

Často bývala kladena otázka (zejména v počátcích měření hypnability pomocí psychometricky konstruovaných škál), zda je *rozdíl mezi individuálním a skupinovým testováním*. Bentler a Roberts (1963) uvažovali o tom, že by se při hromadném testování mohly vyskytnout dva efekty, které by mohly způsobit nepřesnost určení hypnability, a to: 1. úzkost vyplývající z přítomnosti mnoha lidí, která by mohla snížit hypnabilitu, a 2. humorné chápání testové situace, takže smích jedné pokusné osoby by mohl narušit proces indukce. Autoři testovali 52 dobrovolníků a 39 „přinucených dobrovolníků“ Harvardskou škálou a potom Stanfordskou škálou formy C. Srovnání dosažené hypnability ve skupinovém a v individuálním sezení přineslo tyto výsledky: korelace pro vysoce hypnabilní subjekty byla 0,72, u „přinucených dobrovolníků“ byla 0,67. Výzkum prokázal, že se při hromadném testování neobjevují žádné efekty masového odporu nebo přesvědčovací efekty v daných podmínkách. Skupinové testování hypnability je tedy zcela rovnocenné s individuálním. Při stanovování norem Harvardské skupinové škály porovnávali Shor a Ornová (1963) výsledky u čtyř referenčních vzorků pokusných osob (celkem 470 osob), jímž byla škála administrována hromadně i jednotlivě a zjistili, že skupinově podávaná verze dosahuje norem kongruentních s individuálně administrovanou škálou.

Ke stejnému závěru dospěli Bentler a Hilgard (1963), kteří zkoumání skupinové a individuální administrace kombinovali ještě s *porožením objektivního skórování se sebeskórováním*. Na vzorku 45 dobrovolníků dosáhli korelace mezi skóry pozorovatelů a sebeskórováním 0,83 pro skupinové sezení a 0,89 pro individuální sezení. Průměrná hypnabilita ve skupinovém sezení byla 6,8 zatímco u individuálního sezení 6,67; rozdíl není významný. Experimentální srovnání sebeskórování a objektivního posouzení provedl O'Connell (1964). Obě posouzení spolu korelovala 0,90, což je vynikající výsledek. Nalezl některé chyby skórování, které jsou funkcí hypnability: subjekty s nízkou hypnabilitou mají tendenci podceňovat svůj výkon, subjekty s vysokou hypnabilitou jej přeceňují.

Ze studie Hoskovce, Svorada a Lance (1963) a Barbera a Calverleye (1964 b) vyplývá, že je lhostejné, je-li škála podávána *ústně nebo z magnetofonu*. Tento fakt je mnoha výzkumy i klinickými zkušenostmi pokládán za jednoznačně prokázaný. Většina experimentálních výzkumů používá nahraných škál z důvodů standardnosti a proto, aby byla zamezena možnost vzniku experimentátorových „bias“ (předpojatostí, které mohou způsobit nekontrolovanou výchylku nezávisle proměnných). U ležících pacientů trpících velkými bolestmi zjistili Johnson a Wiese (1979) vyšší hypnabilitu u živé procedury než u nahrávky, avšak korelace mezi oběma způsoby byla 0,66. Autoři vyvozují závěr, že u takovýchto pacientů nelze považovat oba typy procedur za rovnocenné.

Často se v souvislosti s používáním škál diskutuje otázka *obtížnosti jednotlivých položek*, z nichž je škála sestavena. Předpokládalo se, že jednotlivé položky mají být seřazeny podle vzrůstající obtížnosti, v opačném případě je vyšší možnost selhání. Tuto hypotézu ověřovali Field, Evans a Orne (1965), kteří administrovali Harvardskou skupinovou škálu dvěma skupinám pokusných osob. Jedné skupině byla škála předložena obvyklým způsobem, druhé skupině byly testovací položky sugerovány v opačném pořadí. Rozdíly mezi oběma skupinami byly sugerovány v opačném pořadí. Rozdíly mezi oběma sku-