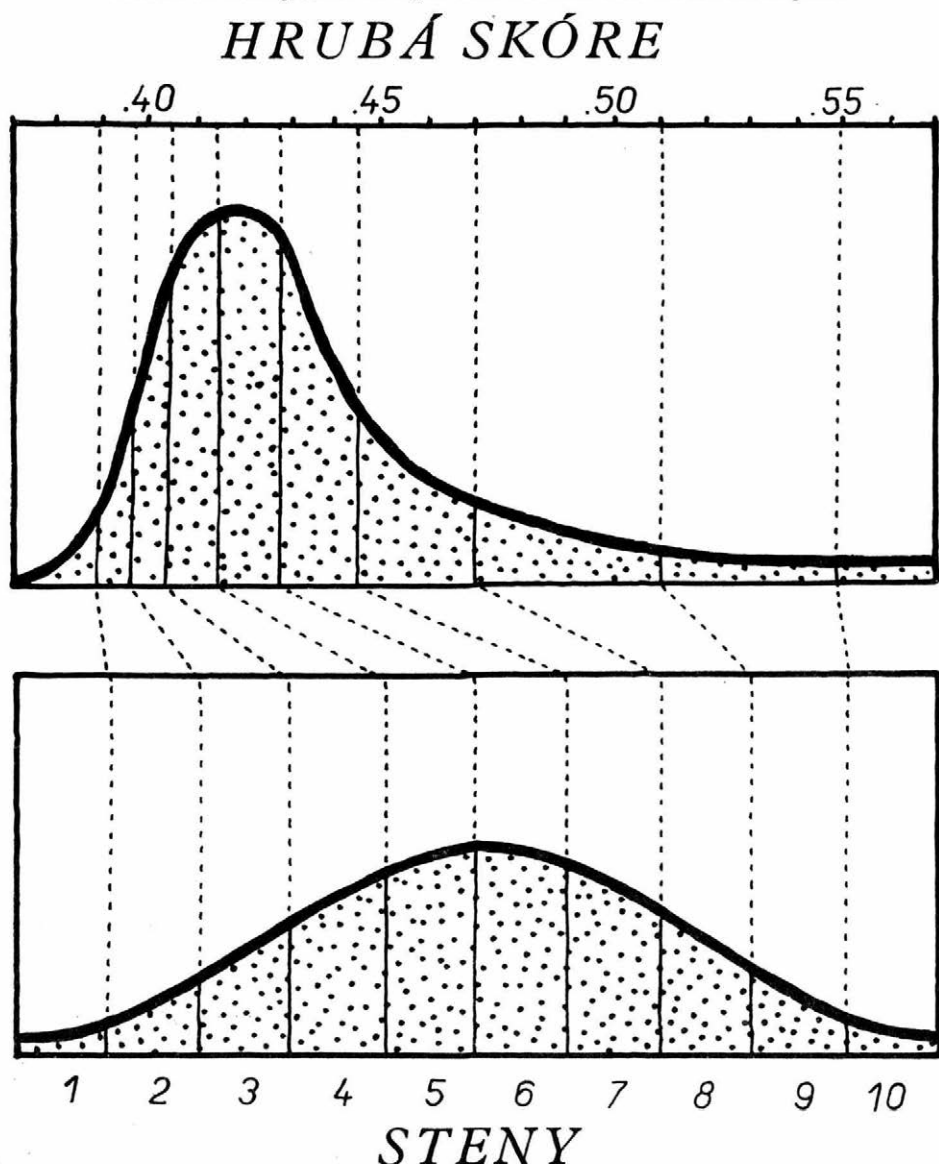
IV. Cejchování a normování testu

Pro další etapu vývoje testu byla velmi důležitá analýza rozdělení hrubých skóre. Bylo nezbytně nutné provést takovou analýzu na výběrovém souboru, který je reprezentativní pro celkovou populaci, na níž má být test v budoucnosti používán. K tomuto účelu byl Modrý test aplikován u 1632 osmnáctiletých chlapců ve dvou okresech v Čechách, v jednom okrese na Moravě a v jednom na Slovensku. Bylo pečlivě zajištěno, aby tento výběrový soubor byl reprezentativní pro celou československou mužskou populaci ročníku na-

rození 1949. Přitom byla sledována hlediska národnosti, vzdělání, profesionálních skupin, velikosti bydliště, rodinného stavu, dřívějšího potrestání a základní kombinace těchto proměnných.

rozdělení (Gaussova). S tímto cílem byla provedena plošná transformace empirického rozdělení tak, aby bylo dosaženo quasinnormálního rozdělení. Přitom jsme podle McCalla nejprve vypočetli centily empi-

Obrázek 1. Transformace hrubých skóre Modrého testu do stenové stupnice



Hrubá skóre tohoto výběrového souboru byly pak podkladem pro cejchování a normování Modrého testu. Rozdělení hrubých skóre v reprezentativním souboru je uvedeno na obr. 1 nahoře. Již z pouhého pohledu na tvar empirického rozdělení je patrna jeho asymetričnost. Pro naše účely však bylo žádoucí dosáhnout normálního

rozdělení a dosud lineární stupnici hrubých skóre jsme transformovali tak, aby příslušná procenta transformovaného rozdělení odpovídala požadavkům ideální Gaussovy křivky. Intervaly mezi hrubými skóre se tímto způsobem buď zvětšily, nebo zmenšily, podle toho, v které části rozdělení se hrubé skóre nalézalo.

Konečné výsledky Modrého testu nejsou udávány v hrubých skóre, ale v tzv. stenech. Je tomu tak proto, že hrubá skóre jednoho testu nejsou srovnatelná s údaji jiných testů a mají celou řadu dalších nevýhod omezujících použití testu. Desetibodová stenová stupnice (sten = standard ten) v soudobé psychodiagnostice stále více vytlačuje jiné způsoby diferenciacie jedinců podle změřené intenzity určité vlastnosti (tak např. hodnotu intelektových schopností uvádějí novější pomůcky již ve stenech namísto tradičního IQ).

Hrubá skóre byla převedena na steny tím způsobem, že se od aritmetického průměru rozdělení stanovily oboustranné intervaly v rozsahu poloviční standardní odchylky normálního rozdělení. První a desátý sten pak zachycuje jedince, kteří jsou od populačního průměru vzdáleni více než činí dvě standardní odchylky. Převod hrubých skóre na steny má především tu výhodu, že steny je možno přičítat, odčítat, počítat průměry atd., tedy pracovat s nimi jako s běžnými číselnými údaji. To nelze dělat ani s hrubými skóre ani s jinými jednotkami, které ve svých intervalech plně nerespektují rozptyl a standardní odchylku měřených hodnot.

Výsledky Modrého testu tedy tiskne počítač ve stenech. Přitom sten č. 5 znamená průměrného jedince s inklinací k lepší polovině populace, sten č. 6 průměrného jedince inklinujícího k hůře adaptované části populace. Čím vyšší sten, tím intenzivnější stupeň sociální maladaptace a naopak. Devátý a zejména pak desátý sten, jak ještě uvidíme, zachycuje jedince s výrazně delikventními dispozicemi.

V. Reliabilita a validita testu

Spolehlivost testu byla ověřena opakovaným sejmutím Modrého testu u téhož souboru 412 respondentů po uplynutí osmi měsíců. Hodnoty stenů naměřených při obou těchto příležitostech jsou patrný z obr. 2. Z nich byla vypočtena reliabilita Modrého testu $r = 0,893$. Vyvinutá pomůcka tedy v tomto směru splňuje požadavky na zvlášť kvalifikované testy. Z obr. 2 je současně patrné, že v oblasti vyšších stenů je reliabilita testu ještě vyšší než je průměr udávaný korelačním koeficientem. To je pozitivní rys, neboť právě tyto vysoké steny zachycují predelikventní populaci.

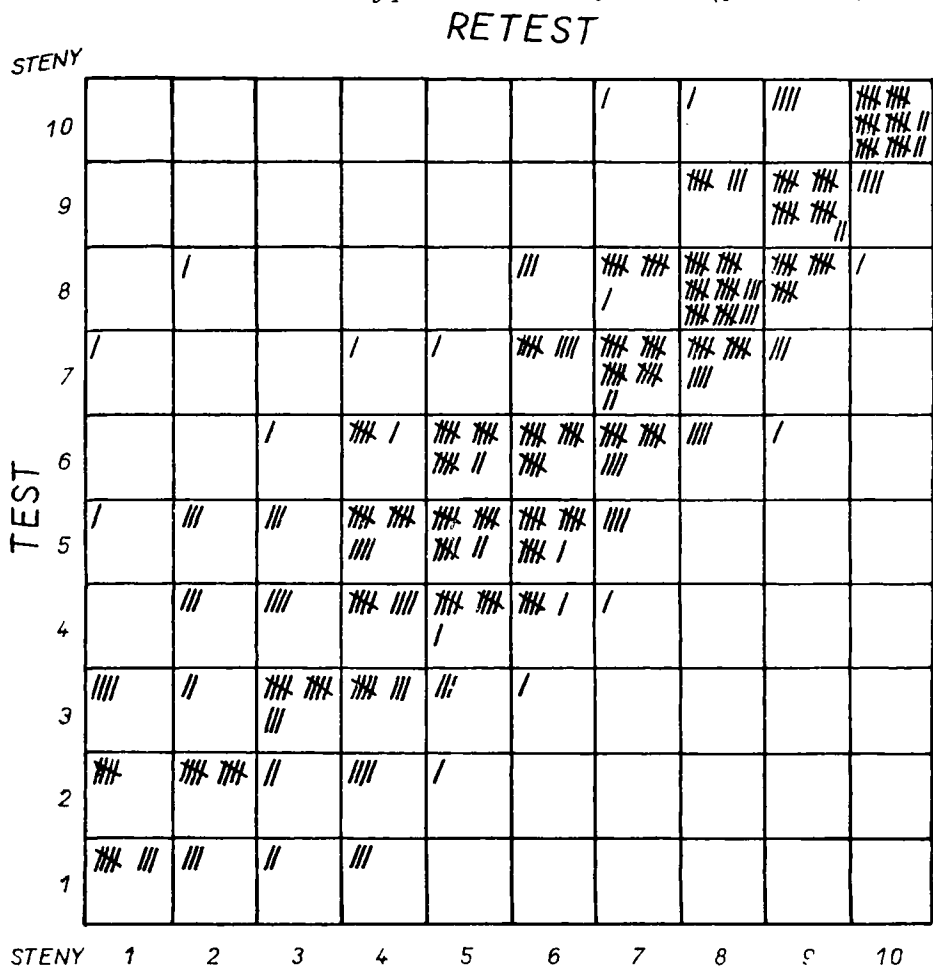
Vysoce shodné výsledky dosažené v časově dosti vzdálených momentech ukazují, že test je dostatečně stálý v čase. Současně jsou dokladem toho, že při jeho správné aplikaci není třeba mít přehnané obavy z rušivých faktorů, které bývají spojovány s dotazníkovou technikou (nepravdivé údaje, neschopnost rozpomenout se, neschopnost správně kategorizovat odpovědi apod.). Po osmi měsících může respondent odpovědět stejně jen za jediného předpokladu: že poprvé i podruhé uvedl vše podstatné podle pravdy.

V části III bylo řečeno, že zvláštním aspektem reliability testu je jeho stabilita. Za předpokladu, že testem měřené znaky jsou dostatečně stálé a nemění se v čase, má být stabilita testu rovna jeho reliabilitě. Tento důkaz se u Modrého testu podařilo podat. Z indexů spolehlivosti položek byla vypočítána [16, 156] stabilita testu $r = 0,886$. Rozdíl mezi empiricky zjištěnou retestovou reliabilitou a na základě položkové analýzy vypočtenou stabilitou Modrého testu činí pouze 0,007. Stabilita Modrého testu je tedy rovna jeho reliabilitě.

Bylo již řečeno, že o validitě testu lze hovořit vždy jen se zřetelem k určitému definovanému kritériu. Jeden a týž test může být tedy charakterizován více údaji různého typu validity. Z obecného hlediska přicházejí u testů osobnosti v úvahu především tyto tři typy validity: 1. souběžná validita (concurrent validity); 2. predikční validita (predictive validity); 3. konstrukční validita (construct validity [1, 135; 5, 106]).

Vztah mezi testovými skóre a vnějším kritériem, měřený přibližně ve stejnou dobu, je označován jako *souběžná validita*. Uvedeným způsobem byla i při tvorbě Modrého testu nejprve ověřována jeho platnost na původním souboru 500 delikventů a 500 příslušníků elity. Takto změřená souběžná validita, vyjádřená koeficientem bodově biseriální korelace, činí 0,824. Údaje o validitě, zpětně získané na souboru, na němž byl test vyvinut, nebývají však zcela objektivní. K jejich ověření se vyžaduje tak zvaná křížová validizace, totiž zkoušky na dalším výběrovém souboru, který s prvním nesouvisí. Bylo k tomu využito tzv. výběrového souboru GAMA [23, 196]. Tento soubor 551 respondentů byl vytvořen metodou záměr-

Obrázek 2: Stenové hodnoty prvního a druhého sejmání testu (po 8 měsících)



Tabulka 1: Stenové hodnoty normální populace a delikventních souborů

Soubor	Počet osob	Sten číslo										Arit-met. prům.
		1	2	3	4	5	6	7	8	9	10	
Normální populace	1632 100%	39 2,4	70 4,3	151 9,2	247 15,1	309 18,9	310 19,0	247 15,1	149 9,1	72 4,4	38 2,3	5,50
Nepodmíněně odsouzení (BETA)	500 100%	—	—	—	—	—	3 0,6	22 4,4	47 9,4	105 21,0	323 64,6	9,44
Dříve soudně trestaní (GAMA)	198 100%	1 0,5	1 0,5	—	2 1,0	3 1,5	7 3,5	22 11,1	41 20,7	41 20,7	80 40,4	8,68
Později soudně trestaní (OMEGA)	73 100%	—	1 1,3	—	—	2 2,7	5 6,8	8 10,9	15 20,5	17 23,3	25 34,2	8,50

Absolutní četnosti jsou uvedeny v horním řádku, procenta v dolním řádku. Podrobnější údaje o souborech viz časopis Čs. kriminallistika 3, 1969.

ného výběru a byl na něm aplikován též Modrý test. Z respondentů dříve soudně trestaných, jichž bylo 198, byla vytvořena jedna skupina. Druhá skupina 200 osob byla sestavena krokovou metodou ze souboru 3541 branců otestovaných tímto testem. Koeficient bodové biseřální korelace činí v tomto případě 0,659.

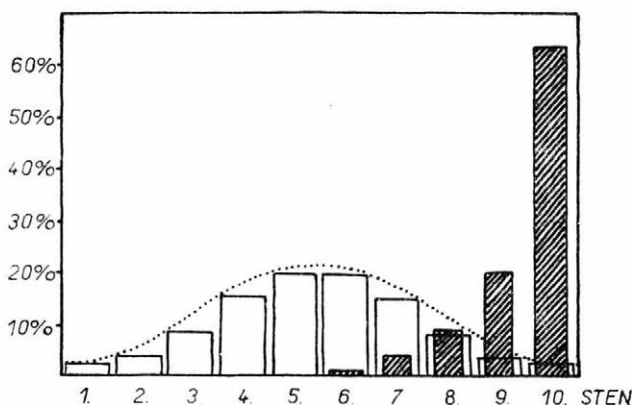
Z toho, co bylo řečeno v části III vyplývá, že je to validita dosti vysoká.

Přehled o stenových hodnotách naměřených v obou zmíněných delikventních souborech je uveden v tab. 1 a graficky je znázorněn na obr. 3.

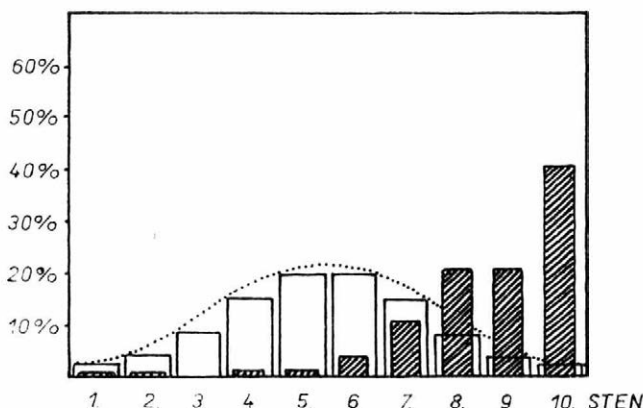
V praxi se často užívá souběžné validity i pro účely prognózy a predikce. Bývá

Obrázek 3: Stenové hodnoty normální populace a delikventních souborů (v procentech)

NEPODMÍNĚNĚ
ODSOUZENÍ
(N = 500)

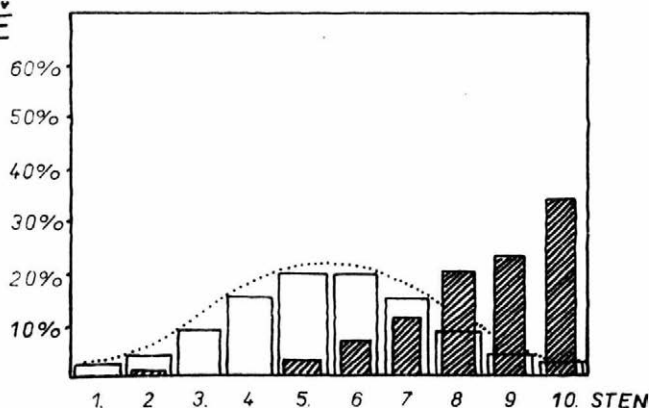


DŘÍVE SOUDNĚ
TRESTANÍ
(N = 198)



POZDĚJI SOUDNĚ
TRESTANÍ
(N = 73)

□ NORMÁLNÍ
POPULACE
▨ DELIKVENTNÍ
POPULACE



to nejčastěji tam, kde je nutno čekat celou řadu let na to, až bude znám skutečný výsledek stanovené prognózy. Jestliže jistí američtí kriminologové učinili experiment s prognózou delikventních dispozic již u dětí ve věku 6—7 let, pak museli čekat 10—15 let na to, než se jejich prognóza praxí prověří. Po tuto dobu slouží souběžná validita zpravidla jako substitut validity predikční. V našem případě jsme nemuseli čekat tak dlouho, neboť jde o test určený pro adolescentní populaci. Přesto však doba 2 let, kterou jsme měli k dispozici, je spíše minimální než optimální pro posouzení predikční hodnoty testu.

Schopnost testu předvídat určité chování je označována jako *predikční validita*. Je vyjádřena korelací mezi původními skóre testu a kritériem pozdějšího chování, měřeným často až po řadě let. Predikční validita je nejlepším základem různých metodik personálního výběru. V rámci výzkumu, o němž je řeč, jsme učinili též pokus o zjištění predikční validity Modrého testu na kritériu pozdější trestné činnosti respondentů. Setkali jsme se však se dvěma těžkostmi. První spočívala v již zmíněném krátkém časovém odstupu dvou let, vněmž se trestná činnost měla projevit. Druhá vyplývala z technických obtíží otestovat již v etapě tvorby a vývoje testu tak velký vzorek populace, abychom z něho po dvou letech dostali dostatečně rozsáhlou skupinu čerstvých delikventů. Podařilo se nám otestovat a sledovat celkem 998 osob (soubory GAMA a EPSILON; viz [23]). Z nich se však po otestování až do uzavření výzkumu dopustilo trestné činnosti pouze 73 jedinců, a to přesto, že jsme záměrným výběrem respondentů chtěli podchytit tu část populace, která podle běžných laických hledisek má k delikvenci nejbližší. Nechceme zde z těchto důvodů činit závěry o predikční validitě Modrého testu. Stenové hodnoty zmíněných 73 později soudně trestaných respondentů jsou uvedeny v tab. 1 a na obr. 3 dole. Vyplývá z nich dosti přesvědčivě, že jsou dostatečné důvody k tomu, aby souběžná validita Modrého testu zatím fungovala jako substitut jeho validity predikční.

Modrý test byl koncipován jako diagnostická pomůcka k depistáží predelikventních jedinců. Delikventní chování

však představuje až nejzazší úsek konfliktu se společností vyžadovanými normami. Vzhledem k tomu, že ve výběrovém souboru nebylo možno očekávat příliš velké množství jedinců s takto extrémně nekonformním chováním, učinili jsme pokus zachytit škálu skutečného chování respondentů a použít jí jako kritéria validity. Způsob měření chování je popsán na jiném místě [23, 197—198]. Bylo k tomu využito jednak objektivních dokladů o kázeňských odměnách a trestech sledovaných vojáků, jednak navzájem nezávislého hodnocení respondenta třemi jeho nadřízenými na devítibodové stupnici kázně. Podle výsledků faktorové analýzy byly tyto čtyři dílčí skóre chování zváženy a souhrnné skóre bylo převedeno do stenové stupnice. Korelace Modrého testu s takto měřeným pozdějším chováním vojáka činí 0,411.

Již v části III byla zmínka o tom, že u některých psychodiagnostických pomůcek nebývá uváděna jejich validita, a to zejména tehdy, je-li vnější kritérium platnosti méně spolehlivé než test sám. Byla tam také zmínka o faktorech, které snižují spolehlivost soudního potrestání jako dostatečného kritéria platnosti. Ani poměrně složitá metoda měření chování, kterou jsme uplatnili v našem výzkumu, však nedosahuje reliability Modrého testu. Chování vojáků bylo během dvouleté vojenské služby měřeno celkem třikrát, vždy po osmi měsících. Mezi prvním a druhým měřením byla zjištěna korelace $r = 0,603$; mezi prvním a třetím $r = 0,628$; mezi druhým a třetím měřením chování bylo pak vypočteno $r = 0,708$. Jak bylo již uvedeno, činí reliabilita Modrého testu 0,893. Je tedy řádově vyšší než reliabilita tohoto kritéria.

V případech, kdy je test reliabilnější než samo vnější kritérium platnosti, uplatňuje se pojem *konstrukční validity* testu. Testová skóre sama o sobě jsou v tomto případě teoretickým základem z nich odvozaného rysu. Výrazně se to např. uplatňuje u testů inteligence, při nichž se někdy položertem i polovážně uvádí, že inteligence je to, co je měřeno pomocí testů inteligence.

Ve výzkumu, o němž zde byla řeč, byla mezi jiným uplatněna i konstrukční validita Modrého testu. Jeho skóre představovala kvantifikované dispozice k neukázněnému a delikventnímu chování. Na

pozadí této známé veličiny byl pak zkoumán vliv nezávisle proměnných, působících na výběrový soubor ze sociálního prostředí a vyvolávajících změny závisle proměnné, totiž výsledného chování zkoumaných osob. Pozoruhodné výsledky, které byly dosaženy na takto koncipované validitě Modrého testu, se však již vymykají z rozsahu této stati.

VI. Použití testu

Před závěrečnými úvahami o použitelnosti Modrého testu si musíme nejprve uvědomit hranice jeho použití, které nelze překročit bez rizika vážného omylu. Test nelze bez dalšího aplikovat na jinou populaci, než na které byl vyzkoušen. Nelze jej aplikovat zejména u populace ženské. Není rovněž použitelný u dětí a mladistvých do 16–17 let, jakož i u věkových skupin nad 30 let. Zůstává tedy omezen na kategorii dospívajících a mladých dospělých mužů ve věkovém rozmezí 18–29 let. Podle statistiky za r. 1968 [21] tvoří však tato kategorie plných 43 procent všech stíhaných osob v ČSSR. Také podle výsledků řady zahraničních výzkumů je tato populační kategorie kriminálně nejproduktivnější, a to jak co do kvantity, tak co do kvalitativního složení trestné činnosti.

Aplikovat test na jinou věkovou kategorii, na ženskou populaci, příp. na jinou kulturní a jazykovou oblast by bylo možno až po provedení jeho restandardizace na nové populaci.

Již za současných podmínek lze vysokých diferenciacních schopností Modrého testu využít v kriminologickém, psychologickém a penologickém výzkumu delikventních osob. Z tab. 1 a z obr. 3 je názorně patrné, jak výrazně se liší delikventní populace od populace normální. Modrý test však diferencuje i mezi jednotlivými delikventními skupinami. Z tab. 1 vyplývá, že stenový průměr osob odsouzených soudy (k podmíněným i nepodmíněným trestům odnětí svobody i k jiným mírnějším trestům) se pohybuje okolo 8,5–8,7. Osoby odsouzené k nepodmíněným trestům odnětí svobody, tedy buď pachatelé vážnějších činů nebo osoby po názoru soudu hlouběji narušené, mají stenový průměr 9,4. Ale i v této skupině lze dále diferencovat na základě počtu dřívějších odsouzení, jak ukazuje tab. 2. Pachatelé, kteří dříve trestáni nebyli a před soudem se ocitli poprvé, mají stenový průměr 9,06. S přibývajícím počtem recidiv roste také stenový průměr. Mnohonásobní recidivisté spadají z 90 procent do desátého stenu. Skóre Modrého testu mohou tedy zde sloužit jako kvantitativní vy-

Tabulka 2: Stenové hodnoty 500 delikventů podle počtu předchozích odsouzení

Dříve trestáni	Počet osob	Sten číslo										Aritmet. prům.
		1	2	3	4	5	6	7	8	9	10	
nebyli	162 100%	—	—	—	—	—	3 2	12 7	28 17	48 30	71 44	9,06
byli jednou	131 100%	—	—	—	—	—	—	7 5	9 7	29 22	86 66	9,48
byli dvakrát	94 100%	—	—	—	—	—	—	3 3	7 7	14 15	70 75	9,61
byli třikrát	62 100%	—	—	—	—	—	—	—	2 3	9 15	51 82	9,79
byli čtyřikrát a vícekrát	51 100%	—	—	—	—	—	—	—	1 2	5 10	45 88	9,86

Absolutní četnosti jsou uvedeny v horním řádku, procenta v dolním řádku.

Číslo položky	Text položky	χ^2	Položkové indexy			Váha K
			obtížnosti	rozlišnosti	validity	
3	Jaké máš vzpomínky na dětský věk? Převážně hezké Převážně smutné Již mi vymizely z paměti	191,5	27,1	,576	,711	,365 ,901 ,708
8	Dávali tě rodiče do jeslí? Ano Ne Již se nepamatuji	22,6	11,8	,185	,059	,654 ,476 ,698
11	Měl jsi někdy zhoršenou známku z chování (z mravů)? Neměl Měl jsem někdy dvojku (ale nikdy trojku) Měl jsem někdy trojku	259,3	55,1	,563	,695	,227 ,630 ,925
15	Kolikrát jsi změnil zaměstnání (podnik)? Nezměnil jsem Změnil jsem jednou Změnil jsem dvakrát Změnil jsem třikrát Změnil jsem čtyřikrát (vícekrát)	292,0	41,9	,606	,735	,275 ,647 ,876 ,890 1,000
20	Jsi tetován? Ano Ne	278,2	36,6	,628	,776	,843 ,292
21	Jsi (byl jsi) členem mládežnické organizace? Ano Ne	35,3	22,8	,337	,370	,441 ,679
23	Tvoji vlastní rodiče: Spolu trvale žijí Spolu nežijí od doby, kdy ti bylo: méně než 3 roky 3— 6 let 7— 14 let více než 14 let	89,6	23,6	,538	,539	,410 ,757 ,898 ,750 ,660
29	Jaký vztah měl k tobě tvůj otec? Nepřátelský — neměl mne rád Lhostejný — přlíš se o mne nezajímal Vřelý — cítil jsem, že mne má rád Nežil jsem s otcem	115,3	24,3	,554	,571	,854 ,730 ,406 ,829
33	Kolikrát jsi utekl z domova? Nikdy jsem neutekl Utekl jsem jednou nebo dvakrát Utekl jsem více než dvakrát	253,5	26,8	,650	,833	,339 ,839 ,989
36	Prodělal jsi někdy zápal mozkových blan? Ano Ne	8,9	4,6	,127	,360	,711 ,486

jádrění delikventních předpokladů, ať již tyto předpoklady označíme jako sociální maladaptaci, morální narušenost či jinak.

Podobné diferenční schopnosti prokázal Modrý test i u jednotlivých typů a skupin trestných činů. O možnosti jeho použití při studiu exogenních faktorů byla již zmínka.

Vlastní smysl Modrého testu je však v oblasti hromadné prevence trestné činnosti. Test umožňuje organizované vyhledávání sociálně ohrožených osob a tím hospodárné soustředění sociální péče na ty, kteří ji nejvíce potřebují. Test ovšem řeší jen technické aspekty této depistáže: politické, organizační, právní, etické a ekonomické předpoklady takto chápané prevence zločinnosti jdou daleko za rámec této stati.

Z hlediska současných podmínek se nabízí praktické využití testu především v těchto oblastech:

1. *Depistáž sociálně maladaptovaných branců.* V rámci základního psychologického výběru před odvodem nebo po nástupu vojenské základní služby je možno při zařazování branců k různým jednotkám v armádě přihlížet též ke stupni jejich sociální adaptace. Diferencovaným zařazením a diferencovaným výchovným působením lze příznivě působit na dispoziční k trestné činnosti prakticky v celé mužské dospívající populaci.

2. *Diferenciace odsouzených v nápravně výchovných ústavech.* Pro diferencované zařazení a diferencovaný nápravně výchovný postup se nabízí perspektiva postupného uplatňování kvantitativních metod jako jednoho z faktorů komplexního hodnocení odsouzeného, jeho chování i možností další nápravy.

3. *Zjišťování potřeby kurátorské péče u podmíněně odsouzených a podmíněně propuštěných osob.* Uvažovaná kurátorská péče může mít dostatečné výsledky jenom tehdy, jestliže nebude rozptylovat omezené síly a hmotné prostředky na všechny osoby podmíněně odsouzené a podmíněně propuštěné, ale bude-li se koncentrovat na osoby, u nichž je nejvíce potřebná.

4. *Zjišťování potřeby ochranného dozoru nad recidivisty.* O uvažovaném ochranném dozoru nad kriminálními recidivisty platí totéž, co bylo řečeno o kurátorské péči. Má-li ochranný dozor respektovat reálné

potřeby prevence recidivy, pak bude tím účinnější, čím více se podaří v jeho rámci využít též nejnovějších metod kriminologické prognózy.

Literatura

- [1] A. Anastazi: *Psychological Testing*, 2nd Ed., Macmillan, New York 1961.
- [2] A. Anastazi: *Fields of Applied Psychology*, Mc Graw-Hill, New York 1964.
- [3] O. K. Buros, Ed.: *The Fifth Mental Measurement Yearbook*, Gryphon Press, Highland Park, N. J. 1959.
- [4] R. B. Cattell: *The Scientific Analysis of Personality*, Penguin Books, Harmondsworth 1965.
- [5] L. J. Cronbach: *Essentials of Psychological Testing*, 2nd Ed., Harper and Row, London 1961.
- [6] H. J. Eysenck: *Crime and Personality*, Routledge and Kegan Paul, London 1964.
- [7] O. A. Gavrilov, G. M. Sobko: *Metody kriminologického prognózování*, sborník: Voprosy naučného prognózování, AN SSSR, Moskva 1968.
- [8] S. and E. Glueck: *Unraveling Juvenile Delinquency*, Harvard University Press, 2nd Ed., Cambridge 1951.
- [9] H. Gulliksen: *Theory of Mental Tests*, Wiley and Sons, New York 1950.
- [10] S. R. Hathaway, E. D. Monachesi, Ed.: *Analyzing and Predicting Juvenile Delinquency with the MMPI*, University of Minnesota Press, Minneapolis 1953.
- [11] D. Chan-Magomedov: *Primenenije ščotno-vyčislitel'noj tehniki i matematičeskich metodov v social'no-pravovyh issledovanijach*, Sovetskaja justicija, 1, 1968.
- [12] V. N. Kudrjavcev: *Socialnoje i individualnoje prognózování v kriminologii*, sborník: Voprosy naučného prognózování, AN SSSR, Moskva 1968.
- [13] V. N. Kudrjavcev: *Pričinnost v kriminologii*, Juridičeskaja literatura, Moskva 1968.
- [14] N. F. Kuzněcova: *Někotoryje voprosy prognózování prestupnosti*, sborník: Voprosy naučného prognózování, AN SSSR, Moskva 1968.
- [15] N. Janoušková, Z. Heinzl, M. Veverka: *Základní psychologický výběr na úrovni branců*, Československá psychologie, 6, 1968.
- [16] G. A. Lienert: *Testaufbau und Testanalyse*, Julius Beltz, Weinheim 1961.
- [17] Z. Ostřihanska: *Prognóza recydywy u nieletnich przestępców oraz wyniki badań prognostycznych 180 recydywistów w wieku 15—16 lat*, Archiwum kryminologii, 3, 1965.
- [18] V. V. Pankratov: *O prognózování prestupnosti*, sborník: Statističeskije metody v kriminologii i kriminalistike, Moskva 1966.
- [19] V. V. Pankratov: *Naučno-techničeskij progress i kriminologičeskoe prognózování*, sborník: Voprosy naučného prognózování, AN SSSR, Moskva 1968.

- [20] R. A. Safarov: *Prognozirovanije i juri-dičeskaja nauka*, Sovetskoje gosudarstvo i pravo, 3, 1969.
- [21] *Statistická ročenka kriminality*, Generální prokuratura ČSSR, Praha 1969.
- [22] M. Veverka: *Kritéria S. a E. Glueckových v našem kriminologickém výzkumu*, Sociologický časopis, 1, 1969.
- [23] M. Veverka, V. Šedivý: *Výzkum trestné činnosti a hrubé nekázně v armádě (1966 až 1968)*, Československá kriminalistika, 3, 1969.
- [24] K. Vitek: *Problematika delikvence mládeže a jejího včasného rozpoznávání v pojetí S. a E. Glueckových*, Sociologický časopis, 5, 1968.
- [25] P. Zakrzewski: *Zagadnienie prognozy kryminologicznej*, Wydawnictwo prawnicze, Warszawa 1964.

Резюме

Веверка М.: Голубой тест — орудие криминологического прогнозирования

Голубой тест — это новое пособие, созданное для выявления потенциальных преступников. Он исходит из традиций, представляемых в этой области произведениями супругов Глюк. При составлении этого теста был применен в большей мере опыт современной психометрии в области развития и анализа психо-диагностических тестов. Большое количество с этим связанных вычислительных операций было выполнено при помощи саморегулируемой вычислительной машины.

Тест предназначен для мужчин в возрасте от 18 до 29 лет, значит для той части популяции, которая является самой продуктивной в области преступности. Тест представляет собой анкету, содержащую 46 вопросов. Тест выполняется групповым порядком в течение 15—20 минут. Работу с тестом может вести ознакомленный ассистент, руководствующийся точной инструкцией. Респондент записывает ответы на заменяемую вкладку. Выполненные вкладки сдаются в вычислительный пункт и целый последующий процесс автоматизирован. Оценку тестов проводит саморегулируемая вычислительная машина, печатающая списки респондентов с указанием правдоподобности их преступного поведения в будущем. Результаты теста выражены в стеновых нормах. В распоряжение предоставлены тесты на чешском, словацком и венгерском языках.

Кроме Голубого существует Зеленый тест, имеющий подобную функцию, но основанный на других началах.

Голубой тест был разработан на материале изучения пяти сот арестованных мужчин в возрасте, составляющем в среднем 20,2 лет, и на материале пяти сот дисциплинированных солдат. Он был стандартизирован на репрезентативном образце 1632 восемнадцатилетних парней. Его ретестовая надежность после 8 месяцев составляет 0,893. Сходящаяся надежность (concurrent v.), измеряемая на группе в 1000 индивидов, на которых тест был разработан, равна $r = 0,824$. Путем перекрестной валидации (cross-validation) у иной группы

198 раньше наказанных лиц и 200 нормальных лиц было определено $r = 0,659$. До сих пор незаконченные эксперименты для определения предикационной валидности (predictive v.) дали замечательные результаты на основе наблюдения за 998 опрошенными в течение 2 лет (смотри табл. 3). К 73 индивидам, совершившим преступную деятельность до истечения срока 2 лет после выполнения теста, относится такое же распределение стеновых норм, как ожидалось.

Голубой тест представляет таким образом диагностическое пособие с высокой надежностью и валидностью. Это очень экономный тест, который не преподносит больших требований ни с точки зрения времени нужного для выполнения, ни с точки зрения израсходуемых материалов, он легко применяется, он приспособлен к групповой аппликации и при помощи саморегулируемой вычислительной машины можно провести быстро и легко его оценку.

Большие дифференциационные свойства Голубого теста можно использовать в криминологическом, психологическом и пенологическом исследованиях. Собственное значение Голубого теста состоит однако в области массовой профилактики преступлений. Он подходит например для дифференциации арестованных в местах заключения. В случае условного наказания и условного освобождения он может служить средством для разыскания лиц, нуждающихся в наибольшей мере в социальной заботе. Также он может служить хорошим пособием для профилактики криминального рецидивизма. Его можно применить при отборе призывников и при их определении в военной службе.

Summary

Miroslav Veverka: The Blue Test — an Instrument of Criminological Prediction

The Blue Test is a newly produced aid in spotting potential delinquents, proceeding from the traditions represented in this area by the work of the Gluecks. In the course of its elaboration, the experience of contemporary psychometrics as regards the development and analysis of psychodiagnostical tests was extensively applied. A computer was used to perform a great number of computing operations connected with it.

The test is designed for men aged 18—29 years — i.e. that part of the population which, as far as criminality is concerned, is most productive. The test has the form of a questionnaire containing forty-six questions. Its application is performed collectively and lasts 15—20 minutes. It may be carried out by a trained assistant proceeding in accordance with exact instructions. The respondent writes his answers on an exchangeable inset. The filled-in insets are handed over to the computing station and all further procedures are automatized. The valuation of the tests is carried out by a computer printing lists of the respondents and giving the probability of their future delinquent behaviour.

The test results are presented in sten norms. The test is available in Czech, Slovak and Hungarian.

The opposite of the Blue Test is the Green Test which has a similar function but is based on different principles.

The Blue Test was developed on a set of 500 imprisoned men whose average age was 20.2 years, and on a set of well disciplined soldiers. It was standardized on a representative sample of 1632 eighteen-year-old boys. Its retest reliability after eight months is 0.893. The concurrent validity measured on a set of the 1000 individuals on whom the test was developed is $r = 0.824$.

By a cross-validation on another set of 198 former convicts and 200 normal individuals, $r = 0.659$ was ascertained. After following up 998 tested persons for two years, the so far uncompleted attempts at determining the predictive validity brought remarkable results (see Fig. 3, at the bottom). 73 individuals who committed a criminal offence in the im-

mediately following two years reveal the expected distribution of the sten norms.

The Blue Test thus represents a diagnostical aid with a high reliability and validity. It is a most economical test, since its realization requires little time and little material; it is easily applicable and adapted to a collective application and, with the use of a computer, it can be quickly and easily evaluated.

The high differentiation abilities of the Blue Test may be utilized in the criminological, psychological and penological research. The true sense of the Blue Test, however, is in the area of mass prevention of criminal activity. It is useful e.g. for the differentiation of convicts in the prisons. As concerns probation and paroles, it may be used for finding out persons most in need of social care. It may be a good aid in the prevention of criminal recidivism and used in selecting and classifying recruits in military service, etc.