

2024 年度 AI カウンセリングの効果検証結果報告書

作成: 2025 年 3 月 17 日
ユースキャリア研究所 高橋浩

【序】

2024 年度に企業で試験導入を行った際の AI カウンセリングおよび対面カウンセリングの効果検証について報告する。

分析対象: 参加企業 3 社のうち 2 社からの有志。第 1 回(9~14 名)、第 2 回(11~19 名)

なお、残り 1 については、ベースライン、第 1 回後、第 2 回後のすべてが揃った回答なかったため分析対象外となった。

実施期間: 2024 年 7 月~12 月

実施内容: 実験群 A は AI カウンセリングを 2 回、実験群 B は AI と対面のカウンセリングを各 1 回実施。統制群はカウンセリング未実施

調査実施: カウンセリング前のベースライン、第 1 回後、第 2 回後に調査を実施

なお未回答者が多数であったため、当初予定の 3 群 × 3 時点の効果検証を断念し、3 群におけるベースラインから第 1 回、およびベースラインから第 2 回の変化を検証した。

【分析結果】

実験群 A のベースラインから第 2 回にかけて(すなわち 2 回の AI カウンセリングの実施で)、精神的不健康度の改善および自己効力感の向上が認められた。また、第 1 回(AI)の実験群 A において自己効力感、ジョブ・クラフティングの下位項目¹が、実験群 B(AI)においては精神的不健康度、ウェルビーイング、ジョブ・クラフティングの下位項目が有意に改善されていた。これらも AI カウンセリングの効果と考えられる。

なお、第 2 回の実験群 B(対面カウンセリング)ではジョブ・クラフティングの下位項目が向上されていた。

【考察】

ベースラインから第 2 回にかけて、実験群 A(AI)の精神不健康度と自己効力感の向上が認められたことにより、AI カウンセリングに効果があることが示唆された。その効果は、主に第 1 回での有用な情報提供であり、これが第 2 回に遅延効果として発現した可能性が考えられた。なお、効果指標の下位項目での変化に注目すれば、第 1 回、第 2 回の実験群 A, B それぞれにおいてカウンセリング効果が認められた。特に、実験群 B の第 2 回の対面カウンセリングでは、自己理解の促進があり、AI カウンセリングとの違いが示された。

なお、第 1 回は実験群 A, B とも AI カウンセリングであるにもかかわらず、実験群 B に効果が認められなかった。この点については、今回のデータだけでは説明が困難であった。実験群 A, B の間に何らかの質的差異や外的環境の違いがあると予想される。

詳細な分析結果と考察について次頁以降を参照されたい。

¹ 下位項目: 効果指標を構成する小項目のこと

分析結果の詳細

対象者

ベースラインから第2回までの全3回の調査に未回答者が多数存在するため、今回は「ベースラインと第1回の比較」、「ベースラインと第2回の比較」をすることとした。そのため、前者と後者で対象者数が異なっている。表1、表2に対象者数とその構成を示す。

表1 第1回の対象者数

性別	統制群		実験群A (AI-AI)		実験群B (AI-人)		
	n=9		n=11		n=14		
	n	%	n	%	n	%	
女	5	55.6	6	54.5	8	57.1	
男	4	44.4	5	45.5	6	42.9	
20代	4	44.4	7	63.6	9	64.3	
30代	5	55.6	3	27.3	4	28.6	
40代	0	0.0	0	0.0	0	0.0	
50代	0	0.0	1	9.1	1	7.1	
婚姻	未婚	3	33.3	5	45.5	8	57.1
	既婚	6	66.7	6	54.5	6	42.9
独居	独居	1	11.1	3	27.3	5	35.7
	非独居	8	88.9	8	72.7	9	64.3
企業	C社	9	100.0	8	72.7	8	57.1
	I社	0	0.0	3	27.3	6	42.9

表2 第2回の対象者

性別	統制群		実験群A (AI-AI)		実験群B (AI-人)		
	n=19		n=11		n=13		
	n	%	n	%	n	%	
女	6	31.6	8	72.7	7	53.8	
男	13	68.4	3	27.3	6	46.2	
20代	6	31.6	7	63.6	9	69.2	
30代	6	31.6	4	36.4	3	23.1	
40代	4	21.1	0	0.0	0	0.0	
50代	3	15.8	0	0.0	1	7.7	
婚姻	未婚	10	52.6	6	54.5	8	61.5
	既婚	9	47.4	5	45.5	5	38.5
独居	独居	6	31.6	2	18.2	6	46.2
	非独居	13	68.4	9	81.8	7	53.8
企業	C社	7	36.8	10	90.9	8	61.5
	I社	12	63.2	1	9.1	5	38.5

効果指標

カウンセリングの効果指標として、精神的不健康度(6項目)、ウェルビーイング(10項目)、自己効力感(9項目)、ジョブ・クラフティング(8項目)、従業員エンゲージメント(6項目)を用い、対象者には7件法(1点～7点)での回答を求めた。なお、精神的不健康度のみ、得点が高いほど不健康であり、それ以外の効果指標は得点が高いほど良好であることを示す。

精神的不健康度およびウェルビーイングは心理的な状態を示すものであり、カウンセリングによる直接的な効果を示すことが想定される。

ジョブ・クラフティングとは「仕事がより快適になる本人が行う調整力」、従業員エンゲージメントとは「本人と組織の一体感やつながり感」を示す指標である。これらの効果指標は、カウンセリングの間接的な効果を示すことが想定されている。

なお、質問項目数を最小限に抑えるため、効果指標は本来の測定尺度から項目を抜粋して使用している。

測定方法

3社のご協力を得て、対象者への説明を実施した後、ベースライン調査を実施。同時に、カウンセリングへの協力可否を確認し、協力不可としたものを統制群に、協力可としたものを実験群とした。実験群はさらにA群とB群に分割した。第1回と第2回ともにAIカウンセリングを実施する実

験群 A と、第1回は AI で第2回を対面でおこなう実験群 B へとランダムに振り分けた。

第1回の AI カウンセリングはチャットによる双方向のコミュニケーションである。操作に慣れてもうるために、AIとの日時の会話を数日行ってもらった。次に、対象者の相談を入力してもらい AI とのやり取りを実施した。なお、AI のやり取りではシステムの都合上 12~24 時間後に AI から回答される場合がある。

第2回のカウンセリングは、実験群 A においては第1回の最終対話から約一か月後に AI カウンセリングを実施した。実験群 B においては、同じく約一か月後に対面カウンセリングを実施した。

評価測定は、第1回と第2回の終了直後(目安として1週間以内)に実施した。

ベースラインからの変化(効果指標の総合得点)

5つの効果指標について、ベースラインから第1回および第2回への変化を確認した(図1~5、表3)。その結果、実験群 A においては第2回で精神的不健康度の有意な改善と自己効力感の有意傾向の改善が認められた。実験群 B においては第2回で精神的不健康度の有意傾向の悪化が認められた。(図1~図5の注意事項: 第1回と第2回では分析対象人数が異なるため、ベースラインの値は異なっている。図1~図5の第2回の値は、第1回のベースラインと同値になるように値を調整した。表3の値とは異なるので注意されたい。)

実験群 B の精神的不健康度の悪化については、統制群も第1回から第2回で悪化しているため、カウンセリング以外の環境的要因による影響と考えられる。次に、実験群 A の精神的不健康度は、統制群が悪化する中での改善が認められていることから、この改善がカウンセリングによる効果と考えられる。また、実験群 A の自己効力感の向上についても、統制群が低下傾向である中の向上なので、カウンセリングによる効果と考えられる。

図1 精神的不健康度の変化

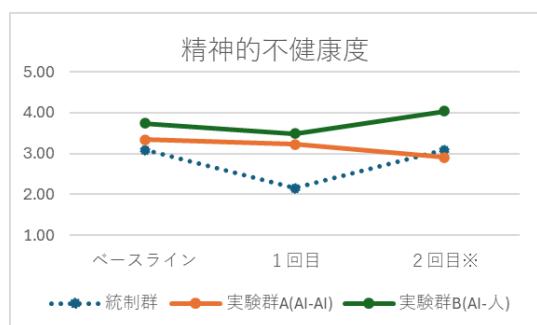


図2 ウェルビーイングの変化

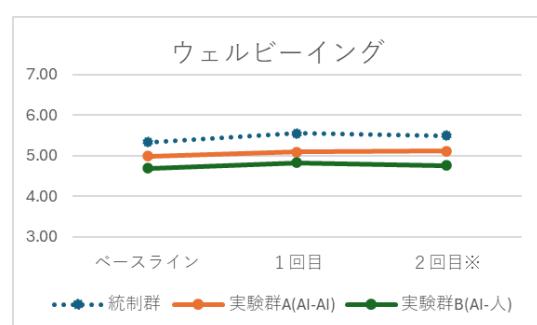


図3 自己効力感の変化

図4 ジョブ・クラフティングの変化

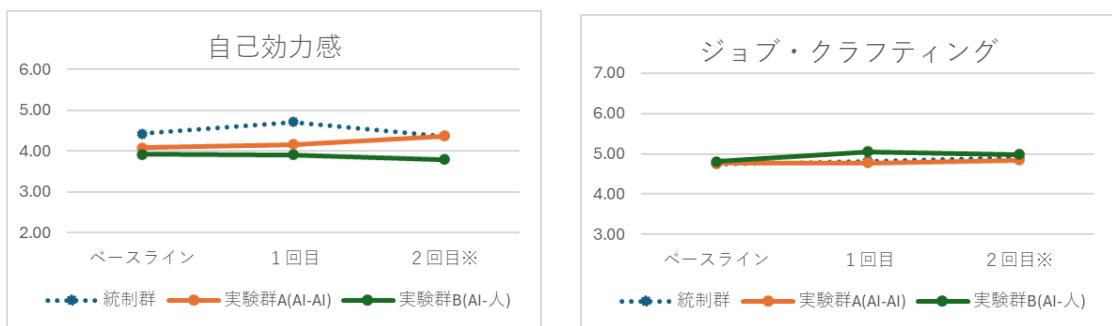
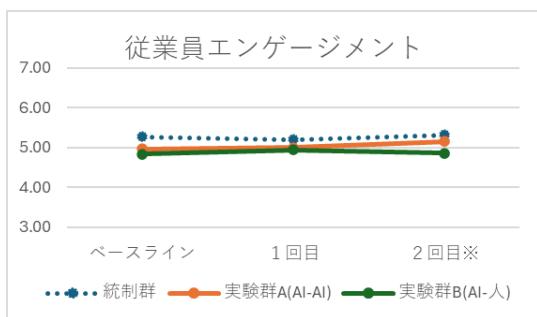


図5 従業員エンゲージメントの変化



ベースラインからの変化(効果指標の下位項目得点)

効果指標の下位項目の得点に注目した場合のベースラインから第1回および第2回への変化を確認した。その結果、有意な変化が確認された(表 4)。総合得点では見られなかった第1回の変化を確認することができた。

まず第1回での変化については、実験群 A では自己効力感の「嫌な体験の回想(反転項目)」「仕事の失敗を感じる(反転項目)」、ジョブ・クラフティングの「フィードバックの希求」が有意に緩和されていた。AI カウンセリングによって、過去の嫌な体験に悩まされることや、仕事での失敗を感じることが軽減され、さらにより良く仕事をするために自らフィードバックをもらうような行動変容が生じたといえる。

また、実験群 B では、精神的不健康度の「睡眠時間の減少」「日常的緊張」、ウェルビーイングの「自己目標の達成」「周囲の援助の獲得」、ジョブ・クラフティングの「思考力消耗の抑制」が有意に改善されていた。AI カウンセリングによって、睡眠時間の確保や日常的な安心感が得られ、さらにより良く仕事をするために過剰に思考力が消耗されないようにする行動が変容したといえる。

次に、第2回での変化については、実験群 A では精神的不健康度の「日常的緊張」とウェルビーイングの「人生の価値の実感」、自己効力感の「仕事の失敗を感じる」が有意に改善されており、これらが総合得点の向上にも寄与していると考えられる。継続的な AI カウンセリングによって、日常的緊張の緩和や失敗感の軽減、さらには人生の価値の実感につながっているといえる。

また、実験群 B は、ジョブ・クラフティングの「上司指導の希求」で有意な向上が認められた。対面カウンセリングによって積極的に上司の支援をえる行動変容が生じたといえる。実験群 B の第1

回の AI カウンセリングで得られた効果は、第2回の対面カウンセリングでは確認ができなかった。AI と対面で取り上げられる問題に違いがあるのかもしれない。

表3 効果指標(総合得点)のベースラインからの変化

評価指標	群	n	ベース ライン		t 値	ベース ライン		t 値
			平均	平均		平均	平均	
			(SD)	(SD)		(SD)	(SD)	
精神的 不健康度	統制群	9	3.09	2.15	2.073 *	2.62	2.61	.042
			(1.167)	(0.699)		(1.202)	(1.097)	
	実験群A (AI-AI)	11	3.33	3.23	.647	3.23	2.80	1.813 *
			(1.169)	(1.550)		(0.987)	(1.390)	
ウェル ビーイング	実験群B (AI-人)	14	3.74	3.49	1.302	3.60	3.90	-1.364 †
			(1.600)	(1.641)		(1.496)	(1.268)	
	統制群	9	5.33	5.56	-1.315	5.28	5.44	-1.233
			(0.477)	(0.667)		(0.691)	(0.873)	
自己効力感	実験群A (AI-AI)	11	4.99	5.10	-.546	5.15	5.28	-.797
			(0.952)	(1.209)		(0.938)	(1.068)	
	実験群B (AI-人)	14	4.69	4.84	-1.055	4.75	4.82	.325
			(1.528)	(1.426)		(1.106)	(0.969)	
ジョブ・ クラフ ティング	統制群	9	4.42	4.70	-1.360	4.69	4.64	.372
			(0.855)	(0.935)		(1.009)	(0.953)	
	実験群A (AI-AI)	11	4.07	4.15	-.505	4.12	4.41	-1.452 †
			(1.064)	(1.076)		(1.095)	(0.962)	
従業員 エンゲージ メント	実験群B (AI-人)	14	3.91	3.90	.079	3.79	3.67	.105
			(0.977)	(1.166)		(0.958)	(1.050)	
	統制群	9	4.75	4.82	-.375	4.66	4.82	-1.042
			(0.576)	(0.673)		(0.539)	(0.709)	
実験群A (AI-AI)	11	4.76	4.77	-.115	4.84	4.92	-.502	
			(0.621)	(0.884)		(0.791)	(1.022)	
	実験群B (AI-人)	14	4.81	5.05	-1.140	4.69	4.87	.149
			(0.761)	(0.808)		(0.801)	(0.687)	
実験群A (AI-AI)	9	5.28	5.20	.471	5.13	5.18	-.427	
			(0.702)	(0.725)		(0.814)	(0.781)	
	実験群B (AI-人)	11	4.97	5.00	-.173	5.24	5.42	-1.174
			(0.927)	(1.169)		(0.761)	(1.009)	
実験群B (AI-人)	14	4.83	4.94	-.751	4.78	4.81	.427	
			(1.068)	(0.802)		(1.033)	(1.116)	

*:5%有意水準で有意、†:10%有意水準で有意傾向

表4 効果指標(下位項目得点)のベースラインからの変化

群	n	評価指標	ベース		t 値	n	ベース		t 値
			ライン 平均 (SD)	第1回 平均 (SD)			ライン2 平均 (SD)	第2回 平均 (SD)	
精神的不健康度									
実験群A (AI-AI)	11	日常的緊張					4.09 (1.375)	3.45 (1.572)	1.399 *
実験群B (AI-人)	14	睡眠時間減少	3.50 (1.990)	2.86 (1.791)	2.590 *				
		日常的緊張	4.07 (1.492)	3.14 (1.657)	2.738 *				
ウェルビーイング									
実験群A (AI-AI)	11	人生の価値の実感					4.91 (1.578)	5.55 (10.368)	-3.130 *
実験群B (AI-人)	14	自己目標の達成	4.00 (1.710)	4.50 (1.345)	-2.188 *				
		周囲の援助の獲得	4.93 (1.685)	5.5 (1.506)	-1.847 *				
自己効力感									
実験群A (AI-AI)	11	嫌な体験の回想R	3.27 (1.794)	4.18 (1.940)	-2.193 *				
		仕事の失敗を感じるR					4.36 (1.206)	5.18 (1.079)	-2.324 *
ジョブ・クラフティング									
実験群A (AI-AI)	11	フィードバックの希求	4.64 (1.120)	5.00 (1.265)	-2.390 *				
実験群B (AI-人)	14	思考力消耗の抑制	2.93 (1.072)	4.07 (1.592)	-2.332 *				
		上司指導の希求					4.62 (1.502)	5.31 (0.947)	-1.812 *

*:5%有意水準で有意

カウンセリング指標と効果指標との関連

効果測定において、カウンセリングに対する対象者の捉え方やかかわり方(以下、カウンセリング指標とする)について8項目のデータを得ている。具体的には、面談への期待度(面談期待度)、面談内容に対する納得度(面談納得度)、面談自体やカウンセラーに対する信頼感(面談信頼性)、面談への積極的な関わり(面談積極性)、面談や役立ち感(面談有用性)、面談による自己理解の程度(自己理解度)、面談による今後の行動の明確化(行動明確化)、面談自体の受けやすさ(面談容易性)の8つである。いずれも、先行研究を参考に、カウンセリング効果に影響する項目となっている。

これらのカウンセリングと効果指標との相関分析を実施した。つまり、カウンセリングにおける対象者の捉え方やかかわり方がどのような効果指標に影響を及ぼすかを確認した。

まず実験群Aは、第1回のAIカウンセリングにおいて、面談有用性が自己効力感、ジョブ・クラフティング、従業員エンゲージメントと有意な相関が認められた(表5)。このことから、実験群Aの場合、第1回のAIカウンセリングで問題解決に役立つ情報を獲得した人ほど各種の効果が得ら

れたといえる。このことは、第1回の下位項目別の効果に寄与していると考えられる。

第2回(AI)では、面談期待度が精神的不健康度および自己効力感と有意な関連が認められた(表6)。このことは、第1回で面談期待度が高まり、そのことが第2回に影響していると考えられる。さらに、第1回のカウンセリング指標と第2回の効果指標との間では、面談有用性が精神的不健康度と有意な相関が認められた(表7)。表6と表7を合わせると、第1回で得られた情報が「遅延効果」として第2回の総合得点に現れたと考えられる。

なお、第2回で面談有用性と効果指標に関連が見られないのは、第1回で有用な情報が十分に入手できたため、第2回ではこれ以上必要な情報が無くなつたためと推察する。

表5 実験群A 第1回(AI)のカウンセリング指標と効果指標の相関分析(spearman)

カウンセリング指標	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント	次回面談の期待度
面談期待度	.368	-.100	.027	-.213	.416	.292
面談納得感	.164	-.208	-.175	-.133	.056	.009
面談信頼性	-.278	.182	.338	.258	.346	.566*
面談積極性	.112	-.109	.026	.007	.097	.044
面談有用性	-.452	.198	.575*	.548*	.530*	.767**
自己理解度	-.102	-.231	.308	.222	.197	.639**
行動明確化	-.102	-.377	.078	.375	.187	.585*
面談容易性	-.275	-.018	.291	.116	.327	.365

*:5%有意水準で有意、**:1%有意水準で有意

表6 実験群A 第2回(AI)のカウンセリング指標と効果指標の相関分析(spearman)

カウンセリング指標	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント	次回面談の期待度
面談期待度	-.607*	.593	.627*	.516	.428	.527
面談納得感	-.074	.117	.236	.173	.215	.251
面談信頼性	-.359	.330	.293	.313	.356	.325
面談積極性	.121	-.072	-.270	-.055	.133	.370
面談有用性	-.161	.014	.199	.294	.354	.833**
自己理解度	-.273	-.116	-.219	.245	-.041	.444
行動明確化	-.027	-.161	.001	.202	.175	.713**
面談容易性	-.487	.218	.126	.435	.226	.307

*:5%有意水準で有意、**:1%有意水準で有意

表7 実験群A 第1回(AI)のカウンセリング指標と第2回(AI)の効果指標の相関分析(spearman)

カウンセリング指標	第1回					第2回				
	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント
面談期待度	-.055	-.065	-.595	-.010	-.040					
面談納得感	.032	.144	-.032	.022	.257					
面談信頼性	-.266	.363	.143	.176	.374					
面談積極性	.161	-.038	-.137	-.099	.179					
面談有用性	-.622*	.568	.587	.512	.447					
自己理解度	-.219	.101	.416	.421	.334					
行動明確化	-.248	.083	.470	.486	.536					
面談容易性	-.029	.048	.169	.314	.489					

*:5%有意水準で有意

次に実験群Bは、第1回のAIカウンセリングでの面談納得感が従業員エンゲージメントとの有意な負の相関が認められた(表8)。負相関であった原因については、面談内容を考慮しなければ検討が困難であるため、ここでは不明である。

表8 実験群B 第1回(AI)のカウンセリング指標と効果指標の相関分析(spearman)

カウンセリング指標	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント	次回面談の期待度
面談期待度	.000	.063	.080	-.087	-.033	.490
面談納得感	-.047	-.266	-.006	-.169	-.444*	.355
面談信頼性	.111	-.002	-.074	.034	-.328	.673**
面談積極性	.105	-.359	-.148	.023	-.127	.002
面談有用性	.172	.039	-.166	-.136	-.190	.828**
自己理解度	.142	.050	-.101	-.051	-.027	.697**
行動明確化	.432	-.055	-.290	-.074	-.181	.646**
面談容易性	.026	-.030	-.067	-.151	-.335	.785**

*:5%有意水準で有意、**:1%有意水準で有意

表9 実験群B 第2回(対面)のカウンセリング指標と効果指標の相関分析(spearman)

カウンセリング指標	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント	次回面談の期待度
面談期待度	-.046	.155	-.019	.016	.331	.528*
面談納得感	.345	.053	-.167	.139	.081	.432
面談信頼性	.135	.299	-.046	.335	.417	.659**
面談積極性	-.135	.193	.094	.273	.258	.331
面談有用性	.043	.385	.056	.278	.442	.812**
自己理解度	-.132	.699**	.237	.576**	.476*	.575*
行動明確化	.115	.433	-.006	.191	.558*	.842**
面談容易性	.221	.150	-.033	.281	.406	.478*

*:5%有意水準で有意、**:1%有意水準で有意

表10 実験群B 第1回(AI)のカウンセリング指標と第2回(対面)の

効果指標の相関分析(spearman)

カウンセリング指標	第1回		第2回		
	精神的不健康度	ウェルビーイング	自己効力感	ジョブクラフティング	従業員エンゲージメント
面談期待度	.052	-.016	.049	.242	.114
面談納得感	-.074	-.113	-.005	-.073	-.407
面談信頼性	.090	.116	-.172	.276	.010
面談積極性	-.140	.019	.141	-.181	-.041
面談有用性	.038	.186	-.135	.081	.168
自己理解度	.055	.091	-.093	.140	.218
行動明確化	.332	-.141	-.387	-.090	.189
面談容易性	.100	-.050	-.155	-.069	.180

第2回の対面カウンセリングでの自己理解度がウェルビーイング、ジョブ・クラフティング、従業員エンゲージメントとの有意な相関が認められた(表9)。対面カウンセリングによって自己理解を深化させた効果だと考えられる。ここにAIカウンセリングとの違いが見られる。

また、第1回(AI)のカウンセリング指標と第2回(対面)の効果指標との間に有意な相関は認められなかった(表 10)。つまり、実験群 A で見られたような第1回から第2回への遅延効果は見られなかった。AIカウンセリングを2回連続で利用するからこそ遅延効果が現れる可能性がある。

ところで、実験群 A と B の第1回は、どちらも AI カウンセリングであるが、実験群 A のみに効果が現れた点は疑問である。その理由の1つとしては、両群が同質ではない可能性がある。少なくとも、実験群 B は実験群 A よりも未婚で独居者が多く、また、企業構成比も異なる。残念ながら、現状のデータからは、実験群 A と B の AI カウンセリングの違いを説明することができなかった。

また、実験群 B は、総合得点ではベースラインからの有意な変化は認められなかった。つまり、対面カウンセリングの効果が十分示されなかつことになる。しかし、一方で第2回(対面)のカウンセリング指標と効果指標との間には有意な相関が認められた(表9)。このことは、対面カウンセリングをしっかりと実施していけば、いずれはその効果が現れる可能性を示している。対面カウンセリングを複数回実施することによって、ベースラインからの向上が現れるものと予想される。

結論

本効果検証では、AIC にて実施する AI カウンセリングあるいは対面カウンセリングの効果を検証することが目的であった。当初、医療等でも用いられる RCT(ランダム化比較試験)を用いる予定であったが、回答データが不足したため、ベースラインから第1回カウンセリング、およびベースラインと第2回カウンセリングの変化について検証を行った。その結果、以下が明らかになった。

- ① AI カウンセリングでは、主に問題解決に有用な知識やハウツーの活用が効果に寄与している(表5)。これに対して対面カウンセリングは、利用者の自己理解の促進が効果に寄与している(表9)。
- ② 連続してAIカウンセリングを利用した場合、第2回において精神的不健康度の改善および自

自己効力感の向上が認められた(図1、図3、表3)。

- ③ AIカウンセリングによる有用情報の獲得は、第1回でなされ、第2回への期待感で維持され、第2回で効果が発現するという遅延効果が考えらえる(表6、表7)。
- ④ 効果指標の下位項目で捉えた場合、第1回のAIカウンセリングにおいてもウェルビーイングや自己効力感、ジョブ・クラフティングへの部分的な効果が確認できた。
- ⑤ AIカウンセリングから対面カウンセリングへの切り替えによって効果を抑制される可能性がある。AIカウンセリングでの有用情報の獲得から対面による自己理解へと、利用者の意識の方向が外から内へと大きく転換が迫られる。これにより、AIカウンセリングでの効果を保持されず、対面による効果が発現する前に終結している可能性がある(表4、上記①、③)。

今後の課題

より精度の高い効果検証の為の課題を列記する。

- ① 統制群を含め、群分けはランダムに行う。今回、統制群と実験群が協力可否で分けたために、統制群と実験群に質的な違いが生じ、一部、同一条件による比較ができなかつた。
- ② 全3回の調査に回答することの徹底。1回でも回答しないと分析対象から外れることになる。今回かなり多くのデータが分析対象から除外される結果となつた。
- ③ 調査項目に群種、業務負荷状況(ストレス)の追加。群種は、今回は実験群A、Bが混在されていたのを識別するため。業務負荷状況は効果指標に大きく影響をする環境要因あり、分析時に考慮していきたい。
- ④ 実験群A、Bの実施内容の変更。実験群AをAIカウンセリング2回、実験群Bを対面カウンセリング2回とした方が、遅延効果も制御してAIと対面の違いを明確にしやすくなる。

上記課題をクリアし、より精度の高い効果検証を行っていく。

以上