تئوری آزمون مدرن حول محور مدل سازی داده های طبقه بندی شده با استفاده از مدل ویژگی نهفته غیر خطی تعبیر می شود، این تئوری در مراجع آزمون های آزمایشی «تئوری پاسخ آیتم[[1]](#footnote-1)» (IRT) نامیده می شود. به نظر می رسد آنچه که اکنون به استاندارد مدل سازی IRT تبدیل شده، بر مدل های تک بعدی تخمین زده شده با استفاده از برآورد بیشترین احتمال اطلاعات-کامل، تمرکز کرده است (برای مثال، لرد 1980). رود مک دونالد، در طول فعالیت های خود، جنبه های مختلف این موضوع را بررسی کرد (برای مثال، مک دونالد، 1999). او از یک طرف بر مدل های چند بعدی تاکید داشت (برای مثال، مک دونالد، 2000)؛ و از طرفی دیگر، از تخمین اطلاعات-محدود و روش های آزمایش مطابق با تئوری استفاده کرده است (برای مثال، مک دونالد، 1985؛ مک دونالد و موک 1995). شاید بیشترین سهم مک دونالد در این زمینه برنامه NOHARM باشد (فریزر و مک دونالد، 1988) که قادر به تطبیق دادند مدل طاقدیسی[[2]](#footnote-2) نرمال چند بعدی به تعداد آیتم های تقریبا نامحدود همراه با تعداد تقریبا نامحدود از ابعاد ویژگی نهفته با استفاده از روش اطلاعات محدود می باشد.

در فصل اول این بخش، **تری آکرمن** نمونه های گرافیکی از اینکه چگونه IRT چند بعدی را می توان در نمونه های اولیه برای درک بهتر داده های پاسخ تست مورد استفاده قرار داد، تهیه کرد. این فصل با مرور مدل های IRT از دو دیدگاه فاکتور تحلیلی مک دونالد( 1999) و رویکرد سنتی IRT لرد(1980)آغاز شده است. این فصل با بحث مختصری درباره روش های نوین مدل سازی IRT چند بعدی در آینده به پایان می رسد. در فصل دوم، **دانیل بولت** اصل اطلاعات-محدود و همچنین رویکردهای اطلاعات-کامل را در برآورد IRT بررسی کرد و با استفاده از زنجیره ی مارکوف مونت کارلو، نمایش رویکردهای تخمین اطلاعات کامل و اطلاعات محدود را نتیجه گیری نمود.

**آلبرت میدهو الیوارز** رابطه ی بین مدل فاکتور مشترک و مدل های IRT را مورد بررسی قرار داد. مک دونالد (1999) نشان داد که در حالت داده های دو دویی، هنگامی که تخمین از اطلاعات دو متغیره استفاده می کند، مدل فاکتور و مدل IRT خطی معادل یک دیگر هستند. در این فصل، میدهو الیواز تخمین اطلاعات کامل حاصل از مدل IRT خطی را در نظر گرفت. وی نشان داد که نیازی به هیچ توزیعی از پراکندگی ویژگی های پنهان برای ارزیابی این مدل وجود ندارد، و در بسیاری از ممان های خصوصیات پنهان را می توان تخمین زد. همچنین او خاطر نشان کرد که آزمون پیشنهاد شده توسط بروون (1984) می تواند برای مقایسه نیکویی برازش مدل های فاکتور و مدل های IRT غیر خطی (مانند مدل های لجستیک) در هر مجموعه داده ارائه شده مورد استفاده قرار گیرد.

در پایان، **جان هورن** گامی به عقب برداشت و درباره ی موضوعات اندازه گیری اولیه (اساسی و پایه) بحث کرد، اگرچه آنها در نوشته های اخیر مربوط به اندازه گیری بیان شده اند، اما در تحقیق روزانه دانشمندان رفتارگرا مورد بررسی قرار نگرفته اند. بنابراین، این فصل درباره ی این موضوع بحث می کند که چگونه عملیات های اندازه گیری شکل پراکندگی ها و دیگر ویژگی های مهم اندازه گیری را تعیین می کند، چگونه افتراق پذیری ( تبعیض) به شکل توزیع و پراکندگی، مربوط می شود و چگونه این موضوعات به استفاده از آمارهای موجود در تحقیق علمی رفتاری مربوط می گردند.

1. item response theory [↑](#footnote-ref-1)
2. ogive [↑](#footnote-ref-2)